Epigenetic PU.1 silencing in myeloid leukemia by mimicrying a T cell specific chromatin loop

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät I

der Humboldt-Universität zu Berlin

von

Diplom-Biologin Chiara Perrod

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz
Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Stefan Hecht

Gutachter: 1. Dr. Prof. Achim Leutz
2. Dr. Prof. Harald Saumweber
3. Dr. Prof. Matthias Selbach

eingereicht: 06.03.2013

Datum der Promotion: 02.07.2013

Table of Contents

I Abstract

Alterations in the local chromatin structure orchestrate the dynamic regulation of differentiation promoting genes. PU.1 is a master transcription factor in hematopoiesis. PU.1 gene must be tightly regulated to achieve lineage specific expression pattern. High levels of PU.1 are required for myeloid commitment: it is expressed at intermediate level in B-cells and must be actively silenced to permit T cell development from early multipotent progenitors. However, little is known of how PU.1 is regulated in T-cells. Moreover, aberrant PU.1 expressions have been observed in multiple leukemias.

Using a genome-wide chromatin interaction screen we identified a cis-repressor with insulating capacity that undergoes long-distant chromatin looping to block PU.1 promoter activity in T cells but not myeloid or B cells. Looping and repression requires binding of the chromatin regulator protein CTCF. In contrast to normal myeloid cells, we found that cancer cells from myeloid leukemia patients adopt the T cell specific repressive chromatin structure bringing the insulator into spatial contact with the PU.1 promoter. These results identify CTCF controlled long-distant insulator looping as a novel mechanism to silence lineage-opposing transcription factor expression, and reveal that cancer cells can mimic the chromatin confirmation of another lineage to block expression of differentiation driving genes.

Keywords: PU.1, hematopoiesis, chromatin conformation, gene regulation, leukemia

II Zusammenfassung

Veränderungen in der lokalen Chromatinstruktur beeinflussen die dynamische Regulation von Genen, welche für die Differenzierung notwendig sind. PU.1 ist ein Master-Transkriptionsfaktor in der Hämatopoese und wird streng reguliert, um ein zelllinienspezifisches Expressionsmuster zu erzielen. Hohe Konzentrationen von PU.1 sind für myeloische Differenzierung erforderlich. In B-Zellen wird PU.1 mittelstark exprimiert und muss aktiv runterreguliert werden, um eine Ausdifferenzierung der multipotenten Vorläuferzellen zu T-Zellen zu ermöglichen. Derzeit ist wenig über die Regulierung von PU.1 in T-Zellen bekannt. Darüber hinaus wurde eine abnormale Expression von PU.1 in verschiedenen Leukämieerkrankungen beobachtet.

Mittels eines genome-wide Chromatin-Interaktions-Screens konnten wir einen cis-Repressor mit insulierender Kapazität identifizieren, welcher mittels eines Chromatinloops die Promotoraktivität von PU.1 in T-Zellen, jedoch nicht in myeloischen oder B-Zellen blockiert. Sowie Looping als auch Insulation erfordern die Bindung des Chromatin-Regulatorprotein CTCF. Im Gegensatz zu normalen myeloischen Zellen finden wir, dass Krebszellen aus myeloischen Leukämie Patienten diese T-Zell-spezifische repressive Chromatinstruktur aufweisen, was einen räumlichen Kontakt des Insulator mit dem PU.1 Promotor ermöglicht.

Die Ergebnisse dieser Arbeit beschrieben das CTCF gesteuerte „long distance looping“ als ein neuer molekularer epigenetischer Mechanismus, um Transkriptionsfaktor PU.1 in T-Zellen runterzuregulieren, und zeigen zum ersten mal, dass Krebszellen die Chromatinstruktur anderer Zelllinien imitieren können, um die Expression von Differenzierungsgenen zu blockieren.

Schlagwörter: PU.1, Hämatopoiesis, Chromatin Struktur, Generegulation, Läukemie

1 Introduction

1.1 Gene regulation

After the sequencing of the human genome had been essentially completed in 2001 (Lander et al., 2001; McPherson et al., 2001), the gene-centric paradigm shifted eventually to a genome-wide investigation, finally confirming the presence of gene networks, which define the effective transcriptome of specific cell populations (McPherson et al., 2001; Zhou et al., 2011). Moreover, comparative genome analysis revealed a surprisingly constancy in genetic content: vertebrate genomes contain only about twice the number of genes compared to invertebrate ones. Hence, evolutionary diversity can not be explained by the number of genes. Emerging evidences suggest that organisms' complexity developed due to a progressively elaborated gene regulation.

Hence, transcriptional regulation of genes is a tightly coordinate process which involves trans-acting proteins (transcription factors), binding at different genomic regions to modulate the transcriptional activity of a certain gene, binding on promoters and cis-regulatory elements in a chromatin-dependent context (discussed in the next chapter).

1.1.1 Trans-acting factors

Gene regulation is controlled by cis and trans- acting factors.

An enlightening example in this perspective is an increased absolute number of transcription factors (TF) (around 3000 in human, 300 in yeast) as well in the ratio of transcription factors per genome (1 per 10 genes in human; 1 per 20 in yeast) and their combinatorial activity on gene regulation (Levine and Tjian, 2003). This underlies the importance of trans-acting transcription factors as one of the driving force of the transcriptional decision-making. Not surprisingly, TF are involved in transcriptional regulation of genes coding for protein involved in every biological process. TF rarely act alone for a single cellular response; their function is rather a coordinate action of TF networks.

Perturbation of these networks can lead to neoplastic transformation resulting in cancer. Next-generation sequencing (NGS) methods allow nowadays to obtain an overview of gene expression on the genome wide level, however it is still challenging to analyze binding and function of multiple TF in different cell type (cancer cells compared to normal controls), due to huge bioinformatics work and lack of appropriate controls and algorithms. Therefore, it is

important to dissect gene regulation of a single TF by genetic manipulation integrated with molecular biology to understand in detail the contribution of a specific TF into different networks (Wilson et al., 2011a).

1.1.2 Cis-acting factors

1.1.2.1 Promoters

The promoter of a gene consists in general of two regions: the core-promoter and the proximal-promoter. The core-promoter is canonically characterized by the presence of the TATA box, located upstream of the transcriptional start site (TSS) Basal transcription machinery, including the TF-II family transcription factors, binds to the TATA box recruiting the RNA-polymerase II, which transcribes the gene (Smale and Kadonaga, 2003). However, only 32% of potential core-promoters in the human genome contain a standard TATA-box, the majority of which are housekeeping genes (Suzuki et al., 2001). The rest of promoters are regulated by the proximal region. The proximal promoter sequences are adjacent to the eventual core-promoter and display multiple TF binding sites which coordinate the recruitment of the transcriptional factory (RNAPII and related cofactors).

1.1.2.2 Enhancers and silencers

Enhancers are cis-elements which influence the transcription of specific genes, by recruiting the transcriptional machinery (RNAPII and cofactors), transcription factors and chromatin regulators to distal promoters, therefore inducing gene expression. Enhancers can be located upstream or downstream the gene promoter, in intronic regions or non-coding ones; hence, their function is independent from their orientation and spatial location (Suzuki et al., 2001; Atchison, 1988). Enhancers do not contain particular DNA sequences as signature, but binding sites of chromatin remodelers and transcription factors determine the specificity of their action on a particular gene and in a particular cell context (Majumder et al., 1997). Histone modification profiles (see next chapter) have been very useful to globally characterize enhancers. Recently, by genome-wide screening of histone modifications (see next chapter), Heintzman and colleagues (Heintzman et al., 2007) identified H3K4me1 (monomethylated lysine 4 on histon 3) as chromatin signature of enhancers in human and mouse genomes.

It remains unclear how enhancers function. There are two current theories: contact and no-contact models. The last one hypothesizes that enhancers indirectly modify the state of the respective promoter by changing the accessibility of promoter sequences to transcriptional

factories. The contact model proposes a physical interaction between promoter and enhancer, resulting in a chromatin loop (Zhou et al., 2011).

Silencers are cis-acting element as well, and their function and properties are very similar to enhancers, with the difference of the final outcome: when they are activated by transcription factors, the transcriptional activity of the relevant gene is reduced or blocked.

Silencers and enhancers are often located in a proximity to each other, and sometimes they overlap, assuming different functions in diferent cell type context. A clear example is represented by the cluster of regulatory element upstream the promoter of the transcription factor PU.1, as explained in chapter 1.4.1. In vertebrate, the majority of genes is regulated by a balance of different elements. This complex regulation is even more profound in genes encoding for transcription factors, which role is crucial in early differentiation programs, as the best known triad oct4, Sox2, Nanog (Wernig et al., 2007) for embryonic stem cells, or the complex network of hematopoietic factors (Rosenbauer and Tenen, 2007).

1.1.2.3 Insulators

A third category of regulatory element is represented by insulators. Their discovery is very recent compared to the other cis-acting elements: in 1999 Bell and colleagues described, and named (Bell et al., 1999), the first insulator. Insulators delimit independent transcriptional domains in eukaryotic genome (Bell and Felsenfeld, 2000; Bell et al., 2001). They are traditionally divided in two categories: barriers insulators and enhancer blocking ones (Bell et al., 1999; Gaszner and Felsenfeld, 2006). Barrier insulators are proposed to shield active genes from heterochromatin spreading (and consequent transcriptional silencing), by recruiting protein complexes that create a sterical obstruction (Felsenfeld et al., 2004; Litt et al., 2001). Enhancer blocking insulators shield a promoter from an enhancer's action. Importantly, these insulators are required to be physically located between the respective promoter and enhancers, in other words, they are position dependent. In eukaryotes, both types of insulator are prevalently bound by CTCF protein (see next chapter), which mediate their function in cooperation with multiple partner (Wallace and Felsenfeld, 2007).

The molecular mechanism underlying insulators' function remains elusive. Possible mechanisms might include sterical effect(Bell et al., 1999), a role in modifying chromatin by recruiting chromatin remodelers (Ishihara et al., 2006; Han et al., 2008) and also involvement of no coding RNA in gene transcription (Grimaud et al., 2006; Lei and Corces, 2006).

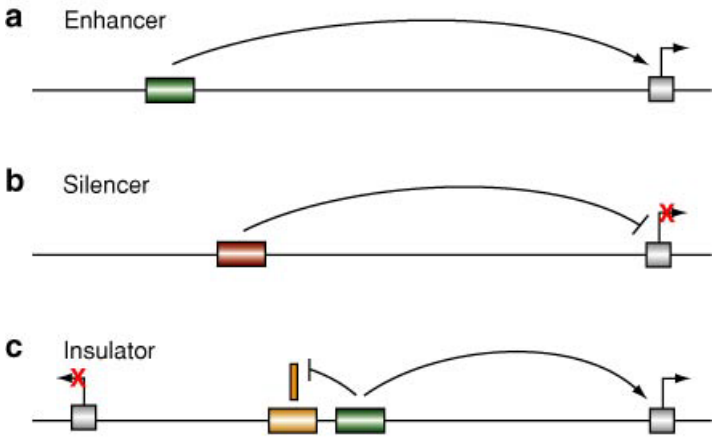


Figure 1: Cis-acting elements regulate transcriptional activity of respective genes. The picture illustrated the three categories of regulatory elements mentioned in the main text, a) enhancer, b) silencer, c) insulator. As the words suggest, the enhancer affects positively the transcription of the respective gene; the silencer plays the opposite role; the insulator constitutes a barrier between different gene loci so that regulatory elements specific for one gene don't affect also the neighbor's ones.

1.2 Epigenetic: chromatin structure and their organizers

The term "epigenetics" was first coined by Conrad H. Waddington in 1942 as a fusion of the words genetics and epigenesis describing "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being" (Waddington, 2012). In other words, it is a field that describes changes in a cellular phenotype that are independent of alterations in the DNA sequence, thus representing a global connection from genotype to phenotype. In the last 50 years, several processes has been explained by epigenetic mechanism: for instance imprinting genes (Ohlsson, 2007; McGrath and Solter, 1984; LYON, 1961) and X chromosome inactivation in vertebrates, which will be not further commented since they are out of scope of this thesis. Hereafter a summary of chromatin organization will be given, with a particular focus on chromatin organizers involved in this study.

1.2.1 Chromatin organization

In eukaryotic cells DNA is packaged repetitively in nucleosomes by interaction with histone proteins. One nucleosome is an octamer of histones (twice each H2A, H2B, H3 and H4) wrapped by 147bp of DNA (Olins and Olins, 1974; Luger et al., 1997). Nucleosome is a fundamental unit of chromatin. Its positioning along with DNA modifications and histone tail variants refers to the primary structure of chromatin. A higher level of chromatin organization

is the high-order structure and nuclear compartmentalization, together leading to the definition of three-dimensional structure of the nucleus (Olins and Olins, 1974; Richmond and Davey, 2003).

1.2.2 DNA modifications

DNA methylation, apart from the just recently discovered 5-hydroxymethylcytosine (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009), is the only known covalent modification of DNA in mammals. DNA methylation is achieved by the addiction of a methyl group to the C5 of the cytosine – it is a very general mechanism and that is found in prokaryotes and eukaryotes (Bird, 1980). In eukaryotes, DNA methylation occurs prevalently on the cytosine in the context of CpG dinucleotide. This DNA modification is usually associated with inhibition of transcription, therefore is considered to be a repressive mark. Its function is achieved by displacement of transcription factor from its binding site in case of methylation-sensitive DNA binding or affecting the chromatin status around the promoter of a gene. Interestingly, in cancer DNA methylation pattern is not randomly altered (Costello et al., 2000), contributing to a tumor-type specific aberrant transcriptional program.

1.2.3 Histon modifications

In 1964 Allfrey proposed that post-translational modifications of histones play a role in transcriptional regulation of genes (ALLFREY et al., 1964). However, this theory was reconsidered only 20 years after (Weintraub, 1984); nowadays, it is one of the most investigated topics in the epigenetic field. Current efforts are mainly based on next generation sequencing (NGS) technique coupled with immunoprecipitation (ChIP-seq) (Barski et al., 2007; Kouzarides, 2007)(for review see (Bell et al., 2011; Zhou et al., 2011)) in order to unequivocally define by a combinations of histone variants DNA sequences in their cellular roles. For example, the a.m. H3K4me1 is a signature for enhancers, or H3K4me3 (trimethylated lysine 4 of histone 3) for promoters, H3K36me3 (trimethylated lysine 36 of histone 3) covers the 3' end of genes, H3K27me3 correlates with gene repression, methylation of H3K9 has been implicated in heterochromatin formation, whereas its acetylation remarks active transcription, or the dynamic histone combination within a gene body guide the elongation of RNAPII (Heintzman et al., 2009; Barski et al., 2007; Boyer et al., 2006; Brookes and Pombo, 2009). Histone modifications represent therefore an important layer of transcriptional regulation.

1.2.4 Nucleosome positioning and remodeling

The nucleosome structure is ubiquitous, though histone post-translational modifications and nucleosome positioning affect the chromatin structure, contributing to cell type specific gene regulation. Nucleosome positioning can affect gene regulation by masking (or protecting, for example preventing methylation on CpG) or exposing DNA sequences, thus allowing transcription factors or transcriptional complexes to bind. It has been shown that an actively transcribed gene displays a well-positioned nucleosome pattern around the first exon and a nucleosome free region just upstream the TSS (Schones et al., 2008),. Interestingly Andersson and colleagues could observe that nucleosomes tend to localize more often with internal exons than with intronic region (Andersson et al., 2009), supporting the emerging concept that splicing, transcription and chromatin depend on each other (Neugebauer, 2002). On this conceptual line, there are more and more evidences of an additional layer of regulation represented by intergenic transcript (Ebralidze et al., 2008; Kim et al., 2007; Orom et al., 2010; Whitehouse and Tsukiyama, 2009).

1.2.5 High-order structure

The major chromatin unit in eukaryotes is the chromosome. Eukaryotic genomes are not randomly organized into the nuclei, but they occupy precise chromosomal territories (CT) (Cremer et al., 1982; Cremer et al., 2006; Hochstrasser et al., 1986; Misteli, 2008). A CT is the unit to measure the nuclear position occupied by a certain chromosome; the distribution of which can be described as radial position (Croft et al., 1999). This concept is very important because it allows quantification of the varying distribution across different cell type. Usually, CT correlates with the state of expression, where actively transcribed gene or cluster of genes are located in the interior and silenced regions are often observed in the periphery (Takizawa et al., 2008). However, CTs do not have rigid or physical boundaries, thus they are very flexible. The emerging idea is that clustering of genes in transcriptional hotspots contributes to their efficient and coordinated expression, like the HOX gene or T cytokines cluster (Cai and Kohwi-Shigematsu, 1999; Soshnikova and Duboule, 2009; Bergman et al., 2003) and to trans-interaction. This occurs not only for chromosomal translocation (Meaburn et al., 2007; Okuno et al., 2001) or different replication timing process, but also interchromosomal reversible interaction, which are important for gene regulation (Ling et al., 2006).

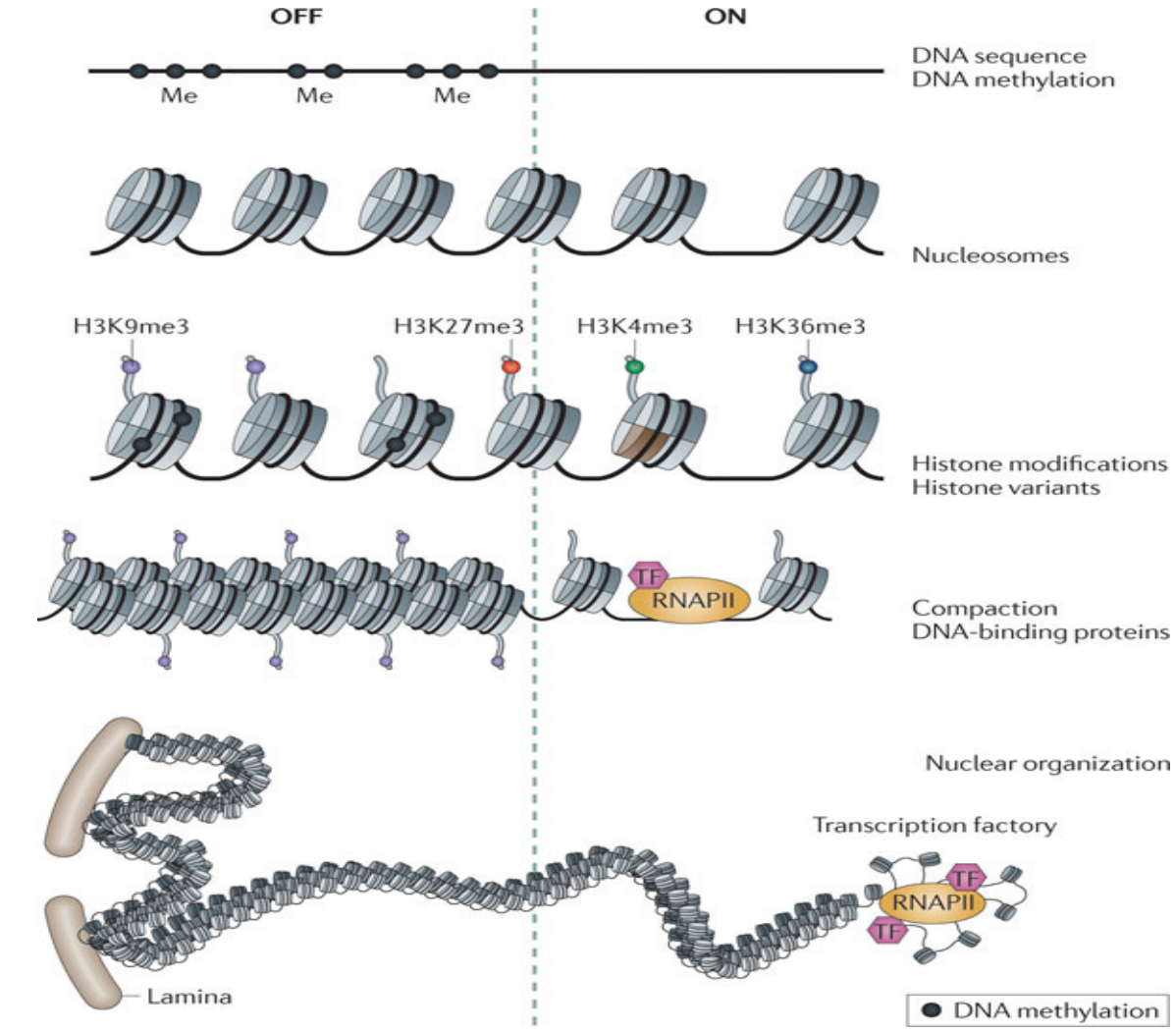


Figure 2: Different layers of epigenetic mechanisms involved in the chromatin organization. Chromatin organization in higher eukaryotes is achieved at different step, from modification in the primary sequences (here is depicted the example of cytosine methylation), to the nucleosomes levels, which controls the DNA accessibility. Histone tails post-translational modifications represent an ulterior layer, and they comprise active and repressive marks, few of which are indicated and discussed in the main text. Formation of heterochromatin is the next higher order structure, compacting chromosome regions which must be silenced. The last level of organization is the spatial organization of the chromosomes in the nuclear space. On the left part of the scheme repressive epigenetic mechanisms are depicted, contrasting the active ones (Zhou et al., 2011).

In order to establish and maintain the nuclear architecture in a specific cell type context, chromatin organizer proteins have been identified during the last decades. In the last two paragraph of this chapter, two very important architects will be discussed, since they are involved in this thesis.

1.2.5.1 SatB1: architect in T cells

SatB1 (special AT-rich binding protein 1) was originally identified in a complex with matrix attachment DNA regions (MARs) (Dickinson et al., 1992). MAR sequences were postulated to modify high order chromatin structure by mediating attachment of chromatin to topologically independent loop domains (Gasser and Laemmli, 1986). An example of the essential role of SatB1 mediated by MAR sequences is the binding of the L2a silencer, which regulates CD8a gene, encoding for an important cofactor of the T cell receptor (Banan et al., 1997); another example is the transcriptional regulation of the immunoglobulin μ heavy chain for BCR (B cell receptor) (Forrester et al., 2004; Goebel et al., 2002). Here the high order chromatin structure affects transcriptional activity by promoting epigenetic changes, which turn the locus in an active status. This results in generation of specific domains accessible for binding of transcription factors, and extensive demethylation across the chromatin domain (Forrester et al., 2004; Regha et al., 2007; Villa et al., 2007).

SatB1 is predominantly expressed in T cells. SatB1 knockout mice die after 3-4 weeks. These mice are thinner and smaller in size compared to the wild type littermate, but the only observed phenotype involves T and B cells. In fact, thymus and spleen are drastically smaller. Concentrating on the T cell phenotype (since in this thesis B cells will not be studied) in the thymus there are less immature progenitors but more double positive cells (CD4,CD8+)compared to the wt, indicating that SatB1 is determinant for T cell specification and development at multiple stages. Moreover, T cell activation in the periphery is impaired (Forrester et al., 2004; Alvarez et al., 2000).

SatB1 ablation leads to gene dysregulation affecting hundred of genes (at least 2% of total gene) (Yasui et al., 2002). This global control on gene regulation can be achieved by the chromatin architecture established by SatB1: it has a unique cage-like distribution in thymocytes nuclei, establishing a unique intranuclear architecture by anchoring DNA sequences and exposing specific gene loci to tissue specific activation or repression by gene regulators as trans-acting transcription factor (SatB1 favors or inhibits their binding by the means of DNA accessibility) or cis acting mechanisms (as histone tail modifications). In other words, this unique structure differentiates between euchromatin and heterochromatin (Cai et al., 2003; Notani et al., 2010). With a such spatial organization SatB1 compacts the cluster of cytokine genes in loop-dense structure, allowing their coordinated transcription during T cell specification and differentiation (Cai et al., 2006; Galande et al., 2001).

All these studies point to a essential role of SatB1 in spatial organization especially in T cells, which contributes to regulate the multistep development of T cell at transcriptional level.

1.2.5.2 CTCF: master weaver of the genome

CTCF (CCCTC-binding factor) is the most widely studied insulator-binding protein in vertebrate. CTCF was originally discovered as a transcriptional inhibitor of chicken myc and lysozyme genes (Baniahmad et al., 1990; Lobanenkov et al., 1990). CTCF protein contains 11 zing fingers, by which can mediate multiple DNA interaction. The same structure is maintained across vertebrates, being conserved from human to Drosophila (Moon et al., 2005). Mapping CTCF binding site (BS) at genome wide level by deep sequencing ChIP technology had shown that many of them remain invariant across different human cell types (> 13.000; (Kim et al., 2007)), and are conserved among human, mouse and chicken CTCF BS (>200 BS are syntenic, meaning located in equivalent genomic position regardless of the sequence conservation; (Martin et al., 2011)). This suggests a crucial role of CTCF genome wide transcriptional regulation.

It is an accepted paradigm that CTCF exerts its functions via establishing long-range chromatin interaction (Ling et al., 2006; Splinter et al., 2006). For this reason CTCF has been called ‘a master weaver of the genome’ (Phillips and Corces, 2009). CTCF’s functions can be summarized in four categories: it acts as 1) a barrier insulator by blocking heterochromatin spreading into active loci 2) an enhancer blocker 3) a three dimensional genomic organizer 4) a transcriptional enhancer. CTCF was found to block the spreading of repressive telomeric heterochromatin in S. Cerevisiae (Bell et al., 1999; Defossez and Gilson, 2002; Bowers et al., 2009). By binding an insulator region CTCF coordinates transcriptional activity of two genes, IL-3 (interleukin-3) and GM-CSF (Granulocytic-Macrophage Colony-Stimulating Factor), located in the same locus, but differently expressed within hematopoietic cell types. Multiple DNA methylation sensitive CTCF BS are located in the Xci (X chromosome Inactivation Center), suggesting a direct role of CTCF in X chromosome inactivation (Chao et al., 2002; Donohoe et al., 2009; Navarro et al., 2006). The imprinted locus Igf2 (Insulin-llike Growth Factor-2)/H19 had been studied in molecular detail: these two genes share an enhancer, whose activity induces H19 expression in maternal allele and Igf2 in the paternal one. This regulation is achieved by DNA-methylation sensitive CTCF binding to an enhancer blocking insulator located between the two genes (ICR=Imprinted Control Region). In the maternal allele CTCF binds to non-methylated ICR and blocks the enhancer function on the

Igf2 gene and vice versa in the paternal allele (Bell et al., 2001; Murrell et al., 2004; Fedoriw et al., 2004).

Multiple evidences suggest a major role for CTCF as gene regulator by long-range interactions, both as repressor (Lutz et al., 2000) or activator (Hadjur et al., 2009). Few examples of the importance of CTCF in coordinating transcriptional regulation are the following: CTCF protects p53 gene in humans by maintaining the locus in an open state and avoiding its repression by incorporation of repressive histone marks in tumor cells (Soto-Reyes and Recillas-Targa, 2010). CTCF plays a functional role in lymphoid pathway, both in T cell and B cell development: during T cell maturation, CTCF regulates cell cycle progression of specific subset of thymocytes (Heath et al., 2008). During B lymphoopoiesis, it is involved in the spatial organization of the Igh locus for a correct BCR rearrangement (Degner et al., 2011). CTCF binds also in the first intron of the key hematopoietic regulator Myb, forming an Active Chromatin Hub (ACH) which undergoes destabilization and consequent disruption during erythroid differentiation (Stadhouders et al., 2012). Moreover, CTCF-mediated loops in the ß-globin gene determine the chromatin conformation of the locus directly involving active and repressive chromatin mark throughout erythroid differentiation (Tolhuis et al., 2002; Splinter et al., 2006; Palstra et al., 2003; Dekker et al., 2002).

CTCF is a crucial organizer of the high-order structure of chromatin, but plays also essential roles at local structure. These studies point to an emerging concept of epigenetic landscape, that controls and modulates trans-acting factors (as transcription factors) activity on gene transcription.

1.3 Hematopoiesis

Hematopoiesis is a developmental process of blood cells of all lineages. The main cellular components of blood are erythrocytes, platelets and leukocytes (white blood cells). Erythrocytes transport oxygen from the lungs to all tissues through the circulation stream, platelets are responsible for blood coagulation and leukocytes, which are represented by myeloid and lymphoid cells, protect individuals from infection forming the immune system. Myeloid cells are monocytes, the precursors of tissue macrophage, and granulocytes; lymphoid cells are mainly T cells and B cells, and a small percentage of natural killer cells (Kawamoto et al., 2010).

Apart from some rare so called memory lymphoid cells, all mature blood cells have a short lifespan ranging from hours for granulocytes to couple of months for erythrocytes. The

turnover of this system is constant throughout life, in man amounting to millions of cells per second in a steady state. This rate can rapidly increase even ten fold when challenged, e.g. by infection or bleeding, in order to compensate stress and maintain the system homeostasis (Ogawa, 1993; Kaushansky, 2006). To sustain this rapid turnover some progenitor cells exist in blood system. Such cells are Hematopoietic Stem Cells (HSC) (Adolfsson et al., 2005; Wilson and Trumpp, 2006) which differ from the more differentiated counterparts in possessing a combination of two properties: they can generate more HSCs (a process called self-renew) and are pluripotent, meaning they have the potential to stepwise differentiate into all mature blood cells (Orkin, 2000). In adult hematopoiesis (after birth), HSCs reside in the bone marrow in specialized cellular structures, essential for maintenance of adult HSCs (Wilson et al., 2009). All blood cells are generated in bone marrow, except T cell lineage, which develops in the thymus from progenitors derived from these organ (Weissman et al., 2001; Rothenberg et al., 2008).

Hematopoiesis is usually depicted in a hierachical fashion, starting from HSC differentiation first to progenitors and then to precursors with varying commitments to multiple or single pathway (Orkin, 2000). Differentiation is defined as a sequence of events through which immature precursor becomes mature effector cells. During differentiation stem cells give rise to progeny that progressively lose self-renewal capacity and become restricted to specific lineage (Akashi et al., 2000); this lineage restriction is also referred to as “cell fate decision” or “lineage commitment”. In the last decades several models have been proposed to describe a hierarchy of the hematopoietic system (Kondo et al., 1997; Reya et al., 2001), and a simplified version of the hierarchy which aim to merge different proposed models (Adolfsson et al., 2005; Kawamoto et al., 2010; Wilson and Trumpp, 2006) is depicted in Figure 1.

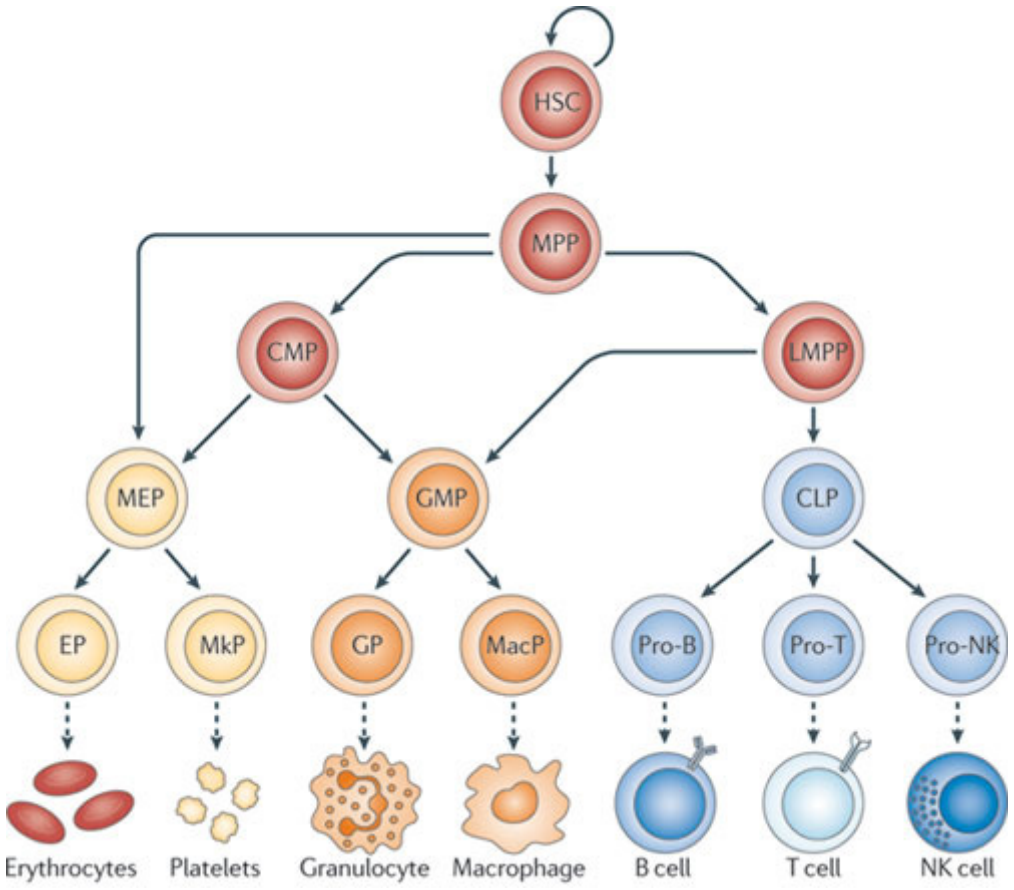


Figure 3: schematic view of the hierarchy of hematopoietic populations. All hematopoietic cells derive from the HSCs, which differentiate into an intermediate MPP (multipotent progenitor) stage, which again can differentiate in all lineages but lost the stem cell property of long-term self-renewal. The model postulates that the binary decision between myeloid and lymphoid pathway occurs at this level. However, as underlined from severale cross-lineage arrows, based on transdifferentiation assays, the illustrated hierarchy is an hypothetic flexible system. Lymphoid and myeloid progenitors (CMP-GMP and CLP) are at this point primed for a specific lineage commitment, which will end up in effector cells through stepwise maturation processes (Cedar and Bergman, 2011).

The hematopoietic system is intensively studied for regulation of cell fate decision, since it provides a valuable model for examining genetic programs, established and executed in vertebrates, and alteration of blood homeostasis in pathological processes, as leukemia. First of all, it is relatively easy to isolate hematopoietic cells from mice. Hematopoietic cells are found in several organs of the body: the bone marrow (BM) contains adult HSCs, myeloerythroid and B-lymphoid progenitors and consists largely of mature blood cells. In the spleen and lymph nodes are found the majority of mature B and T cells, but also macrophages and granulocytes. The thymus is mostly populated by T cells carrying the T cell receptor

(TCR) and still expressing both CD4 and CD8 (double positive, DP cells) and in a little extent by T cell progenitors. Mature hematopoietic cells are also present in the peritoneum and in the peripheral blood. Secondly, many hematopoietic populations with different lineage capabilities and at various stage of differentiation were defined by isolation though the using monoclonal antibodies that recognize specific surface markers (CD = cell determinants). Therefore, it is possible to stain isolated tissue with different combination of antibodies and characterize distinct hematopoietic cell lineages by FACS (Fluorescence Activated Cell Sorting) at a single cell level. By this technique is not only possible to analyze a distribution of different lineages in a tissue, but also to sort specific populations for further functional assays (Forsberg et al., 2006).

1.3.1 Transcription factors regulate hematopoietic development

In hematopoietic development, blood cells acquire defined phenotypes as a result of coordinated regulation of cell-specific molecular pathways, involving cytokines receptors and transcription factors (Kondo et al., 2000; Shivdasani and Orkin, 1996; Zhu and Emerson, 2002). It is clear that transcriptional factors play a key role in the determination of cell fate programs by promoting or repressing lineage-specific genes. However, there is no single "master" transcription factor, which is responsible for pivotal steps of differentiation; rather, gene expression is controlled by combinatorial function of several transcription factors, resulting in cell type specific network (Rosenbauer and Tenen, 2007; Rosmarin et al., 2005). To define these networks of lineage-specific expression profiles on a large scale, genome wide technologies as microarrays and next generation deep-sequencing analysis are carried out on different blood cells populations in the last decade (Forsberg et al., 2006). A bottom-up approach investigating a role of specific transcription factors, especially in very rare haematopoietic populations (difficult to be purifying and studied), consists in genetic strategies, primarily gene-targeting (knock-out, knock-in and transgenic approach) and consequent gene overexpression or downregulation using retroviral vectors (Zhu and Emerson, 2002; Laiosa et al., 2006). Mouse models lacking a specific transcription factor do not only play an important role in understanding major changes in the haematopoietic development, but provide a reproducible biological material to perform genome-wide analysis on different cell populations. Up to now, the role of distinct hematopoietic transcription factors in lineage commitment and cell fate decision has been widely investigated in last

decades, predominantly by a loss or gain-of- function mouse models: However, there are still many aspects, which remain elusive (Bonadies et al., 2011).

