From Department of Biosciences and Nutrition
Karolinska Institutet, Stockholm, Sweden

EPIGENETIC REGULATION OF TRANSCRIPTION AND CELLULAR DEVELOPMENT

Farzaneh Shahin Varnoosfaderani

Karolinska Institutet

Stockholm 2020

Published by Karolinska Institutet.
Printed by Arkitektkopia AB, 2020
Cover illustration by Farzaneh Shahin Varnoosfaderani

ISBN 978-91-7831-745-5

Epigenetic regulation of transcription and cellular development

THESIS FOR DOCTORAL DEGREE (Ph.D.)

The thesis will be defended Wednesday the 1st of April 2020 at 10.00
Location: GENE (Room 5108), NEO, plan 5, Blickagången 16, 14152 Huddinge

By

Farzaneh Shahin Varnoosfaderani

Principal Supervisor:

Associate Professor Andreas Lennartsson
Karolinska Institutet
Department of Biosciences and Nutrition

Co-supervisor(s):

Associate Professor Peter Svensson
Karolinska Institutet
Department of Biosciences and Nutrition

Professor Sören Lehmann
Karolinska Institutet
Department of Medicine
Center for Hematology and Regenerative Medicine (HERM)

Anna Palau de Miguel
Karolinska Institutet
Department of Biosciences and Nutrition

Opponent:

Esteban Ballestar
Josep Carreras Leukaemia Research Institute
Cancer and Leukemia Epigenetics and Biology Program PEBCL

Examination Board:

Assistant Professor Robert Månsson
Karolinska Institutet
Department of Medicine
Center for Hematology and Regenerative Medicine (HERM)

Professor Peter Zaphiropoulos
Karolinska Institutet
Department of Biosciences and Nutrition

Professor Neus Visa
Stockholm Universitet
Department of Molecular Biosciences

To my FAMILY
&
To those who NEVER GIVE UP

"Experience is what you get when you didn't get what you wanted.
And experience is often the most valuable thing you have to offer."

Randy Pausch, The Last Lecture

ABSTRACT

Epigenetic machinery can regulate different biological processes via different mechanisms. In this thesis, we explore the effects of the epigenetic system on transcription and how it can differ during cellular development in different human cell lines models, with focus on hematopoiesis.

Paper I, aimed to identify new roles for different epigenetic regulators in myeloid differentiation. We performed a CRISPR-Cas9 screen that targeted 1092 epigenetic factors in a model for myeloid differentiation, with the objective to uncover novel roles for regulatory factors that are important for differentiation in hematopoiesis. In our analysis, the chromodomain helicase DNA-binding 2 (CHD2) showed a crucial impact on megakaryocytic differentiation in the K-562 cell line model.

In paper II, our aim was to identify the roles of different PHC subunits in Polycomb repressive complex 1 during hematopoiesis. Data mining from publicly available datasets showed opposite expression pattern between each PHC subunit. *PHC1* is higher expressed in early stages of myelopoiesis that is opposite to *PHC2*, and *PHC3*, which expression increasing with differentiation. *PHC1-3* was knocked down individually, using siRNA in the myeloblast cell line KG-1. RNA-sequencing analysis after knock down for each specific *PHC* subunit, showed how *PHC1*, *2* and *3* play different roles during development and myeloid differentiation.

In paper III, we used the FANTOM5 database for transcription start sites (TSS) in a wide variety of primary cells. The study mapped the usage of alternative TSS that leads to exclusion of coding sequence, and exclusion of annotated protein domains. We demonstrated a dynamic usage of alternative TSS and their potential regulatory roles in different cell lineages and development stages. We investigated the role of alternative TSSs for *KDM2B* in the Jurkat T-cell lineage and their potential functional consequences.

In paper IV, our aim was to study the dynamics of 3D chromatin structure in relation to the circadian rhythm. We demonstrated that chromosomal fiber interactions are organized by PARP1- CTCF activity. We showed how the 3D genome structure can influence circadian rhythm machinery and how the transcription activation and silencing are under oscillation.