1.3.2 Leukemia: when hematopoiesis is perturbed

The hematopoietic system has been studied in detail in order to understand the etiology of leukemia from macroscopic changes up to molecular details. Leukemia refers in general to all blood cancers, including myeloma, myeloid leukemia and myelodysplastic syndromes. The word 'leukemia' derives from ancient Greek: λευκό/lephko means white and αίμα/hema is the suffix pertaining to blood: leukemia is in fact a malignant disease marked by altered proliferation and development of leukocytes and/or their precursors in blood-forming organs, affecting bone marrow, blood cells, lymph nodes and other part of the lymphatic system. Since this thesis investigates acute myeloid leukemia (AML), this chapter will focus only on this particular type of leukemia.

AML is a predominant type of acute leukemia in adults. The median age at diagnosis is 67 years (Estey, 2012a). Leukemia, mainly AML and T-ALL (T cells acute lymphoblastic leukemia) are together the main cause of death by cancer in childhood (Deneberg, 2012).. AML is characterized by an accumulation of granulocytic or monocytic precursor in the bone marrow. It is a very heterogeneous cancer, therefore complicated to classify. The most widely used classification is still the French American British scheme (FAB), which was first proposed in 1976. It relies on morphologic, cytochemical, and immunophenotypic features of the neoplastic cells to establish their lineage and degree of maturation, taking into account genetic lesions. Following these criteria, AML can be divided in 7 subtypes, named M0-M7 according with increasing degree of differentiation of the cancer blasts. Incidence and prognosis are also approximately indicated in the table below.

Current treatment options include chemotherapy, allogenic stem cell transplantation of bone marrow, and other drug therapies. Current efforts are directed to identify novel genetic and epigenetic markers or combination of them in order to design specific drugs for specific AML (see for review, (Estey, 2012b)).

FAB subtype	Name	% of adult AML patients	Prognosis compared to average for AML
M0	Undifferentiated acute myeloblastic leukemia	5%	worse
M1	Acute myeloblastic leukemia with minimal maturation	15%	average
M2	Acute myeloblastic leukemia with	25%	better
M3	Acute promyelocytic leukemia (APL)	10%	best
M4	Acute myelomonocytic leukemia	20%	average
M4 eos	Acute myelomonocytic leukemia with eosinophilia	5%	better
M5	Acute monocytic leukemia	10%	average
M6	Acute erythroid leukemia	5%	worse
M7	Acute megakaryoblastic leukemia	5%	worse

Table 1: French-American-British (FAB) classification. Incidence frequency and prognosis predictions are also indicated (Jemal et al., 2002).

The role of genes involved in leukemogenesis has not been yet completely understood, leaving open questions on gene regulation and pivotal transcription factors. Ideally, such a molecular approach has to be integrated with clinical research in a translational medicine perspective. So far investigations of AML pathogenesis mainly focus on the expression analysis of oncogenes (Mueller and Pabst, 2010; Pabst and Mueller, 2007) and tumor suppressor genes that regulate cell proliferation and differentiation.This thesis focuses on master transcription factor PU.1 and investigates its role in regulation of hematopoiesis and involvement in leukemogenesis.

1.4 Master hematopoietic transcription factor PU.1

PU.1, a widely studied member of the large family of Ets transcription factors, plays a crucial role in all hematopoietic cell types, since it is a primary lineage determinant from the early progenitor state. PU.1 is required for myeloid and lymphoid lineages development for orchestrating a transcriptional network specific for each of the lineages. PU.1 is expressed in a lineage specific manner throughout the hematopoietic system; it can be detected at low levels in HSCs and in LMPP, low and constant levels during B-cell development, is upregulated in

the myeloid compartment, and becomes absent in the erythroid and in T-cell lineages(Scott et al., 1994; Klemsz et al., 1990; Iwasaki et al., 2005).

The Pu.1 gene consists of 5 exons and is located on chromosome 2 (2.E3) in mouse and 11 (p11.22) in human. The homology between the proteins is 85% (264 amino acids of the human and 266 or 272 amino acids of the murine one, depending on the initial transcription codon); the DNA exon sequences are evolutionary conserved between the two species by meaning of more than 75% (Gupta et al., 2009). The PU.1 protein consists of four different domains: 1) the ETS domain at the carboxyl terminus, which mediates both protein interactions (with partner such as C/EBP, c-Jun, GATA-1 and Runx1) as well as binding to the DNA by a loop-helix-loop structure, 2) the PEST domain which is essential for protein-protein interactions (such as IRFs family members) 3) a glutamine-rich domain, which interact with GATA-1, GATA-2 and RB (retinoblastoma protein) and 4) an acidic domain, which can recruit components of the basal transcriptional machinery (such as TFIID and TBP); therefore the last two domains are together considered as trans-activating domain for the direct role of recruiting essential factor of the transcriptional machinery for the initiation of transcription and the recruitment of RNA polymerase II.

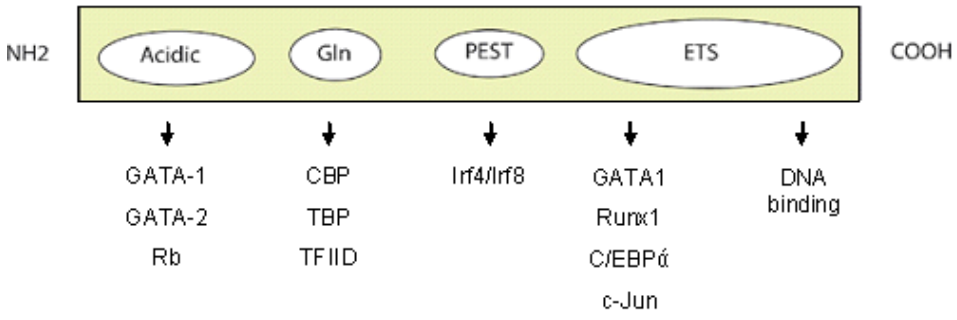


Figure 4: schematic representation of the PU.1 protein and relative domains. Several proteins, mostly transcription factors, interact with PU.1 promtoer. The ETS domain is the best conserved within the ETS-transcription factors.

The interaction with protein partners may result in positive (c-Jun) or negative (GATA-1), or synergic (C/EBPα) effects on PU.1 function: the stechiometric balance between PU.1 and GATA-1 determines myeloid (higher PU.1 expression) against megakaryocyte-erythroid cell fate (higher GATA-1 expression) by reciprocal displacement at DNA binding site of target genes (Chou et al., 2009; Nerlov et al., 2000; Rekhtman et al., 1999; Stopka et al., 2005). c-Jun interacts with PU.1 promoting myeloid transcriptional program, and associated binding to

the DNA of C/EBPα and PU.1 marks specificly monocyte target gene after myeloid commitment (Dahl et al., 2003; Heinz et al., 2010; Laslo et al., 2006; Behre et al., 1999; Grondin et al., 2007).

The PU.1 binding site recognized by the EST domain consists of a purine-rich sequence core (5'GGAA), which is present in the promoter region or in distal sequence of its target genes. Through the trans-activation DNA domain PU.1 can bind upstream the promoter, interact with the TATA binding protein and universal cofactors like Sp1 or hematopoietic cofactors like AML-1, inducing the initiation of transcription. Most of PU.1 targets are the key factors for myeloid and lymphoid differentiation like receptors for lineage specific cytokines. Some examples are macrophage, granulocyte or both colony stimulating factor receptor (M-CSFR; G-CSFR; GM-CSFR) for the myeloid lineage, or IL7Rα, immunoglobulin (Ig) light- (κ and λ) and heavy- (μ) chain gene, required for a correct B cell development (Friedman, 2007).

Since different levels of PU.1 modulate blood cell fate decision and deregulation of PU.1 leads to abnormal hematopoiesis and might be involved in leukemogenesis, it is important to understand how Pu.1 gene is regulated in distinct population types.

1.4.1 PU.1 role in hematopoiesis by *in vivo* mouse models

PU.1 has emerged as a central regulator of hematopoiesis and elicits its function in an orchestrated and coordinated manner with lineage specific transcription factors networks. For this reason it is important to investigate its role using in vivo models, which allows studying the consequence of the PU.1 loss or overexpression in the whole hematopoietic system. Several mouse models have been developed in the past 20 years and some of them as listed in the table below (Table 2). Since PU.1 knockout is embryonic or neonatal lethal, conditional strains were generated, to understand the consequence of PU.1 loss of function in a specific lineage. Another approach involves gene regulation studies, with consequent ablation of regulatory elements or indispensable genomic region for PU.1 gene transcription. These models aim to perturb PU.1 expression pattern and observe the phenotype resulting from lower than physiological (Houston et al., 2007; Rosenbauer et al., 2004) or higher (Anderson et al., 2002; Moreau-Gachelin et al., 1996) PU.1 levels.

Type of mutation	Genetic modification	Main phenotype	Reference
Overexpression	PU.1 transgene driven by LTR from SFFV virus	Erythroleukemia	Moreau-Gachelin et al. (1996)
Null allele	Inserion of neomycin cassette into exon 5, disrupting the Ets domain	Late fetal / neonatal death. No fetal B,T, and myeloid cells	Scott et al. (1994) McKercher et al. (1996)
B cell specific deletion	Deletion of floxed exon 5 with CD19-cre	Development of functional B cells	Polli et al. (2005)
T cell specific deletion	Deletion of floxed exon 5 with Lck-cre	Impaired response for allergic inflammation (TH9 phenotype)	Chang et al. (2010)
Inducible deletion	Poly-IC induced deletion of floxed exon 5 with Mx1-cre in aduts	Granulocytic expansion and myeloid leukemia	Dakic et al. (2005) Metcalf et al. (2006)
Deletion of -14Kb URE	PU.1 kd/kd	Block in T cell differentiation, T cell lymphoma in 70% of mice. Block of B cell differentiation in BM and expansion of B1 cells. Late development of myeloid leukemia (AML)	Rosenbauer et al. (2006)

Table 2: List of the main mouse models generated to study PU.1 role in hematopoiesis. Several models have been developed, and here only a summary of those is given, focusing on the main phenotype as representative example of the investigated lineages.

Summarizing the outcome of these models, it has been proved that PU.1 is necessary for myeloid pathway development, commitment and differentiation, although its high levels in mature macrophages are maintained by default; PU.1 is dispensable for B lymphopoiesis but not for lymphoid commitment (Polli et al., 2005). PU.1 must be downregulated during T cell development and for erythroid lineage (Rothenberg et al., 1999; Moreau-Gachelin et al., 1988; Rao et al., 1997); even if expressed at lower level, PU.1 plays a role in periphery activation of TH9 lineage, involved in allergic inflammation, by regulating interleukin-9 production (Chang et al., 2010). Importantly, PU.1 perturbation is sufficient to lead to leukemic transformation in vivo, enlightening its role as tumor suppressor (Rosenbauer et al., 2006).

After summarizing PU.1 role in hematopoietic development in vivo, in the next chapter PU.1 gene regulation will be discussed.

1.4.2 PU.1 gene regulation

In vivo mouse models represent a powerful approach to investigate the role of a gene in physiological onset, as well as loss and gain of function effects in different lineages. However, mouse models actively perturbing PU.1 expression are not useful for understanding the molecular mechanisms involved in establishing PU.1 transcriptional activity. To address this question, multiple studies have been performed to figure out how PU.1 gene is regulated. PU.1 gene doesn't contain the standard TATA box, and PU.1 promoter is not sufficient to

drive alone luciferase expression in reporter assay (Li et al., 2001). Instead, PU.1 gene regulation is orchestrated by multiple regulatory elements not in the promoter region, responsible for its graded and lineage-specific expression: it has been shown, there is a cluster of discrete regulatory regions upstream the PU.1 promoter, with different dynamics throughout hematopoietic cell types (Leddin et al., 2011; Zarnegar et al., 2010). Among these, the URE (upstream regulatory element: located -14Kb or -17Kb upstream the PU.1 promoter, in mouse and human, respectively) has been widely investigated because of its primary role in PU.1 gene regulation. The URE function is necessary for PU.1 expression already in the stem cell compartment (Steidl et al., 2006). Different studies support the hypothesis that the URE is crucial for opening the PU.1 locus and therefore stepwise conditioning its expression at progenitors level, both myeloid or in lymphoid pathway. Several pivotal transcription factors have been found to bind the URE region: RUNX1, a master hematopoietic transcription factor essential for lineage commitment in the stem cell compartment and throughout hematopoietic development (Hoogenkamp et al., 2009; Huang et al., 2008), Ikaros, priming stem cell for lymphoid commitment (Zarnegar and Rothenberg, 2012), SatB1, a chromatin organizer and major repressor during T cell development (Steidl et al., 2007), member of the LEF/TCF family, downstream target of Wnt signalling (Rosenbauer et al., 2006), Egr family members, involved in establishing myeloid transcriptional network together with PU.1 (Hoogenkamp et al., 2007), NF-κB, an activator of PU.1 transcription (Bonadies et al., 2010) and PU.1 itself, via an autoregulation transcriptional loop (Okuno et al., 2005).

Besides the URE, another element (-12Kb upstream the PU.1 promoter) has been identified as an enhancer in myeloid cells but not in B cells, suggesting that the threshold of PU.1 expression maintained by an additional autoregulatory loop can discriminate between high and intermediate PU.1 expression (Leddin et al., 2011).

Rothenberg and Zarnegar (Zarnegar et al., 2010) dissected the whole cluster of conserved region after identification of two different additional elements in a proximity to the URE region. Function analysis of the elements was performed in vitro, by comparing myeloid cell with immature T cell lines, and Runx1 dosage-dependent PU.1 regulation model was proposed, where levels of Runx1 establish its binding through mutliple discrete elements in a cell-type specific manner, favoring myeloid factors to enhance PU.1 transcription or cooperating with T cell specific factors as Ikaros strongly silence PU.1 locus. The two models below represent current hypothesis of PU.1 gene regulation. It appears clear though, that many components and their interactions should be investigated in detail.

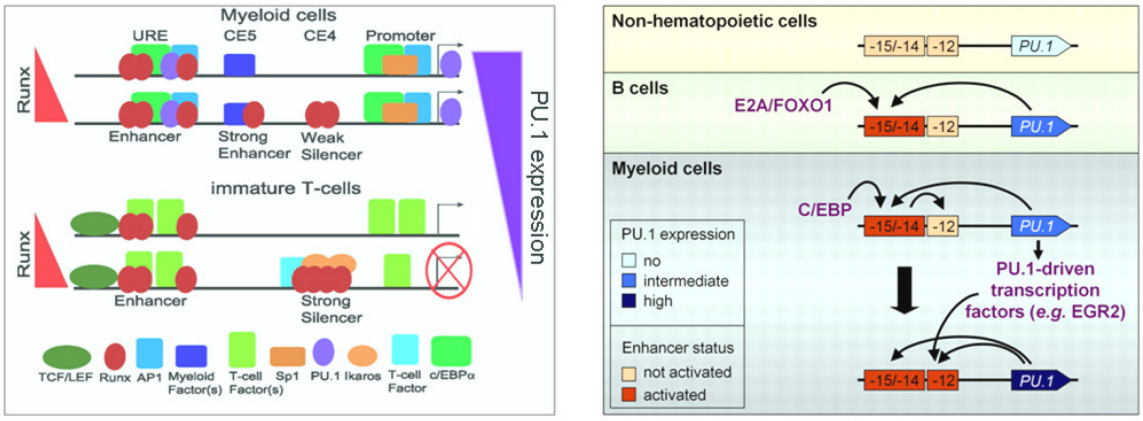


Figure 5: Most recent models for PU.1 gene regulation. Several studies investigate PU.1 gene regulation. Depending on the approach, there are several proposed models, but all of them have some shadowed aspects which remain unknown. For instance, the model of Rothenberg and colleagues (on the left) (Zarnegar et al., 2010) is based only on in vitro gel shift assay modulating Runx1 levels; for that reason, a lot of complexes could not been identified, in fact there are named as "myeloid factor" or "T-cell factor". However, this model provided the first evidence of the existence of a cluster of discrete elements, in addition to the previously described URE. The model on the right was proposed by a previous work of our lab (Leddin et al., 2011): here the focus is directed on the differential activation state of the PU.1 gene from an intermediate state (B-.cells) toward a very active state (macrophages), explained by the function of a novel enhancer.

1.4.3 PU.1 as tumor suppressor: lessons from AML

Given the importance of PU.1 for normal hematopoiesis, PU.1 represents an obvious target for disruption in AML. Indeed, mutations of the PU.1 gene were detected in 7% of patients with AML. The mutations decreased PU.1 ability to interact with the other partner protein, such as RUNX1 or c-Jun, therefore diminishing their synergistic effect on transcriptional regulation of target genes, responsible for myeloid differentiation (Mueller et al., 2003). However, these point mutations are too rare to explain the high frequency of PU.1 dysregulation in AML patients (Mueller et al., 2003; Mueller et al., 2006b; Zhu et al., 2012). In fact, PU.1 expression and function are often altered by other mechanisms. PU.1 is known to be a target of three fusion products, frequently generated in leukemias due to chromosomal translocations.. AML1-ETO, a fusion product generated by (8;21) translocation, is found in 10% of AML. This oncogene can ablate the function of RUNX1 (AML1), leading to leukemia. PU.1 function is down-regulaed by AML1-ETO through protein-protein interactions and consequent displacing of the PU.1 transcriptional coactivator (Vangala et al., 2003), c-Jun. FLT3-ITD (tyrosine receptor kinase FLT-3 with internal tandem duplications,

which lead to constitutive activation) mutation occurs in 25% of AML. This oncogene represses PU.1 transcription (Mizuki et al., 2003). PML-RARα, resulting from the (15;17) translocation, is the genetic hallmark of APL M3 leukemia (98% cases). As FLT3-ITD, also PML-RARα, inhibits PU.1 at transcriptional level (Cook et al., 2004; Grignani et al., 1998; Mueller and Pabst, 2006; Walter et al., 2005). Importantly, treating APL with ATRA (all-trans retinoic acid), which neutralize the function of the fusion oncogene, results in PU.1 upregulation and differentiation of leukemic blasts (Huang et al., 2008).

Inactivation of PU.1 in adult mice is sufficient for leukemia development. And URE knockout mice, where PU.1 expression is downregulated up to 20%, develop AML. Additionally, Steidl and colleagues (Steidl et al., 2007) identified a SNP (single nucleotide polymorphism) within the URE region, which is associated more frequently with complex karyotype leukemia (pour prognosis).

1.5 Aim of the thesis

Transcription factor PU.1 plays a major role in hematopoietic development as lineage priming determinant. PU.1 is expressed in a dynamic pattern in different lineages and alteration of PU.1 levels can lead often to leukemia or proliferative disorders. It is therefore necessary to understand the molecular mechanisms behind PU.1 cell-type specific gene regulation. In the last decades PU.1 gene regulation has been widely investigated in the myeloid and B-cell compartment. However, little is known about PU.1 regulation in T-cells. PU.1 must be in fact actively silenced during T cell lymphopoiesis, and lack of its downregulation leads to T-cell leukemia.

On the other hand, high levels of PU.1 are required for myeloid differentiation, and PU.1 expression block at early myeloid progenitors has been observed in human acute myeloid leukemia.

Thus, this thesis aims to understand how PU.1 is downregulated in T-cells and blocked in acute myeloid leukemia by investigating the 3 dimensional chromatin structure of the PU.1 locus.

2 Materials and Methods

2.1 Materials

2.1.1 General equipment

Mastercycler Gradient	Eppendorf
7300 Real Time PCR System	Applied Biosystems
Multicentrifuge 3 S-R	Heraeus
Geldoc 2000	Biorad
Thermomixer	Eppendorf
Agarose gel chambers	Biosteps
Power supply EV231	Consort
Flow hood	BDK
Microscope DMIL	Leica
Nanodrop	PeqLab
FACS LSRII	BD
Incubator	Binder
Luminometer	Berthold Technologies
Stratalinker 2400	Stratagene
PDVF membrabe	Pall Corporation
XAR film	Kodak
Hypercassette	Amersham
Phosphoimager cassette FLA-3000	FUJI
Fuji RGII photographic system	FUJI
7300 Real Time PCR System	Applied Biosystem
CFX96 Real Time System	Biorad

2.1.2 Cell culture equipment

Cell culture dishes and flasks, sterile	TPP or Falcon
Centrifuge tubes, sterile, different sizes	TPP or Falcon
Serological pipettes	Falcon
Neubauer cell-counter chamber	Superior Marienfeld
Needles for single-use, sterile, different sizes	Neoject
Polystyrene tubes, 5mL	BD Falcon
Cryotubes, sterile, 1.5mL	Nunc
Cell strainer, sterile, different sizes	BD

2.1.3 Mouse dissection equipment

Dissecting board and pins	
Forcipes and scissors of different sizes	Brand
Scalpels, sterile	Cutfix

2.1.4 Chemicals and reagents

1Kb ladder	Fermentas
100bp ladder	Fermentas
Agarose	Roth
Ammonium persulfate (APS)	Roth
ATP	Roth
[α-32P]dCTP	Amersham
Bromophenol blue	Roth
BSA	Roth
$CaCl_2$	Roth
Chlorophorm/Isoamylalcohol	Roth
dNTPs 10mM each	Fermentas
DTT	Roth
Ethanol absolute	MDC Lager
Ethidium bromide	Roth
Formaldehyde 37%	Roth
Giemsa stain	Roth
Glacial acetic acid	Roth
Glycerol	Roth
Glycine	Roth
Glycogen	Roth
HEPES	Roth
Hybri-Quick	Roth
Isopropanol	Roth
KCl	Roth
KH_2PO_4	Roth
IPTG	Roth
KCl	Roth
LB medium	Roth
LB agar	Roth
LiCl	Roth
Methanol	Roth
$MgCl_2$	Roth

NaCl	Roth
Powder milk	Roth
NaCl	Roth
$NaHCO_3$	Roth
NaOH	Roth
Nonidet P-40	Sigma-Aldrich
Pageruler prestained protein ladder	Fermentas
Phenol	Roth
Polyacrilammide 37.5:1	Roth
Propidium iodide	Sigma-Aldrich
Proteinase K (stock 10mg/mL)	Roth
Proteinase inhibitor cocktail	Roche
RNase free water	Fermentas
Sodium citrate	Roth
Sodium butyrate	Roth
Sodium dodecyl sulphate	Roth
TEMED	Roth
Trichlormethan (Chloroform)	Roth
Triton X-100	Roth
Tris base	Roth
Tris-HCl	Roth
Xylene cyanol	Roth

2.1.5 Buffers and solutions

ACK (red blood cells lysis buffer)	0.15 M NH4Cl, 10 mM KHCO3, 0.1 mM EDTA, pH 7.3 in water
ChIP cell lysis buffer	10 mM Tris pH 8.0, 10 mM NaCl, 0.2 % NP40, protease inhibitors
ChIP nuclei lysis buffer^	50 mM Tris pH 8.1, 10 mM EDTA, 1 % SDS, protease inhibitors
ChIP IP dilution buffer	20 mM Tris pH 8.1, 2 mM EDTA, 150 mM NaCl, 1 % Triton X-100, 0.01 % SDS, protease inhibitors

ChIP IP wash buffer 1	20 mM Tris pH 8.1, 2 mM EDTA, 50 mM NaCl, 1 % Triton X-100, 0.1 % SDS, protease inhibitors
ChIP IP wash buffer 2	10 mM Tris pH 8.1, 1 mM EDTA, 0.25 M LiCl, 1 % NP40, 1 % deoxycholic acid, protease inhibitors
ChIP elution buffer	100 mM NaHCO3, 1 % SDS
Nuclear extract buffer A	10 mM HEPES pH 7.9, 1.5 mM MgCl2, 10 mM NaCl, protease inhibitors
Nuclear extract buffer C	20 mM HEPES pH 7.9, 25 % glycerol, 0.42 M NaCl, 1.5 mM MgCl2, 0.2 mM EDTA
FACS buffer	2 % (v/v) FCS, 2 mM EDTA, in PBS
Freezing medium	50 % (v/v) medium, 40 % FCS (v/v), 10 % (v/v) DMSO
Phosphate buffered saline (PBS)	137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 2 mM KH2PO4, pH 7.4
Southern blot wash solution I	100 ml 20 x SSC, 10 ml 20 % SDS, in 890 ml water
Southern blot wash solution II	10 ml 20 x SSC, 10 ml 20 % SDS, in 980 ml water
Tail digestion buffer	10 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 50 mM NaCl, 0.5 % (v/v) SDS,
TE buffer	10 mM Tris, 1 mM EDTA, pH 7.5
3C cell lysis buffer	10 mM Tris pH 8.0, 10 mM NaCl, 0.2 % NP40, protease inhibitors

1x EMSA binding buffer	20 mM Hepes, pH 7.6, 10 % glycerol, 2 mM EDTA, 2 mM DTT, 10 mM MgCl2, 100 mM KCl in water
1x SDS sample buffer	0.5 M Tris-HCl pH 6.8, 5 % glycerol, 2 % SDS and 100 mM DTT
1x TAE	40 mM Tris acetate, 2 mM EDTA, pH 8.5
1x TBE	10.8 g Tris base, 5.5 g boric acid, 0.93 g EDTA add 1000 ml water
1x transfer buffer	3.4 g Tris-base, 14.4 g glycine, 200 ml methanol, add 1000 ml water
1x western running buffer	3.4 g Tris-base, 14.4 g glycine, 5 ml 20 % SDS, add 1000 ml water
4x Tris/SDS pH 6.8	18.6 g Tris-base, 6 ml 20 % SDS, pH 6.8 with HCl, add 300 ml water
4x Tris/SDS pH 8.8	91 g Tris base, 10 ml 20 % SDS, pH 8.8 with HCl, add 500 ml water
6x Loading buffer for agarose gel electrophoresis	0.25 % (w/v) bromphenol blue, 0.26 % xylene cyanol (w/v), 30 % glycerol (v/v), in water
20x SSC	3.0 M NaCl, 0.3 M sodium citrate, 1 mM EDTA

2.1.6 Cell culture media and reagents

Dulbecco's modified Eagle's Medium (DMEM)	PAA
RPMI 1640, with HEPES	PAA
RPMI 1640 with HEPES, sodium carbonate, glucose	Gibco-Invitrogen
Fetal Calf Serum (FCS)	Biochrom
Penicillin/streptomycin, 100 x concentrate	PAA
Stable glutamine, 200mM concentrate	PAA
Trypsin EDTA (1:250), 1 x Concentrate	PAA

ß-mercaptoethanol 50mM	Gibco-Invitrogen
N,N-dimethylsulfoxide (DMSO)	Roth
Polybrene or Hexadimethrine bromide	Sigma-Aldrich
G418 sulphate	Invitrogen
Puromycin	Roth
Doxocyclin	Sigma
5-AZA-cytidine	Sigma

2.1.7 Enzymes and appending buffers

Restriction endonucleases 10.000-50.000 unit/mL	Fermentas
10x red, green, blue, orange, tango, EcorI, BamHI, SacI buffers	Fermentas
ApoI 10.000 unit/mL and buffer 3	NEB
BglII 50.000 unit/mL and buffer 3	NEB
MspI and buffer 3	NEB
Klenow fragment and 10x buffer	Fermentas
Taq polymerase and 10x KCl buffer	Fermentas
Pfx proof reading Taq polymerase and 10x buffer	Invitrogen
GoTaq Taq polymerase and 5x buffer	Promega
2x Mastermix SYBR Green	Applied Biosystems
2x Mastermix TaqMan	Applied Biosystem
T4 Ligase 500u/µL and 10x buffer	NEB
T4 Ligase high concentrated 2000u/µL	NEB
Retrotranscriptase	Fermentas
DNase	Invitrogen
RNase out	Fermentas
RNase I	Fermentas
X-Galatossidase	Roth

2.1.8 Kits

Invisorb Spin DNA extraction kit	Invitek
Invisorb Spin Plasmid mini two	Invitek
Purelink HiPure plasmid Maxiprep kit	Invitrogen
Purelink HiPure plasmid Miniprep kit	Invitrogen
MSB Spin Rapace kit	Invitek
QIAquick Nncleotide Removal Kit	Qiagen
ECL Western Blot	Invitrogen
RNease Micro kit	Qiagen
Calcium Phosphate Transfection kit	Invitrogen
TurboFect Transfection Reagent	Fermentas

Amaxa Cell Line Nucleofector Kit L	Lonza
Dual-Luciferase Reporter Assay System	Promega
LowCell ChIP kit	Diagenode

2.1.9 Antibodies

ANTIBODY AGAINST	APPLICATION	SUPPLIER
acetyl histone H3(H3K9Ac)	ChIP	Santa Cruz Biotechnology
trimethyl histon H3 (H3K27me3)	ChIP	Santa Cruz Biotechnology
SatB1	ChIP/western/EMSA	Santa Cruz Biotechnology
SatB1	ChIP (mouse)	Kohwi-Shigematsu lab
CTCF (human, mouse)	ChIP/western	Millipore/Upstate
Rabbit IgG	ChIP/EMSA	Millipore/Upstate
Goat IgG	ChIP/EMSA	Sigma
ß-TUBULIN	western	Sigma
ß-actin	western	Sigma
anti-mouse IgG-HRP	western (II)	Santa Cruz Biotechnology
anti-rabbit IgG-HRP	western (II)	Santa Cruz Biotechnology
anti-goat IgG-HRP	western (II)	Santa Cruz Biotechnology

2.1.10 Micro Beads and FACS antibodies

mouse CD19 microbeads	Miltenyi
mouse CD117 microbeads	Miltenyi
human CD3 microbeads	Miltenyi
human CD19 microbeads	Miltenyi
human CD34 microbeads	Miltenyi
mouse CD19-FITC	Miltenyi
mouse CD117-APC	Miltenyi
human CD3-FITC	Miltenyi
human CD19-PE	Miltenyi
human CD34-FITC	Miltenyi
CD117-APC	BD
CD11b / Mac1-APC	Biologend
Sca1-PE	Biolegend

2.1.11 Cell lines and mouse strains

CELL / MOUSE	REF: / ATCC NR
B6 wild type mice	Taconic

Transgenic human PU.1 PU.1$^{-/-}$ background	(Leddin et al., 2011)
HEK-293T	CRL-11268
HeLa	CCL-2
PlatE	(Morita et al., 2000)
EL-4	TIB-39
416B	
Jurkat	TIB-152
HL-60	CCL-240
U937	CRL-1593.2
Inducible Jurkat cells	(Ngo et al., 2006)

2.1.12 Oligonucleotides

Oligos were purchased by Metabion, MWG, Invitrogen or Biotez. Taqman probes were provided by ABI.
In the following tables, m means oligo specific for mouse genome, h for human; F stays for forward and R for reverse. In the reference column is indicate the origin of the primers, either from a previous published work or from the internal database of the Rosenbauer research team (the OR nnumbers)

GENOTYPING PCRs	**SEQUENCE 5'-3'**	**REFERENCE**
m PU.1 KO 1	ctt cac tgc cca ttc att ggc tca tca	(Scott et al., 1994)
m PU.1 KO 2	gct ggg gac aag gtt tga taa ggg aa	(Scott et al., 1994)
m PU.1 KO 3	caa ccg gat cta gac tcg agg a	(Scott et al., 1994)
h BAC PCR1 F	ggc aat atg tca ggg agg tg	(Leddin et al., 2011)
h BAC PCR1 R	atg act gga tgg gac tgg ag	(Leddin et al., 2011)
h BAC PCR2 F	cat ctg atc ccc tcc aga ga	(Leddin et al., 2011)
h BAC PCR2 R	cag caa agg ctt ttg aga cc	(Leddin et al., 2011)
GFP F	ctg acc tac ggc gtg cag tg	OR25
GFP R	gtt ctg ctg gta gtg gtc gg	OR24
PSK T7 F	taa tac gac tca cta tag g	OR460
PSK M13 rev	cac aca gga aac agc tat gac ca	OR461
m/h PU.1 promoter F	atc agg aac ttg tgc tgg ccc tgc	OR41

CLONING FOR LUCIFERASE	**SEQUENCE 5'-3'**	**REFERENCE**
m -25kb A F BamHI	cgg gat ccc gac caa ggc agg ccc tcg c	OR684
m -25kb A R HindIII	ccc aag ctt ggg caa gtc att tgt gtg gga cca	OR642
m -25kb B F BamHI	cgg gat ccc gat gtg ccc tcc tac atc cag t	OR686
m -25kb B R HindIII	ccc aag ctt ggg cat gtg gac cta cca tgc ct	OR644
m +71Kb A F BamHI	cgg gat ccc gat gaa aat gag gat gtc aca	OR690
m +71Kb A F HindIII	ccc aag ctt ggg aga tgc cag tga agt gga gat	OR691

m +71Kb B F HindIII	ccc aag ctt gag aga cca aga ggt aca gac c	OR692
m +71Kb B R HindIII	ccc aag ctt ggg cct tcc agc gga ttc aca aca	OR693
m +71Kb CL F HindIII	ccc aag ctt ggc ttg gag atg cct gct gtg gc	OR694
m +71Kb CL R HindIII	ccc aag ctt ggg tta aat tta cta agg atg acc	OR695
m +71Kb C F BstBI	gcg ttc gaa ctt gga gat gcc tgc tgt ggc	OR845
m +71Kb C R NheI	cta gct agc tag tta aat tta cta agg atg acc	OR846
m +71Kb CS F HindIII	ccc aag ctt ggg caa gcc aac agc tgc agt aa	OR843
m +71Kb CS R HindIII	ccc aag ctt ggg tga ggg agt tgg ttc tct cc	OR844
m +71 CTCF del F	gat atc cta agt gct cac gta cag agt g	OR1291
m +71 CTCF del R	gat atc ttt agg atg tgg gcc taa act g	OR1292

EMSA	**SEQUENCE 5'-3'**	**REFERENCE**
m SatB1 +71Kb BS F	agt cca aaa taa taa aat aat aaa taa ata aat	OR1271
m SatB1 +71Kb BS R	agt cat tta ttt att tat tat ttt att att ttg	OR1272
m CTCF +71Kb BS F	agt cgc cag cct tca gca ggt ggc act gtt ggg	OR1277
m CTCF +71Kb BS R	agt ccc caa cag tgc cac ctg ctg aag gct ggc	OR1278
m SatB1 pos BS (MAR) F	agt cga cct att agt aat aac gac cta tta gta ata ac	OR1140
m SatB1 pos BS (MAR) R	agt cgt tat tac taa tag gtc gtt att act aat agg tc	OR1142

EXPRESSION	**SEQUENCE 5'-3'**	**REFERENCE**
m PU.1 F	aga agc tga tgg ctt gga gc	OR1295
m PU.1 R	gcg aat ctt ttt ctt gct gcc	OR1296
Taqman mPU.1	FAM -tgg gcc agg tct tct gca cgg - TAMRA	ABI
m ßactin endogenous control	4352341E	ABI
m Actin F SYBR	tga cat c cg taa aga cct cta	OR926
m Actin R SYBR	cag gag gag caa tga tct tga	OR927
m gapdh F	aag ggc tca tga cca cag tc	OR141
m gapdh R	cac att ggg ggt agg aac ac	OR142
m CTCF F	caa ttg cac ctg tat tct gat c	OR1327
m CTCF R	gag gag gag gag gag cct gc	OR1328
m SatB1 F	ctt tgg agc agc aag ttt cc	OR1451
M SatB1 R	ctc tca gtg gca agg gta gc	OR1452
m/h MADD F	tgg gta gca ctt ctg cat tg	OR1864
m/h MADD R	cct gcc agc ttc ttc ctc ta	OR1862
h PU.1 F	gga gag cca tag cga cca tta c	OR727
h PU.1 R	cgg cga agc tct cga act c	OR728
Taqman hPU.1	NED - ctt cca ccc cca cca cgt gca - MGB	ABI
h ßactin endogenous control	4326315E	ABI
h gapdh F	atg ctg gcg ctg agt ac	OR721

h gapdh R	tga gtc ctt cca cga tac	OR722
h CTCF F	ttg gca aaa aga ccc aga cta t	OR1556
h CTCF R	gac acg tgt aac tgc aaa gct c	OR1557
h SatB1 F	att gcc act gaa agg aat gg	OR1453
h SatB1 R	ttc gga tca tgg aga ggt tc	OR1454