LIST OF SCIENTIFIC PAPERS

I. **A regulator role for CHD2 in myelopoiesis.**

Shahin Varnoosfaderani F, Palau A, Dong W, Persson J, Durand-Dubief M, Svensson JP, Lennartsson A. Epigenetics. 2020 Jan 10 :1-13.

II. **Distinct roles for Polycomb repressive complex 1 subunits PHC1, PHC2 and PHC3 in myeloid differentiation.**

Palau Anna, **Shahin Varnoosfaderani Farzaneh** and Lennartsson Andreas. Manuscript.

III. **Investigation of protein coding sequence exclusion by alternative transcription start site usage across the human body.**

Wenbo Dong*, Berit Lilje*, **Farzaneh Shahin Varnoosfaderani**, Erik Arner, The FANTOM consortium, Andreas Lennartsson*, Albin Sandelin* Manuscript.

*Authors contributed equally to this study

IV. **PARP1- and CTCF-Mediated Interactions between Active and Repressed Chromatin at the Lamina Promote Oscillating Transcription.**

Zhao H*, Sifakis EG*, Sumida N*, Millán-Ariño L*, Scholz BA, Svensson JP, Chen X, Ronnegren AL, Mallet de Lima CD, **Varnoosfaderani FS**, Shi C, Loseva O, Yammine S, Israelsson M, Rathje LS, Németi B, Fredlund E, Helleday T, Imreh MP, Göndör A. Mol Cell, 2015. 59(6): p. 984-97.

*Authors contributed equally to this study

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

2-OG	2-oxoglutarate
3C	Chromatin conformation capture
3D	Three-dimensional
4C	Circular chromatin conformation capture
5caC	5- carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AML	Acute Myeloid Leukemia
Ash1	Absent, small and homeotic discs 1
ASXL1	Additional sex-comb like-1
BAC	Bacterial artificial chromosome/clone
bHLH	Basic helix-loop-helix
BMAL1	Brain and Muscle ARNT-like 1
bp	Base pair
BRD4	Bromodomain-containing 4
CAGE	Cap analysis of gene expression
CAS9	CRISPR Associated Protein 9
CBX	Chromobox homolog
CCG	Clock-controlled gene
CD	Cluster of Differentiation
CFU	Colony Forming Unit
CGI	CpG islands
CHD2	Chromodomain Helicase DNA Binding Protein 2
ChrISP	Chromatin in situ proximity analysis
CLOCK	Circadian Locomotor Output Cycles Kaput
CLP	Common Lymphoid Progenitor
CMP	Common Myeloid Progenitor
Cpf1	CRISPR from *Prevotella* and *Francisella* 1
CRISPR	Clustered Regularly Interspaced Short Palindrome Repeats
crRNA	CRISPR RNA
CRY	Cryptochrome
CT	Chromosomal Territory

CTCF	CCCTC-binding factor
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferases enzyme
DSB	Double-strand break
EED	Embryonic Ectoderm Development
EPO	Erythropoietin
EZH2	Enhancer of zeste homolog 2
FISH	Flourescence in situ hybridization
GFP	Green fluorescent protein
GM	Granulocytic-macrophage
GM-CSF	Granulocyte-macrophage colony-stimulating factor
G-SCF	Granulocyte colony-stimulating factor
H3K4	Lys4 of histone H3
H3K9me2	Di-methylation of Histone H3 lysine 9
H3K27me3	Tri-methylation of lysine 27 on histone H3
HAT	Histone acetyltransferases
HDAC	Histone deacetylases
HDR	Homology-directed repair
HEBs	Human embryoid bodies
HESCs	Human embryonic stem cells
HP1	Heterochromatin protein 1
HSC	Hematopoietic Stem Cell
ICR	Imprinted control region
IDH1/2	Isocitrate dehydrogenase 1 and 2
IL-3	Interleukin 3
isPLA	*in situ* proximity ligation assay
ISWI	Imitation switch
LAD	Lamina-associated domain
lncRNA	Long non-coding RNA
LOCK	Large organized chromatin lysine modification
LSD1	Lysine specific histone demethylase 1
MegE	Megakaryocytic-erythroid
MLL1	Methyltransferase mixed lineage leukemia 1
ncRNA	non-coding RNA
NHEJ	Nonhomologous end-joining