ACT and 3C	**SEQUENCE 5'-3'**	**REFERENCE**
ACT linker MspI	gct gac cct gaa ttc gca cgt gcc tgt cgt tag cgg aca cag ggc gat tca c	OR1549, (Ling et al., 2006)
ACT linker short	cgg tga atc	OR1550, (Ling et al., 2006)
ACT primer ext	gct gac cct gaa ttc gca cgt gcc t	OR1551, (Ling et al., 2006)
ACT primer int	gtc gtt agc gga cac agg gcg att c	OR1552, (Ling et al., 2006)
3C PU.1 pm down F outside	aac agc tat tta tag atg ggt tca gtg	OR1419
3C PU.1 pm down F inside	ggc agg gtt ctg ggt tga aga t	OR1420
3C PU.1+481 mm	att ctg ggt tga aga tgg cga gg	OR264
PU.1 1st intron	gtt cag tgg cag gcc gaa ctc tag	OR671
mmTaq_PUP5_R_undig	tgt tgc ctc cag aga ctc ct	OR1775
PU1mm Prom.out	ata ggg gga gaa tgg tct ggg atg tg	OR373
PU1mm Prom. In	gaa cca ggc ccc aac tgt tat ttt tg	OR374
m GAPDH2	aca cag gca aaa tac caa tg	OR1661, (Spilianakis et al., 2005)
m GAPDH3	ctg cgc ctc aga atc ctg	OR1662, (Spilianakis et al., 2005)
m GAPDH5	caa aac tcc tgg gtg caa g	OR1664, (Spilianakis et al., 2005)
m URE up 1 3C locus	tcc tgg aac tca ctc tga aga c	OR1657
3C PU.1-14134 mm	tca ggg tgg gca aag tgt tat ctg	OR267
m –11.8kb-F	ctc tgc ccg ctc tta acc tt	OR1731
3C PU.1-8286 mm	ttg cct gct acc agg gag gtt g	OR266
PU.1 mm +22kb	cag gct cct caa aag gtc aag t	OR802
PU.1 mm +37kb	ctc ctg acc ttg tgt ctg gga	OR803
PU.1 mm +50 kb	cca aag act tgt tgg agt cat agt	OR806
PU.1 mm +71kb	gct atg atc caa ggt ctt agg	OR807
m 71C ApoI 5` out	aat ggc cca gct aag ggt ccc tca act	OR1106
m 71 in MspI R	cta gag ttc ggc ctg cca ct	OR1585
m PUP 3C F2	tgc aga cac act tgg gac tc	OR1713
m PUP 3C F3	agg ccc ctc tct gag gtt ag	OR1714
h CC2	cct gcc ttc aga gag ttc ctg atc	OR1
h CC4	cct atg aca cgg atc tat acc aac	OR2

h CC6	ccc aca cat taa aga gaa gga tgc	OR4
PU.1P H F ApoI	tcc tgg ggt ctg gag aga gaa a	OR1314
PU.1P H RR ApoI	tgt tgt tgt ttt tct ggc cca gtg a	OR1311
+94 H 5`ApoI out	agc cct aat cca gtc att tgg gtg	OR1170
h 3C Apo1 +94kb R	ggg aca gat tag tca tta tgc ctg g	OR1318
h 3C-TUBLa	cac tac acc att ggc aag ga	OR1773
h 3C-TUBLb	cat atg tgg cca gag gga ag	OR1774
h tub4	ctc aaa gca agc att ggt ga	OR1889
hsTaq_PUP5'	FAM – tct ctg gtg agt ccc ctc tgc tg – TAMRA	ABI

ChIP	**SEQUENCE 5'-3'**	**REFERENCE**
m/h PU.1 promoter F	atc agg aac ttg tgc tgg ccc tgc	OR41
m/h PU.1 promoter R	cca ttt tgc acg cct gta aca tcc	OR42
m PU.1 promoter endo F	gat cca cgc tct cgc tat tc	OR111
m PU.1 promoter endo R	gcg cta cag gaa gtc tct gg	OR112
m URE F	ttt ctc ctg ctt cga ctt cc	OR91
m URE R	gag gcc tgt gtt cct tca ac	OR92
m +71 F	caa gcc aac agc tgc agt aa	OR939
m +71 R	agc acc gtt gtc cac agt	OR940
m +37 F (neg ctrl)	gcc ttc tag gta cta aca ggc t	OR804
m +37 R (neg ctrl)	tct ctg tat tgc cat cat ctt g	OR1788
m CTCF F (pos ctrl)	ccc tcc tca ttc agg aaa ca	OR1681, (Weth and Renkawitz, 2011)
m CTCF R (pos ctrl)	ggc ggg agc aca tag act aa	OR1682, (Weth and Renkawitz, 2011)
h PU.1 promoter F	gga agg aag aaa gca gca cta	OR1316
h PU.1 promoter R	ctc tga tct agc agg cta tca	OR1317
h URE F	cta ggc ctg aag aga gat ctg g	OR1736
h URE R	gga cag caa gga aaa gag aag a	OR1737
h +94 F	gcc cat gaa gag gaa gat gag aa	OR1320
h +94 R	cag aca gtg aga act ctc agg a	OR1321
h +67 F (neg ctrl)	ggg gtc tat gtt ccc cag ac	OR1781
h +67 R (neg ctrl)	gca ccc agc caa cag tat tt	OR1782
h CTCF F (pos ctrl)	cct gac cta tat ctg gca gga c	OR1685, (Weth and Renkawitz, 2011)
h CTCF R (pos ctrl)	gca ccc acc ttc aat caa aa	OR1686, (Weth and Renkawitz, 2011)

2.1.13 Vectors

pBluescript 2SK (+)	Stratagene
enhancer blocking luciferase plasmid	(Ishihara et al., 2006)
pcDNA3.1-GFP	Promega
pPAC4 carrying the murine PU.1 locus	(Frengen et al., 2000)
BAC, carrying the human PU.1 locus	RP11-125
sh-retrovirus	(Ngo et al., 2006)

2.1.14 Software

BLAST (NCBI)
BLAST (Genomatics)
Transcription factors binding site prediction (Genomatics)
Primer 3.0 free online, primer design program
7300 System SDS software (Applied Biosystems) real time PCR analyzing software
CFX96 RealTime System BioRad
Vector NTI (Invitrogen)
Finch TV (Geospiza)
Ensemble (genome database)
UCSC browser (genome database and mapping of genome wide databases)

2.2 Methods

2.2.1 Mice

Daily animal care, breeding and offspring separation was carried out in collaboration with the animal core facility of the Max-Delbrueck-Center for Molecular Medicine, Berlin, Germany. All mice were housed and bred in specific pathogen-free animal facilities. All animals experiments were approved by the local authorities according to the German Federal Animal Protection Act.

2.2.1.1 Description of used mouse strains

The transgenic hsPU.1 strain was generated with the help of the transgenic facility of the Max-Delbrueck-Center for Molecular Medicine, Berlin, Germany. Briefly, the BAC vector containing the human PU.1 locus and adjacent regions (total 160Kb, as illustrated in fig.) was microinjected into fertilized oocytes of FVB/N mice and implanted *in uteri* of pseudopregnant FVB/N. 4 weeks after birth founder mice were identified by PCR and Southern blotting. Two

different lines (#1 and #55) were crossed with the *PU.1* $^{-/-}$ strain (Klemsz et al., 1990) to generate *hsPU.1*$^{+}$ */ PU.1* $^{-/-}$ mice. These mice were used for organ collection.

2.2.1.2 Genotyping

Tagging of the offspring was carried out according to standard protocols. Mice were genotyped by locus specific polymerase chain reaction (PCR) on genomic DNA extracted from tail issue (see 2.2.3.1).

2.2.1.3 Isolation of mouse organs material

After euthanizing the mouse with CO_2, the animal was rinsed in ethanol and pinned down on a dissecting board with the belly facing up. Mice were opened and the legs and arms were dissected with scissors and forceps to isolate bone marrow (BM) cells. Spleen and thymus were also collected. All organs were kept on ice in cold PBS until preparation. Single cell suspensions were generated by cut the organs in small pieces and subsequently filter them through a cell strainer. BM was isolated by flushing sterile the bones with PBS.

Single suspensions were treated with ACK buffer for 5 minutes to destroy red blood cells. The reaction was interrupted by add PBS.

CD19 positive cells were extracted from the single suspension cell of the spleen by CD19 microbeads and subsequent purification by MACS columns, according to Miltenyi manufacture's instruction.

Mouse hematopoietic stem cells were extracted from the single cell suspension of the bone marrow by sorting for CD117 (c-kit) positive, Sca-1 positive and Lineage cocktail negative cells.

Total bone marrow cells were seeded on not coated plates with conditioned medium to promote macrophage differentiation (DMEM supplemented with 20% FCS and 10mg/mL M-CSF). Primary mature cells will die, whereas stem cells and progenitors will live as suspension cells. The only cells, which can adhere under these conditions, are *in vitro* differentiated macrophages. After 10 days suspension cells were removed, and only adherent cells were collected.

For murine embryonic fibroblast (MEF), embryos were collected at 13.5 dpc from a C57/bl pregnant female. With curved forceps "red" tissue were removed from the body cavity (heart, most of blood vessels and fetal liver). Then every embryo was minced with razor blaze and transfered to 15mL conical tube and incubated o/n in trypsin 1X at 4°C. The next day trypsin embryo tissue was spin down by centrifuging, and the pellet was resuspended in complete

DMEM medium, pass through a cell strainer to remove not digested tissue fragments and seeded in 15cm cell culture tissues. MEF grow very fast and need to be split every 2-3 day to keep maximum confluence at 70-80%.

2.2.2 Cell Culture and primary human cells

2.2.2.1 Cell lines

All indicated cells were grown in indicated media supplemented with fetal calf serum (FCS), glutamine and antibiotics (penicillin/streptomycin; ciprofloxacin, gentamycin and fungizone when needed) in an incubator at 37°C with 5% CO_2.

Human Embryonic Kydney cell lines (HEK293T), Plat_E and HeLa cells are adherent cells, were cultured in dishes in supplemented DMEM (10% FCS) and split every 2-3 days.

EL-4 T-lymphoblastoma cells is a suspension cell line, was cultured in flask in RPMI supplemented with 10% FCS, glutamine, 4,5g/L of glucose and 50µM of ß-mercaptoethanol to prevent differentiation.

Human lymphoblastom Jurkat-eco cells are Jurkat cells stable transfected with an ecotropic receptor, which allows infecting them with murine viruses. They were cultured in RPMI supplemented with 10%FCS and selection antibiotics (150ng/mL G418 for selection of the ecotropic receptor; 10µg/mL puromycin for selection of the plasmid containing the sh sequence; 20µg/mL of doxocyclin for induction of the inducible promoter)

416B cells are a murine myeloid progenitor cell line (Dexter et al., 1979); it was cultured in supplemented DMEM medium (20% FCS). HL60 and U937 are human acute myeloid leukemia cell lines and were cultured in supplemented RPMI medium (10% FCS).

All suspension cells were counted every 2-3 days and splitted to be maintained between 1E4 and 1E6 cells per mL.

2.2.2.2 Thawing, cultivating and freezing of cells

Cells were thawed quickly in a 37°C water bath and transferred drop-wise to a tube with 10mL medium After centrifugation (standard condition: 1200 rpm, 5 minutes, room temperature) supernatant was discarded and pellet was gently resuspended in 1 mL medium. Cells were plated in appropriate cell number on culture dishes/flasks.

For passaging of adherent cells, medium was aspired, cells were washed once with PBS and incubated with 0,05% trypsin-EDTA for 5 minutes at 37°C (2mL trypsin per 10cm dish). Once detached, cells were harvested by adding 10mL of medium and centrifuged; supernatant

was discarded and pellet was resuspended in medium. Cell suspension was transferred in different dilutions to new tissue culture dishes. For suspension cells, they were centrifuged and resuspended in fresh medium.

For cryo-preservation, cells were transferred in a Falcon tube, medium was removed and pellet resuspended in 1mL of freezing medium (90% FCS, 10% DMSO) and transferred to a cryo tube. Cryo tubes were stored in polystyrene boxes at -80°C for at least 24 hours to allow a gradual freezing process and to avoid crystal formation, after which they were placed in liquid nitrogen or at -80°C for short-term storage.

2.2.2.3 Transient and stable transfection of cells

HEK293T cells were transfected for transient assay according to the Turbofect *in vitro* transfection kit's instructions (Fermentas). Briefly, 5E4 per well cells were seeded on 24-wells plate. After approximately 24 hours 70-80% confluent wells were co-transfected with 0,4µg of plasmid.

EL-4 cells were transfected by electroporation according to Amaxa Cell Line Nucleofector Kit L (Lonza) manufacture. Cells were counted, 2E6 cells were centrifuged at 900 rpm for 10 minutes at room temperature and the pellet was resuspended in 100µL of Cell Line Nuclefector Solution L. The cell suspension was gently mixed with 2µg of DNA (1,5µg of linearized construct to be tested and 0,5µg of pcDNA3.1-GFP plasmid, carrying neomycin resistance cassette) transferred into certified cuvettes (provided by the kit) and electroporated with Nucleofector Program C-009 (C-09 for Nucleofector I Device). 500µL of pre-warmed medium was added to the cuvettes and the suspension was transferred into the prepared 12-well plate, containing 1mL of pre-warmed medium (final volume 1,5mL medium). To generate stable cell lines, cells were transferred to a flask in 10mL medium and selected for G418 sulfate resistant colonies after 48h of electroporation. Selection was performed 10 days long, after that viable and GFP positive cells were sorted by FACS.

2.2.2.4 Virus production and transduction of cells

Viral supernatants were used to stable transfect Jurkat-eco cells. Murine retroviral supernatants were produced by co-transfection of Plat-E cells with the retroviral construct (10µg), gag-pol and ecotropic env constructs as support for viral particle assemble (10 and 2 µg respectively). Transfection was performed with Calcium Phosphate Transfection Kit according to manifacture. Viral supernatant were collected at 72 and 96 h after transfection, filtered through 0,45µm filters and either used directly or frozen and stored at -80 °C.

For retroviral transduction of Jurkat-Eco cells, 1E6 cells were seeded in 24-well plate in 500µL of RMPI medium supplemented with 20% FCS and mixed with 500 µL of filtered virus supernatant in the presence of 8µg/mL polybrene and cultured for 24 hours. The next day fresh medium was added and cells were expanded.

2.2.2.5 Patients cells and healthy donors

AML (acute myeloid leukemia) bone marrow aspirates were collected according to ethic rules. All the patients presented blast infiltration of more than 80% of bone marrow total cells. Cells were treated with ACK for erythrolysis; this step was repeated 1 to 3 times, until when the cell pellet was white. Then the sample was depleted from lymphocytes by first CD3 (T cells) and second CD19 (B cells) microbeads according to Miltenyi manufacture's instruction: briefly, cells were stained with proper microbeads, then the flow though was collected for further purification step, whereas the eluate (CD3 or CD19 positive cells, respectively) was used for testing the purity of lymphocyte depletion. The blasts content afterwards was higher than 95%.

Hematopoietic stem cells were purchased from Lonza manufacture as adult bone marrow CD34 purified cells.

Primary monocytes were collected from blood of healthy donors. The cells were processed as the AML: first erythrolysed by ACK treatment and then lymphocyte-depleted by magnetic beads separation. The final content of cell was mostly mature myeloid cells.

2.2.3 Molecular Biology

2.2.3.1 Preparation of genomic DNA: Phenol/Chloroform extraction

DNA was extracted following standard protocols. Briefly, cells or tissue was digested with tail digestion buffer supplemented freshly with 10µg/mL proteinase K at 56°C o/n. Genomic DNA was purified by adding 1 volume of Phenol/Chloroform/Isolamylalcohol mixture (in 25:24:1 ratio), mixed by vortexing or inverting the tubes multiple times, and centrifuged at high speed (13000g, 5 minutes, 4°C). The supernatant was then collected and subsequently mixed with 1 volume of Chloroform/Isomamylalcohol mixture (24:1) and centrifuged as above. The supernatant which contains nucleic acids (mostly DNA) was transferred to a new tube and DNA was precipitated by adding 2,5 volumes of ethanol 100%, supplemented with 0,1 volume of NaOAc 3M pH5,2 and mixed thoroughly. After 20 minutes incubation at -20°C followed by high speed centrifugation (13000 g, 20minutes, 4°C) supernatant was discard and

pellet DNA was washed with 1mL 70% cold ethanol. After centrifugation (13000 g, 5 minutes, 4°C), DNA pellet was air-dried until complete evaporation of ethanol. DNA was dissolved in water or TE buffer.

2.2.3.2 RNA extraction

Depending on cells numbers 2 different methods were used: for small amount of cells (less than 5E4) RNA was extracted using RNeasy Micro Kit according to manufacture. For bigger amount of cells triazol method was performed as followed.

First, cell pellet was resuspended in 500µL Triazol by passing the solution through a 16 gauge needles several times. 100µL pure Chloroform was added to the tube and the solution was vortexed for 30 seconds and incubate 10 minutes at RT (room temperature). After centrifugation (13000 g, 5 minutes, 4°C) two phases can be separated: the lower phase is organic, meaning contains proteins and lipids, the interphase contains DNA (which can be separated from RNA based on the pH), and the upper phase is inorganic and contains RNA. So the upper phase was carefully transferred to a new tube and precipitated with 1mL Ethanol 100% supplemented by 50µL of NaOAc 3M pH 5,2 (for less than 2E6 cells as starting material, also 4g of glycogen were added to visualize the pellet afterwards). After centrifugation (13000rpm, 20 minutes, 4°C) the supernatant was discarded and the pellet wash with 1mL of 70% cold Ethanol. After centrifugation (13000 rpm, 10 minutes, 4°C) supernatant was discarded and pellet was air-dried until complete evaporation of Ethanol. The pellet was resuspended in DEPC water and incubated 10 minutes at 55°C for better resuspension. RNA was stored at -80°C. All buffers, solutions and tips were RNase free.

2.2.3.3 cDNA synthesis

cDNA synthesis was performed according to Fermentas Revertaid First strand cDNA synthesis kit instructions, after DNase treatment (Invitrogen). Briefly, maximum of 1 µg of RNA was treated with DNase in the presence of RNase inhibitors for 15 minutes at room temperature. The reaction was stopped by vortexing, adding 1.25mM of EDTA (final concentration) and 10 minutes incubation at 65°C.

Random primers provided by Fermentas BLABLA kit were annealed by 5 minutes incubation at 65°C, afterwards quickly chilled on ice. Then a mastermix containing proper buffer, 1mM dNTP, revert transcriptase and RNase inhibitors was added, and the cDNA synthesis reaction occurred at 42°C 1h long, followed by enzymes heat inactivation (10 minutes at 72°C).

Dilution and amount of cDNA used per reaction was optimized for every gene and every cell context.

2.2.3.4 Polymerase chain reaction and Real Time PCR

PCR was carried out with standard procedure. For all reactions PCRs 1,5mM $MgCl_2$, 1x KCl buffer, 400nM forward and reverse primers, 400nM dNTPs and 2.5 units Taq polymerase were used (25µL final volume). For semi-quantitative PCR the reaction was carried out in 50µL final volume and 12µL were taken 3-4 times; 10µL for each cycle were loaded on an agarose gel. For nested PCR 1µL of the first round was used as template for the second round. Oligonucleotides for SYBR green Real Time RT PCR were design spanning different exons to avoid detecting DNA contamination in the cDNA sample. Amplified PCR product were separated and visualized on 1-2% agarose gels containing 1µg/mL ethidiumbromide.

For Real Time PCR 2x SYBR Green or 2x Taqman mix (Applied Biosystem) was used according to instructions. 7300 System SDS Software or CFX96 Real-Time System were used to analyze results.

Protocols:

All endpoint PCR: 94°C 10'', 55-59 20'', 72C 15''/500bp

All SYBR green: 94°C 10'', 60°C 30'' (minimum time for fluorescence reading), 72°C 10-15''.

Taqman probe for expression (mm and hs actin, mm and hs PU.1): as SYBR green

Taqman reactions for q3C (hs): 94°C 10'', 60°C 30'', 64°C 10'', 72°C 10''

Sybr green at Biorad machine when II temperature read was needed: 94°C 10'', 60°C 15'', 72°C 10'', II read (variable temperature according to the melting curve in order to distinguish SYBR grren signal of amplified product versus primer dimers or background signals which are usually lower) 2''.

2.2.3.5 Agarose gel electrophoresis

Agarose was dissolved in 1x TAE at the desired concentration and boiled in a microwave until the solution was completely dissolved. Ethidiumbromide was added in the appropriate concentration when the temperature of the agarose gel was around 50°C. After casting the gel, hardening occurred within 10-30 minutes. Gel chamber was filled with 1x TAE and aliquots of samples supplemented with 6x loading buffer were loaded into the wells. To determine length of the fragments, 1Kb or 100bp ladder was used. The gel was run at 80-130V

depending on further applications. DNA was visualized by UV-light. Picture of the agarose gel was taken by BioDoc Analyzer and printed or saved as file.

2.2.3.6 Cloning

All used vector are listed in 2.1.13

As general process, PCR-amplified products were purified by MSB Spin Rapace Kit and blunt by Klenow fragment for 1h at 37°C; blunt-ends DNA was then ligated into appropriate plasmids. DNA fragments deriving from plasmids were digested with specific restriction enzymes, gel-extracted according to Invisorb Spin DNA extraction Kit and blunt if necessary. Ligation was performed with T4 Ligase (NEB) using 50-100ng of the vector and 10fold of the insert for 2 h at RT. The ligation reaction was transformed into DH10b chemically competent bacteria according to heat shock standard procedure: briefly DNA was incubated with bacteria on ice for 20 minutes, placed in a 42°C 90 seconds long, then on ice for 2 minutes. Then, 1mL of pre-warmed LB medium was added to the bacteria, placed them at 37°C for 1h and then plated on LB-agar plates supplemented with appropriate antibiotic for selecting resistant colonies. Plates were incubated o/n at 37°C. The following day correct cloned vectors were screen by directional colony-PCR or directly by extracting the DNA according to Invisorb Plasmid Mini Two Kit.

For regular subcloning of PCR products, PBluescript was EcorV digested.

The PU.1 promoter sequence was PCR-amplified and inserted as XhoI-HindIII fragment into the PBluescript vector for the experiment of chapter 3.1.4.

For luciferase enhancing blocking assay vector was used as backbone plasmid for further modification: briefly, all the DNA fragment to be tested were PCR-amplified with primers containing BamHI and HindIII overhangs respectively. The PCR products were cloned between the luciferase-encoding gene and the SV40 enhancers. The mutant fragments were inserted by triple ligation using BamHI-SacI-HindIII endonuclease.

To test different short hairpin against CTCF a retroviral vector was used: 12 oligos as putative shRNA against CTCF mRNA were cloned into the retrovirus, and only the colonies resistant to ampicillin but not anymore to kanamycin carried the desired oligo inserted. Oligo direction was validated by sequencing.

2.2.3.7 Luciferase reporter assay

Luciferase constructs were used as reporter assay to test enhancing blocking function of desired DNA fragments. For this assay the DNA fragment to be tested is inserted between an

enhancer and a promoter, which drives the luciferase gene. Comparison between the "empty" vector (where no putative insulator is inserted) and the "insulator" vector provides a functional prove of the DNA tested. The plasmid used as control and as backbone for further cloning has already been described (see 2.1.13)

Briefly, HEK293T cells were co-transfected (see 2.2.2.3) with 400ng of the luciferase construct and 10ng of Renilla plasmid (provided along with the kit). The next day medium was removed and cells were washed with PBS and then lysed according to the Dual Luciferase Reporter Assay System. 20% of the lysate was then analyzed with the luminometer. The firefly luciferase and the Renilla luciferase catalyze two different reactions. The light produced by the different enzyme can be red at two different wavelengths (560nm and 480nm respectively). The Renilla plasmid is used as internal control for transfection efficiency, so that the ratio between luciferase and Renilla activity guarantees for comparing same amount of transfected cells.

2.2.3.8 Total and nuclear protein extract

For Western Blot analysis was sufficient to extract total proteins (cytoplasmatic and nuclear, although transcriptional regulation occurs only in the nucleus). Proper number of cells (if possible 5E6) was spin at standard condition, washed once with PBS and centrifuged again. The cell pellet was resuspended with 40-400µ (depending on cell numbers: usually 2 volumes of cell pellet were used) RIPA lysis buffer. Lysis was performed by shaking 30 minutes at 4°C; for bigger number of cells was necessary to sonicate 3 minutes (medium power, 30 second on / off) to ensure all membranes to break.

The lysate was centrifuged (13000 g, 5 minutes, 4°C), the supernatant containing all proteins was collected and diluted in 1:2 ratio with 2x Laemmli buffer. Proteins were boiled 5 minutes and then analyzed by Western Blot or stored at -20°C. Laemmli buffer is highly reducing and breaks disulphite bounds; by boiling complete denaturation of protein was achieved.

Nuclear extracts were made by harvesting 1.0-3.0*E07 cells. The cells were washed twice in ice cold PBS (14000 rpm, 2 min, 4°C) and incubated in 400µl ice-cold buffer A (15 min, 4°C, shaking) in the presence of protease inhibitors (Sigma). 25µl 10 % NP 40 were added and mixed thoroughly. After centrifugation (14000rpm, 30sec, 4°C) the pellet was resuspended in 50-100µl buffer C and incubated (15min, 4°C, shaking). After centrifugation the supernatant was flash-frozen in liquid nitrogen and stored at -80°C.

Whole cell lysates were done employing the TCA (1,1,1-trichloroethane) method. Briefly, cells were washed twice in ice cold PBS (14000rpm, 2min, 4°C), resuspended in 500µl PBS and mixed with 10% (v/v) TCA (15-30 min, 4°C). The lysate was centrifuged (14000rpm, 2min, 4°C) and resuspended in 1 x SDS sample buffer and stored at -20°C.

2.2.3.9 Western Blot Analysis

To analysis specific protein level western blot technique was used. First of all, a polyacrilammide gel was made mixing liquid polyacrilammide with *polimerazing* agents (APS and TEMED) to obtain desired % of gel (6% for CTCF, 12% for tubulin or ßactin and PU.1). A Western Blot gel consists in 2 phases: the stacking gel, which contain glycin which align all the protein at the same line and the separating gel, well proteins start to run specifically according to their size.

Samples were loaded in the pockets and the gel was run for 1-2 h at 90V for the stacking gel and 120V when proteins reach the separating phase. Then the chamber was disassembled and the gel was capillary-blotted onto a previously methanol-activated PVDF membrane in the transfer buffer. Protein transfer was controlled staining with Ponceau solution, which binds to proteins; the solution can be easily removed washing with PBS the membrane several times. To prevent unspecific binding of the antibody to the membrane, it was incubate with 4% of milk (dissolved in PBS) and 0,05% of Tween 20 for 1h at RT. Milk has a highly proteic concentration, which will cover by ionic interaction the surface of the membrane still protein-free (everything apart from the transferred lanes). Then the blocked membrane was incubated with the desired antibody (here, polyclonal antibody against CTCF and ß-actin) o/n at 4°C. The next day the membrane was washed 3 times with PBS and then incubate 2 hours with a HPR conjugated secondary antibody, which recognizes the constant part of the first antibody according to the species were this was raised (here, Ab α CTCF was raised in rabbit, α ß-actin in mouse; therefore the second antibody was (goat)-anti-rabbit and (rabbit)-anti-mouse respectively; all of them were conjugated with HPR, horseradish peroxidase; see 2.1.9). Afterward, the membrane was washed again with PBS to remove all free antibody in the solution and HPR activity was measured by ECL method, according to Invitrogen kit instructions, and developed by X-ray film.

2.2.3.10 Electrophoresis mobility shift assay (EMSA)

By gel shift or EMSA assay is possible to investigate *in vitro* DNA-protein interaction. Nuclear protein Extracts (NE) were mixed with oligonucleotides containing putative binding

site for a specific protein, loaded on a polyacrilammide gel and separated by electrophoresis. Due to size and charge, free DNA runs fast, whereas a protein/DNA complex is slower which produces the so-called "gel shift". Moreover by adding an antibody against a presumably protein only, the mobility is further decreased, producing the commonly known "super-shift". By radioactive labelling of DNA oligos is possible to detect different bands by autoradiography of the gel. Moreover, by competing the radioactive (hot) oligo with an excess of a not radioactive cold mutant oligo (where the putative binding sites was deleted) is possible to prove which base pairs are involved in the binding.

First of all, a not denaturating polyacrilammide gel was assembled and polymerized o/n. Olgonucleotides were annealed producing ATGC overhangs (10 minutes at 94, 75, 50, 25, 4°C) followed by labelling with Klenow fragment using α-32P-dCTP and 800nM of each dATP, dGTP, dTTP. The solution was purified by Qiagen kit to remove all free radioactive nucleotides and radioactivity's yield was measured with a scintillator counter. 30000cpm (counts per minutes) were used per reaction. Equal aliquots of nuclear protein extracts (NE) were tested for binding different combinations of hot and not cold probes and with or without specific antibody as a.m for 30 minutes. All reactions were performed in 2x shift buffer on ice-cold racks. Then were loaded without loading buffer (to prevent denaturation) and separated by electrophoresis. The gel ran 4h long, 230V, approximately 26mA. Afterwards the gel was dried at 80°C for 1h. Dried gel was placed in an Amersham Hypercassete, with an x-ray photographic fil on top. The film was developed in a Fuji RGII photographic developing system.

2.2.3.11 Chromosome conformation capture (3C)

This technique allows the detection of two interacting chromatin fragment *in vivo*. 1E7 cells were resuspended in 10mL of proper complete medium, fixed with 2% formaldehyde (540µL of 37% formaldehyde) and mixed by tumbling at 4°C for 10 minutes. Crosslinking reaction was stopped by adding 630µL of 2M glycine (final concentration 0,1375M) and mixed by tumbling at 4°C for 5 minutes. Cells were pellet, washed in 10mL of cold PBS, centrifuged at standard conditions and the pellet was resuspended in 5mL of cold lysis buffer with freshly added protease inhibitors. Lysis was performed on ice for 20 minutes, afterwards cells were centrifuged at 600g at 4°C for 5 minutes. During lysis only cell membranes were dissolved, whereas the nuclei remained intact: in the following steps a digestion in intact nuclei was performed, in order not to destroy chromatin-protein complexes. The pellet was washed in

500µL of 1.2x digestion buffer (in this case, 60µL of NEB buffer 3 and 440µL of water), spun again at 600g, 4°C, 5 minutes. Pellet was resuspended in 500µL 1.2x digestion buffer and 7.5 µL of SDS 20% were added. The solution was incubated in a thermomixer shaking at 900 rpm at 37°C for an exact hour to permeate the nuclear membrane; then 50µL of preheated 20% Triton X-100 were added to quench SDS activity and the solution was shaken for one more hour. By adding 800 units of restriction enzyme (in this work, BglII or ApoI enzyme were used) the digestion reaction was incubated at optimize temperature (37°C for BglII, 50°C for ApoI) shaking at 900rpm.

The next day digestion was stopped adding 40µL of SDS 20% and 20 minutes incubation at 65°C. Then 80µL of the digestion solution was diluted in 3,7mL of 1x ligation buffer supplemented with 187,5µL of 20% Triton X-100 and incubated for 1 hour at 37°C (the rest of digestion reaction was stored at -20°C). Digested material was highly diluted in order to prevent intermolecular ligation and let possible only intramolecular reactions (between two region which really interact *in vivo*) The reaction was then chilled at 16°C, and 1mL aliquot was taken as control of not ligated material. Then 16µL of high concentrated NEB ligase was added in the 3mL solution and ligation was performed o/n.

The following day 10µL/mL of proteinase K were added (also to the "no ligation" control aliquot) and the solution was incubated at least 4 hours @ 67°C shaking at 900rpm to decrosslink protein complexes from DNA and to denaturate proteins. Afterwards DNA was purified by phenol/chloroform extraction, and precipitated by ethanol. Pellet was resuspended in 50µL of TE buffer and 50ng were used for each PCR reaction. DNA material was stored at -20°C.

2.2.3.12 Associated chromosome trap (ACT)

Associated Chromosome Trap (ACT) is a modification of the 3C technique (see 2.2.3.13): by this method is also possible to detect inter- and intra-chromosomal interactions and co-localization. By ACT DNA is digested first with one restriction enzyme which cuts with an average of 3-10Kb in the genome: ligation of these products will produce DNA circles consisting of two interacting DNA sequence. By a second digestion with a restriction enzyme which cut the genome with a high frequency is possible to linearized the previous DNA circles and producing protruding ends, which can be used to anneal linker oligonucleotides. By designing a proper primer on the query DNA sequence (in this work, PU.1 promoter) and the other primer on the linker is possible to amplify unknown interacting DNA sequences.

PCR products were afterwards subcloned into the PBluescript vector and sequenced. Sequence results were BLASTed and mapped in the genome.

Briefly, after standard 3C assay, 1µg of DNA was further digested with 20unit of MspI for 6 hours and then the solution was by MSB Rapace Spin Kit and eluted in 50µL. 1µg of MspI-treated DNA with mixed with linker oligonucleotides (2µL of Long Linker 10µM and 8µL of Short Linker 10µM), 2µL of 10x ligase buffer (NEB). The solution was denatured at 50 °C for 1min, and it was then allowed to cool down gradually to 10°C in a 0,5°C/min gradient. One µL (400 units) of T4 DNA ligase (NEB) was then added, and the mixture was incubated at 15°C overnight. The ligated DNA was purified by MSB Rapace Kit and eluted in 50L of bidistilled water.

One µL of eluted DNA was used to perform the first PCR reaction using the PU.1 promoter specific primer and a linker specific primer. The PCR protocol was performed as following: 2 min at 94 °C, 35 cycles of 94°C for 20 sec, 65°C for 40 sec, 72°C for 1 min, then 72°C for 5 min. The first PCR product was purified by MSB Rapace Spin Kit and eluted in 50 µL of bidistilled water. 1µL was used for a second PCR reaction using nested primers with the same protocols a.m. PCR product were visualized by agarose gel, extracted by Invisorb Spin DNA extraction Kit, and cloned in pBluescript vector. White recombinant bacterial colonies were screened by colony PCR using universal primers flanking the multiple cloning site, then DNA was extracted by Invisorb Spin Mini Two Kit and subjected to sequence analysis.

Sequencing results were analyzed by Finch TV software and mapped to the murine genome by NCBI BLAST.

2.2.3.13 Chromatin immunoprecipitation (ChIP)

To investigate *in vivo* binding of proteins to DNA Chromatin ImmunoPrecipitation (ChIP) was performed. In this assay proteins are covalently bound to the DNA by formaldehyde and then immunoprecipitated by antibody recognizing the desired protein. Then the DNA-protein complexes are dissembled by heat and DNA is purified and analyzed by PCR.

ChIP assay was performed according to standard procedure. 2E7 cells were resuspended in 10mL of complete medium, fixed with 1% formaldehyde (270µL of 37% formaldehyde) and mixed by tumbling at room temperature (RT) for 10 minutes. Crosslinking reaction was stopped by adding 630µL of 2M of glycine (final concentration 0,1375M) and mixed by tumbling at RT for 5 minutes. Cells were pellet, washed in 1mL of cold PBS, centrifuged at standard conditions and the pellet was resuspended in 400µL of ChIP cell lysis buffer with

freshly added protease inhibitors (1:500 diluted) and 10μM of Sodium Butyrrate (histone deacetylated inhibitor). The mixture was incubated on ice for 10 min, spun at 600g at 4°C for 5min and supernatant was discarded. The lysate pellet was resuspended in 400μL of ChIP nuclei lysis buffer with freshly added protease inhibitors and sodium butyrate as a.m., and incubated for 10min on ice. 5μL aliquot was taken prior sonication. The DNA was sheared in eppendorf tubes in a Bioruptor Sonicator for 15 minutes by pulsing sonication 15sec on/off. The chromatin was submitted to a preclear process, in which rabbit serum and agarose beads were incubated with the chromatin at 4°C at least 1h long. This step allows cleaning the background of chromatin fragments binding aspecifically the F_c (constant part) of any immunoglobulin or the beads themselves. By spinning and taking the supernatant, cleared chromatin was then incubated o/n with specific antibodies as well as aspecific IgG control antibody by rotating on wheel at 4°C. An aliquot of 10% in volume was kept for the input control. The next day, beads were washed 2 times with wash buffer I, once with wash buffer II and twice with TE: all these steps were performed at 4°C and every washing step was carried out by 10 minutes incubation on a rotating wheel, spinning down the beads, remove supernatant. Chromatin was then eluted from the beads by elution buffer, thoroughly vortex, spin, and carefully take the supernatant without touching the beads. The elution step was repeated twice, then IP samples and inputs (diluted with same amount of elution buffer) were incubate o/n at 65°C after adding of 250mM filtered NaCl (final concentration) to decrosslink protein complexes from DNA. The next day, after 2h incubation at 45°C in the presence of 5 μg of Proteinase K, all the samples were phenol/chloroform purified, followed by ethanol precipitation.

When analyzing the *in vivo* occupancy starting from rare population (less than 1E5 cells), ChIP was performed using the LowCell ChIP kit of Diagenode according to manufacture's instructions. In this thesis, this kit was used for sorted population and human primary material.

3 Results

3.1 Identification and characterization of a novel PU.1 regulatory element with insulating capacity

3.1.1 PU.1 promoter interacts with lineage-specific elements

The PU.1 promoter region is not sufficient to drive PU.1 expression alone, but relies on the cooperation with distinct regulatory elements. Several of these have been identified and characterized so far, most of which act in a lineage specific manner, (Okuno et al., 2005; Rosenbauer et al., 2004; Leddin et al., 2011; Zarnegar et al., 2010). However, the molecular mechanisms by which the PU.1 gene is maintained silenced in T-cells are still to be understood. In order to discover putative new regulatory elements we performed ACT (association chromosome trap) (Ling et al., 2006), a modification of standard chromosomal conformation capture (3C) (Dekker et al., 2002).

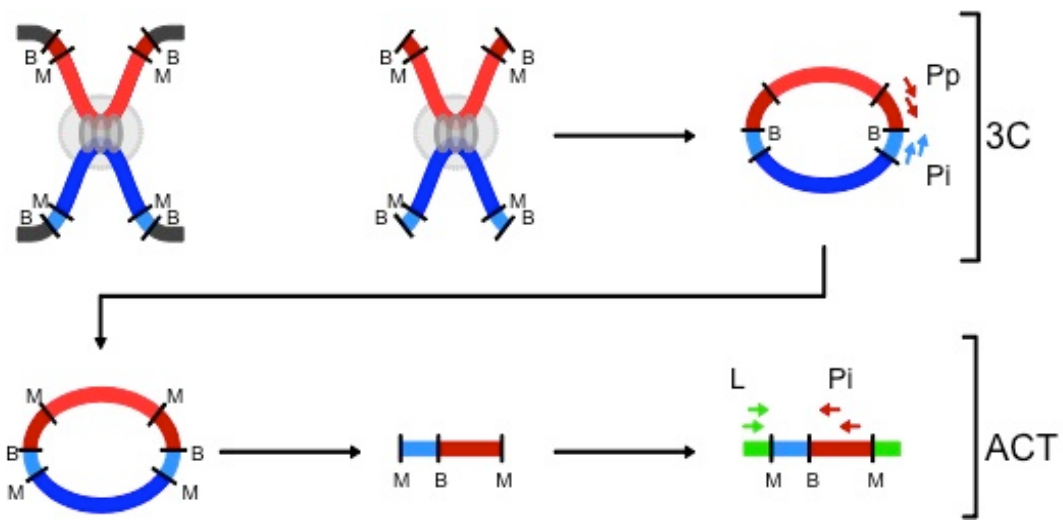


Figure 6: schematic representation of Chromosome Conformation Capture (3C, upper panel) and Associated Chromosome Trap (ACT, lower panel). Grey ovals represent proteins, the grey circle delimits the conceptual area of DNA-protein complex. B=BglII site; M=MspI; Pp=primer on PU.1 promoter; Pi=primer on known interacting fragment; L=linker-based primer.

As depicted in Fig. 6, by this technique it is possible to detect long-range chromatin interactions between a candidate region (here the PU.1 promoter) and unknown DNA

fragments. By subsequent linker-mediated PCR and sequencing, the new regions can be mapped to the genome.

We analyzed thymocytes (T cells), where PU.1 gene is actively silenced, bone-marrow derived macrophages (Macs) as highly PU.1 expressing cell type and murine embryonic fibroblast (MEF) as non hematopoietic cells. PU.1 expression was measured at mRNA level in these cells, as shown in Fig.7.

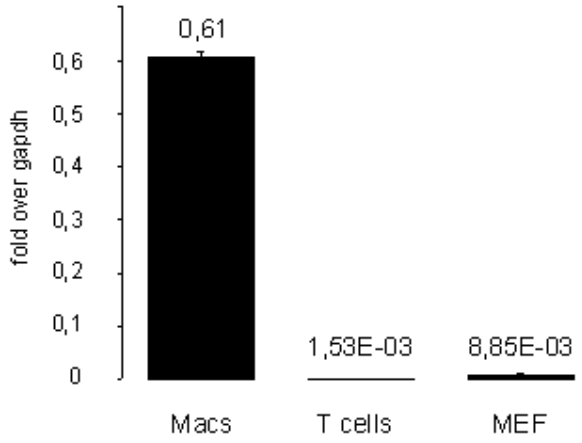


Figure 7: qRT-PCR showing Pu.1 mRNA level. Bars indicate values ± s.d. (standard deviation); data were normalized to the gapdh expression levels. Macs = bone marrow derived macrophage; T cells = thymocytes; MEF = murine embryonic fibroblast

ACT assay is based on crosslinking protein complexes to DNA by formaldehyde covalent binding; to test the outcome of this initial reaction a not-crosslinked control sample was carried out for all three cell types. As Figure 8B shows, in not-crosslinked controls no interaction could be observed; on the contrary (Figure 8A), crosslinking DNA to protein complexes allows the detection of chromatin interactions, confirming that the chromatin organization is achieved with protein complexes which modulate the DNA structure. Interestingly, Macs and T cells show a different pattern of amplified products, whereas in MEF only a basic smear was detected, suggesting that no specific other genomic loci spatially interact with the PU.1 promoter in non-hematopoietic cells.

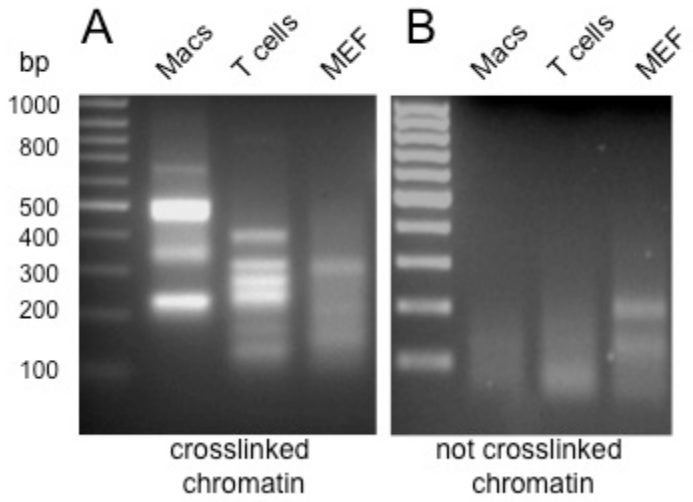


Figure 8: endpopint PCR on linked-based PCR on ACT material. A) crosslinked reaction with indicated cell types B) not-crosslinked controls. All the reactions were loaded on 2% agarose gel.

The hematopoietic cells, represented by two different lineages, show a different interplay of the promoter region with other chromatin partners. This is a sign of active regulation, either to keep to locus active for macrophages (Macs) or repressed for T cells. In MEF cells PU.1 expression is absent; here we didn't detect any interacting region with the PU.1 promoter, suggesting that the locus is closed by different mechanisms (e.g., spread of heterochromatin).
All amplified PCR products were extracted from the agarose gel, subcloned in the PGEM-T-easy vector, sequenced and subsequently mapped to the genome using NCBI blast browser. The sequences were considered only if they met the following criteria: first, they had to contain the PU.1 promoter sequence; second, the mapping to the genome region had to start with a MspI site (in both directions); third, BglII site had to separate the PU.1 promoter sequence from the interacting region; fourth, BglII region flanking the PU.1 promoter were considered as background due to insufficient spatial distance, and the cut off for this artefact interaction was set to 2 BglII sites distance.
In table 3 newly identified fragments are indicated for hematopoietic lineages. The bands we analyzed from MEF didn't follow the a.m. criteria, therefore were background noise and were excluded. The murine PU.1 gene is located on chromosome 2, and the coding sequence's (CDS) genomic coordinates are chr2:90,936,954-90,955,913 (according to USCS browser, assembly on July 2007; NCBI37/mm9). In macrophages, an interaction between the PU.1 promoter and a known enhancer located 14Kb upstream of the transcription start site (TSS) as

well as another interaction within the same chromosome were observed. In T cells the intronic part of Ppard (peroxisome proliferator-activated receptor delta) gene, a metabolic gene, was found in physical proximity with the PU.1 promoter; moreover, two distinct chromatin interactions were detected in the PU.1 locus, -25 and +71Kb relative to PU.1 TSS.

cells	chr.	BglII start	BglII end	region
Macs	1	120.483.644	120.493.282	intronic: CLIP-ass. protein 1 isoform 1
Macs	2	91.158.402	91.158.505	hypothetical protein LOC228356 (+263 kb)
Macs	2	90.922.792	90.925.236	Intergenic (-14 kb, URE)
T cells	2	90.911.726	90.912.086	Intergenic (-25 kb)
T cells	17	28.399.612	28.401.964	intronic: Ppard
T cells	2	91.005.810	91.010.557	intronic: Madd (+71 kb)

Table 3: genomic location of ACT-identified product. Cell type and BglII genomic coordinates are indicated.

In this thesis we will focus on intra-chromosomal interaction in T cells, so we further confirm and characterize the elements found to interact with the PU.1 promoter. We first confirmed the -25Kb and the +71Kb element by 3C, to be sure that their detection was not a background PCR artefact. For this purpose we designed primers spanning the BglII site one on the PU.1 promoter (bait) and the other on the candidate region. As control we use the bait primers together with primers located on the Gapdh promoter. As Fig. 9 illustrates, we could confirm both interaction, compared to the control locus (Spilianakis et al., 2005), where no amplified product was observed.

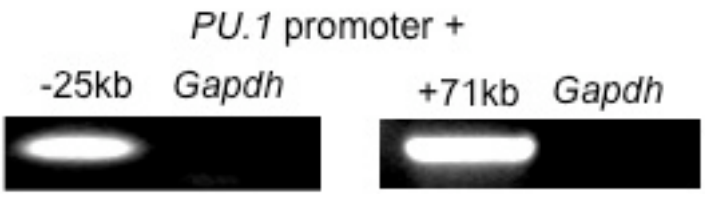


Figure 9: 3C standard assay on T cells with indicated primers.

3.1.2 +71Kb, but not -25Kb element possesses insulting activity *in vitro*

PU.1 gene must be actively silenced during physiological T cell development. Based on this knowledge, we investigated whether the two interacting elements could be responsible for PU.1 gene down-regulation. We thus tested them in reporter assay for insulator or repressor function, by the use of a vector containing the H19 promoter driving the firefly luciferase gene and a strong universal enhancer (SV40). In this assay the putative insulator should decrease the luciferase acitivity, when being cloned between the H19 promoter and the SV40 enhancers.

The -25Kb BglII fragment is almost 500bp long, whereas the +71Kb is around 4Kb long. In order to compare these two regions to each other and with the control insulator (Ishihara et al., 2006), fragments on the same Kb scale were PCR amplified, spanning the -25Kb region or subdividing the +71Kb fragment.

Two regions spanning the -25Kb BglII fragment in both directions were cloned in both orientations between enhancer and promoter to test for insulating function in vitro (Fig. 10). By monitoring luciferase activity there was no significant difference between negative control (non-insulating-control) and -25Kb fragments were observed. We therefore concluded that this region does not have insulator function and we excluded it from further studies.

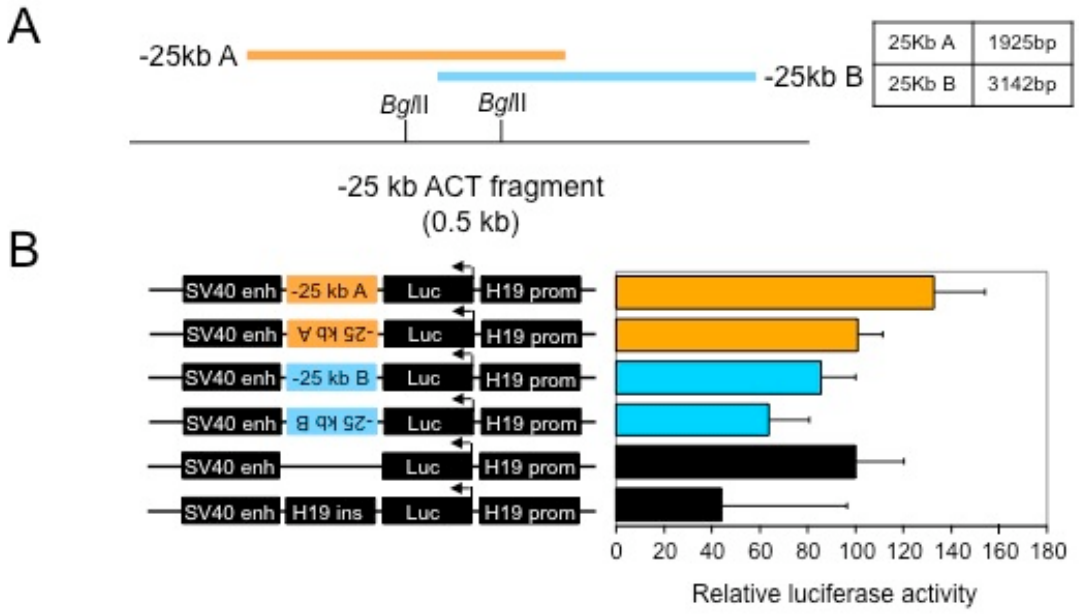


Figure 10: enhancer blocking luciferase reporter assay. A) schematic representation and size of cloned fragment spanning the -25Kb BglII fragment. B) luciferase activity was measured after 24 hours of transfection in HEK23T cells and was normalized to the Renilla codifying plasmid used as internal control for transfection efficiency. Values of sv40 enhancer-H19 promoter construct were set to 100% and served as reference luciferase activity. Error bars represent S.D. of one out of at least three independet experiments

The +71Kb BglII ACT fragment was too big to be cloned in one plasmid, so it was split in different regions and cloned separately: first in 3 fragments (A,B,CL) with the same size (around 2Kb), then in 2 overlapping ones (AC, BC), and finally narrowing down the region with the major insulating activity (from 71CL to 71C and 71CS).

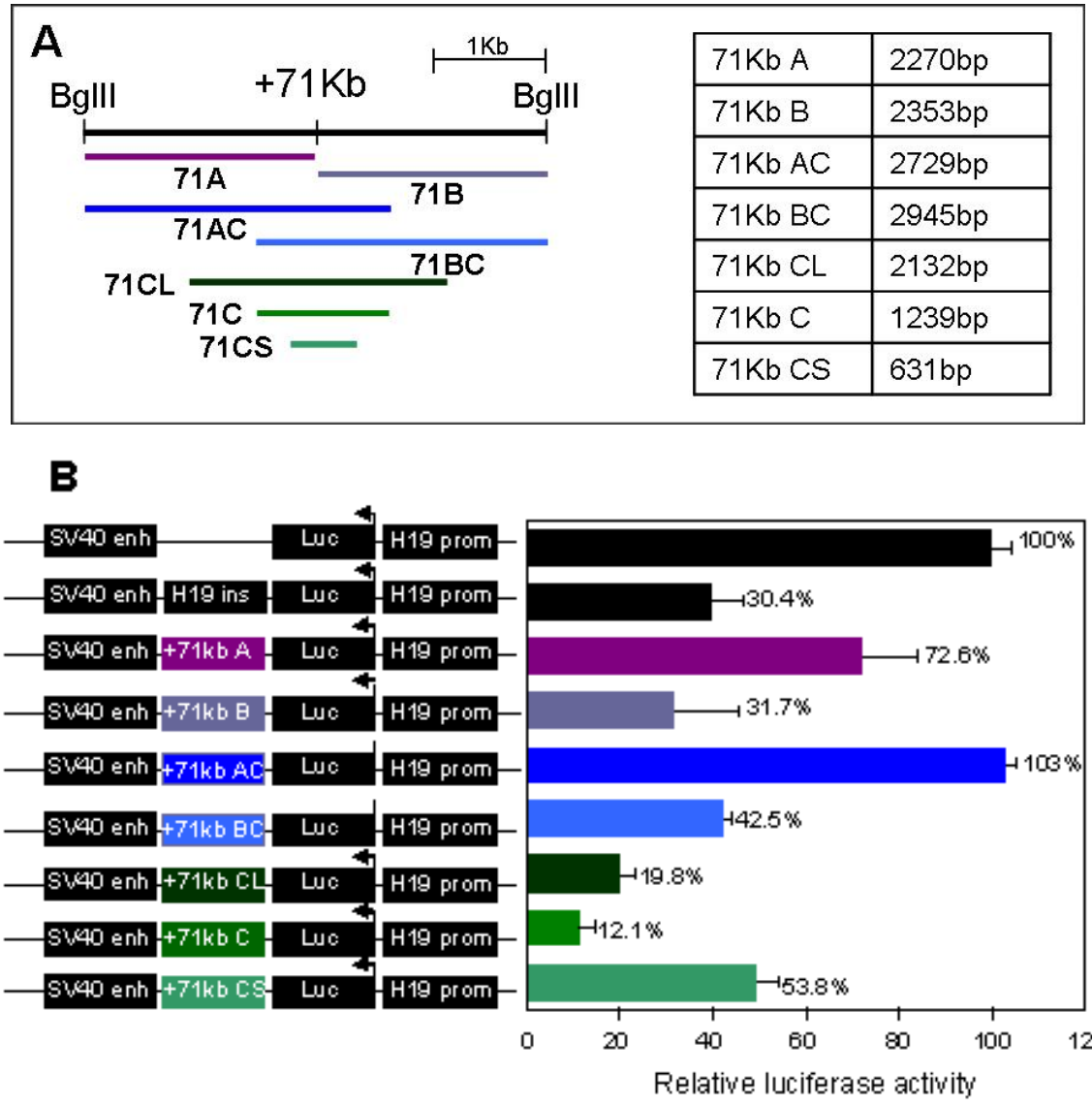


Figure 11: enhancer blocking luciferase reporter assay. A) schematic representation of tested fragments and relative sizes. B) relative luciferase activity of depicted constructs. Error bars represent SEM (mean standard deviation) of at least three experiments.

As illustrated in Figure 11, the fragment with best insulating activity on the SV40 universal enhancer is a 1.2Kb region located in the middle of the +71 BglII fragment. Importantly, the enhancer blocking activity was not size dependent, since fragments with similar sizes (e.g. A, B, CL or AC and BC) show different outcome.

We next investigated in more detail the in vitro properties of the most important insulator candidate (+71C), testing it for orientation and position effect.

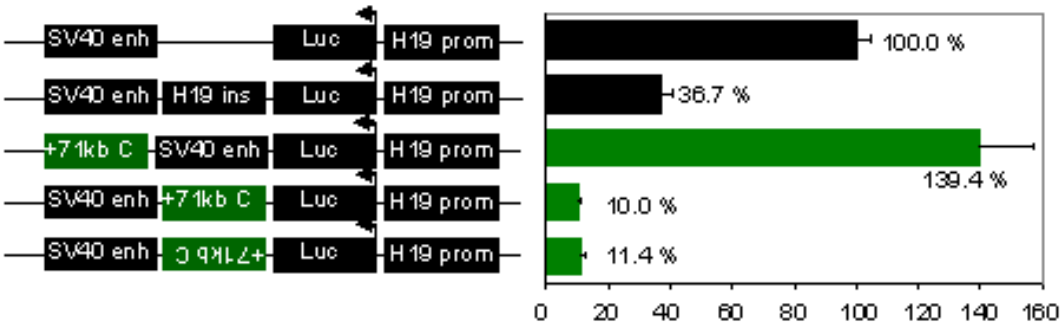


Figure 12: enhancer blocking luciferase reporter assay testing orientation and position effect of the +71 element. Error bars represent S.D. of at least three independent experiments.

Despite the -25Kb fragment, the +71Kb element is able to block enhancer activity on the promoter. Moreover, it shows two fundamental properties of an enhancer blocking insulator: 1) is orientation independent 2) is position dependent. Both properties are illustrated in Figure 12: when the core of the putative insulator (+71C) is cloned after the SV40 enhancer, luciferase activity is not repressed, indicating that in this position the +71 element has no activity. Whereas, when it is cloned between promoter and enhancer, it can block the inducing function of the enhancer on the promoter regardless its orientation. In conclusion, this in vitro assay suggests that the +71 element is an enhancer-blocking insulator.

3.1.3 PU.1 locus chromatin interaction screen to dissect linage-specific local chromatin structure

In order to investigate the nature of the interaction between the PU.1 promoter and the +71 element in vivo we screened the whole PU.1 locus for chromatin interaction taking the promoter region as bait. This approach was chosen first to determine whether the physical

proximity detected by ACT and confirmed by 3C (see previous chapters) is a specific chromatin loop or just a proximity due to a closed structure along the whole locus. In second instance, such method gives more detailed information about the start and the end of the interacting fragments. Hereafter there is a summary of the experimental procedure and correspondent explanation of this technique; this chapter should be taken as reference and example for further experiments using the same technique.

The first problem about 3C-based technologies is the missing positive control; for this reason every sample is related to a PAC carrying the genomic sequence of the required locus. In other words the reference is not standard, therefore the entity of one interaction assumes a meaning only when compared to other interactions or to other samples. As discussed below, we decided to analyze more samples and more interactions (different primer fractioning the PU.1 locus).

To investigate the local chromatin structure behavior in accordance to PU.1 transcriptional activity, we applied the 3C technique followed by qPCR (q3C) and screened for chromatin interaction along the PU.1 locus in several hematopoietic populations with different PU.1 expression levels: T cells as PU.1low lineage, B cells as PU.1intermediate=int lineage and macrophages as PU.1high population (see Figure 13).

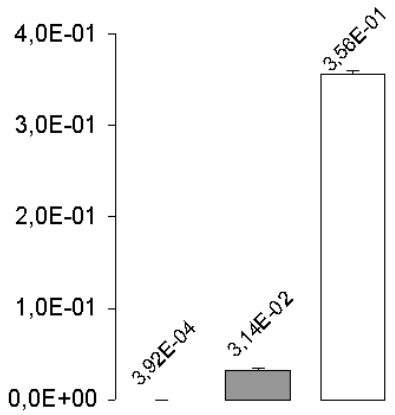


Figure 13: PU.1 expression of indicated population is indicated as cDNA fold over βactin housekeeping gene. T cells = whole thymus extract; B cells = CD19+ splenic B cells, purified by microbeads and magnetic separation; BM-Macs = bone marrow derived macrophages.

In order to compare different 3C samples quantitatively by RealTime PCR every step was controlled, according to (Hagege et al., 2007).

1) crosslinking efficiency: the efficiency of this reaction is practically impossible to be checked. A not-crosslinking control was run in parallel for the pilot experiment for all three lineages: no interaction was detected, confirming that without formation of DNA-protein complexes it is not possible to analyze chromatin structure (data not shown).

2) digestion efficiency: it is very important to check that digestion efficiency is close to 100%. Moreover, different samples should be digested with very similar efficiency in order to compare them. Digestion efficiency was calculated by RealTime PCR as ΔCt between amplified product spanning BglII restriction site (PCR reaction on restriction enzyme site: PCR^{RE}) in relation with a PCR product within 2 BglII sites (PCR reaction on genomic DNA not spanning the chosen endonuclease: PCR^{gDNA}). These values were then related to a not-digested sample used also for standard curve. Assuming that the primer efficiency of both of the two primer pairs is 2, the applied formula was:

$$[2^{\wedge} (Ct_{standard}^{gDNA} - Ct_{standard}^{RE})] / 100 = [2^{\wedge}(CtquerygDNA - CtqueryRE)] / x$$

where x is the undigested percentage of the query (3C sample).

Digestion efficiency is then

$y = 100 - x$

% undigested	% digested	sample
1,94	98,06	T cells
2,09	97,91	
1,40	98,60	
2,67	97,33	Macs
0,69	99,31	
1,66	98,34	
1,40	98,60	B cells
3,08	96,92	
2,09	97,91	
11,61	88,39	neg. ex.
0,29	99,71	PAC ctrl

Table: 4 independent samples for every indicated population were analyzed. PAC ctrl = PU.1-PAC BglII digested and religated. neg. ex. = negative example.

3) ligation efficiency: the ligation efficiency cannot be quantified within the same samples for every locus in the genome. To overcome this problem, an aliquot for each sample was run in parallel but without ligase treatment, which allowed for a yes or no answer, whether the ligase worked at all. All samples (with and without the ligase) were tested for the circle formation reaction, meaning the ligation of digested fragments within themselves, which also represent theoretically the most probable ligation product.

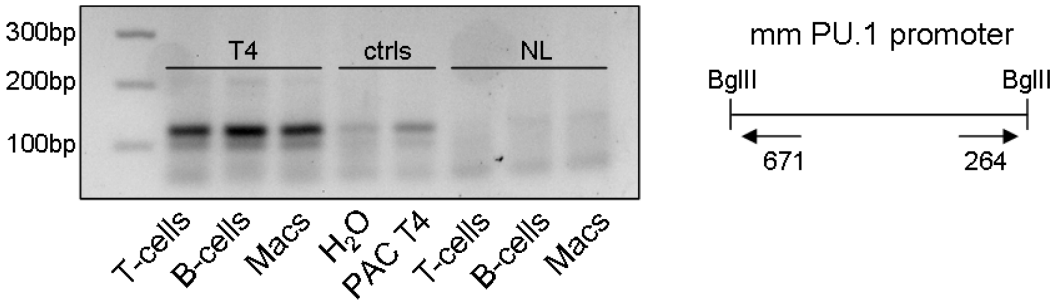


Figure 13: PU.1 promoter circle formation verifies the activity of the ligase. On the right, agarose gel showing PCR-amplified product from 3C samples, each representative for one lineage. T4 = ligated samples; NL = not.ligate samples; PAC T4 = positive control, here highly diluted (1pg), see bullet point 6 for explanation. On the left, essential graphic representation of the BglII fragment spanning the PU.1 promoter region; the primers used for circle PCR are indicated.

4) 3C assay test: 3C samples were first tested by endpoint PCR for chromatin loops which were previously reporter for the Gapdh locus. This interaction is required for GAPDH expression, therefore can be considered as "housekeeping loop":

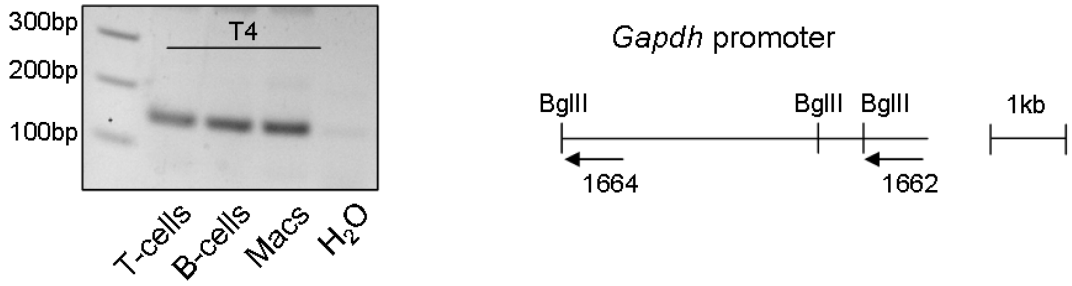


Figure 14: Long-range interation within Gapdh promoter as housekeeping loop. On the right, agarose gel showing PCR-amplified product from 3C samples, each representative for one lineage. T4 = ligated samples. On the left, graphic representation of the BglII interacting fragments on the Gapdh locus; the primers used for the PCR are indicated.

5) DNA content and salt effect: it is known that salt concentration in the DNA pellet at the last DNA purification step of the 3C assay can affect downstream RealTime PCR and interfere with spectrometer quantification of DNA content. To exclude this inhibitory effect in the PCR reaction and to compare samples with very similar DNA amount, samples were normalized by PCR for gDNA amount by serial dilution of every sample using primers not spanning sites of the chosen restriction enzyme.When by dilution of 10 fold the sample correlates with the standard curve, then no salt effect is observed (data not shown).

6) primer efficiency: to calculate primer efficiency and to compare Ct values of different PCR reactions, a PAC carrying the PU.1 locus was digested with BglII followed by ligation reaction in small volume. This sample was used as positive control assuming that all possible ligation products between the BglII fragments were represented. The same sample was used to run a standard curve for every primer pair, by diluting the PAC control samples over 3 Log folds. Every efficiency was extracted by plotting the Ct versus dilution, and interpolating the data by adding a logarithmic trendline:

$y = mx + q$

converting the slope from log (base 10) scale to ln (base e) one:

$M= m/\log_{10}e$,

where e is the Euler's number;

and finally

Eff= 10^(1/M)

Where Eff. is the primer efficiency

Pu.1 promoter primer +	efficiency
-21Kb	1.88
-17Kb	1.80
-14,5Kb	1.88
-11,8Kb	1.83
+22Kb	1.83
+37Kb	1.92
+55Kb	1.79
+71Kb	1.89

Table 5: set up of primer pairs for the PU.1 PAC control digested and religated (designed as PAC T4). A) schematic representation of primers closed to BglII restriction site delimitating the region tested by 3C. B) agarose gel shows the specificity of every primer pair used. PAC T4: positive control; PAC: negative control as genomic DNA not BglII and ligase treated; w: water. C) an example how to calculate the primer efficiency. For all PCR reactions the correlation coefficient was > 0,95. D) the table indicates primer efficiency of every primer pair, calculated by a standard curve diluting the PAC T4 control over 3 Log folds.

As depicted in Figure 15, the three indicated lineages show different local chromatin structure spanning 100Kb around the PU.1 locus. When PU.1 gene is transcribed as in B cells and macrophages (Macs), we observe an interaction of the upstream cluster of enhancers, within which the URE shows the highest crosslinking frequency. On the contrary, the locus is strikingly differently organized in T cells where PU.1 is silenced: here the only prominent interaction is indeed the one between the PU.1 promoter and the +71 element. Interestingly, there is no or very low interaction in the region between these two fragments, demonstrating that the nature of this chromatin interaction is a loop conformation.

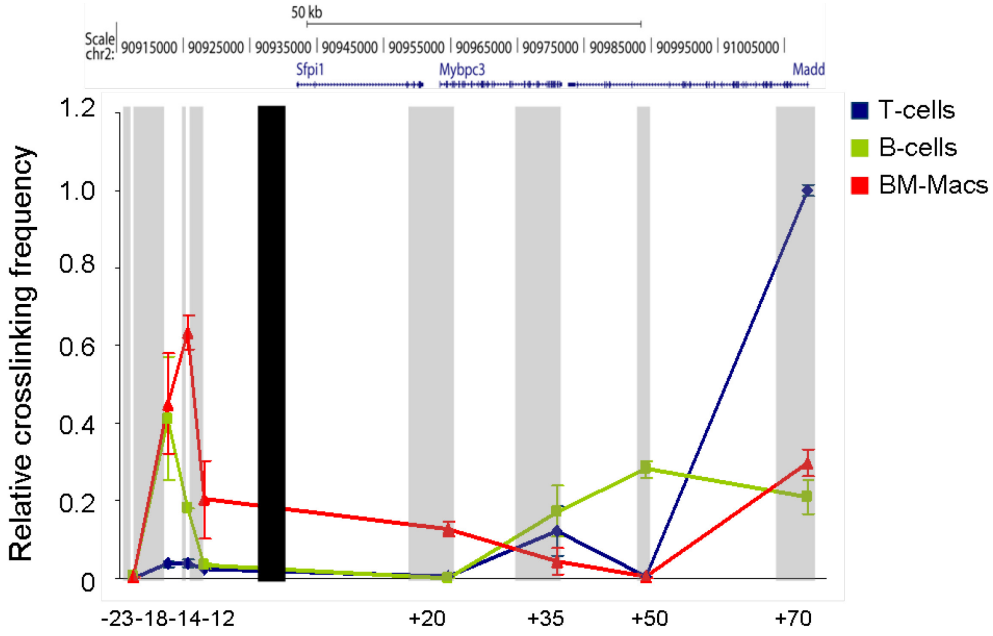


Figure 15: q3C on PU.1 locus and surrounding genomic regions. On top of the graph genomic coordinates and gene bodies are depicted. Crosslinking frequency refers to the ΔCt between the PAC control and the query: then the highest detected value was set to 1 to obtain the relative crosslinking frequency. Cell types are depicted with different colors. Grey bars represent the analyzed BglII fragments and the black the PU.1 promoter one, used as bait. The X coordinates indicate the genomic distance of the primers in relation to the primer on the PU.1 promoter.

This whole-locus screen confirms that the enhancer cluster, including the URE, is in physical proximity with the promoter in PU.1 expressing cells, whereas this interaction is absent in T cells, where the promoter loops with the +71 element.

Are the two main interactions of the promoter with the upstream regulatory cluster URC (observed in myeloid and B-cells) or with the +71 element (observed in T-cells) actively formed during lineage commitment or are these default conformations from the stem cell status? To address this question, LSK (lin- Sca1- Kit+) stem cells were sorted from mouse bone marrow and their chromatin analyzed for interactions. Figure 16 illustrates the 2 chromatin loops in stem cells, T-cells, B-cells and BM-Macs and MEFs as not hematopoietic cell line. In the stem cell compartment none of the loops could be detected in a prominent manner; confirming that local structure of the PU.1 locus is organized in different ways in various hematopoietic lineages.

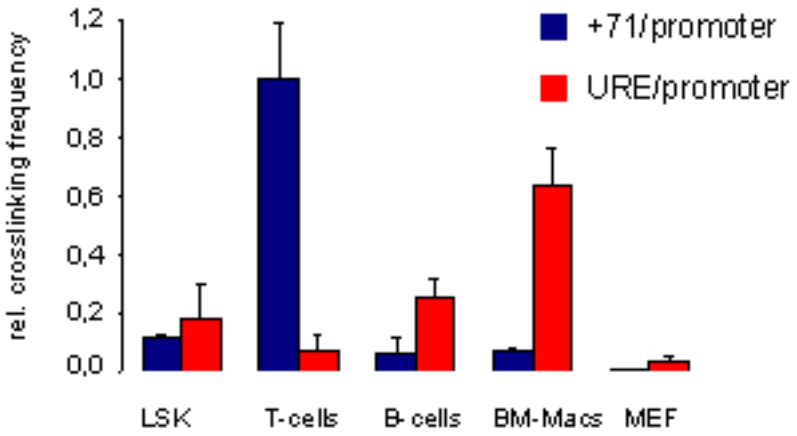


Figure 16: Two chromatin loops characterize specifically opposite lineages. In blue is depicted the interaction of the +71 element with the PU.1 promoter, in red of the URE. Error bars represent at least 2 independent experiments. Crosslinking frequency was calculated as explained in the previous figure. The highest value was set to 1.

3.1.4 Is the PU.1 promoter / +71 chromatin interaction sequence-specific?

In the previous chapters we analyzed the +71 element interaction with the PU.1 promoter in the endogenous locus. We next asked whether the DNA sequences are sufficient to form the a.m. loop also when not located on the same chromosome. For this reason we used a mouse strain previously created in our lab: we inserted the human PU.1 locus by injecting a BAC (bacterial artificial chromosome) into murine oocytes (Fig. 17), resulting in transgenic mice

which carry both murine and human PU.1 copies. By crossing these offspring into the PU.1 knockout background, we could demonstrate that the human locus can rescue the PU.1 lethality of PU.1KO mice (Leddin et al., 2011).