NK	Natural killer
oxi-mC	Oxidized methylcytosine
PAM	Protospacer adjacent motif
PAR	Poly(ADP-ribose)
PcG	Polycomb group
PCGF	PcG RING finger protein
PER	Period
PHC	Polyhomeotic homolog proteins
PHD	Plant homeodimain
PMA	Phorbol 12-myristate 13-acetate
PRC1	Polycomb Repressive Complex 1
PTM	Post translational modification
RAWUL	Ring finger and WD40 associated Ubiquitin-Like
RYBP	RING1- and YY1-binding protein
SCF	Stem cell factor
SCN	Suprachiasmatic nucleus
sgRNA	single guide RNA
SSC	Sodium salt citrate
SUZ12	Suppressor of zeste 12
SWI/SNF	Switching/sucrose non fermenting
TAD	Topological associate domain
TALENs	Transcription activator-like effector nucleases
TC	Tag clusters
TCH	Terminal conserved hairpin
TET	Ten eleven translocation
tracrRNA	Transactivating crRNA
TrxG	Trithorax group
TSS	Transcription start site
ZFNs	Zinc-finger nucleases

1 INTRODUCTION

1.1 EPIGENETICS

The “epigenetics” term was used in 1942 by Conard Waddington to describe events that genetic principles could not explain those [1], and later various inexplicable biological phenomena added to the categories of epigenetics [2]. The “epigenetics” comes from Greek, and it means “outside conventional genetics” [1,3], which is a bridge between genotype and phenotype- a series of events that change the final consequence of a locus or chromosome without DNA sequence alteration during development [2]. In other words, the genetic information of an organism can express differentially in both time and space without directly affecting the sequence of DNA, and this can only happen with the help of epigenetic regulators [3,4].

1.1.1 Chromatin

As the DNA in a eukaryotic cell pictured in figure 1, it organized in chromatin fibers with the nucleosome as a repeating unit [5,6]. In each nucleosome, 145-147 base pairs (bp) of DNA are wrapped around a nucleosome core: two copies of histone proteins H2A, H2B, H3, and H4 [5]. The linker histone H1 can assemble the nucleosome cores into higher-order structures and compact linear DNA by approximately 30–40 folds. Nucleosome core, linker DNA, and H1 form the nucleosome. The nucleosome is the main factor for DNA condensation within the nucleus, and DNA accessibility [5].

Euchromatin and heterochromatin are two well-defined states of chromatin that considered to be active and repressive, respectively. Facultative heterochromatin is a region that can switch between two states of transcription: activation and repression. Heterochromatin regions divided into different domains based on their modification and position [7].