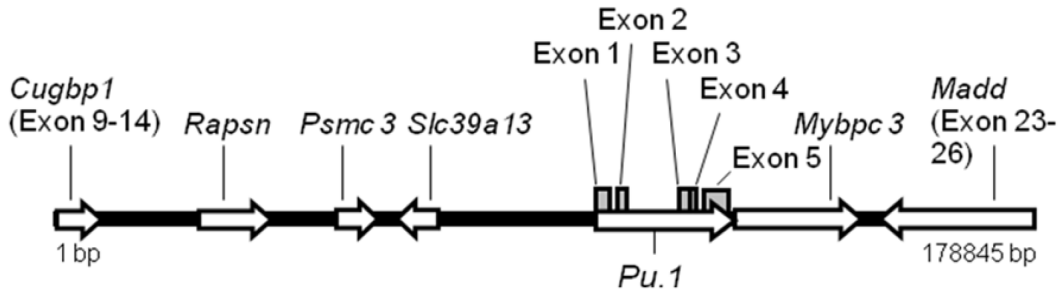
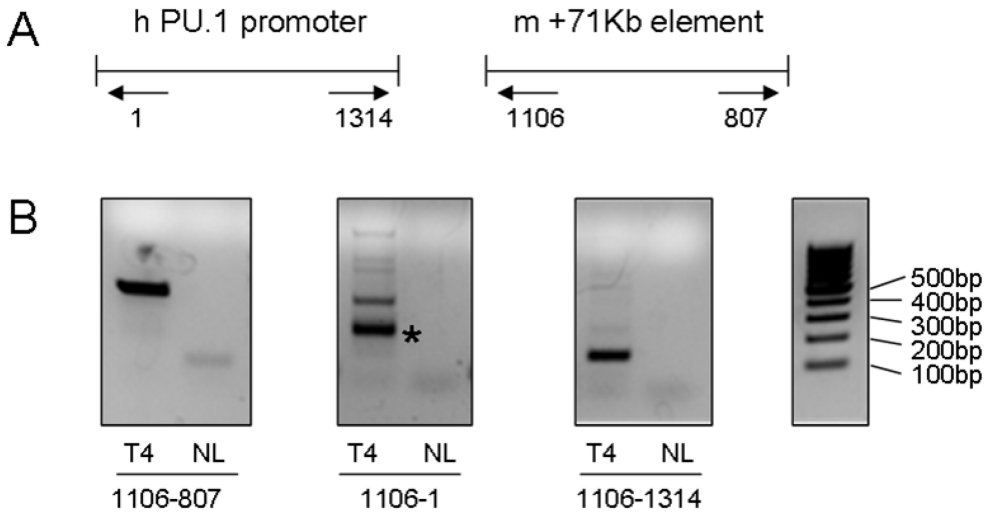


Figure 17: genomic region contained in the BAC injected in oocytes. The total length is around 160Kb and the integrity of integrated BAC was confirmed by several Southern Blot probes (data not shown).

The human PU.1 protein resembled qualitatively the tissue-speci fic expression of the murine counterpart. The transgenic BAC doesn't contain the region homologous to the +71 element. In the human genome this element is located +94Kb downstream the PU.1 TSS. We therefore hypothesized that the expression of the human PU.1 could be silenced during T cells development by forming the previously identified loop between the human copies and the endogenous +71. For that purpose, we performed 3C on thymocytes on this mouse strain, as it is possible to distinguish the endogenous (murine) from the transgenic (human) sequences by using a different restriction enzyme, suitable for the human promoter (ApoI rather than BglII as in the previous assays).

This result demonstrates that the chromatin interaction takes place regardless of the genomic position, and that the DNA sequences by themselves contain all the information needed to form the chromatin loop.



C

primers	size	location
1106-807	458bp	m71circle
1106-1	239bp	hPUp_5'-m71_5'
1106-1314	126bp	hPUP_3'-m71_5'

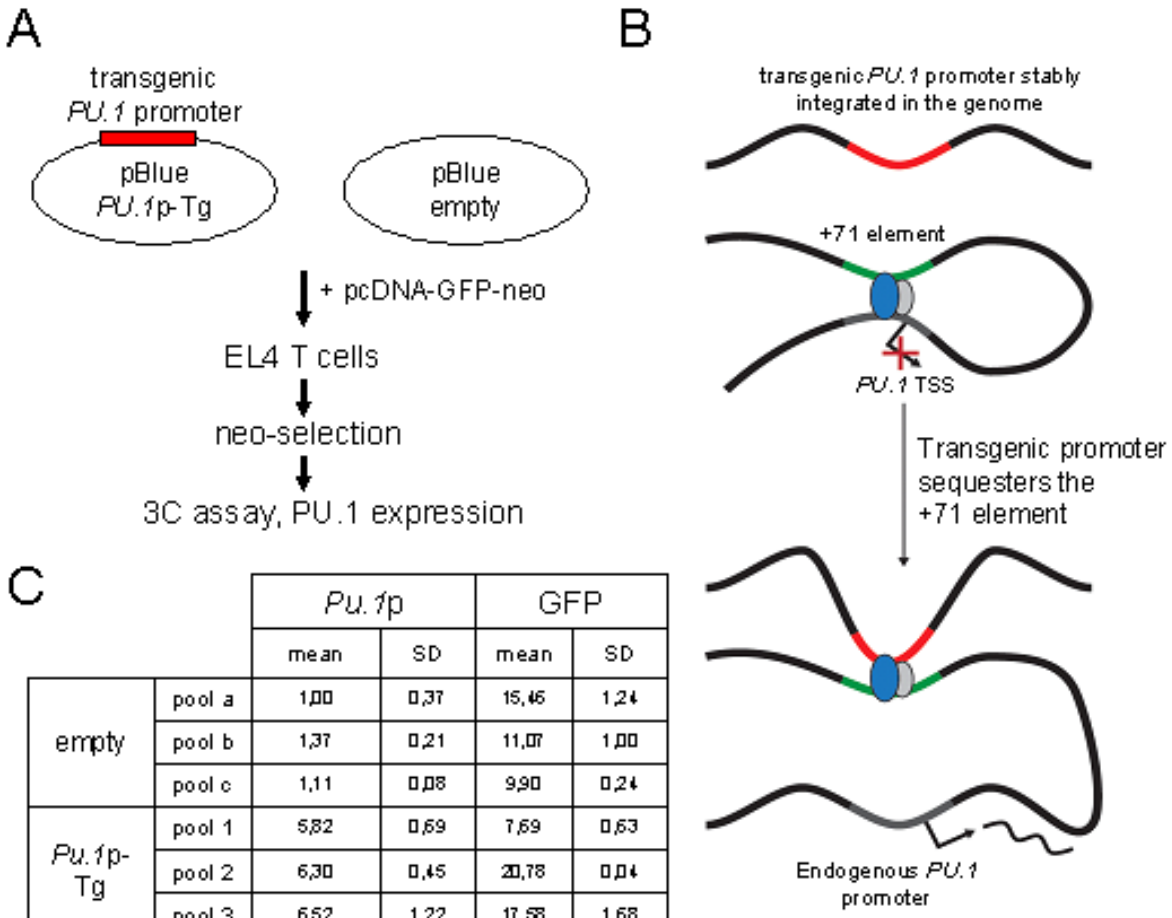
Figure 18: Interaction between human PU.1 promoter and murine +71 element. A) schematic representation of the analyzed ApoI-fragments with the primers used for PCR on 3C material. B) agarose gel showing PCR products detecting, from left to right, the murine +71 circle, the interaction between human PU.1 promoter and murine +71 element in two genomic direction. T4=3C sample; NL=No Ligase control. The ladder is indicated. The star on the second panel indicates the specific product. C) the table summarizes the displayed amplified products with their relative size and genomic location. All PCR products were verified by sequencing.

3.1.5 The PU.1 promoter / +71 element loop regulates PU.1 expression

Enhancer blocking insulators are usually located between an enhancer and a proximal promoter to control the functional interaction between both elements (West et al., 2002). One of the current models explains enhancer-blocking activity by a direct interaction of an insulator with an enhancer or a promoter, leading to disruption of the promoter-enhancer interaction and gene inactivation (reviewed in (Raab and Kamakaka, 2010)). Here, we identified a genomic element with insulator properties that is located far outside of the PU.1 enhancer/promoter cassette, but can loop to the promoter in a cell-type specific manner.

We tested the hypothesis that the interaction of the insulator with the promoter blocks endogenous PU.1 expression. For this aim, we generated stable T cell lines carrying

transgenic PU.1 promoter fragments. The idea behind this approach consists in overexpressing exogenous copies of the PU.1 promoter which will interact with the +71 insulator, disrupting its interaction with the endogenous PU.1 promoter (Figure 19).



C

		Pu.1p		GFP	
		mean	SD	mean	SD
empty	pool a	1,00	0,37	15,46	1,24
	pool b	1,37	0,21	11,07	1,00
	pool c	1,11	0,08	9,90	0,24
Pu.1p-Tg	pool 1	5,82	0,69	7,69	0,63
	pool 2	6,30	0,45	20,78	0,04
	pool 3	6,52	1,22	17,58	1,68

Figure 19:.Concept of sequestering the endogenous loop approach. A) Flow chart of generation of stable cell lines. B) schematic conformational change of loop formation from endogenous to exogenous PU.1 promoter. in red: exogenous promoter, in grey: endogenous promoter, in green: +71 endogenous insulator. Blue and grey ovals: protein complex forming the chromatin interaction.C) copy number of different pools by qPCR shows the heterogeneity of the pools. For GFP, a standard curve was performed related to a transgenic mouse model containing only one copy. For the PU.1 promoter, the values were normalized to 1 related to normal genomic DNA content (2n=1)

Lymphoblastoma EL-4 cells were employed as T cell line, because they express low amount of PU.1. EL-4 cells were transfected with a plasmid carrying 2,1Kb of the PU.1 promoter region or the empty vector as control. In order not to have an additional promoter in the same plasmid, a vector containing neomycin resistance cassette and coding for GFP as reporter was co-transfected in a ratio of 1/10 compared to the PBluescript empty or the exogenous PU.1

promoter: this vector allowed us both to select for antibiotic resistance and to follow the selection process by GFP signal in FACS analysis (data not shown). Different pools were generated and tested for heterogeneity by quantitative RealTime PCR, detecting either the GFP coding sequence, present in both, or the PU.1 promoter using primer spanning both endogenous and exogenous copy.

To validate the system, it is required to detect the interaction between the exogenous promoter and the endogenous +71 insulator. It is possible to distinguish between the two promoters by using different restriction enzymes in 3C assay: while BglII is the enzyme of choice for using the endogenous promoter as bait, because this fragment contain the region upstream the PU.1 TSS, ApoI is suitable for the exogenous one, because its restriction site are contained in the backbone of the pBlue plasmid. On the other hand, the +71 element can be analyzed by both digestions, which span over the insulator core with similar size fragments, 4Kb for BglII digestion and 6Kb for ApoI (as ligation efficiency might also depend on length of ligation partners) Figure 20 shows an exemplarily interaction pattern of cell lines containing the PU.1p-Tg (transgenic PU.1 promoter sequence) compared to the control lines. For this endpoint PCR, the PAC control was adapted to the context: the mPU.1-PAC (PAC carrying the PU.1 locus) as well as the two plasmids used to generate the stable cell lines (pBlue PU.1-Tg and pBlue) were digested by ApoI. Then the digested PAC was separately religated in equimolar ratio to every of the two plasmids. In the bottom panel, PCR detecting the circle formation around endogenous promoter was performed as control for genomic DNA content as well as for the 3C procedure: cell lines with the exogenous promoter or the empty control show similar intensity of the amplified product. Importantly, the transgenic promoter copies interact specifically with the endogenous +71 element, depicted in the upper panel. Along the PU.1 locus we also controlled that this interaction is unique and specific, as confirmed by the testing interaction between the exogenous promoter and a BglII fragment located +55Kb downstream the PU.1 promoter (middle panel). In contrast, cell lines carrying the empty vector (pBlue) didn't show any of these interactions. The PAC controls verified that the PCR products were specific: in fact, the control not including the exogenous promoter plasmid didn't show interaction within this fragment. All the PCR amplified products were verified by sequencing.

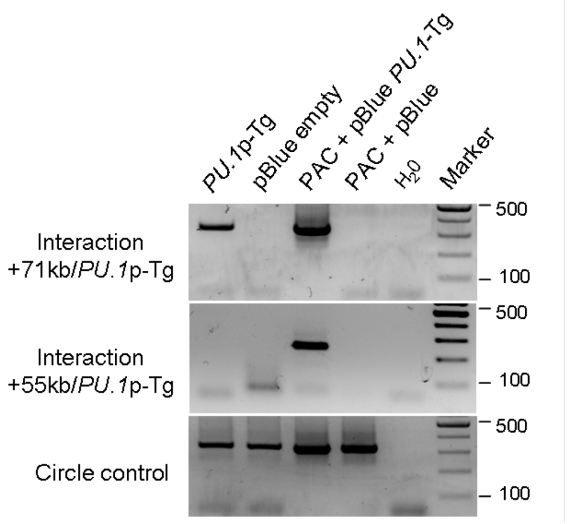


Figure 20: the exogenous PU.1 promoter interacts with the endogenous +71 element. PU.1p-Tg = exogenous promoter carrying cell lines. pBlue empty = control cell lines with empty vector. 3rd lane = control for PU.1p-Tg, 4th lane = control for pBlue empty lines. 3C assays was based on ApoI digestion.

We finally tested the different pools of both genotypes for PU.1 expression by qPCR as well as for the endogenous loop formation by 3C. Interestingly, we could demonstrate that in the transgenic cell lines carrying the PU.1p-Tg integrated into the genome the loop between endogenous promoter and +71 element is significantly reduced to approximately 20% compared to the control cell lines (Fig. 21A). Intriguingly, PU.1 expression is increased by 4-fold upon loop formation reduction (Fig. 21B).

By this transgenic approach, we could show that the exogenous promoter sequence is able to compete with the endogenous promoter locus, interacting and therefore sequestering the +71 element from its physiologic function. The consequent reduction of endogenous loop formation correlates with an enhanced PU.1 transcription. This is shown both by increased histon acetylation of the PU.1 promoter, a hallmark of active transcription, as well by cDNA PU.1 level.

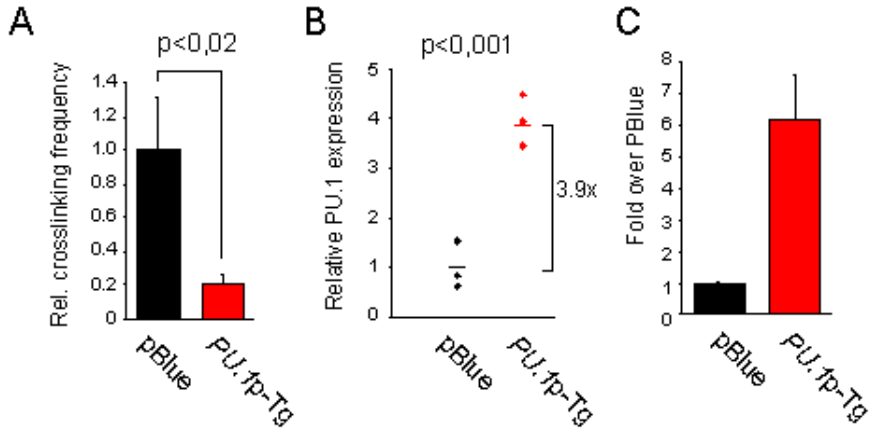


Figure 21: reduced loop formation corresponds to PU.1 upregulation. A) q3C on cell lines containing the empty vector pBlue compared to transgenic promoter PU.1p-Tg. Crosslinking frequency was calculated as ΔCt between every sample and the PU.1-PAC control; average of pBlue cell lines was set to 1 to obtain relative crosslinking frequency. P value is indicated. B) qPCR with exon spanning primers on PU.1 cDNA. pBlue mean expression was set to 1. C) H3K9Ac in vivo binding on the PU.1 promoter. Bars show the average of 3 independent pools each. Every enrichment was first calculated as fold over correspondent control IgG antibody, then pBlue average was set to 1.

Taken together, these data provide strong evidences for a functional role of the newly discovered genomic insulator that represses PU.1 gene by chromatin looping.

3.2 Chromatin organizer proteins mediate the PU.1 promoter/+71 chromatin loop

3.2.1 CTCF and SatB1 bind to the PU.1 promoter region and to the +71 insulator

In the previous chapter we identified and characterized a novel element responsible for PU.1 gene regulation in T cells. We showed that the +71 element is an insulator, which physically contacts the PU.1 promoter. In this chapter we focus on the proteins that are involved in the formation of this chromatin loop. Bioinformatic screening was performed on the +71 element, using free online web tool as Genomatics and the insulator database. The list of most prominent candidates was obtained by applying the following criteria: 1) the proteins should be expressed in T cells, 2) its binding site should be conserved between mammalian species,

3) the putative protein should dimerize and/or be able to form chromatin loops. The outcome of this screening for transcription factor binding sites revealed CTCF and SatB1 as good candidates. Both are chromatin organizers and remodelers at local level as well as in high order chromatin structures, such as heterochromatin formation or chromosome territories organization. Moreover, both proteins are highly expressed in T cells and their function is required for proper T-cell development, as was demonstrated in loss-of-function models (KO mice).

SatB1 is a 100Da protein with multiple functional domains. Among them, a BUR-binding domain is necessary for DNA recognition sites, an atypical homeodomain and a PDZ domain. All these domains are necessary to confer SatB1 high specificity to the BUR DNA sequences, and to distinguish them from any AT-rich DNA sequences (Burute et al., 2012). CTCF (CCCTC binding factor) is nowadays recognized as the main "chromatin weaver" of the genome, binds the DNA via zinc finger domains (it contains 11 of them) as monomer, dimer, tetramer, heterocomplexes with homologues protein like BORIS (Phillips and Corces, 2009). Predicted binding sites of these nuclear factors need to be confirmed experimentally since the core sequence has not been fully characterized so far. For that purposes, nuclear extracts from murine thymocytes were first tested by gel shift assay for the capability to bind the +71 element.

As depicted in Figure 22, binding of CTCF and SatB1 to the +71 was confirmed. To test whether a band is the specific result of the protein binding to the radioactively labelled oligos representing the putative binding site, different not labeled oligos were co-incubated in a molar excess ratio. The cold probe contains the same sequence as the labelled oligo, so here all bands disappear. To exclude unspecific bands known binding sites for CTCF or SatB1 were used as competing oligos. Thus, in these lanes the specific band could not been detected. Importantly, incubating nuclear extracts with the labeled binding site oligo and the antibody targeting the candidate protein (here only displayed in the SatB1 EMSA), a super shift was observed, which could occur only in the case that the ternary complex (oligo+protein+antibody) has formed. The specificity of the antibody was compared using an unspecific antibody (IgG) which does not bind to the binding candidate, and therefore does not produce any ternary complex. This is shown in the last two lanes of each gel shift. In conclusion, CTCF and SatB1 are indeed able to bind in vitro the identified binding sites.

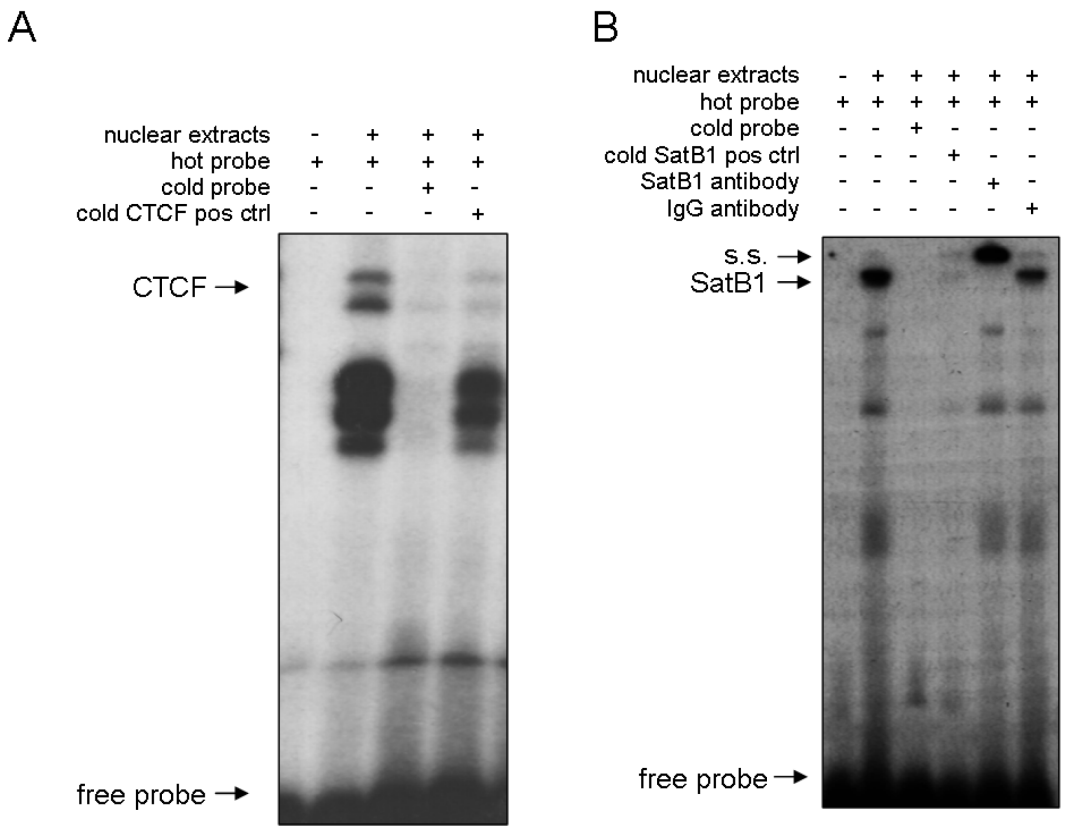


Figure 22: The +71 element is bound by chromatin nuclear factors in vitro. A) EMSA with nuclear extracts from murine thymocytes shows that CTCF can bind the +71 element. – and + represent absence or presence of indicated component in the binding reaction. Free probe, CTCF binding shift and super-shift with a CTCF antibody (s.s) are displayed. B) EMSA for SatB1 binding site on the + 71 element.

Next, we tested the +71 element and the PU.1 promoter region for in vivo binding of CTCF and SatB1: it could happen in fact, that the sequence is bound in vitro, especially by abundant proteins like CTCF and SatB1 in thymocytes nucleus, but not living cells. Chromatin immunoprecipitation (ChIP) assay was thus performed on three different cell types, T cells, B cells and macrophages, which represent different PU.1-expressing lineages (see chapter 3.1.3). In parallel, also the acetylation on lysine 9 of the histone 3 (H3K9Ac) was investigated by ChIP on the PU.1 promoter region: H3K9Ac is a hallmark of active transcription because it is involved in recruiting the RNA polymerase II activated complex to the promoter of a gene. Both CTCF and SatB1 bind their relative binding sites within PU.1 promoter and +71 element. Figure 19 illustrates the occupancy of CTCF (a), SatB1 (b) and H3K9ac (c) in the indicated DNA regions. Interestingly, CTCF and SatB1 bind the PU.1 promoter and the +71 element in a T cell specific fashion, suggesting a nice correlation between their occupancy and the repressive chromatin loop. Histone acetylation confirms at epigenetic level that the

promoter region is transcriptionally active in B cells and macrophage lineages, reflecting the intermediate and high level of PU.1 protein, respectively. CTCF binds the DNA very frequently with very variable affinity, therefore is important to check the threshold between background and reliable enrichments. To control the chromatin immunoprecipitated by CTCF antibody, one known binding site (on the c-MYC gene) was amplified as positive control, whereas a region located +37Kb downstream the PU.1 promoter was used as negative control. The positive control for SatB1 binding (Cai et al., 2003) was run to test the efficiency of the antibody: for this protein, none of the commercially available antibody is working for in vivo binding assay, and here we used a friendly gift of the Kohwi-Shigematsu's lab, who raises their own antibody.

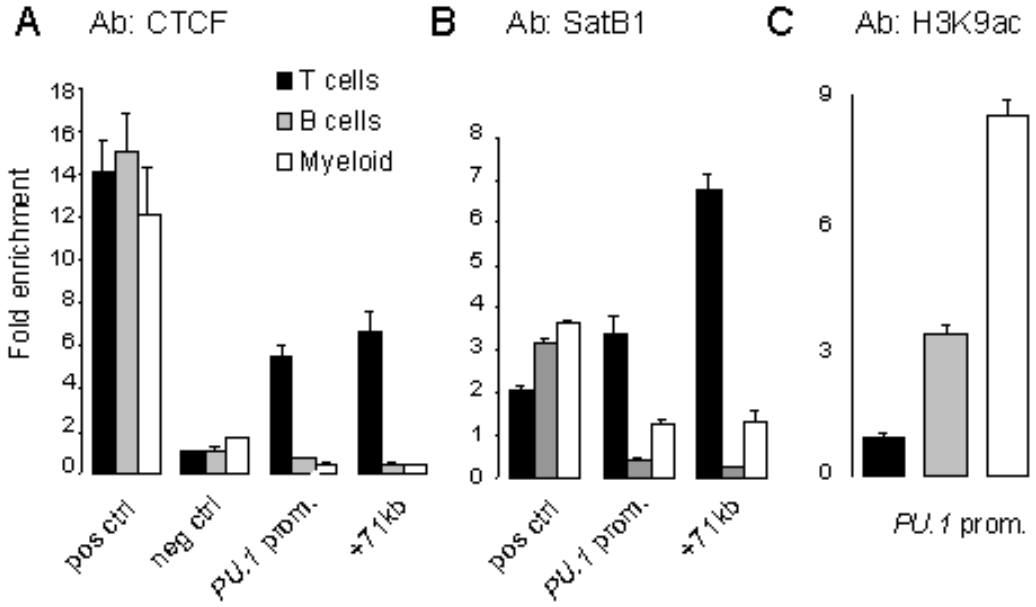


Figure 23: ChIP assay on T cells (black bars), B cells (grey bars) and myeloid cells (white bars). T cells were extracted from thymocytes, B cells were purified by magnetic separation from spleen using a CD19 antibody and myeloid cells were sorted as Ly6c+ MCSFR- from bone marrow suspension. In all panels, enrichment is displayed as fold over IgG control. Crosslinked chromatin was immunoprecipitated with antibody against A) CTCF, B) SatB1 and C) H3K9ac.

3.2.2 CTCF mediates the chromatin loop formation: a loss in function approach

To test whether the binding of CTCF to the insulator and the PU.1 promoter region is involved in the chromatin loop formation between these regions, a knock-down approach was

used. CTCF was shown to be necessary for T cell development (REF.). Since the deletion of CTCF is lethal, an inducible knock-down system is required, an inducible T cell line was used. We therefore engineered Jurkat cells (human lymphoblastoma cell line) by first transfecting with a plasmid encoding for a mouse ecotropic retroviral receptor, increasing the susceptibility of cells to infection with retroviral vectors of ecotropic host range. In this way mouse retroviral virus can easily be produced in vitro and afterwards used to infect Jurkat cells with high transduction efficiency (30-80%); moreover, by this system harder to handle human lentivirus can be avoided. The resulting Jurkat-eco cells were then transduced with a TET-on system (Ngo et al., 2006), and finally with vectors containing shRNA against the CTCF mRNA.

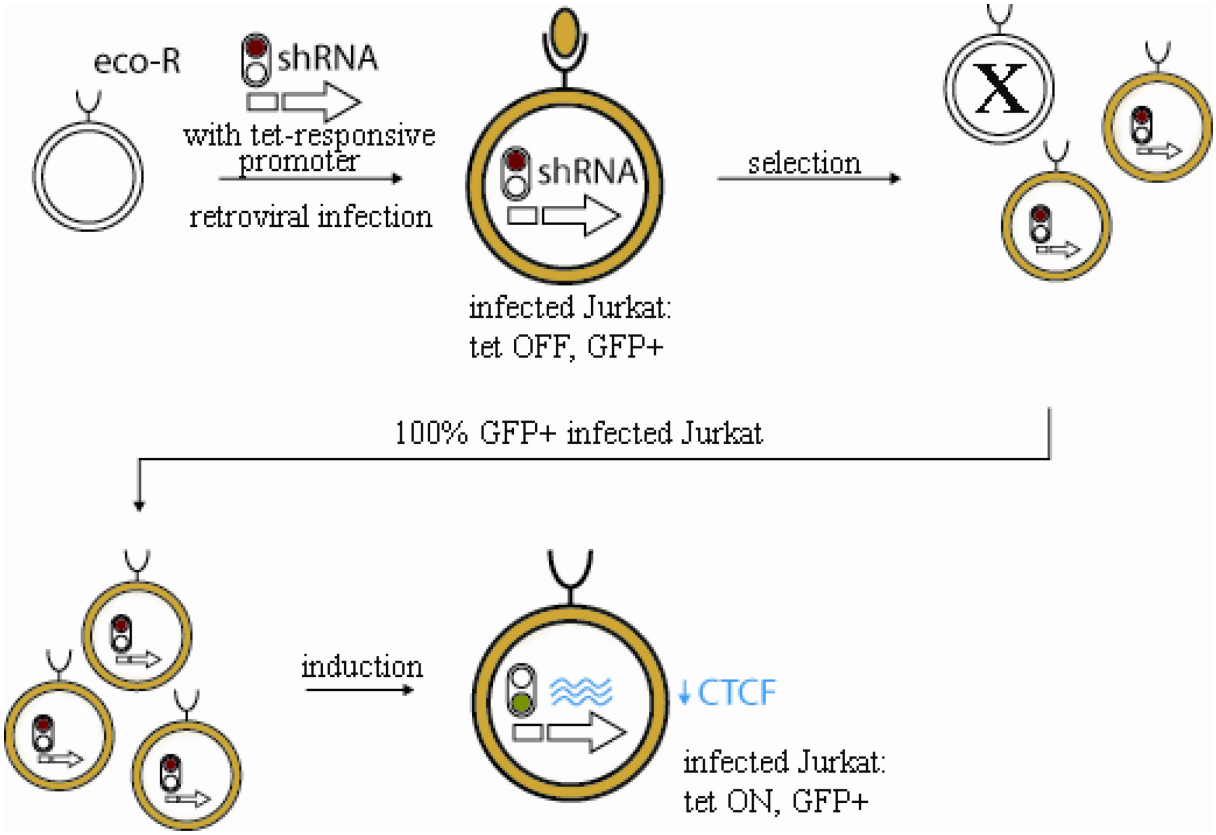


Figure 24: flow chart of TET on/off for inducible cell line. EcoR: ecotropic receptor.

Since Jurkat cells are human cells, it was first necessary to test the homologue region to the murine +71 element had insulating properties. For that, the luciferase enhancer blocking assay was performed and indeed the +94Kb element (human insulator, located 94Kb downstream the human PU.1 promoter) was ablating the SV40 enhancer function on the promoter.

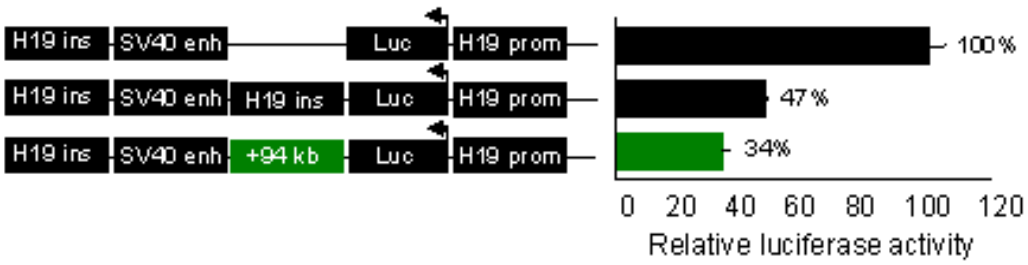


Figure 25: +94 element acts as insulator in vitro. Enhancer blocking assay indicates that the +94 element leads to a reduction of luciferase activity. Thus the +94 element is an insulator like the murine +71 homologue region.

It has been shown for several discrete regulatory elements (Wilson et al., 2011a) as well as at genome wide scale (Schmidt et al., 2012a)that conserved regions are coding for more vital function. So it is interesting that the +71 (mouse) or +94 (human) insulators exert the same functionality: this allows to switch easily between the two species for technical reasons, and also suggests that the insulating effect of these region located downstream the PU.1 promoter represents an evolutionary conserved mechanism to regulate the PU.1 gene in T cells.

After screening 12 different shRNA constructs (data not shown) for CTCF downregulation, lines #2 and #12 were chosen for the best knock down (kd) of CTCF after 72 hours of induction. Importantly, these lines did not show increased cell death in 72 hours compared to the scramble (sc4) control line: we therefore assume, that the following results are indeed CTCF-level dependent rather than apoptosis effects. In Figure 26, we confirmed the downregulation of CTCF at protein level. In these induced cells, the loop between PU.1 promoter and +94 element was measured by quantitative 3C assay, as explained in chapter 3.1.3. The chromatin loop formation decreased down to 20-40% in a reproducible way (in three biological independent experiments), strongly correlating the level of CTCF and the insulating loop.

Since CTCF plays distinct and crucial cellular functions, it is hard to distinguish whether this effect is direct or indirect. The scramble cell lines controls only for side effect of the shRNA vector, but not for a general knock-down of CTCF. In other words, is the reduction of the insulating loop formation a specific effect or are all chromatin interactions affected upon CTCF downregulation? Since “positive” and “negative” controls are impossible to be defined, this issue was addressed by measuring a loop of a housekeeping gene as positive control, in this case in the promoter of the TUBULIN gene (Hou et al., 2010), which entity didn’t

significantly change, and which protein level also doesn't change as shown in the western blot (Fig. 26A, third panel). On the other side, a region between the +94 element and the PU.1 promoter, which doesn't loop with the promoter, was chosen as negative control. Upon downregulation of CTCF this chromatin interaction was still not detectable (data not shown). Hence, the reduction of the +94/PU.1 promoter loop depends on CTCF levels, suggesting that CTCF physically mediates this chromatin interaction.

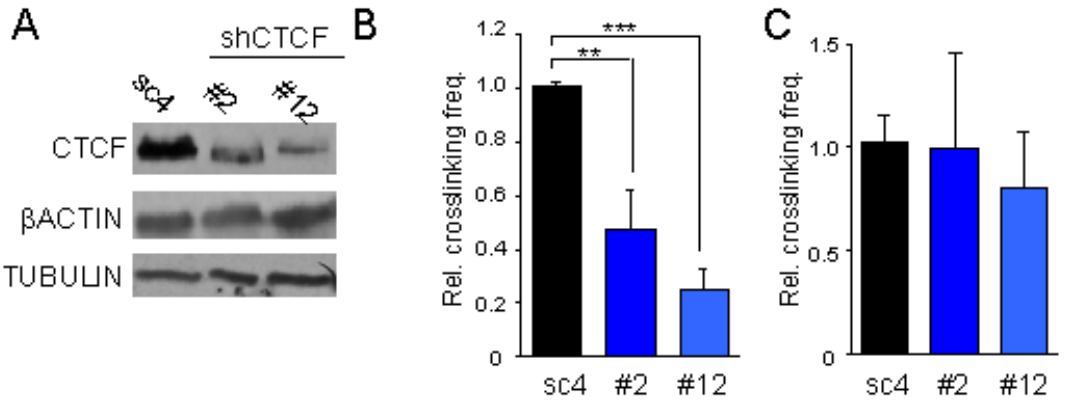


Figure 22: Loop formation in kdCTCF human cell lines. A) western blot on total extracts confirm the downregulation of CTCF at protein level compared to the tubulin loading control. B) Upon kd of CTCF, q3C on PU.1 promoter / +94 element loop is reduced. On the contrary, housekeeping chromatin loop did not vary in C).

To investigate whether the PU.1 gene is directly regulated by the chromatin loop, PU.1 expression was measured at transcript level; but, as shown in Figure 27A, mRNA level of PU.1 is not affected upon CTCF downregulation. Thus, the disruption of the insulating loop is not sufficient to activate the PU.1 locus in Jurkat cells; but several epigenetic mechanisms could occur concomitantly, and affecting only one of these (the high order chromatin structure at local level) might be not sufficient to reprogram the regulation of the gene. Supporting this hypothesis, histon acetylation did indeed increase in kdCTCF lines (Figure 27B) compared to scramble Jurkat cell line, indicating that the PU.1 promoter status is prone to be activated. Another epigenetic mechanism, which is known to maintain PU.1 gene silenced, is DNA methylation at the promoter and upstream regions. It has been shown that upon treatment of a demethylating agent (5-AZA) PU.1 promoter loses its methylated status resulting in upregulation of the PU.1 transcript (Amaravadi and Klemsz, 1999). Therefore kdCTCF lines and scramble control were treated with low concentration of 5-AZA for 36 hours, and both CTCF and PU.1 levels were measured at mRNA level (Fig. 27C and 27D): while CTCF levels were not affected, PU.1 expression rose proportionally more in the kdCTCF lines. This

result indicates that the loop formation is mediating by CTCF, that the PU.1 locus is silenced by different and synergistic mechanisms and that upon demethylation of the promoter the silencing effect of the chromatin loop in kdCTCF lines was reduced.

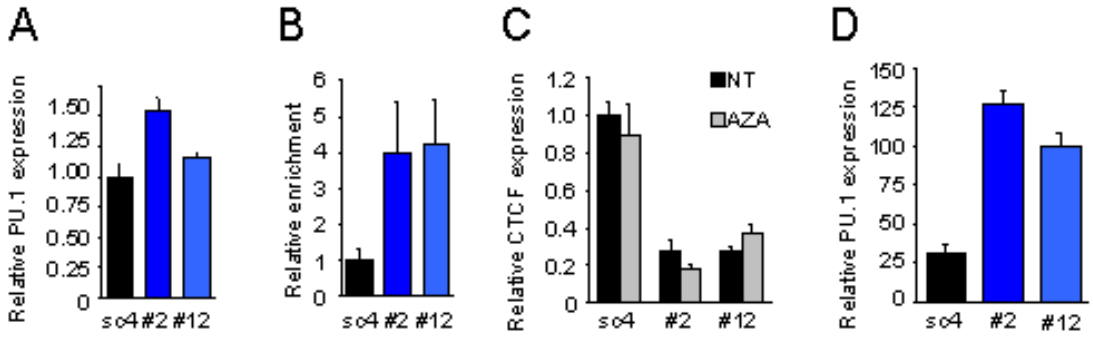


Figure 27: PU.1 expression upon CTCF downregulation. All 4 panels' results are derived from the same representative experiment, which was repeated 3 independent times. In all panels, sc4=scramble cell line, #2 and #12 are kdCTCF cell lines. A) PU.1 transcripts, relative to gapdh expression, were normalized to the scramble control. B) ChIP versus H3K9ac in indicated cell lines. Enrichment was calculated as fold over IgG control antibody. Data were afterwards normalized for the scramble's one. C) CTCF expression normalized for the not treated scramble line. NT= not treated. AZA= 36h treatment with 1μM of demethylating reagent 5_AZA. D) PU.1 mRNA normalized for the scramble cell line, not treated with AZA, as shown in panel A.

Does CTCF change its occupancy when downregulated? Figure 28A illustrates CTCF ChIP assay on kdCTCF lines. The positive control PCR demonstrated that is technically possible to detect very low level of CTCF (even in the mentioned cell lines, where CTCF is drastically downregulated); in fact, the enrichment doesn't significantly varying between scramble and kd. Notably both the PU.1 promoter and the +94 insulator are bound by CTCF in a level-fashion way. Still, the results are hard to interpret since CTCF could exploit its function by binding with different affinity, meaning that the observed outcome could have biological functions or not (Splinter et al., 2006). To support our hypothesis that the loop is directly mediated by CTCF to silence PU.1 gene, we decided to employ the previous established EL-4 cell lines for the sequestering experiment. In these cell lines we observed a reduction of the loop without changing CTCF expression; investigating CTCF binding in this setting will thus provide the direct correlation between CTCF occupancy and loop formation. We designed primers, which distinguished the endogenous promoter region from primers spanning both endogenous and transgenic copy. As control we monitor the enrichment on the mouse +71 element. Interestingly, while the binding of CTCF on the insulator remained invariant, CTCF occupancy didn't vary at the insulator level, whereas it increases in the transgenic cell lines at

the overall PU.1 promoter (endogenous plus transgenic) compared to the specific PU.1 promoter (endogenous only). These results suggest that CTCF occupancy shifted from the endogenous to the transgenic promoter region (Fig. 28B).

Taken together, in this chapter we demonstrated that the loop is mediated by CTCF, which binds to the insulator and promoter region. This chromatin loop contributes to silence PU.1 gene in T cells.

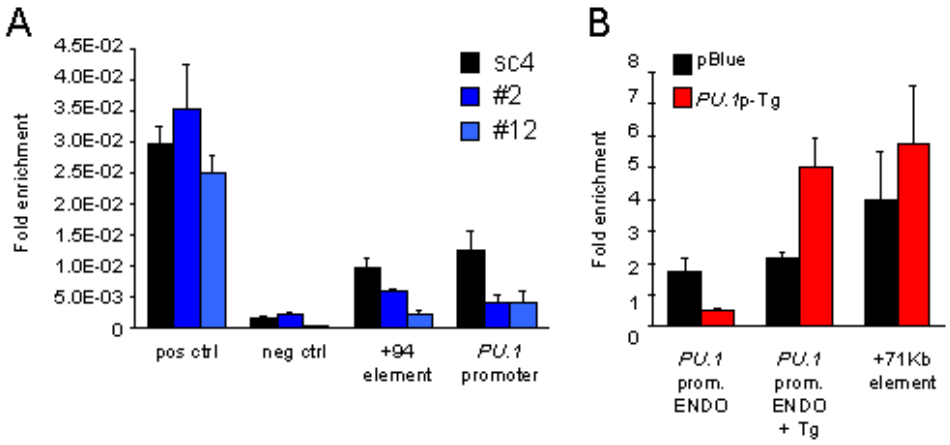


Figure 28: CTCF binding correlates with the insulating loop formation. A) CTCF ChIP on kdCTCF cell lines compared to the scramble control. The enrichment of CTCF binding is shown as percentage of the input. Error bars indicate S.D. of one representative experiment. Experiments were performed at least on three biological independent experiments. Positive and negative controls are on the MYC promoter gene and +55Kb downstream the PU.1 promoter, respectively. +94 element and promoter region are indicated. B) CTCF ChIP on mouse EL-4 cell lines comparing the mock control versus the stable cell lines containing transgenic copies of the PU.1 promoter. Enrichment is illustrated as fold over IgG negative antibody control, error bars represent S.D. between three pools each genotype of one representative experiment. Genomic region amplified by PCR are indicated: in the first lane, primers spanning only the endogenous PU.1 promoter region are used, whereas in the second primers amplifying a common region within the 2,1 Kb of he transgenic promoter and the endogenous one are indicated. The third lane represents enrichment at the murine +71 element.

3.2.3 Chromatin architect SatB1 does not mediate the insulating loop

In chapter 3.2.1 it was shown that SatB1 binds both at promoter and insulator level. SatB1 is essential for proper chromatin architecture, especially for establishing chromosome territories. Moreover, SatB1 is crucial for T cell development; its loss compromises T cell differentiation at the level of gene regulation and at nucleus architecture level (Galande et al., 2007). During T cell differentiation SatB1 undergoes to gene downregulation, so it would be interesting to follow the loop formation during T cell development. This experiment would require too

many undifferentiated cells; thus, the same inducible loss in function approach as illustrated in the previous chapter was applied for SatB1. To efficiently knock down SatB1, 18 different shRNA constructs were tested in Jurkat cells and transcript level was measured after 48 hours (data not shown). The best two cell lines were selected for further analysis, named #15 and #18, as illustrated in the western blot in Figure 29A. Upon SatB1 downregulation, confirmed both at mRNA and at protein level, PU.1 expression did not rise (Figure 29B). Also histon acetylation did not change, suggesting that knocking down SatB1 did not modulate the histon code at the promoter level (Figure 29C). For that reason, 5-AZA treatment effect is not shown here, since no differences in PU.1 levels was observed between scramble and kdSatB1 lines.

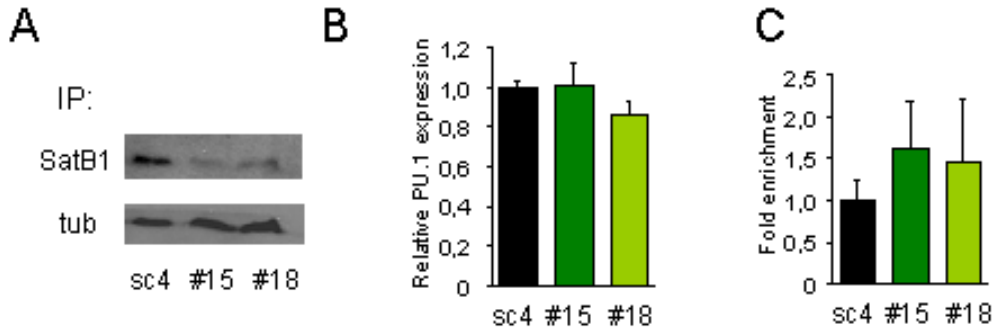


Figure 29: Validation of inducible Jurkat lines to knock down SatB1. A) Downregulation of SatB1 at protein level compared to loading control. In the upper lane IP against SatB1 is shown; in the lower one, tubulin control. Sc4 = scramble; #15 and #18 are the most effective shRNA used. B) PU.1 mRNA level in indicated cell lines. Values are calculated as ΔCt related to GAPDH levels, then normalized for the scramble control. C) ChIP against H3K9ac: enrichment is depicted as fold over IgG and then normalized to the scramble control.

Interestingly, insulating loop formation increased upon knock down of SatB1 (Figure 29A). Chromatin loop at the TUBULIN promoter was used as housekeeping control and did not significantly change, indicating that the differences observed in the +94 / PU.1 promoter loop were not an artefact of downregulation of SatB1. As mentioned before, SatB1 downregulation occurs during T cell maturation; the increase of the chromatin loop could be therefore explained as a maturation of Jurkat cells with lower level of SatB1. Since CTCF levels were also increasing upon kd of SatB1 (Figure 30C), the increased chromatin loop could be also be a consequence of higher CTCF expression. Hence, in order to understand whether SatB1 directly mediates the loop, in vivo studies would be necessary; for that purpose, a valid SatB1

antibody would be required. For all these reasons, we did not investigate the role of SatB1 further in this thesis.

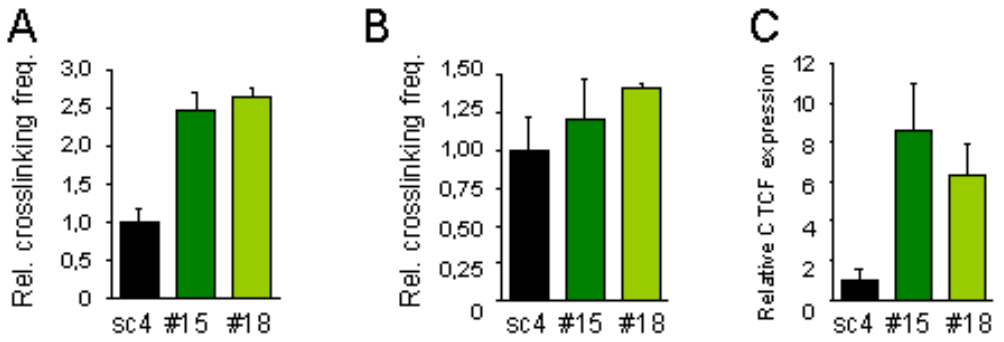


Figure 30: Loop formation upon kdSatB1. Relative crosslinking frequency is depicted for q3C assay detecting in A) the chromatin loop between +94 element and PU.1 promoter and in B) the housekeeping loop on the TUBULIN promoter in indicated cell lines. C) CTCF expression was measured at transcript level and represented as fold of the scramble control.

3.3 Acute myeloid leukemia displays the insulating loop between insulator and PU.1 promoter

In the previous chapters a lineage specific chromatin loop was described to silence the PU.1 gene in T cells, but not in B cells or myeloid cells where PU.1 is expressed at higher levels; the T cell specificity of the loop was investigated in physiological hematopoietic populations, where PU.1 is required for myeloid differentiation and its silencing is necessary for T cell development.

But what happens to the local chromatin structure in a leukemic context? It has been shown in mouse models that altering PU.1 levels in distinct lineages can lead to different leukemia. In human patients, PU.1 dysregulation occurs very often in Acute Myeloid Leukemia (AML). AMLs harbor a block in myeloid differentiation, in which leukemic blasts display lower levels of PU.1 in respect to mature myeloid cells.

This chapter investigates the AML context of PU.1 locus chromatin structure.

3.3.1 Insulating T cell specific loop adopted by AML cell lines

To test the hypothesis whether PU.1 downregulation in AML can be accomplished by the insulating chromatin loop, two AML cell lines were tested for the loop formation: HL-60 and U937, promyelocytic blasts and histiocytic lymphoma with monocyte morphology, respectively. Two epithelial cell lines (non hematopoietic) were chosen as negative controls, HeLa cells and HEK293T cells, deriving from an adenocarcinoma and embryonic kidney, respectively. As illustrated in Figure 31, PU.1 expression, insulating loop formation and in vivo CTCF binding were investigated. This were in line with the experiments of previous chapters, where we showed a strong correlation between the transcriptional silencing of PU.1 gene, the formation of the CTCF-dependent insulating loop in mature T cells.

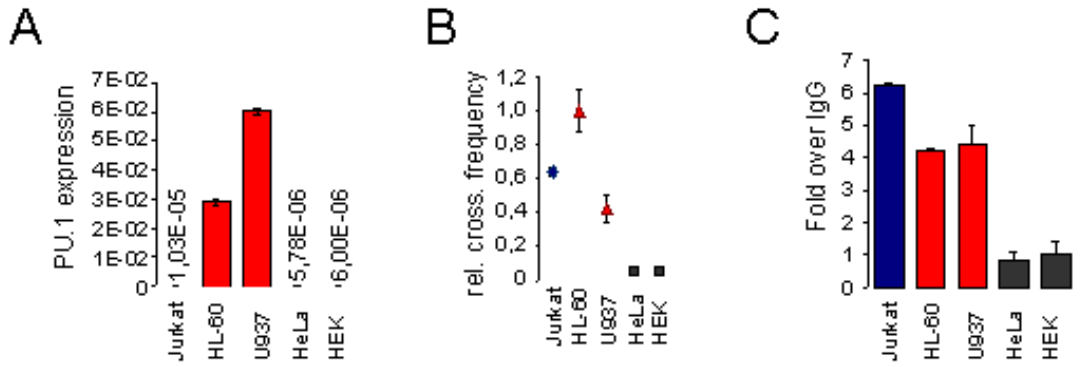


Figure 31: Investigating the insulating loop in AML cell lines. A) PU.1 mRNA levels as fold of GAPDH expression. B) q3C detecting the +94/PU.1 promoter loop on indicated cell lines depicted as crosslinking frequency. All the values were normalized for the highest level, here in HL-60 cells. C) ChIP assay against CTCF: enrichment was calculated as fold over the control antibody IgG. Primers are spanning the human insulator region.

In tested AML cell lines PU.1 expression is very high compared to not hematopoietic cell types or T cell line, but is still lower compared to peripheral blood myeloid cells (there are no macrophages in the blood, just monocytes), which is at least 10 fold higher (Fig. 31A, data not shown). Notably, the chromatin loop could be detected also in AML lines, where PU.1 regulation is altered (Fig. 31B); and interestingly, CTCF binds to the insulator region as in T cells (Fig. 31C). From these results it could be observed that in an AML context the local chromatin structure can mimic conformations specific for another lineage.

As conclusion, AML lines were a suitable model for further experiments, and within them HL-60 cells were selected as a model cell, since they harbor the chromatin loop with higher frequency and lower PU.1 levels.

3.3.2 Chromatin conformation of the PU.1 locus shifts from inactive to an active status upon induced differentiation

It has been reported that the HL-60 cell line can be differentiated into mature monocytes upon several stimuli, like phorbol myristic acid (PMA, TPA), retinoic acid and butyrate, hypoxanthine, dimethylsulfoxide (DMSO, 1% to 1.5%), actinomycin D. Here we made use of this knowledge to track the chromatin conformational change upon induced differentiation. This approach would let us monitor the local structure of the PU.1 locus when an AML cell line is forced toward terminal differentiation and therefore forced to escape from the malignant context of leukemia. HL-60 cells were thus treated for 72 hours with TPA, then only adherent cells were collected and analyzed for expression, and chromatin conformation.

Both CTCF and PU.1 level were measured as transcript (Fig. 32 A and B): while CTCF didn't vary, PU.1 transcript increased compared to not-treated cells. Thus we can conclude that the during the induced differentiation PU.1 gene regulation. In this set up, the local chromatin structure of the PU.1 gene turned from a repressive state to a more active one (Fig. 31 C and D). While the insulating chromatin loop between the insulator and promoter drastically decreased, the loop of the promoter with the upstream regulatory cluster (URC) increased.

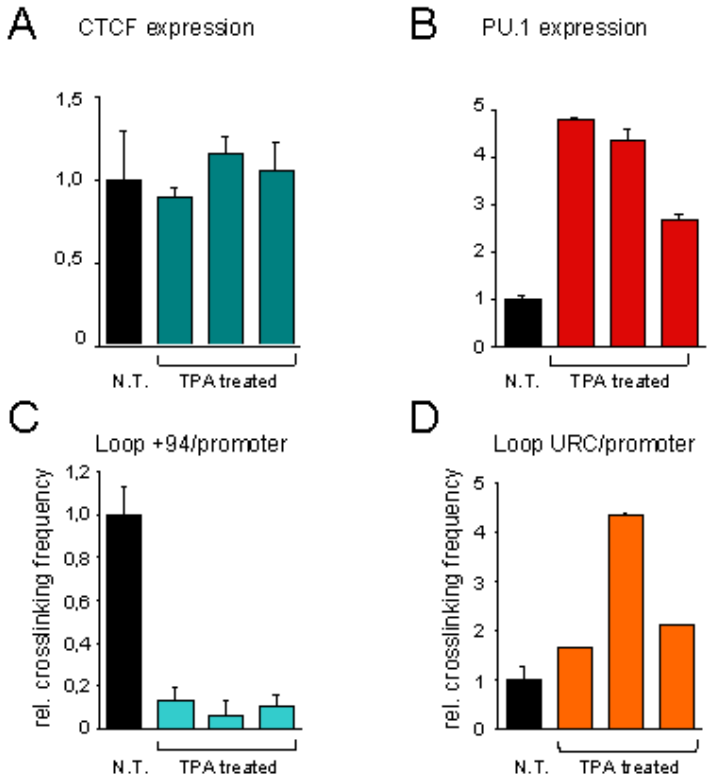


Figure 32: Chromatin characterization of HL-60 upon induced monocytic differentiation. A) cytospin stained with May/Grünewald-Giemsa after 72 hours of TPA treatment shows macrophage and granulocyte morphology compared to the solvent not-treated control (N.T.). Cells were differentiated in 3 independent pools. B) CTCF and C) PU.1 transcript as fold over GAPDH expression. Not-treated control was set to 1. D, E) q3C assay in not-treated cells and 3 independent pool of HL-60 TPA treated: on the left the repressive loop between +94 insulator and PU.1 promoter is depicted, on the right between the same promoter and the upstream cluster of enhancers (URC). Crosslinking frequency was calculated as explained in chapter 3.1.3 and not-treated control value was set to 1.

3.3.3 Chromatin insulating loop in AML patients

In the previous chapter it was shown that AML cell lines display the +94 / promoter loop, found before only in T cells. This result was a prerequisite before analyzing primary cells from AML patients. As already mentioned in chapter 1.3.2, acute myeloid leukemia are extremely variable in regard to differentiation stage block, dysregulation of oncogene/tumor suppressor genes, heterogeneous surface markers' pattern, etc. It is therefore challenging to purify only the leukemic blast from bone marrow aspirates: they do not only contain several cell types, but even more strikingly their distribution could dramatically vary from patient to patient. To overcome this issue, a negative selection approach was applied, trying to deplete the bulk of bone marrow cells from undesired lineages: red blood cells were eliminated by

basic osmolitic buffer (ACK buffer), lymphocytes were excluded by magnetic beads coupled with antibodies against pan T and B cell determinants (CD3 and CD19, respectively). The remaining cells comprised cancer blasts, but also progenitors and stem cells, even if in a very minor percentage. In order to compare bone marrow with similar distribution of blasts, only patients harboring blasts infiltration > 90% (more than 90% of the bone marrow are leukemic cells) were analyzed. To exclude false positive signals deriving from the not-leukemic population (progenitors and stem cells), stem cell were employed as undifferentiated control. As already demonstrated in chapter 3.1.3, the insulating loop doesn't take place in the stem cell compartment in the mouse species; in this chapter the human counterpart was included for the same reason. Human stem cells were purchased by Lonza manufacture as CD133+ cells: CD133 (prominin-1) is a hematopoietic stem cells marker, even more stem cell specific than CD34, which is also present at progenitors stages; cells were expanded in vitro and CD34+ cells were collected. Mature monocytes were collected from blood after erythrolysis and lymphocytes depletion; they were chosen as control for the myeloid pathway; blood was friendly donated from healthy donors from our institute.

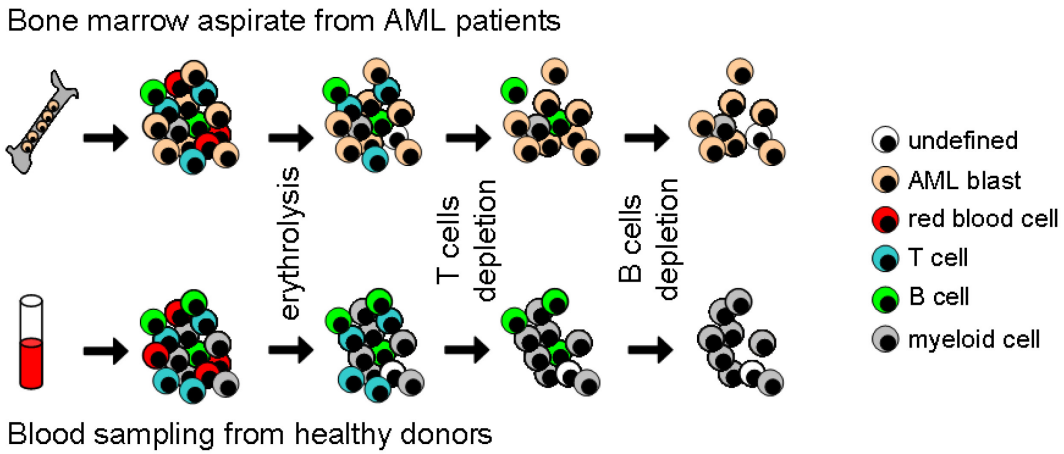


Figure 33: Graphic representation of cell collection's approach. In the upper lane AML bone marrow aspirates are depicted, in the lower blood sampling for healthy donors. At the side are illustrated the code colors reflecting different lineage. Both bone marrow and blood were treated to deplete not-myeloid lineage. Note that the percentage of blast is very high in bone marrow, but is not 100%: there could be some myeloid contamination, differentiated and not, or stem cells and progenitors. On the contrary the blood cell content from healthy humans after depletion of red blood cells and lymphocytes consists predominantly in circulating monocytes.

All collected samples were tested for two chromatin interactions, one insulating (+94 / promoter loop) and the other enhancing (URE / promoter loop) PU.1 gene by quantitative 3C; all samples underwent the initial controls as explained in chapter 3.1.3. Notably, the T cell specific loop, already detected in AML cell lines, took place also in primary AML blasts. Considering the variability of different acute myeloid leukemia, and also the heterogeneity of the leukemic population within the same patient, the distribution of all 11 analyzed samples was broad, as illustrated in Figure 34A, but strikingly higher than physiological cell types at earlier stages during hematopoietic development (stem cells,CD34+ cells) or later ones (mature myeloid cells). Since acute myeloid leukemia is characterized by a differentiation block during myeloid development, both these controls are required to state that the chromatin loop is not related to differentiation, but indeed only occurs in cancer situation. As additive control also the enhancing loop was investigated (Fig. 34B): as expected, mature myeloid cells harbor the highest frequency, CD34 cells the lowest level and AML blast an intermediate average within the patients. The reduction of the enhancing loop in AML blasts is significantly lower than the mature myeloid counterpart. Importantly, there is no correlation between the differentiation stage of leukemia subtypes and the insulating loop formation; for instance, 4 patients classified as M0-2 (undifferentiated blasts) harbored very different chromatin loop frequencies.

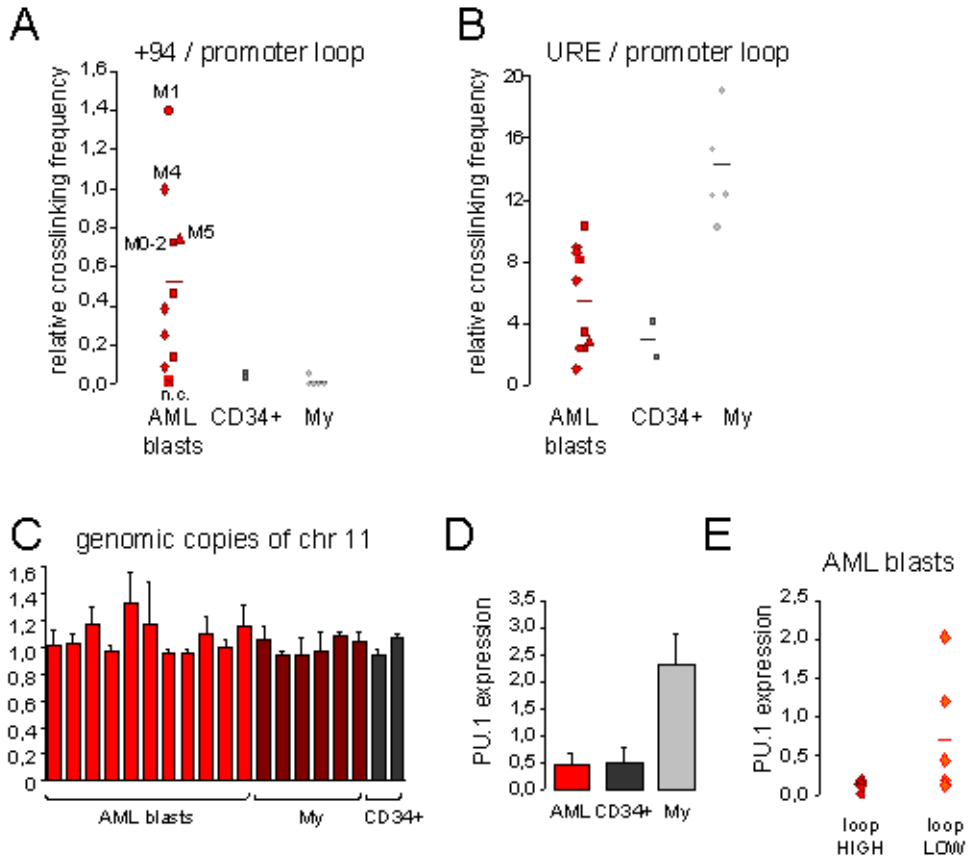


Figure 34: Insulating chromatin loop was found in AML blasts. q3C assay detecting in A) the insulating loop between +94 element and promoter and in B) the enhancing loop between the URE and the promoter. The first analyzed AML treated sample was set to 1 in both cases. 13 AML samples are displayed with different symbols for indicated subtype of leukemia (according to FAB classification). CD34+ cells were bought as CD133+, shortly expanded in vitro, and then collected for further analysis. My = myeloid cells, purified from blood sampling of healthy donors as mentioned in Figure 29. Horizontal bars represent S.D. for CD34 and S.E.M. for the other groups, where n > 5. In C) quantitative PCR was employed to assure that the chromosome 11 at PU.1 locus was not duplicated or deleted. Genomic DNA was tested with three different primer pairs, two located in the PU.1 locus on human chromosome 11 and the third on chromosome 17. ΔCt was indicated as ratio between chr. 17 signal and each of the chr. 11 – PU.1 locus signal. Therefore the value 1 represent normal diploid context (2n, two alleles per gene). DNA was extracted from all samples, patients and donors. Error bars represent S.D. D,E) PU.1 levels were measured as fold over GAPDH by RealTime PCR on cDNA. While in D) three different populations are shown (P value between AML blasts and myeloid cells is >0,001), in E) the AML blasts were separated according to the loop formation entity in loophigh (n=5) or looplow (n=6) in respect with the mean in panel A). Horizontal bars represent the mean within each group.

To exclude that the previous results were not artifacts due to macroscopic chromosomal aberrations (e.g. duplication of the PU.1 locus), copy number of PU.1 locus located on human chromosome 11 was compared to a random region on chromosome 17. Figure 34C shows that

all the patient samples as well as the healthy donors contained normal (2n) chromosome copies in the PU.1 locus.

Next, to link PU.1 expression to the chromatin conformation of the locus, PU.1 transcript levels were measured (Figure 34D); as expected, mRNA differences between mature myeloid cells and more immature cells (both blast and CD34+) are statistically significant using t student method. But, dividing the AML samples in two groups according to the insulating loop frequency (loophigh and looplow), PU.1 mRNA didn't significantly change (Figure 34E; p value = 0,148); however, in patients with high insulating loop frequency, PU.1 expression is downregulated more than in patients with less loop formation. More detailed statistical analysis with more patient samples are required to define whether the anti-correlation between loop and mRNA is significant.

Last but not least, whenever the number of collected cells allowed to perform additional assays, CTCF binding and histone acetylation status at the promoter level were investigated. In Figure 35 illustrates the outcome with the first two analyzed patients. On the left a summary of previous q3C results is depicted, in order to show the chromatin conformation of the only two patients which could be processed also for ChIP. The patient 1# manifests a M4 subtype of leukemia (acute myelomonocytic leukemia), whereas patient 9* a M0-2 (minimal differentiated acute myeloblast leukemia). Interestingly CTCF binds the PU.1 promoter in the AML which harbors the insulating loop and not viceversa; on the contrary, acetylation of H3K9 reflects the lack of the insulating loop. CD34+ cells were used as not differentiated but not leukemic cells: the PU.1 promoter is slightly acetylated and CTCF doesn't bind. These initial but very promising results suggested again that the insulating loop is present not only in a specific lineage as T cells to silence PU.1, but this confirmation is also adopted in myeloid leukemic onset, where PU.1 gene is downregulated.

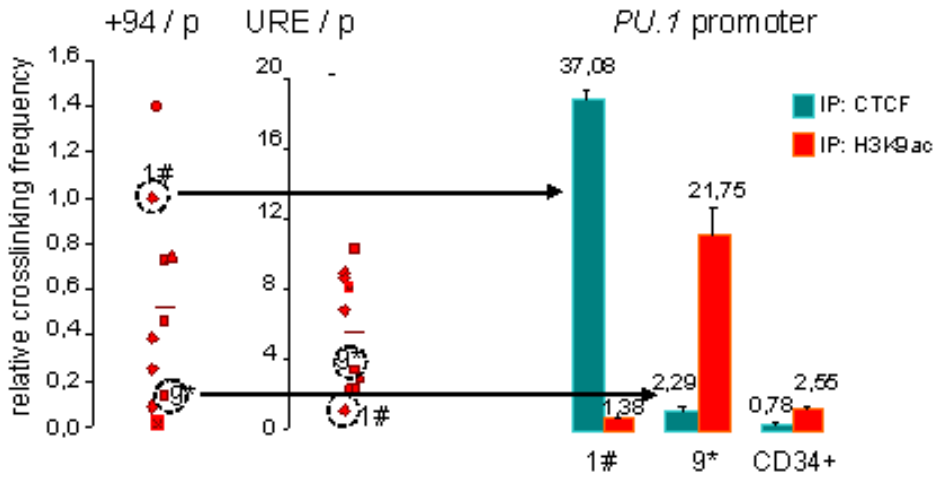


Figure 35: CTCF binding and histon acetylation was measured in two AML blasts and in CD34+ as negative control. On the right side the summary of the chromatin loops is illustrated and which samples were processed for ChIP assay. On the left enrichment of ChIP using CTCF and H3K9acetylated antibody is depicted as fold over IgG negative control. Values are indicated, error bars represent S.D.

4 Discussion

4.1 A novel insulator regulates the PU.1 gene in T-cells

The ETS-transcription factor PU.1 is a crucial lineage determinant in hematopoietic development. PU.1 is already expressed in hematopoietic stem cells at a sub-threshold low level, allowing a multilineage expression network. After reaching a critical threshold, for example by overexpression in myeloid pathway or downregulation during T-cell development, PU.1 determines commitment of multipotent progenitors to a specific lineage (Rosenbauer and Tenen, 2007; Burda et al., 2010). Thus, giving its crucial role in cell fate decision, its gene regulation has been intensively studied since its discovery (Klemsz et al., 1990). However, many aspects of how PU.1 is differently regulated in different lineages remain an open question. One of this, and thus aim of this thesis, is to clarify how *PU.1* gene is silenced during the T-cell lineage, which is a prerequisite of early thymic progenitors for correct progression of the multistep T cell differentiation (Rothenberg et al., 2008).

4.1.1 Local high-order chromatin structure of the PU.1 gene in different hematopoietic lineages

Cell-type PU.1 gene regulation is achieved by an orchestrated action of several discrete regulatory elements (Leddin et al., 2011; Zarnegar et al., 2010). However, up to now none of them could exhaustively explain PU.1 regulation in T-cells. So far, these regions were mainly identified based on two approaches: the evolutionary conservation of DNA sequences and chromatin structure features like histone modifications or DNA accessibility, interpolated by next generation sequencing coupled with either ChIP or DNase hypersensitivity (Wilson et al., 2011b; Zhang et al., 2012). Here, we chose a different approach to investigate uncovered layers of epigenetic regulation: the PU.1 promoter was screened for chromatin interactions in T-cells by genome-wide assay It is now a common perspective, that long-range interactions are directly involved in gene regulation, by which cis-acting elements can be recruited to their respective promoters (as enhancers, silencers or insulator, see 1.1.2). As shown in Table 3, intra-chromosomal interactions were detected with higher frequency (de and de, 2012), supporting the theory that most chromatin contacts occur within the same chromosome. It is worthwhile to notice, however, that also the inter-chromosomal interactions were differently located in T-cells (on chromosome 17) in respect to macrophages (chromosome 1), suggesting

that chromosomal territories are distributed in a cell-type specific fashion (Felsenfeld and Dekker, 2012). Importantly, we could confirm that the URE contacts the PU.1 promoter in macrophages, confirming by a genome-wide approach that this region is the dominant enhancer in the myeloid pathway (Ebralidze et al., 2008). MEF here served only as negative control: PU.1 expression is restricted within the hematopoietic system, therefore is absent in MEF. The fact that no interactions could be detected within the PU.1 locus in MEF suggested that in no hematopoietic cell type context PU.1 locus is most likely silenced by more global high-order epigenetic mechanisms, as the heterochromatin spreading. In this case the locus would be inaccessible for a proper digestion during the 3C assay, thus not forming any long-range interactions.

Next we analyzed the chromatin structure in committed cells in more detail, in which PU.1 is differently expressed: T-cells, B-cells and Macs representing low, intermediate and high level of PU.1, respectively. We further confirmed the putative chromatin interaction along the PU.1 locus in a quantitative fashion by quantitative 3C (q3C), dissecting 100Kb around the PU.1 gene body and analyzing also intermediate genomic location. As illustrated in Fig. 15, the local high-order chromatin structure of PU.1 reflects its differential expression, showing chromatin proximity of the URE and neighbor enhancers where PU.1 is expressed (Macs and B-cells), whereas the only interaction with the PU.1 promoter in T-cells involves the +71 element. Only this interaction could be detected along the PU.1 locus, suggesting that the chromatin assume a loop conformation.

PU.1 must be actively silenced in thymic progenitors, meaning that it has to be downregulated from a basal expression level in HSCs. Thus, to prove that the +71/PU.1 promoter loop is an active epigenetic mechanism, it is important to exclude its presence in more undifferentiated stages. Since early thymic progenitors are a rare population, the required number of cells for q3C could not be obtained. Here, we decided to take HSCs as multipotent cell population, which is even more undifferentiated than thymic progenitors. We only analyzed the two most prominent loops, the URE and the +71 with the promoter:indeed HSCs did not show high frequency of any above mentioned loops, supporting the idea that the loop observed in mature T-cells is established during differentiation. Curiously, in HSCs the loop formation was detected at basal level, while in MEF it was absent. It is therefore tempting to speculate that in the hematopoietic stem cell compartment PU.1 chromatin structure is bivalent, where both enhancing and repressing loop are present at basal level. It would be very interesting to

analyze the chromatin structure in CLPs (common lymphoid progenitors) as intermediate differentiation stage between HSCs and thymocytes. Taken into account that the PU.1 gene is transcribed at low level in HSCs, this suggests that not only histone modifications, but also chromatin conformation could determine the multilineage potential of a priming transcription factor.

4.1.2 The +71 insulator silences PU.1 expression by chromatin looping

In order to evaluate whether the *PU.1 promoter*-interacting DNA fragments have a functional role in *PU.1* gene regulation, we performed *in vitro* reporter assay. We chose two regions which undergo intra-chromosomal interaction, locating both in proximity with the PU.1 promoter (-25 Kb and +71 element respectively). Since our goal was to identify regulatory element of PU.1 gene in T-cells, we tested the above mentioned elements for insulating or repressive function. For that purpose we made use of the reporter assay previously established by Ishihara and colleagues (Ishihara et al., 2006). The -25Kb region was excluded from further investigation, because it couldn't affect the luciferase transcription in the tested reporter assay; long-range interactions can in fact be just structural, hence chromatin interactions studies have always to be carefully interpretated and supported by functional assay (de and de, 2012; Splinter et al., 2006). The +71 element showed instead canonical properties of insulator, being position dependent but orientation independent. We therefore demonstrated that the +71 element is a potential enhancer blocking insulator (Bell et al., 2001).