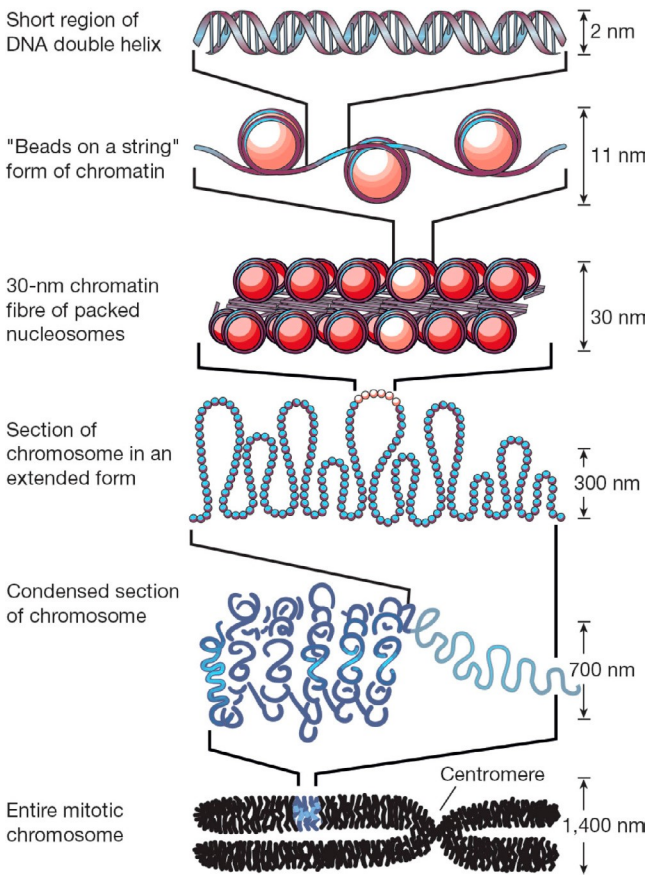

Figure 1. The organization of DNA within the chromatin structure. The nucleosome is the lowest level for DNA organization and folded approximately 50-fold into 30 nm fiber. The details structures of folding are still unclear. Figure reprinted with permission from the publisher [8].

1.1.2 Epigenetic Mechanisms

Nowadays, all inherited changes that can alter gene expression without changes in primary nucleotide sequences defined as epigenetics [2,9] and there are different mechanisms for these alterations such as methylation of cytosine in the CpG dinucleotide in the DNA [10], covalent modifications of the N-terminal histone tails in the nucleosomes [11], remodeling of nucleosomes [12], and transcriptional or post-transcriptional gene silencing through the small regulatory non-coding antisense microRNA [13], or long non-coding RNA (lncRNA) [14]. We will discuss some of these mechanisms in more detail in this thesis.

1.1.2.1 DNA Methylation

As mentioned before, DNA methylation is one of the epigenetic mechanisms. It is responsible for inactivation status in one of the X-chromosome in female cells, and because of that, it has suggested being an epigenetic mechanism of imprinting [15,16]. DNA methylation is a dynamic epigenetic mark and mostly happens at the five positions of the cytosine position in CpG dinucleotides [17]. The prevalence of CpG methylation (approximately 70-80% in mammalian genomes) occurs in specific regions called CpG islands (CGIs), which are rich for GC sequences [18]. CGIs are around 1 kb in length and are, to a high degree, non-methylated in germ cells, in the early embryo, and most somatic tissues. GCI's promoters found in around 60% of human genes [19]. As it shows in figure 2, there are two other regions with different levels of methylation: located up to 2 kb away from a CGI named shores and within 2 – 4 kb of a CGI as shelves [20]. There are open sea areas with more distance from CGI, which are not in figure 2.

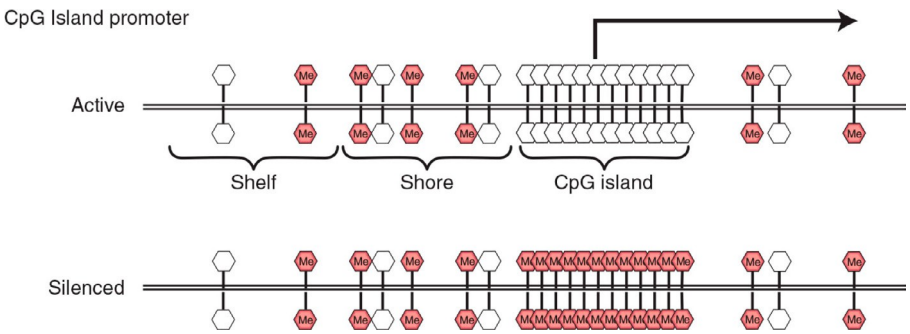

Figure 2. CpG iIsland (CGI) promoter. CGI, shore, and shelf pictured in the figure for the active and silenced situation. Figure reprinted with permission from the publisher [19]. Copyright to Cold Spring Harbor Laboratory Press.

There are three different DNA methyltransferases enzymes: DNMT1, DNMT3A, and DNMT3B. DNMT1 is maintaining DNA methylation patterns during mitosis, DNMT3A and DNMT3B are *de novo* methyltransferases. In normal cells, approximately 3-6% of cytosines are methylated. In the cancer cells, the aberrant DNA methylation can take place. Hypomethylation in the genome of cancer cells can cause genome instability [21]. There are two pathways for DNA demethylation: active and passive [22]. In the active pathway, the ten-eleven translocation (Tet) proteins family (TET1, TET2, and TET3) can change 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) [22,23]. It followed by the next steps in which

5hmC, undergoes further oxidation into 5-formylcytosine (5fC) and 5- carboxylcytosine (5caC), which finally converted to unmethylated cytosine. In order to facilitate their functions, TET proteins need ferrous iron (Fe^{2+}) as an essential cofactor and 2-oxoglutarate (2-OG) as an obligatory co-substrate [22]. Oxidized methylcytosines (oxi-mCs) enriched at promoters, enhancers, and gene bodies which, can have effects on gene expression [24]. Passive demethylation happens through replication during cell division. This can happen in the absence or inhibition of DNMT1 [19].

1.1.2.2 Histone Post Translational Modifications

As mentioned earlier, DNA packed with the help of histones proteins into chromatin. Histones can obtain different post-translational modifications on their N-terminal tails, such as methylation, acetylation, sumoylation, phosphorylation, ubiquitination, and ADP-ribosylation [25,26]. In table 1, some of these modifications and their functions summarized.

Different Classes of Modifications Identified on Histones

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	**K**-ac	Transcription, Repair, Replication, Condensation
Methylation	**K**-me1 **K**-me2 **K**-me3	Transcription, Repair
Methylation	**R**-me1 **R**-me2a **R**-me2s	Transcription
Phosphorylation	**S**-ph **T**-ph	Transcription, Repair, Condensation
Ubiquitylation	**K**-ub	Transcription, Repair
Sumoylation	**K**-su	Transcription
ADP ribosylation	**E**-ar	Transcription
Deimination	**R>Cit**	Transcription
Proline Isomerization	**P-Cis> P-trans**	Transcription

Table 1. Different classes of modifications identified on the core histones and their modified residues. The table adapted with permission from the publisher [26].

Histone modifications can classify into two different groups according to their effects on transcription: activators or repressors. However, some modifications can act as an activator or repressor under different situations, for example, methylation at lysine 9 at histone H3 has a negative effect on the promoter and positive in the coding region [27,28].

1.1.2.3 Histone-Modifying Enzymes

Most of the histone modifications are dynamic processes. In table 2, some of the histone-modifying enzymes and their target residues summarized.

Histone-Modifying Enzymes

Enzymes that Modify Histones	Residues Modified
Acetyltransferases	
HAT1	H4 (K5, K12)
CBP/P300	H3 (K14, K18) H4 (K5, K8) H2A (K5) H2B (K12, K15)
PCAF/GCN5	H3 (K9, K14, K18)
TIP60	H4 (K5, K8, K12, K16) H3 K14
Methyltransferases	
G9a	H3K9
CLL8	H3K9
MLL1	H3K4
SET1A	H3K4
SET1B	H3K4
ASH1	H3K4
EZH2	H3K27
RIZ1	H3K9
Demethylases	
Lsd1/BHC110	H3K4
JHDM1a	H3K36
Deacetylases	
SirT2 (ScSir2)	H4K16

Table 2. Histone-Modifying Enzymes and their modified residues. The table adapted with permission from the publisher [26].

Two main functions considered for histone modifications: implementing global chromatin environments and coordinating DNA-based biological functions. Histone modifications not only can affect each other, but they can also communicate with DNA methylation [26].

Histone acetyltransferase (HAT) enzymes transfer acetyl group from acetyl-CoA to specific lysine residues, which can result in a chromatin structure diffusion, which gives accessibility to transcriptional factors [29]. Based on their cellular localization, HATs classified into two groups [30]. The nuclear localization has seen in type A HATs; there are several transcriptional factors such as p600, CLOCK, and TAF1 among this group [29]. Type B HATs acetylate the newly synthesized histones, and they are localized in the cytoplasm [29]. HATs family members play a role in normal hematopoiesis and malignancies, as pictured in figure 3 [29].