The next question which arose was whether the +71 element represses PU.1 expression *in vivo* by chromatin looping. To address this crucial point, we first demonstrated that the chromatin interaction is DNA-sequence specific regardless the genomic context by making use of a transgenic mouse strain carrying the human *PU.1* locus, but lacking the homologue region corresponding to the +71 element, in an endogenous $PU.1^{KO}$ background. Importantly, we have previously shown that the human counterpart can efficiently rescue the endogenous murine PU.1 loss in the hematopoietic system (Leddin et al., 2011). In this thesis, we were able to prove the murine +71 element to co-localize with the transgenic human *PU.1* promoter. Intriguingly, this finding indicates that both regulatory elements (+71 element and *PU.1* promoter) are able to undergo interactive cross-binding irrespective of their genomic location, a phenomena which has been shown before for regulatory elements controlling expression of odorant receptor genes in neurons or expression of the interferon γ gene in

immune cells (Apostolou and Thanos, 2008; Lomvardas et al., 2006). Next, we generated transgenic T cell lines carrying exogenous copies of the *PU.1* promoter, aiming to disrupt the endogenous loop by sequestering the insulator region by the transgenic promoter copies. First, we validated the system by detecting the long-range interaction between the endogenous insulator and the transgenic copies of the promoter. Then, we were able to detect a reduction of the endogenous loop, concomitantly with an increased PU.1 expression, thus demonstrating that the chromatin loop is responsible for maintaining the *PU.1* locus silenced. By this assay we demonstrated the biological function of the chromatin loop as regulator of *PU.1* transcription.

In conclusion, in a murine cell line context, a forced disruption of the endogenous loop is sufficient to de-repress the *PU.1* gene. The increased transcription was demonstrated by a histone acetylation rise on the promoter as well as at transcript level.

4.2 The +71 insulator function involves binding of chromatin organizers

We discovered a novel insulator silencing PU.1 specifically in T-cells by chromatin looping. We next questioned which factors mediate this loop: several proteins are responsible to establish long-range interactions, determining a controlled high-order chromatin structure, at global level for spatial nuclear organization and at local level regulating gene expression. Within the insulator region the predicted binding sites of two of these chromatin organizer, CTCF and SatB1 (Galande et al., 2007; Phillips and Corces, 2009), were found and further confirmed by gel shift assay. Moreover, *in vivo* ChIP assay showed that both CTCF and SatB1 binding on the +71 element is restricted to T-cells.

4.2.1 CTCF mediates the insulator function of the +71 element

CTCF binding site is located within the insulator region, identified as core by luciferase reporter assay (see chapter 3.1.2). Thus, CTCF binding site was deleted in the construct employed in the enhancer blocking luciferase assay; the mutation of CTCF binding site led to a complete ablation of the repressing function on the luciferase gene, indicating that CTCF binding is necessary to confer to the +71 element the enhancer blocking insulating function. It is important to note that previously reported ChIP-seq experiments detected high frequency

CTCF binding throughout the genome(Barski et al., 2007; Cuddapah et al., 2009; Jothi et al., 2008; Kim et al., 2007; Schmidt et al., 2012b). It is an ongoing challenge in science to understand which binding sites have a direct function on gene-specific regulation and which are involved in maintaining the high order chromatin structure, sometimes being sufficient but not necessary. An striking example in this direction is represented by the ß-globin conformation locus during erythroid differentiation: the ß-globin gene must be induced, and CTCF was found to establish several differentiation-stage specific long-range interactions. Albeit many of them were altered upon loss of CTCF by knockout targeting, only few of them were responsible for the transcriptional regulation (Splinter et al., 2006).

To investigate the biological function of CTCF in the insulation loop formation, we generated a CTCF inducible knock-down system in Jurkat cells. Here, we detected a reduction of the loop formation and a concomitant increase of histone acetalytion (H3K9ac) status on the *PU.1* promoter. H3K9ac is a hallmark of active transcription, but PU.1 transcript's level remained invariant in kdCTCF lines. However, it is known that the *PU.1* promoter is maintained in a silenced state by DNA methylation (Amaravadi and Klemsz, 1999); it had also been shown that the 5-AZA demethylating agent can de-repress PU.1 transcription by erasing its promoter methylation pattern. Taken this knowledge into account, we combined the downregulation of CTCF with the 5-AZA treatment, affecting in this way both chromatin looping and methylation of the *PU.1* promoter. By this approach, PU.1 upregulation in kdCTCF lines was more profound, indicating that the epigenetic regulation is achieved by a coordination of different mechanism, as histone modification, chromatin conformation and DNA methylation. In conclusion, we showed that the loop formation is CTCF-dependent based on a loss-in-function approach.

It had been shown that CTCF binding can be methylation sensitive (Hark et al., 2000), however, CTCF is very versatile in accomplishing different roles, and the methylation inhibition on CTCF binding is not a universal feature: in fact, was mostly restricted to the sequences where CTCF plays a role in regulation of imprinted loci or during X inactivation (Ohlsson et al., 2010). Instead, in the two T-cell lines used in this thesis (EL-4 in chapter 3.1.3 and Jurkat in chapter 3.2.2) this effect could not been observed for two main reasons. First, the insulating loop formation could occur in the physiological cellular context, where the *PU.1* promoter region is methylated. Second, 5-AZA was sufficient for *PU.1* derepression in both lines, and also necessary for Jurkat cells; on the other hand, the reduction of the

insulating loop was sufficient for *PU.1* reactivation in EL-4, but not in Jurkat cells. Thus there is no direct correlation between methylation status of the promoter and CTCF occupancy. Third, treating EL-4 transgenic and mock cells with 5-AZA affected PU.1 expression, but not the entity of the loop disruption (data not shown). Taken together, our data strongly suggested that DNA methylation and CTCF-dependent long-range interactions are two independent mechanisms, whose cooperation result in a coordinated function on silencing *PU.1* expression.

To gain new insights whether the CTCF mediation on this chromatin loop is dependent on direct binding of CTCF to the Pu.1 promoter, we used transgenic cell lines overexpressing the *PU.1* promoter region. In this cellular context CTCF levels don't vary, but the endogenous loop is reduced being sequestered by an exogenous promoter copy. This is the best model in our hands to determine the occupancy change of CTCF. Indeed, CTCF binding was decreased at endogenous promoter level, while remained invariant on the +71 insulator, which strongly correlates CTCF binding and loop formation. Therefore we suggest that an induced disruption of the endogenous insulating loop causes a shift in CTCF binding from the endogenous promoter to the exogenous one.

Taken these findings together, we could demonstrate that CTCF directly mediates the insulating loop, which regulates PU.1 expression in a coordinated fashion with other epigenetic mechanism, as DNA methylation.

4.2.2 SatB1 binds the +71 element but is not necessary for insulation activity

The SatB1 binding site identified on the +71 element does not locate in the core region responsible for insulator function. However, taking into account that SatB1 is not that ubiquitous as CTCF, it was tempting to study whether CTCF and SatB1 could cooperate, and if so, whether this cooperation would confer T-cell specificity of PU.1 gene regulation. On the other hand, SatB1 was shown to bind the URE region in myeloid progenitors, mediating the respective enhancing function (Steidl et al., 2007). In conclusion, little is known about the role of SatB1 in gene regulation in hematopoiesis, motivating us to investigate the function of SatB1 T-cell specific binding in the *PU.1* locus. Thus we first screened the perturbation of the insulating loop by applying the same approach as for CTCF.

Upon SatB1 downregulation, no significant changes were detected neither on PU.1 expression levels, nor on the histone acetylation, and the loop formation significantly increased. Hence, SatB1 is not responsible to directly repress *PU.1* transcription. But these results are controversial to be analyzed: first, CTCF was upregulated, so the increasing of loop formation could be an indirect effect of the increased levels of CTCF; second, SatB1 down-regulation is necessary for T cell maturation, therefore our results could reflect a more mature status of *PU.1* gene. All these open questions are intriguing, but they have not been investigated further in this thesis due to time constrains.

4.3 AML blasts adopt the T-cell specific insulating loop conformation

PU.1 aberrant expression is observed in different hematopoietic malignancies. PU.1 is overexpressed in erythroleukemia (Kosmieder 2005), its downregulation leads to AML in mouse models (DeKoter, 2006?, Rosenbauer, 2004); lack of silencing PU.1 causes T-ALL (Rosenbauer, 2006). Taken together, these studies performed on mouse models indicate that PU.1 role is dosage-dependent and lineage-specific not only as master lineage determinant in normal hematopoietic development, but also as tumor suppressor in leukemia. In human leukemia, point mutations are found in *PU.1* promoter and gene body with very rare probability (Bonadies et al., 2011; Mueller et al., 2002); one SNP in the URE region could be involved in the leukemic progression of complex karyotype AMLs (Steidl, JCI 2007). Rather than mutations in the DNA sequence, PU.1 had been found to be blocked by other mechanisms: it is a target of oncogenic fusion protein in AML (AML1-ETO, Flt3-ITD, PML-RARA) (Gilliland et al., 2004), PU.1 promoter acquires an aberrant methylation pattern in CML, especially evident during the blast crisis.

It is clear that epigenetic plays a major role in gene regulation of master transcription factors as PU.1 in hematopoietic differentiation as well as in leukemia development. Therefore we investigated whether the T-cell specific conformation we identified in this thesis could also be present in the myeloid leukemic onset, where PU.1 is often downregulated. This hypothesis is based on the expression levels of PU.1, which are silenced both in mature thymocytes and myeloid blasts, respectively; this would reveal that the repressing loop between *PU.1* promoter and insulator is a specific mechanism of PU.1 gene regulation, occurring when PU.1 transcription has to be inhibited. Here we show that AML (acute myeloid leukemia) cell lines

harbor the chromatin loop and CTCF occupancy at similar levels found in Jurkat, T-cell lines, suggesting that also in an AML context the insulating loop is CTCF-mediated. As control, two not hematopoietic lines were analyzed and don't show neither the same conformation nor CTCF binding, suggesting, as for MEF in the mouse counterpart, that the *PU.1* locus is here shut down likely by other more permanent epigenetic events, like heterochromatin spreading. To our knowledge, this is the first evidence of a long-range chromatin interaction, which can mark a specific gene for transcriptional silencing and can be adopted in an opposite lineage in a malignant context.

4.3.1 The insulating loop is disrupted during differentiation of leukemic blasts

To test the biological function of the insulating loop found in AML blast line, we decided to induce differentiation of these cells and observe the chromatin changes during this process. HL-60 cells derives from an APL-leukemia (acute promyelocytic leukemia), characterized by the t(15;17) translocation, causing the formation of the PML-RARα oncogene (Rizzo et al., 1998), which traditionally correspond to the FAB subtype M3 (even if Dalton WT Jr and colleagues reporter in 1988 that HL-60 represent an M2 subtype, (Dalton, Jr. et al., 1988; Mueller et al., 2006a)). For our purpose, anyway, the detailed classification is not crucial, since we try here to identify a signature for AMLs unrelated to their differentiation stage; what was relevant for this thesis, is the property of HL-60 to be differentiated in mature monocytes upon phorbol ester (TPA) treatment. Thus HL-60 line represented a model to observe epigenetic changes when the cells escape from the leukemic program to establish a terminal differentiation cell fate. The chromatin conformation switched from a repressive state of PU.1, harboring the insulating loop, to an active state, marked by an increased loop with the URE enhancing region. This result strongly correlates with the previously analyzed populations in 3.1.3, confirming that within the same system PU.1 expression could be reflected by monitoring the two determinant loops of the *PU.1* promoter, the insulating and the enhancing one. The current leading treatment of APL leukemia is *all-trans* retinoic acid (ATRA), which functions as the TPA by displacing the fusion protein PML-RARα from its targets. It had been shown that treating primary APL blast *ex vivo* with ATRA is sufficient to restore PU.1 expression and drive differentiation. Hence we speculate that one of the molecular mechanisms underlying this therapeutic effect could be the modulation of the long-range interactions from the insulator to the URE enhancer.

4.3.2 PU.1 expression block is associated with the insulating loop in AML patients

AML is a very heterogeneous disease, and different subtypes present drastically different prognostic prediction. The FAB classification was mainly based on morphological and cytochemical characteristic, integrating also genetic aberration such as chromosomal translocations. However, morphologic and genetic features don't always correlate, or the molecular defects underlying cannot be identified. It is thus required to consider different levels to provide a more reliable classification, to close to gap between (epi)genome and phenotype. Current efforts are focusing on the application of genome-wide techniques to discovery new markers and tool for prognostication and prediction of response to therapy. In fact, apart from rare cases, such as PML-RARα in M3, where therapeutic drugs could be directed against one main oncogene, usually leukemia are treated by a combination of therapies, as allogenic bone marrow transplant and chemotherapy. Recently, epigenetic drugs are used in clinical treatment, as demethylating agents (citarabine) or inhibitors of histone deacetylase (for instance, tricostatin A), after astonishing results during the clinical trials (Estey, 2012a). It is therefore still challenging to investigate how genes are dysregulated and by which mechanisms, including epigenetic ones, can be involved in leukemogenesis.

We therefore started to analyze primary AML blasts, asking whether the insulating loop we observed *in vitro* in an APL cell lines could also be detect in human patients. We collected blasts from 10 patients and compared them with 2 independent samples of CD34+ as undifferentiated progenitors and 5 samples of human myeloid cells, including monocytes and granulocytes. Intriguingly, the AML blasts assume the T-cell specific chromatin conformation in a wide distribution fashion. Notably, the loop formation doesn't correlate with the FAB classification's differentiation stage; in progenitors and mature myeloid cells, the loop was very low or absent, respectively. On the other hand, the levels of the *PU.1*/URE interaction anti-correlate with the insulating loop in AML patients; this difference is even more strikingly in mature monocytes, where the enhancing loop is very pronounced. We also confirmed that PU.1 transcripts is significantly lower than mature monocytes and is comparable to CD34+ cells, confirming a common knowledge of PU.1 expression in stem cells and overall leukemic blast (Zhu et al., 2012). The number of cells and of patients doesn't allow us to perform statistically significant correlation studies between the chromatin loops and PU.1 expression. For the same reason CTCF occupancy and histone acetylation status of the *PU.1* promoter

could have been investigated only in two patients, one harboring high levels and the other low ones of the insulating loop; the results of this pilot experiment again suggested the role of CTCF as mediator of the loop. On the other hand the acetylation status was inversed represented, indicating that the loop correlates with an inhibited state of transcription. However, taken in account the heterogeneity of the AMLs, these conclusions remain speculations due to limited number of the samples.

In addition to limitations in collecting samples and in blast numbers obtained per patient, it is worthwhile to remind that the blasts must be considered heterogeneous (Dick, 2008) also within the same leukemia and that 3C analysis are performed on the whole population. Therefore these results, as every q3C outcome, have to be interpreted with particular care (de and de, 2012) when aiming of establishing a correlation between different chromatin conformations (here, the insulating and the enhancing loop): it is easy to compare loops levels in different populations when one of them doesn't harbor them, like in the case of mature monocytes. Hence, it was possible to determine the threshold of the insulating loop to 0, and consequently to state that the levels in AML are consistently relevant. But the difficulty increases when samples present intermediate levels of loop formation, such as for the enhancing loop: here we could only assume that the basal level of this loop is represented by CD34+. Albeit these complications, by the analysis of two separate loops within the same locus, which often anti-correlating one to another, our results clearly demonstrated the coexistence of two opposite loops only in AML onset.

In this thesis, the chromatin structure description aimed to enlighten another layer of PU.1 gene regulation, therefore it would be interesting to quantify the association between local structure and PU.1 expression in AML onset and physiological one. In this perspective, it would be risky to correlate each loop with PU.1 expression; more informative would be instead, to first investigate statistically the relations between the two chromatin loops, for which we speculate in an anti-correlation, and subsequently associate the "local chromatin structure" score to the PU.1 transcript level. However, for such statistically based investigations the required number of samples should be incredibly higher. Moreover, the complicated PU.1 gene regulation and its dysregulation in leukemia can only be achieved by a coordination of several epigenetic mechanisms, as histone modification and DNA methylation. Anyway, these results demonstrated the need to include chromatin conformation studies in the epigenetic investigation. In fact, here we provide the first evidence how the

chromatin structure of PU.1 can assume in AML context the long-range interaction of an opposite lineage.

4.4 A conclusive model

The ETS-transcription factor PU.1 is one of the master driving differentiating gene in the hematopoietic system. PU.1 dynamic and lineage-specific expression underlies its versatility in lineage specification and cell fate decision; beyond that, a tight and coordinated transcriptional regulation is required. In this thesis we described an uncovered layer of *PU.1* gene regulation, the local chromatin structure. First, we identified a CTCF-dependent T-cell specific insulator, which contributes to inhibit PU.1 transcription during T-cell commitment and specification. Moreover we proposed that AML blasts silence *PU.1* gene by mimicking the same chromatin loop, perturbing a myeloid-specific chromatin structure.

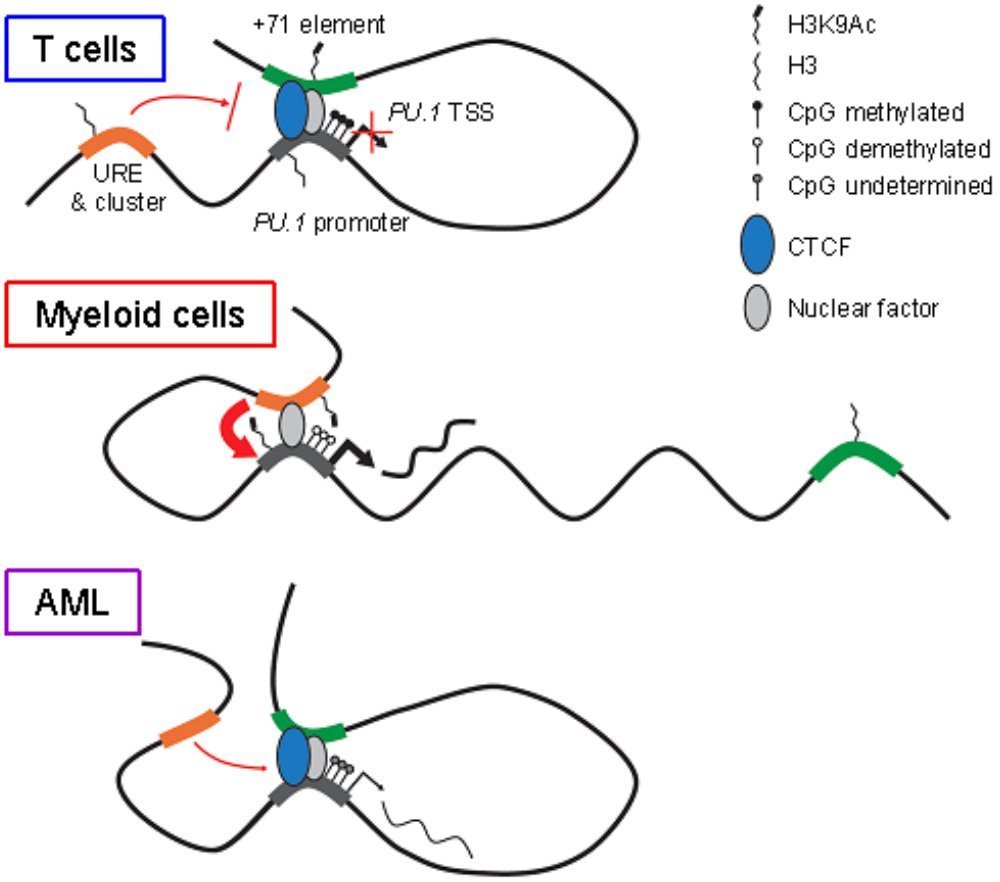


Figure 36: Local chromatin structure of the PU.1 gene in opposite hematopoietic lineages and its perturbation in AML condition. Schematic representation of the proposed role of chromatin conformation on PU.1 gene regulation: H3 states for not-acetylated-histone; CpG are known to be methylated and unmethylated in T cells and macrophages, respectively. In the AML situation are named undetermined, because of the heterogeneity of AML disease. We demonstrated that the loop is mediated by CTCF binding; however, it is not to be excluded, that other factors are involved, directly or indirectly.

4.5 Perspectives

In this thesis a novel molecular mechanism which regulates *PU.1* silencing in T cells was identified; next, the hypothesis whether those could give new insight in PU.1 block in AML was investigated. These findings opened very important consequent questions.

First of all, a subject for an ongoing project consists to validate the CTCF role for the insulating loop formation in AML lines. We therefore planned to apply the same approaches performed in T cell line also in AML context *in vitro*: insulating and enhancing PU.1 loops will be monitored upon knocking-down CTCF and by sequestering the endogenous loop by the mean of an exogenous *PU.1* promoter. In both contexts DNA methylation contribution will be investigated by citarabine treatment.

Moreover, it would be interesting to increase the number of primary AML patients, screening them for *PU.1* locus long-range interaction, PU.1 expression levels and CTCF occupancy; by that, statistical approaches could be performed.

The correlation of these studies could be potentially developed a novel platform to better define leukemia classification, therefore improving prognostic prediction and *ad hoc* therapies. As long term goal, the *PU.1* gene could represent one reference gene for integrating current genome-wide associations (GWAs) studies with epigenetic signatures.

Another aspect which was not conclusively investigated is the role of SatB1 in regulating the insulating loop is the speculating interplay between the chromatin organizers CTCF and SatB1. Their apparent opposite function in mediating and inhibiting the insulating loop leads to speculate on the relevance of the interplay between the two chromatin organizers, both at spatial nuclear architecture level and at local high-order chromatin structure. It would be very informative to then study the perturbation of such potential coordinated function in a neoplastic context. It had been shown that loss in function of CTCF can promote tumor progression (Witcher and Emerson, 2009), whereas it is suggested than an overexpression of SatB1 is associated with different malignancies, as some type of breast or colon cancer

(Kohwi-Shigematsu et al., 2012). But, their role in cancer remains elusive. In a preliminary study, we could show that CTCF and SatB1 co-occupy the insulator. Hence, it is fascinating to investigate whether they interact at protein or DNA level, and if so, if they cooperate in a synergistic, redundant or antagonist manner.

Bibliography

Adolfsson,J., Mansson,R., Buza-Vidas,N., Hultquist,A., Liuba,K., Jensen,C.T., Bryder,D., Yang,L., Borge,O.J., Thoren,L.A., Anderson,K., Sitnicka,E., Sasaki,Y., Sigvardsson,M., and Jacobsen,S.E. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell *121*, 295-306.

Akashi,K., Traver,D., Miyamoto,T., and Weissman,I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature *404*, 193-197.

ALLFREY,V.G., FAULKNER,R., and MIRSKY,A.E. (1964). ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. Proc. Natl. Acad. Sci. U. S. A *51*, 786-794.

Alvarez,J.D., Yasui,D.H., Niida,H., Joh,T., Loh,D.Y., and Kohwi-Shigematsu,T. (2000). The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. Genes Dev. *14*, 521-535.

Amaravadi,L. and Klemsz,M.J. (1999). DNA methylation and chromatin structure regulate PU.1 expression. DNA Cell Biol. *18*, 875-884.

Anderson,M.K., Weiss,A.H., Hernandez-Hoyos,G., Dionne,C.J., and Rothenberg,E.V. (2002). Constitutive expression of PU.1 in fetal hematopoietic progenitors blocks T cell development at the pro-T cell stage. Immunity. *16*, 285-296.

Andersson,R., Enroth,S., Rada-Iglesias,A., Wadelius,C., and Komorowski,J. (2009). Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. *19*, 1732-1741.

Apostolou,E. and Thanos,D. (2008). Virus Infection Induces NF-kappaB-dependent interchromosomal associations mediating monoallelic IFN-beta gene expression. Cell *134*, 85-96.

Atchison,M.L. (1988). Enhancers: mechanisms of action and cell specificity. Annu. Rev. Cell Biol. *4*, 127-153.

Banan,M., Rojas,I.C., Lee,W.H., King,H.L., Harriss,J.V., Kobayashi,R., Webb,C.F., and Gottlieb,P.D. (1997). Interaction of the nuclear matrix-associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene. J. Biol. Chem. *272*, 18440-18452.

Baniahmad,A., Steiner,C., Kohne,A.C., and Renkawitz,R. (1990). Modular structure of a chicken lysozyme silencer: involvement of an unusual thyroid hormone receptor binding site. Cell *61*, 505-514.

Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I., and Zhao,K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Behre,G., Whitmarsh,A.J., Coghlan,M.P., Hoang,T., Carpenter,C.L., Zhang,D.E., Davis,R.J., and Tenen,D.G. (1999). c-Jun is a JNK-independent coactivator of the PU.1 transcription factor. J. Biol. Chem. *274*, 4939-4946.

Bell,A.C. and Felsenfeld,G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482-485.

Bell,A.C., West,A.G., and Felsenfeld,G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.

Bell,A.C., West,A.G., and Felsenfeld,G. (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. Science *291*, 447-450.

Bell,O., Tiwari,V.K., Thoma,N.H., and Schubeler,D. (2011). Determinants and dynamics of genome accessibility. Nat. Rev. Genet. *12*, 554-564.

Bergman,Y., Fisher,A., and Cedar,H. (2003). Epigenetic mechanisms that regulate antigen receptor gene expression. Curr. Opin. Immunol. *15*, 176-181.

Bird,A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. *8*, 1499-1504.

Bonadies,N., Foster,S.D., Chan,W.I., Kvinlaug,B.T., Spensberger,D., Dawson,M.A., Spooncer,E., Whetton,A.D., Bannister,A.J., Huntly,B.J., and Gottgens,B. (2011). Genome-wide analysis of transcriptional reprogramming in mouse models of acute myeloid leukaemia. PLoS. One. *6*, e16330.

Bonadies,N., Neururer,C., Steege,A., Vallabhapurapu,S., Pabst,T., and Mueller,B.U. (2010). PU.1 is regulated by NF-kappaB through a novel binding site in a 17 kb upstream enhancer element. Oncogene *29*, 1062-1072.

Bowers,S.R., Mirabella,F., Calero-Nieto,F.J., Valeaux,S., Hadjur,S., Baxter,E.W., Merkenschlager,M., and Cockerill,P.N. (2009). A conserved insulator that recruits CTCF and cohesin exists between the closely related but divergently regulated interleukin-3 and granulocyte-macrophage colony-stimulating factor genes. Mol. Cell Biol. *29*, 1682-1693.

Boyer,L.A., Plath,K., Zeitlinger,J., Brambrink,T., Medeiros,L.A., Lee,T.I., Levine,S.S., Wernig,M., Tajonar,A., Ray,M.K., Bell,G.W., Otte,A.P., Vidal,M., Gifford,D.K., Young,R.A., and Jaenisch,R. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Brookes,E. and Pombo,A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. EMBO Rep. *10*, 1213-1219.

Burda,P., Laslo,P., and Stopka,T. (2010). The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. Leukemia *24*, 1249-1257.

Burute,M., Gottimukkala,K., and Galande,S. (2012). Chromatin organizer SATB1 is an important determinant of T-cell differentiation. Immunol. Cell Biol.

Cai,S., Han,H.J., and Kohwi-Shigematsu,T. (2003). Tissue-specific nuclear architecture and gene expression regulated by SATB1. Nat. Genet. *34*, 42-51.

Cai,S. and Kohwi-Shigematsu,T. (1999). Intranuclear relocalization of matrix binding sites during T cell activation detected by amplified fluorescence in situ hybridization. Methods *19*, 394-402.

Cai,S., Lee,C.C., and Kohwi-Shigematsu,T. (2006). SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. Nat. Genet. *38*, 1278-1288.

Cedar,H. and Bergman,Y. (2011). Epigenetics of haematopoietic cell development. Nat. Rev. Immunol. *11*, 478-488.

Chang,H.C., Sehra,S., Goswami,R., Yao,W., Yu,Q., Stritesky,G.L., Jabeen,R., McKinley,C., Ahyi,A.N., Han,L., Nguyen,E.T., Robertson,M.J., Perumal,N.B., Tepper,R.S., Nutt,S.L., and Kaplan,M.H. (2010). The transcription factor PU.1 is required for the development of IL-9-producing T cells and allergic inflammation. Nat. Immunol. *11*, 527-534.

Chao,W., Huynh,K.D., Spencer,R.J., Davidow,L.S., and Lee,J.T. (2002). CTCF, a candidate trans-acting factor for X-inactivation choice. Science *295*, 345-347.

Chou,S.T., Khandros,E., Bailey,L.C., Nichols,K.E., Vakoc,C.R., Yao,Y., Huang,Z., Crispino,J.D., Hardison,R.C., Blobel,G.A., and Weiss,M.J. (2009). Graded repression of PU.1/Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate. Blood *114*, 983-994.

Cook,W.D., McCaw,B.J., Herring,C., John,D.L., Foote,S.J., Nutt,S.L., and Adams,J.M. (2004). PU.1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain. Blood *104*, 3437-3444.

Costello,J.F., Fruhwald,M.C., Smiraglia,D.J., Rush,L.J., Robertson,G.P., Gao,X., Wright,F.A., Feramisco,J.D., Peltomaki,P., Lang,J.C., Schuller,D.E., Yu,L., Bloomfield,C.D., Caligiuri,M.A., Yates,A., Nishikawa,R., Su,H.H., Petrelli,N.J., Zhang,X., O'Dorisio,M.S., Held,W.A., Cavenee,W.K., and Plass,C. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nat. Genet. *24*, 132-138.

Cremer,T., Cremer,C., Schneider,T., Baumann,H., Hens,L., and Kirsch-Volders,M. (1982). Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. Hum. Genet. *62*, 201-209.

Cremer,T., Cremer,M., Dietzel,S., Muller,S., Solovei,I., and Fakan,S. (2006). Chromosome territories--a functional nuclear landscape. Curr. Opin. Cell Biol. *18*, 307-316.

Croft,J.A., Bridger,J.M., Boyle,S., Perry,P., Teague,P., and Bickmore,W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. J. Cell Biol. *145*, 1119-1131.

Cuddapah,S., Jothi,R., Schones,D.E., Roh,T.Y., Cui,K., and Zhao,K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. *19*, 24-32.

Dahl,R., Walsh,J.C., Lancki,D., Laslo,P., Iyer,S.R., Singh,H., and Simon,M.C. (2003). Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor. Nat. Immunol. *4*, 1029-1036.

Dalton,W.T., Jr., Ahearn,M.J., McCredie,K.B., Freireich,E.J., Stass,S.A., and Trujillo,J.M. (1988). HL-60 cell line was derived from a patient with FAB-M2 and not FAB-M3. Blood *71*, 242-247.

de,W.E. and de,L.W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev. *26*, 11-24.

Defossez,P.A. and Gilson,E. (2002). The vertebrate protein CTCF functions as an insulator in Saccharomyces cerevisiae. Nucleic Acids Res. *30*, 5136-5141.

Degner,S.C., Verma-Gaur,J., Wong,T.P., Bossen,C., Iverson,G.M., Torkamani,A., Vettermann,C., Lin,Y.C., Ju,Z., Schulz,D., Murre,C.S., Birshtein,B.K., Schork,N.J., Schlissel,M.S., Riblet,R., Murre,C., and Feeney,A.J. (2011). CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. Proc. Natl. Acad. Sci. U. S. A *108*, 9566-9571.

Dekker,J., Rippe,K., Dekker,M., and Kleckner,N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Deneberg,S. (2012). Epigenetics in myeloid malignancies. Methods Mol. Biol. *863*, 119-137.

Dexter,T.M., Allen,T.D., Scott,D., and Teich,N.M. (1979). Isolation and characterisation of a bipotential haematopoietic cell line. Nature *277*, 471-474.

Dick,J.E. (2008). Stem cell concepts renew cancer research. Blood *112*, 4793-4807.

Dickinson,L.A., Joh,T., Kohwi,Y., and Kohwi-Shigematsu,T. (1992). A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. Cell *70*, 631-645.

Donohoe,M.E., Silva,S.S., Pinter,S.F., Xu,N., and Lee,J.T. (2009). The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature *460*, 128-132.

Ebralidze,A.K., Guibal,F.C., Steidl,U., Zhang,P., Lee,S., Bartholdy,B., Jorda,M.A., Petkova,V., Rosenbauer,F., Huang,G., Dayaram,T., Klupp,J., O'Brien,K.B., Will,B., Hoogenkamp,M., Borden,K.L., Bonifer,C., and Tenen,D.G. (2008). PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. Genes Dev. *22*, 2085-2092.

Estey,E.H. (2012a). Acute myeloid leukemia: 2012 update on diagnosis, risk stratification, and management. Am. J. Hematol. *87*, 89-99.

Estey,E.H. (2012b). How to manage high-risk acute myeloid leukemia. Leukemia *26*, 861-869.

Fedoriw,A.M., Stein,P., Svoboda,P., Schultz,R.M., and Bartolomei,M.S. (2004). Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science *303*, 238-240.

Felsenfeld,G., Burgess-Beusse,B., Farrell,C., Gaszner,M., Ghirlando,R., Huang,S., Jin,C., Litt,M., Magdinier,F., Mutskov,V., Nakatani,Y., Tagami,H., West,A., and Yusufzai,T. (2004). Chromatin boundaries and chromatin domains. Cold Spring Harb. Symp. Quant. Biol. *69*, 245-250.

Felsenfeld,G. and Dekker,J. (2012). Genome architecture and expression. Curr. Opin. Genet. Dev. *22*, 59-61.

Forrester,W.C., Kim,C., and Garriga,G. (2004). The Caenorhabditis elegans Ror RTK CAM-1 inhibits EGL-20/Wnt signaling in cell migration. Genetics *168*, 1951-1962.

Forsberg,E.C., Bhattacharya,D., and Weissman,I.L. (2006). Hematopoietic stem cells: expression profiling and beyond. Stem Cell Rev. *2*, 23-30.

Frengen,E., Zhao,B., Howe,S., Weichenhan,D., Osoegawa,K., Gjernes,E., Jessee,J., Prydz,H., Huxley,C., and de Jong,P.J. (2000). Modular bacterial artificial chromosome vectors for transfer of large inserts into mammalian cells. Genomics *68*, 118-126.

Friedman,A.D. (2007). Transcriptional control of granulocyte and monocyte development. Oncogene *26*, 6816-6828.

Galande,S., Dickinson,L.A., Mian,I.S., Sikorska,M., and Kohwi-Shigematsu,T. (2001). SATB1 cleavage by caspase 6 disrupts PDZ domain-mediated dimerization, causing detachment from chromatin early in T-cell apoptosis. Mol. Cell Biol. *21*, 5591-5604.

Galande,S., Purbey,P.K., Notani,D., and Kumar,P.P. (2007). The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. Curr. Opin. Genet. Dev. *17*, 408-414.

Gasser,S.M. and Laemmli,U.K. (1986). The organisation of chromatin loops: characterization of a scaffold attachment site. EMBO J. *5*, 511-518.

Gaszner,M. and Felsenfeld,G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat. Rev. Genet. *7*, 703-713.

Gilliland,D.G., Jordan,C.T., and Felix,C.A. (2004). The molecular basis of leukemia. Hematology. Am. Soc. Hematol. Educ. Program. 80-97.

Goebel,P., Montalbano,A., Ayers,N., Kompfner,E., Dickinson,L., Webb,C.F., and Feeney,A.J. (2002). High frequency of matrix attachment regions and cut-like protein x/CCAAT-displacement protein and B cell regulator of IgH transcription binding sites flanking Ig V region genes. J. Immunol. *169*, 2477-2487.

Grignani,F., De,M.S., Nervi,C., Tomassoni,L., Gelmetti,V., Cioce,M., Fanelli,M., Ruthardt,M., Ferrara,F.F., Zamir,I., Seiser,C., Grignani,F., Lazar,M.A., Minucci,S., and Pelicci,P.G. (1998). Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. Nature *391*, 815-818.

Grimaud,C., Bantignies,F., Pal-Bhadra,M., Ghana,P., Bhadra,U., and Cavalli,G. (2006). RNAi components are required for nuclear clustering of Polycomb group response elements. Cell *124*, 957-971.

Grondin,B., Lefrancois,M., Tremblay,M., Saint-Denis,M., Haman,A., Waga,K., Bedard,A., Tenen,D.G., and Hoang,T. (2007). c-Jun homodimers can function as a context-specific coactivator. Mol. Cell Biol. *27*, 2919-2933.

Gupta,P., Gurudutta,G.U., Saluja,D., and Tripathi,R.P. (2009). PU.1 and partners: regulation of haematopoietic stem cell fate in normal and malignant haematopoiesis. J. Cell Mol. Med. *13*, 4349-4363.

Hadjur,S., Williams,L.M., Ryan,N.K., Cobb,B.S., Sexton,T., Fraser,P., Fisher,A.G., and Merkenschlager,M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature *460*, 410-413.

Hagege,H., Klous,P., Braem,C., Splinter,E., Dekker,J., Cathala,G., de,L.W., and Forne,T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat. Protoc. *2*, 1722-1733.

Han,H.J., Russo,J., Kohwi,Y., and Kohwi-Shigematsu,T. (2008). SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. Nature *452*, 187-193.