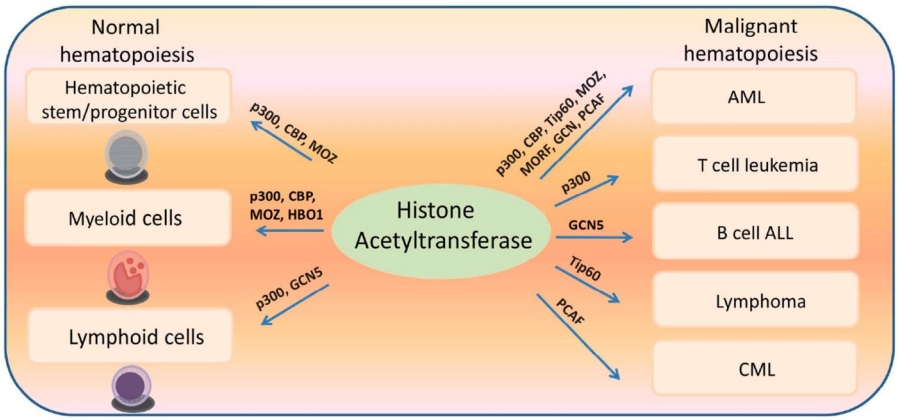

Figure 3. Histone acetyltransferases (HATs) regulate both normal and malignant hematopoiesis. HATs generate H3K27ac in, active enhancers. There is crosstalk between histone acetylation and methylation in hematopoiesis. HDAC inhibitors used for the therapy in malignant hematopoiesis. Figure reprinted with permission from the publisher [29].

Lysine or arginine residues in histones can accept different modifications. They can be methylated. Lysine methylation of histones creates specific signals depending on the residue. Lysine can be acetylated, which can promote gene activation [31-33]. Most of the characterized histone methyltransferases contain a SET domain, and typically they are specific about their targets on histone proteins [25]. Arginine methylation can regulate transcriptional activation [34]. The enzymes associated with adding or removing of histone methylation are key regulators for cell development and linked with human diseases [25]. For example, a lysine-specific histone demethylase 1 (LSD1) is necessary for the differentiation of hematopoietic cell lines in human [35]. The overexpression of several histone demethylases has reported in various cancers [35].

1.1.2.4 Polycomb-group Proteins

Polycomb group (PcG) proteins are transcriptional repressors that can modify histone proteins and their activities [36]. The PcG complexes play critical functions in regulating cell proliferation, self-renewal, and differentiation in several tissues, including blood cells [37]. Polycomb Repressive Complex 1 (PRC1) and 2 (PRC2) belong to nuclear complexes [37,38]. In table 3, each complex and its subunits are listed.

PRC2 complex catalyzes the transcriptionally repressive di-methylation and trimethylation of lysine 27 on histone H3 (H3K27me2/3). The catalytic subunit of PRC2 is the enhancer of zeste homolog 2 (EZH2) subunit. EZH2 binds to suppressor of zeste 12 (SUZ12) and embryonic ectoderm development (EED). All these subunits together constitute the core unit of the enzymatically active PRC2 complex [37,38] which are necessary for PRC2 integrity and PRC2-mediated H3K27 methylation [39]. The other PRC2 subunits are the histone deacetylases HDAC1, HDAC2, histone-binding proteins retinoblastoma-associated protein 46 (RbAp46), and RbAp48, that are not essential for its activity [38].

PRC1		PRC2		TrxG	
Subunit	**Molecular Function**	**Subunit**	**Molecular Function**	**Subunit**	**Molecular Function**
CBX	Chromodomain binds H3K27me3	**EZH2**	SET domain methylates H3K27	**ASH1L**	SET domain methylates H3K36
PCGF	Binds DNA and compacts chromatin	**SUZ12**	Enhances E(z) HMTase activity	**MLL C-ter**	SET domain methylates H3K4
PHC	SAM domain self-associates	**EED**	Enhances E(z) HMTase activity binds H3K27me3	**MLL N-ter**	Required for H3K27 acetylation by CBP
RING1A and RING1B	Ubiquitylates H2AK118 (K119 in vertebrates), compacts chromatin	**RbAp46 and RbAp48**	Binds histones and SU(z)12	**BRD4**	Bromodomains bind acetylated Lys BRD4 phosphorylates Pol II CTD at Ser2

Table 3. Polycomb repressive complex 1 (PRC1), Polycomb repressive complex 2 (PRC2), and Trithorax group (TrxG) proteins subunits and their functions. The table adapted with permission from the publisher [40].