Hark,A.T., Schoenherr,C.J., Katz,D.J., Ingram,R.S., Levorse,J.M., and Tilghman,S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486-489.

Heath,H., Ribeiro de,A.C., Sleutels,F., Dingjan,G., van de Nobelen,S., Jonkers,I., Ling,K.W., Gribnau,J., Renkawitz,R., Grosveld,F., Hendriks,R.W., and Galjart,N. (2008). CTCF regulates cell cycle progression of alphabeta T cells in the thymus. EMBO J. *27*, 2839-2850.

Heintzman,N.D., Hon,G.C., Hawkins,R.D., Kheradpour,P., Stark,A., Harp,L.F., Ye,Z., Lee,L.K., Stuart,R.K., Ching,C.W., Ching,K.A., Antosiewicz-Bourget,J.E., Liu,H., Zhang,X., Green,R.D., Lobanenkov,V.V., Stewart,R., Thomson,J.A., Crawford,G.E., Kellis,M., and Ren,B. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108-112.

Heintzman,N.D., Stuart,R.K., Hon,G., Fu,Y., Ching,C.W., Hawkins,R.D., Barrera,L.O., Van,C.S., Qu,C., Ching,K.A., Wang,W., Weng,Z., Green,R.D., Crawford,G.E., and Ren,B. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. *39*, 311-318.

Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H., and Glass,C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576-589.

Hochstrasser,M., Mathog,D., Gruenbaum,Y., Saumweber,H., and Sedat,J.W. (1986). Spatial organization of chromosomes in the salivary gland nuclei of Drosophila melanogaster. J. Cell Biol. *102*, 112-123.

Hoogenkamp,M., Krysinska,H., Ingram,R., Huang,G., Barlow,R., Clarke,D., Ebralidze,A., Zhang,P., Tagoh,H., Cockerill,P.N., Tenen,D.G., and Bonifer,C. (2007). The Pu.1 locus is differentially regulated at the level of chromatin structure and noncoding transcription by

alternate mechanisms at distinct developmental stages of hematopoiesis. Mol. Cell Biol. *27*, 7425-7438.

Hoogenkamp,M., Lichtinger,M., Krysinska,H., Lancrin,C., Clarke,D., Williamson,A., Mazzarella,L., Ingram,R., Jorgensen,H., Fisher,A., Tenen,D.G., Kouskoff,V., Lacaud,G., and Bonifer,C. (2009). Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. Blood *114*, 299-309.

Hou,C., Dale,R., and Dean,A. (2010). Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proc. Natl. Acad. Sci. U. S. A *107*, 3651-3656.

Houston,I.B., Kamath,M.B., Schweitzer,B.L., Chlon,T.M., and Dekoter,R.P. (2007). Reduction in PU.1 activity results in a block to B-cell development, abnormal myeloid proliferation, and neonatal lethality. Exp. Hematol. *35*, 1056-1068.

Huang,G., Zhang,P., Hirai,H., Elf,S., Yan,X., Chen,Z., Koschmieder,S., Okuno,Y., Dayaram,T., Growney,J.D., Shivdasani,R.A., Gilliland,D.G., Speck,N.A., Nimer,S.D., and Tenen,D.G. (2008). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. Nat. Genet. *40*, 51-60.

Ishihara,K., Oshimura,M., and Nakao,M. (2006). CTCF-dependent chromatin insulator is linked to epigenetic remodeling. Mol. Cell *23*, 733-742.

Iwasaki,H., Somoza,C., Shigematsu,H., Duprez,E.A., Iwasaki-Arai,J., Mizuno,S., Arinobu,Y., Geary,K., Zhang,P., Dayaram,T., Fenyus,M.L., Elf,S., Chan,S., Kastner,P., Huettner,C.S., Murray,R., Tenen,D.G., and Akashi,K. (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood *106*, 1590-1600.

Jemal,A., Thomas,A., Murray,T., and Thun,M. (2002). Cancer statistics, 2002. CA Cancer J. Clin. *52*, 23-47.

Jothi,R., Cuddapah,S., Barski,A., Cui,K., and Zhao,K. (2008). Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. Nucleic Acids Res. *36*, 5221-5231.

Kaushansky,K. (2006). Lineage-specific hematopoietic growth factors. N. Engl. J. Med. *354*, 2034-2045.

Kawamoto,H., Ikawa,T., Masuda,K., Wada,H., and Katsura,Y. (2010). A map for lineage restriction of progenitors during hematopoiesis: the essence of the myeloid-based model. Immunol. Rev. *238*, 23-36.

Kim,T.H., Abdullaev,Z.K., Smith,A.D., Ching,K.A., Loukinov,D.I., Green,R.D., Zhang,M.Q., Lobanenkov,V.V., and Ren,B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231-1245.

Klemsz,M.J., McKercher,S.R., Celada,A., Van,B.C., and Maki,R.A. (1990). The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. Cell *61*, 113-124.

Kohwi-Shigematsu,T., Poterlowicz,K., Ordinario,E., Han,H.J., Botchkarev,V.A., and Kohwi,Y. (2012). Genome organizing function of SATB1 in tumor progression. Semin. Cancer Biol.

Kondo,M., Scherer,D.C., Miyamoto,T., King,A.G., Akashi,K., Sugamura,K., and Weissman,I.L. (2000). Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines. Nature *407*, 383-386.

Kondo,M., Weissman,I.L., and Akashi,K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell *91*, 661-672.

Kouzarides,T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kriaucionis,S. and Heintz,N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Laiosa,C.V., Stadtfeld,M., and Graf,T. (2006). Determinants of lymphoid-myeloid lineage diversification. Annu. Rev. Immunol. *24*, 705-738.

Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W., Funke,R., Gage,D., Harris,K., Heaford,A., Howland,J., Kann,L., Lehoczky,J., LeVine,R., McEwan,P., McKernan,K., Meldrim,J., Mesirov,J.P., Miranda,C., Morris,W., Naylor,J., Raymond,C., Rosetti,M., Santos,R., Sheridan,A., Sougnez,C., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Wyman,D., Rogers,J., Sulston,J., Ainscough,R., Beck,S., Bentley,D., Burton,J., Clee,C., Carter,N., Coulson,A., Deadman,R., Deloukas,P., Dunham,A., Dunham,I., Durbin,R., French,L., Grafham,D., Gregory,S., Hubbard,T., Humphray,S., Hunt,A., Jones,M., Lloyd,C., McMurray,A., Matthews,L., Mercer,S., Milne,S., Mullikin,J.C., Mungall,A., Plumb,R., Ross,M., Shownkeen,R., Sims,S., Waterston,R.H., Wilson,R.K., Hillier,L.W., McPherson,J.D., Marra,M.A., Mardis,E.R., Fulton,L.A., Chinwalla,A.T., Pepin,K.H., Gish,W.R., Chissoe,S.L., Wendl,M.C., Delehaunty,K.D., Miner,T.L., Delehaunty,A., Kramer,J.B., Cook,L.L., Fulton,R.S., Johnson,D.L., Minx,P.J., Clifton,S.W., Hawkins,T., Branscomb,E., Predki,P., Richardson,P., Wenning,S., Slezak,T., Doggett,N., Cheng,J.F., Olsen,A., Lucas,S., Elkin,C., Uberbacher,E., Frazier,M., Gibbs,R.A., Muzny,D.M., Scherer,S.E., Bouck,J.B., Sodergren,E.J., Worley,K.C., Rives,C.M., Gorrell,J.H., Metzker,M.L., Naylor,S.L., Kucherlapati,R.S., Nelson,D.L., Weinstock,G.M., Sakaki,Y., Fujiyama,A., Hattori,M., Yada,T., Toyoda,A., Itoh,T., Kawagoe,C., Watanabe,H., Totoki,Y., Taylor,T., Weissenbach,J., Heilig,R., Saurin,W., Artiguenave,F., Brottier,P., Bruls,T., Pelletier,E., Robert,C., Wincker,P., Smith,D.R., Doucette-Stamm,L., Rubenfield,M., Weinstock,K., Lee,H.M., Dubois,J., Rosenthal,A., Platzer,M., Nyakatura,G., Taudien,S., Rump,A., Yang,H., Yu,J., Wang,J., Huang,G., Gu,J., Hood,L., Rowen,L., Madan,A., Qin,S., Davis,R.W., Federspiel,N.A., Abola,A.P., Proctor,M.J., Myers,R.M., Schmutz,J., Dickson,M., Grimwood,J., Cox,D.R., Olson,M.V., Kaul,R., Raymond,C., Shimizu,N., Kawasaki,K., Minoshima,S., Evans,G.A., Athanasiou,M., Schultz,R., Roe,B.A., Chen,F., Pan,H., Ramser,J., Lehrach,H., Reinhardt,R., McCombie,W.R., de la Bastide,M., Dedhia,N., Blocker,H., Hornischer,K., Nordsiek,G., Agarwala,R., Aravind,L., Bailey,J.A., Bateman,A., Batzoglou,S., Birney,E., Bork,P., Brown,D.G., Burge,C.B., Cerutti,L., Chen,H.C., Church,D., Clamp,M., Copley,R.R., Doerks,T., Eddy,S.R., Eichler,E.E., Furey,T.S., Galagan,J., Gilbert,J.G., Harmon,C., Hayashizaki,Y., Haussler,D., Hermjakob,H., Hokamp,K., Jang,W., Johnson,L.S., Jones,T.A., Kasif,S., Kaspryzk,A., Kennedy,S., Kent,W.J., Kitts,P., Koonin,E.V., Korf,I.,

Kulp,D., Lancet,D., Lowe,T.M., McLysaght,A., Mikkelsen,T., Moran,J.V., Mulder,N., Pollara,V.J., Ponting,C.P., Schuler,G., Schultz,J., Slater,G., Smit,A.F., Stupka,E., Szustakowski,J., Thierry-Mieg,D., Thierry-Mieg,J., Wagner,L., Wallis,J., Wheeler,R., Williams,A., Wolf,Y.I., Wolfe,K.H., Yang,S.P., Yeh,R.F., Collins,F., Guyer,M.S., Peterson,J., Felsenfeld,A., Wetterstrand,K.A., Patrinos,A., Morgan,M.J., de,J.P., Catanese,J.J., Osoegawa,K., Shizuya,H., Choi,S., and Chen,Y.J. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Laslo,P., Spooner,C.J., Warmflash,A., Lancki,D.W., Lee,H.J., Sciammas,R., Gantner,B.N., Dinner,A.R., and Singh,H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. Cell *126*, 755-766.

Leddin,M., Perrod,C., Hoogenkamp,M., Ghani,S., Assi,S., Heinz,S., Wilson,N.K., Follows,G., Schonheit,J., Vockentanz,L., Mosammam,A.M., Chen,W., Tenen,D.G., Westhead,D.R., Gottgens,B., Bonifer,C., and Rosenbauer,F. (2011). Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. Blood *117*, 2827-2838.

Lei,E.P. and Corces,V.G. (2006). RNA interference machinery influences the nuclear organization of a chromatin insulator. Nat. Genet. *38*, 936-941.

Levine,M. and Tjian,R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151.

Li,Y., Okuno,Y., Zhang,P., Radomska,H.S., Chen,H., Iwasaki,H., Akashi,K., Klemsz,M.J., McKercher,S.R., Maki,R.A., and Tenen,D.G. (2001). Regulation of the PU.1 gene by distal elements. Blood *98*, 2958-2965.

Ling,J.Q., Li,T., Hu,J.F., Vu,T.H., Chen,H.L., Qiu,X.W., Cherry,A.M., and Hoffman,A.R. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science *312*, 269-272.

Litt,M.D., Simpson,M., Recillas-Targa,F., Prioleau,M.N., and Felsenfeld,G. (2001). Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. EMBO J. *20*, 2224-2235.

Lobanenkov,V.V., Nicolas,R.H., Adler,V.V., Paterson,H., Klenova,E.M., Polotskaja,A.V., and Goodwin,G.H. (1990). A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene *5*, 1743-1753.

Lomvardas,S., Barnea,G., Pisapia,D.J., Mendelsohn,M., Kirkland,J., and Axel,R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F., and Richmond,T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Lutz,M., Burke,L.J., Barreto,G., Goeman,F., Greb,H., Arnold,R., Schultheiss,H., Brehm,A., Kouzarides,T., Lobanenkov,V., and Renkawitz,R. (2000). Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Res. *28*, 1707-1713.

LYON,M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372-373.

Majumder,S., Zhao,Z., Kaneko,K., and DePamphilis,M.L. (1997). Developmental acquisition of enhancer function requires a unique coactivator activity. EMBO J. *16*, 1721-1731.

Martin,D., Pantoja,C., Fernandez,M.A., Valdes-Quezada,C., Molto,E., Matesanz,F., Bogdanovic,O., Calle-Mustienes,E., Dominguez,O., Taher,L., Furlan-Magaril,M., Alcina,A., Canon,S., Fedetz,M., Blasco,M.A., Pereira,P.S., Ovcharenko,I., Recillas-Targa,F., Montoliu,L., Manzanares,M., Guigo,R., Serrano,M., Casares,F., and Gomez-Skarmeta,J.L. (2011). Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes. Nat. Struct. Mol. Biol. *18*, 708-714.

McGrath,J. and Solter,D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell *37*, 179-183.

McPherson,J.D., Marra,M., Hillier,L., Waterston,R.H., Chinwalla,A., Wallis,J., Sekhon,M., Wylie,K., Mardis,E.R., Wilson,R.K., Fulton,R., Kucaba,T.A., Wagner-McPherson,C., Barbazuk,W.B., Gregory,S.G., Humphray,S.J., French,L., Evans,R.S., Bethel,G., Whittaker,A., Holden,J.L., McCann,O.T., Dunham,A., Soderlund,C., Scott,C.E., Bentley,D.R., Schuler,G., Chen,H.C., Jang,W., Green,E.D., Idol,J.R., Maduro,V.V., Montgomery,K.T., Lee,E., Miller,A., Emerling,S., Kucherlapati, Gibbs,R., Scherer,S., Gorrell,J.H., Sodergren,E., Clerc-Blankenburg,K., Tabor,P., Naylor,S., Garcia,D., de Jong,P.J., Catanese,J.J., Nowak,N., Osoegawa,K., Qin,S., Rowen,L., Madan,A., Dors,M., Hood,L., Trask,B., Friedman,C., Massa,H., Cheung,V.G., Kirsch,I.R., Reid,T., Yonescu,R., Weissenbach,J., Bruls,T., Heilig,R., Branscomb,E., Olsen,A., Doggett,N., Cheng,J.F., Hawkins,T., Myers,R.M., Shang,J., Ramirez,L., Schmutz,J., Velasquez,O., Dixon,K., Stone,N.E., Cox,D.R., Haussler,D., Kent,W.J., Furey,T., Rogic,S., Kennedy,S., Jones,S., Rosenthal,A., Wen,G., Schilhabel,M., Gloeckner,G., Nyakatura,G., Siebert,R., Schlegelberger,B., Korenberg,J., Chen,X.N., Fujiyama,A., Hattori,M., Toyoda,A., Yada,T., Park,H.S., Sakaki,Y., Shimizu,N., Asakawa,S., Kawasaki,K., Sasaki,T., Shintani,A., Shimizu,A., Shibuya,K., Kudoh,J., Minoshima,S., Ramser,J., Seranski,P., Hoff,C., Poustka,A., Reinhardt,R., and Lehrach,H. (2001). A physical map of the human genome. Nature *409*, 934-941.

Meaburn,K.J., Misteli,T., and Soutoglou,E. (2007). Spatial genome organization in the formation of chromosomal translocations. Semin. Cancer Biol. *17*, 80-90.

Misteli,T. (2008). Cell biology: Nuclear order out of chaos. Nature *456*, 333-334.

Mizuki,M., Schwable,J., Steur,C., Choudhary,C., Agrawal,S., Sargin,B., Steffen,B., Matsumura,I., Kanakura,Y., Bohmer,F.D., Muller-Tidow,C., Berdel,W.E., and Serve,H. (2003). Suppression of myeloid transcription factors and induction of STAT response genes by AML-specific Flt3 mutations. Blood *101*, 3164-3173.

Moon,H., Filippova,G., Loukinov,D., Pugacheva,E., Chen,Q., Smith,S.T., Munhall,A., Grewe,B., Bartkuhn,M., Arnold,R., Burke,L.J., Renkawitz-Pohl,R., Ohlsson,R., Zhou,J., Renkawitz,R., and Lobanenkov,V. (2005). CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep. *6*, 165-170.

Moreau-Gachelin,F., Tavitian,A., and Tambourin,P. (1988). Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. Nature *331*, 277-280.

Moreau-Gachelin,F., Wendling,F., Molina,T., Denis,N., Titeux,M., Grimber,G., Briand,P., Vainchenker,W., and Tavitian,A. (1996). Spi-1/PU.1 transgenic mice develop multistep erythroleukemias. Mol. Cell Biol. *16*, 2453-2463.

Morita,S., Kojima,T., and Kitamura,T. (2000). Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther. *7*, 1063-1066.

Mueller,B.U. and Pabst,T. (2006). C/EBPalpha and the pathophysiology of acute myeloid leukemia. Curr. Opin. Hematol. *13*, 7-14.

Mueller,B.U. and Pabst,T. (2010). Lineage-specific transcription factor aberrations in AML. Cancer Treat. Res. *145*, 109-125.

Mueller,B.U., Pabst,T., Fos,J., Petkovic,V., Fey,M.F., Asou,N., Buergi,U., and Tenen,D.G. (2006a). ATRA resolves the differentiation block in t(15;17) acute myeloid leukemia by restoring PU.1 expression. Blood *107*, 3330-3338.

Mueller,B.U., Pabst,T., Hauser,P., Neuberg,D., and Tenen,D.G. (2006b). Mutations of the transcription factor PU.1 are not associated with acute lymphoblastic leukaemia. Br. J. Cancer *94*, 1918-1920.

Mueller,B.U., Pabst,T., Osato,M., Asou,N., Johansen,L.M., Minden,M.D., Behre,G., Hiddemann,W., Ito,Y., and Tenen,D.G. (2002). Heterozygous PU.1 mutations are associated with acute myeloid leukemia. Blood *100*, 998-1007.

Mueller,B.U., Pabst,T., Osato,M., Asou,N., Johansen,L.M., Minden,M.D., Behre,G., Hiddemann,W., Ito,Y., and Tenen,D.G. (2003). Heterozygous PU.1 mutations are associated with acute myeloid leukemia. Blood *101*, 2074.

Murrell,A., Heeson,S., and Reik,W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat. Genet. *36*, 889-893.

Navarro,P., Page,D.R., Avner,P., and Rougeulle,C. (2006). Tsix-mediated epigenetic switch of a CTCF-flanked region of the Xist promoter determines the Xist transcription program. Genes Dev. *20*, 2787-2792.

Nerlov,C., Querfurth,E., Kulessa,H., and Graf,T. (2000). GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. Blood *95*, 2543-2551.

Neugebauer,K.M. (2002). On the importance of being co-transcriptional. J. Cell Sci. *115*, 3865-3871.

Ngo,V.N., Davis,R.E., Lamy,L., Yu,X., Zhao,H., Lenz,G., Lam,L.T., Dave,S., Yang,L., Powell,J., and Staudt,L.M. (2006). A loss-of-function RNA interference screen for molecular targets in cancer. Nature *441*, 106-110.

Notani,D., Gottimukkala,K.P., Jayani,R.S., Limaye,A.S., Damle,M.V., Mehta,S., Purbey,P.K., Joseph,J., and Galande,S. (2010). Global regulator SATB1 recruits beta-catenin and regulates T(H)2 differentiation in Wnt-dependent manner. PLoS. Biol. *8*, e1000296.

Ogawa,M. (1993). Differentiation and proliferation of hematopoietic stem cells. Blood *81*, 2844-2853.

Ohlsson,R. (2007). Genetics. Widespread monoallelic expression. Science *318*, 1077-1078.

Ohlsson,R., Bartkuhn,M., and Renkawitz,R. (2010). CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin. Chromosoma *119*, 351-360.

Okuno,Y., Huang,G., Rosenbauer,F., Evans,E.K., Radomska,H.S., Iwasaki,H., Akashi,K., Moreau-Gachelin,F., Li,Y., Zhang,P., Gottgens,B., and Tenen,D.G. (2005). Potential autoregulation of transcription factor PU.1 by an upstream regulatory element. Mol. Cell Biol. *25*, 2832-2845.

Okuno,Y., McNairn,A.J., den,E.N., Pines,J., and Gilbert,D.M. (2001). Stability, chromatin association and functional activity of mammalian pre-replication complex proteins during the cell cycle. EMBO J. *20*, 4263-4277.

Olins,A.L. and Olins,D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Orkin,S.H. (2000). Stem cell alchemy. Nat. Med. *6*, 1212-1213.

Orom,U.A., Derrien,T., Beringer,M., Gumireddy,K., Gardini,A., Bussotti,G., Lai,F., Zytnicki,M., Notredame,C., Huang,Q., Guigo,R., and Shiekhattar,R. (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46-58.

Pabst,T. and Mueller,B.U. (2007). Transcriptional dysregulation during myeloid transformation in AML. Oncogene *26*, 6829-6837.

Palstra,R.J., Tolhuis,B., Splinter,E., Nijmeijer,R., Grosveld,F., and de,L.W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nat. Genet. *35*, 190-194.

Phillips,J.E. and Corces,V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Polli,M., Dakic,A., Light,A., Wu,L., Tarlinton,D.M., and Nutt,S.L. (2005). The development of functional B lymphocytes in conditional PU.1 knock-out mice. Blood *106*, 2083-2090.

Raab,J.R. and Kamakaka,R.T. (2010). Insulators and promoters: closer than we think. Nat. Rev. Genet. *11*, 439-446.

Rao,G., Rekhtman,N., Cheng,G., Krasikov,T., and Skoultchi,A.I. (1997). Deregulated expression of the PU.1 transcription factor blocks murine erythroleukemia cell terminal differentiation. Oncogene *14*, 123-131.

Regha,K., Sloane,M.A., Huang,R., Pauler,F.M., Warczok,K.E., Melikant,B., Radolf,M., Martens,J.H., Schotta,G., Jenuwein,T., and Barlow,D.P. (2007). Active and repressive chromatin are interspersed without spreading in an imprinted gene cluster in the mammalian genome. Mol. Cell *27*, 353-366.

Rekhtman,N., Radparvar,F., Evans,T., and Skoultchi,A.I. (1999). Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. Genes Dev. *13*, 1398-1411.

Reya,T., Morrison,S.J., Clarke,M.F., and Weissman,I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105-111.

Richmond,T.J. and Davey,C.A. (2003). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Rizzo,M.G., Zepparoni,A., Cristofanelli,B., Scardigli,R., Crescenzi,M., Blandino,G., Giuliacci,S., Ferrari,S., Soddu,S., and Sacchi,A. (1998). Wt-p53 action in human leukaemia cell lines corresponding to different stages of differentiation. Br. J. Cancer *77*, 1429-1438.

Rosenbauer,F., Owens,B.M., Yu,L., Tumang,J.R., Steidl,U., Kutok,J.L., Clayton,L.K., Wagner,K., Scheller,M., Iwasaki,H., Liu,C., Hackanson,B., Akashi,K., Leutz,A., Rothstein,T.L., Plass,C., and Tenen,D.G. (2006). Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. Nat. Genet. *38*, 27-37.

Rosenbauer,F. and Tenen,D.G. (2007). Transcription factors in myeloid development: balancing differentiation with transformation. Nat. Rev. Immunol. *7*, 105-117.

Rosenbauer,F., Wagner,K., Kutok,J.L., Iwasaki,H., Le Beau,M.M., Okuno,Y., Akashi,K., Fiering,S., and Tenen,D.G. (2004). Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. Nat. Genet. *36*, 624-630.

Rosmarin,A.G., Yang,Z., and Resendes,K.K. (2005). Transcriptional regulation in myelopoiesis: Hematopoietic fate choice, myeloid differentiation, and leukemogenesis. Exp. Hematol. *33*, 131-143.

Rothenberg,E.V., Moore,J.E., and Yui,M.A. (2008). Launching the T-cell-lineage developmental programme. Nat. Rev. Immunol. *8*, 9-21.

Rothenberg,E.V., Telfer,J.C., and Anderson,M.K. (1999). Transcriptional regulation of lymphocyte lineage commitment. Bioessays *21*, 726-742.

Schmidt,D., Schwalie,P.C., Wilson,M.D., Ballester,B., Goncalves,A., Kutter,C., Brown,G.D., Marshall,A., Flicek,P., and Odom,D.T. (2012a). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell *148*, 335-348.

Schmidt,D., Schwalie,P.C., Wilson,M.D., Ballester,B., Goncalves,A., Kutter,C., Brown,G.D., Marshall,A., Flicek,P., and Odom,D.T. (2012b). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. Cell *148*, 335-348.

Schones,D.E., Cui,K., Cuddapah,S., Roh,T.Y., Barski,A., Wang,Z., Wei,G., and Zhao,K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887-898.

Scott,E.W., Simon,M.C., Anastasi,J., and Singh,H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science *265*, 1573-1577.

Shivdasani,R.A. and Orkin,S.H. (1996). The transcriptional control of hematopoiesis. Blood *87*, 4025-4039.

Smale,S.T. and Kadonaga,J.T. (2003). The RNA polymerase II core promoter. Annu. Rev. Biochem. *72*, 449-479.

Soshnikova,N. and Duboule,D. (2009). Epigenetic temporal control of mouse Hox genes in vivo. Science *324*, 1320-1323.

Soto-Reyes,E. and Recillas-Targa,F. (2010). Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. Oncogene *29*, 2217-2227.

Spilianakis,C.G., Lalioti,M.D., Town,T., Lee,G.R., and Flavell,R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature *435*, 637-645.

Splinter,E., Heath,H., Kooren,J., Palstra,R.J., Klous,P., Grosveld,F., Galjart,N., and de,L.W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev. *20*, 2349-2354.

Stadhouders,R., Thongjuea,S., Andrieu-Soler,C., Palstra,R.J., Bryne,J.C., van den Heuvel,A., Stevens,M., de,B.E., Kockx,C., van der Sloot,A., van den Hout,M., van,I.W., Eick,D., Lenhard,B., Grosveld,F., and Soler,E. (2012). Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. EMBO J. *31*, 986-999.

Steidl,U., Rosenbauer,F., Verhaak,R.G., Gu,X., Ebralidze,A., Otu,H.H., Klippel,S., Steidl,C., Bruns,I., Costa,D.B., Wagner,K., Aivado,M., Kobbe,G., Valk,P.J., Passegue,E., Libermann,T.A., Delwel,R., and Tenen,D.G. (2006). Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. Nat. Genet. *38*, 1269-1277.

Steidl,U., Steidl,C., Ebralidze,A., Chapuy,B., Han,H.J., Will,B., Rosenbauer,F., Becker,A., Wagner,K., Koschmieder,S., Kobayashi,S., Costa,D.B., Schulz,T., O'Brien,K.B., Verhaak,R.G., Delwel,R., Haase,D., Trumper,L., Krauter,J., Kohwi-Shigematsu,T., Griesinger,F., and Tenen,D.G. (2007). A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. J. Clin. Invest *117*, 2611-2620.

Stopka,T., Amanatullah,D.F., Papetti,M., and Skoultchi,A.I. (2005). PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. EMBO J. *24*, 3712-3723.

Suzuki,Y., Tsunoda,T., Sese,J., Taira,H., Mizushima-Sugano,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Suyama,A., Sakaki,Y., Morishita,S., Okubo,K., and Sugano,S.

(2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Res. *11*, 677-684.

Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L., and Rao,A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Takizawa,T., Meaburn,K.J., and Misteli,T. (2008). The meaning of gene positioning. Cell *135*, 9-13.

Tolhuis,B., Palstra,R.J., Splinter,E., Grosveld,F., and de,L.W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol. Cell *10*, 1453-1465.

Vangala,R.K., Heiss-Neumann,M.S., Rangatia,J.S., Singh,S.M., Schoch,C., Tenen,D.G., Hiddemann,W., and Behre,G. (2003). The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. Blood *101*, 270-277.

Villa,R., Pasini,D., Gutierrez,A., Morey,L., Occhionorelli,M., Vire,E., Nomdedeu,J.F., Jenuwein,T., Pelicci,P.G., Minucci,S., Fuks,F., Helin,K., and Di,C.L. (2007). Role of the polycomb repressive complex 2 in acute promyelocytic leukemia. Cancer Cell *11*, 513-525.

Waddington,C.H. (2012). The epigenotype. 1942. Int. J. Epidemiol. *41*, 10-13.

Wallace,J.A. and Felsenfeld,G. (2007). We gather together: insulators and genome organization. Curr. Opin. Genet. Dev. *17*, 400-407.

Walter,M.J., Park,J.S., Ries,R.E., Lau,S.K., McLellan,M., Jaeger,S., Wilson,R.K., Mardis,E.R., and Ley,T.J. (2005). Reduced PU.1 expression causes myeloid progenitor expansion and increased leukemia penetrance in mice expressing PML-RARalpha. Proc. Natl. Acad. Sci. U. S. A *102*, 12513-12518.

Weintraub,H. (1984). Histone-H1-dependent chromatin superstructures and the suppression of gene activity. Cell *38*, 17-27.

Weissman,I.L., Anderson,D.J., and Gage,F. (2001). Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. Annu. Rev. Cell Dev. Biol. *17*, 387-403.

Wernig,M., Meissner,A., Foreman,R., Brambrink,T., Ku,M., Hochedlinger,K., Bernstein,B.E., and Jaenisch,R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

West,A.G., Gaszner,M., and Felsenfeld,G. (2002). Insulators: many functions, many mechanisms. Genes Dev. *16*, 271-288.

Weth,O. and Renkawitz,R. (2011). CTCF function is modulated by neighboring DNA binding factors. Biochem. Cell Biol. *89*, 459-468.

Whitehouse,I. and Tsukiyama,T. (2009). Opening windows to the genome. Cell *137*, 400-402.

Wilson,A., Laurenti,E., and Trumpp,A. (2009). Balancing dormant and self-renewing hematopoietic stem cells. Curr. Opin. Genet. Dev. *19*, 461-468.

Wilson,A. and Trumpp,A. (2006). Bone-marrow haematopoietic-stem-cell niches. Nat. Rev. Immunol. *6*, 93-106.

Wilson,N.K., Calero-Nieto,F.J., Ferreira,R., and Gottgens,B. (2011a). Transcriptional regulation of haematopoietic transcription factors. Stem Cell Res. Ther. *2*, 6.

Wilson,N.K., Tijssen,M.R., and Gottgens,B. (2011b). Deciphering transcriptional control mechanisms in hematopoiesis:the impact of high-throughput sequencing technologies. Exp. Hematol. *39*, 961-968.

Witcher,M. and Emerson,B.M. (2009). Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. Mol. Cell *34*, 271-284.

Yasui,D., Miyano,M., Cai,S., Varga-Weisz,P., and Kohwi-Shigematsu,T. (2002). SATB1 targets chromatin remodelling to regulate genes over long distances. Nature *419*, 641-645.

Zarnegar,M.A., Chen,J., and Rothenberg,E.V. (2010). Cell-type-specific activation and repression of PU.1 by a complex of discrete, functionally specialized cis-regulatory elements. Mol. Cell Biol. *30*, 4922-4939.

Zarnegar,M.A. and Rothenberg,E.V. (2012). Ikaros represses and activates PU.1 cell-type-specifically through the multifunctional Sfpi1 URE and a myeloid specific enhancer. Oncogene.

Zhang,J.A., Mortazavi,A., Williams,B.A., Wold,B.J., and Rothenberg,E.V. (2012). Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. Cell *149*, 467-482.

Zhou,V.W., Goren,A., and Bernstein,B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7-18.

Zhu,J. and Emerson,S.G. (2002). Hematopoietic cytokines, transcription factors and lineage commitment. Oncogene *21*, 3295-3313.

Zhu,X., Zhang,H., Qian,M., Zhao,X., Yang,W., Wang,P., Zhang,J., and Wang,K. (2012). The significance of low PU.1 expression in patients with acute promyelocytic leukemia. J. Hematol. Oncol. *5*, 22.

Abbreviations

% percentage

°C degree centigrade

3C chromatin conformation capture

AB antibody

a.m. above mentioned

ALL acute lymphoid leukemia

AML acute myeloid leukemia

APC allophycocyanin or adenomatous polyposis *coli*

ATCC american tissue culture collection

BCR B cell receptor

BM bone marrow

BMMΦ bone marrow derived macrophages

bp base pair

BS binding site

BSA bovine serum albumin

CDS coding determining sequence

CLP common lymphoid progenitor

CML chronic myeloid leukaemia

CMP common myeloid progenitor

c-Myc myelocytomatosis viral oncogene homolog

CTCF CCCTC-binding factor

DBD DNA binding domain

DMEM Dulbecco's modified Eagle medium

DMSO dimethylsulfoxid

DNA desoxyribonucleic acid

dNTP deoxynucleotide triphosphate

DTT dithiotreitol

ECL enzymatic chemiluminescence

e.g. exempli gratia

EDTA ethylenediaminetetraacetate

Env envelope protein

ES cells embryonic stem cells

F forward

FACS fluorescent activated cell sorting

FCS fetal calf serum

FITC fluoresceinisothiocyanat

GMP granulocyte-macrophage precursor

Gag-pol group antigen protein - polymerase

GFP green-fluorescent protein

h hour

HCl hydrochloric acid

HRP horseradish peroxidase

HSC hematopoietic stem cell

Ig immunoglobulin

IgG immune globulin gamma

Kb kilobase

kDakilo Dalton

KCl potassium chloride

KH_2PO_4 potassium phosphate monobasic

KO knock out

kd knock down

Lin lineage

LSC leukemia stem cell

m or mm mouse, murine

M molar

mg milligram

MIG MSCV-IRES-GFP

min minutes

ml millilitre

mM milli molar

µg microgram

µl microliter

µM micromolar

mRNA messenger ribonucleic acid

n number

Na_2HPO_4 Sodium phosphate dibasic

NaCl Sodiumchloride

NaOH sodium hydroxide

ng nano gram

nm nano meter

NLS nuclear localization sequence

p probability value

PAC phage artificial chromosome

PBS phosphate buffered saline

PcG polycomb group

PCR polymerase-chain-reaction

PFU *Pyrococcus furiosus* derived polymerase

pH potentia hydrogenii

PI propidium iodide

Pol polymerase

P/S penicillin/streptomycin

R reverse

RNA ribonucleic acid

rpm rotations per minute

RT room teperature

RT-PCR reverse transcription real time polymerase chain reaction

s.d. standard deviation

SDS Sodiumdodecylsulfat

Seq sequencing

SEM standard deviation of the mean

TAE Tris/acetate/EDTA buffer

Taq Taqman DNA polymerase

TE Tris/EDTA

U units

URC upstream regulatory cluster

URE upstream regulatory element

wt wild type

Selbständigkeitserklärung

Ich versichere hiermit, dass die von mir vorgelegte Dissertation eigenständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Ich versichere, dass alle aus anderen Quellen oder indirekt übernommenen Daten und Konzepte, sowie Ergebnisse aus Kooperationsprojekten unter Angabe der Referenz gekennzeichnet sind.
Außerdem versichere ich, dass mir die aktuelle Promotionsordnung bekannt ist und ich mich nicht anderwärts um einen Doktorgrad bewerbe, bzw. noch keinen entsprechenden Doktorgrad besitze.

Berlin, den 26. Januar 2013

Chiara Perrod

Acknowledgements

First, I would like to thank my supervisor Dr. Frank Rosenbauer for providing the opportunity to perform my PhD project in his laboratory, for his financial and permanent scientific support. Furthermore, I want to thank all members of Achim Leutz's group for practical support and constructive criticism in shared seminars and of course Dr. Achim Leutz to accept to be my supervisor.

I am also very thankful to the collaborators who participated in this project: thanks to Sören Wenzel and Georg Lenz for providing cell lines, Saeed Ghani, Wolf-Dietrich Ludwig, Jörg Westermann, Peter Schlenke and Carsten Müller-Tidow to have provided primary leukemic samples.

Within the Rosenbauer's lab a particular thank to Andreas Enns, for having started this project, and Jörg Schönheit, for having tolerated me for all these years as chaotic neighbour in the lab as in the office; and all the Rosis', past and present members, for fruitful scientific discussions and help as well as for extra-scientific wonderful time.

Within the MDC I thank everyone who shared opinions, reagents and relaxing chats during the usual trip in the S-Bahn back to the city; and especially the "Beer Hour community", for not only nice time but friendly scientific support when urgently needed.

And of course all my not-scientific friends, in Berlin and around the world, for their love, and my family, especially my mother, for love and support in any moment.