As the schematic figure 4 shows, in mammalian cells, PRC1 complexes are heterogeneous and classify into two groups based on PcG RING finger proteins (PCGFs): PRC1.2 and PRC1.4 as canonical PRC1 complexes and PRC1.1, PRC1.3, PRC1.5, and PRC1.6 belong to uncanonical category [36,41].

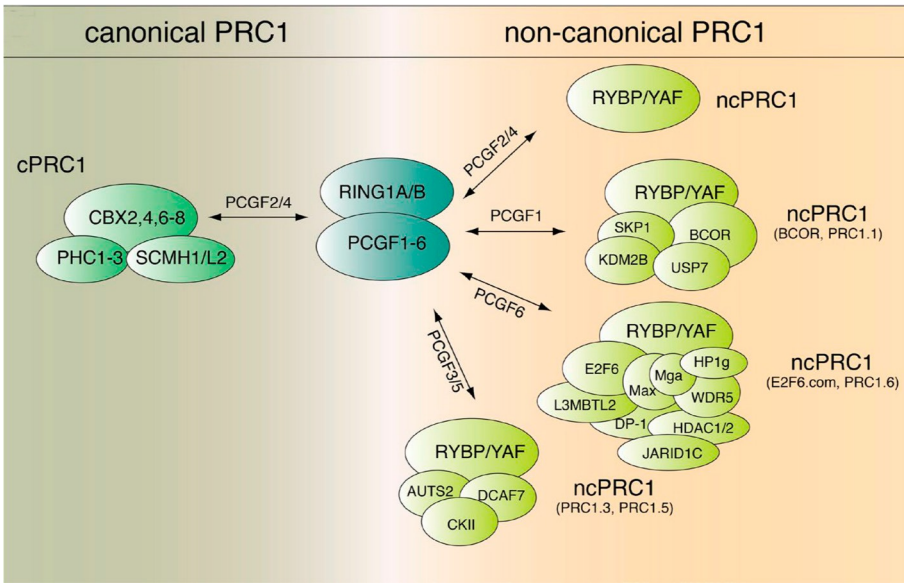

Figure 4. Subunit content in canonical and non-canonical PRC1 complexes. The RING1A/B and PCGF1-6 form the core subunits in PRC1 complexes. In canonical PRC1, one PHC and one CBX protein incorporate. There are incorporations of RYBP/YAF2 and some other subunits in noncanonical PRC1. Figure reprinted with permission from the publisher © 2019, Di Carlo V, et al. Originally published in the Journal of Cell Biology. https://doi.org/10.1083/jcb.201808028

PRC1 catalyzes mono-ubiquitylation on lysine 119 (K119) of histone H2A (H2AUb119), which also is a repressive histone marker [38]. H2AK119Ub1 can promote compaction in chromatin, which leads to inhibition of transcriptional elongation and gene silencing [38].

The core PRC1 complex consists of RING1A/B and PcG ring finger (PCGF) proteins (Figure 4). Each PRC1 complex has a specific PCGF1-6 subunit: NSPC1/PCGF1, MEL-18/PCGF2, PCGF3, BMI-1/PCGF4, PCGF5, or MBLR/PCGF6 [36].

The ubiquitin E3 ligase complex contains different PcG proteins, such as RING1/Ring1A, RING2/Ring1B, and BMI-1/Bmi-1 [42]. The catalytic subunits of PRC1 are RING1A and RING1B [38,42]. At the N-terminal of the RING finger proteins, there is a specific type of Zn^{2+}-binding motif, Cys3HisCys4. There is a Ring finger and WD40 associated Ubiquitin-Like (RAWUL) domain at their C-terminal. The RING finger motifs pair and use E3 ubiquitin ligases and the RAWUL motifs as binding platforms for other PRC1 subunits [41]. The canonical PRC1.2 and PRC1.4 complexes are the only ones with Polyhomeotic homolog proteins (PHC) and chromobox homolog (CBX) proteins [36].

Different mechanisms have suggested for Polycomb complexes recruitment to their specific targets [43]. The PRC1 recruitment to its target can be dependent or independent on pre-existing trimethylation (H3K27me3) marks [44]. The CBX subunit in PRC1 recognizes target sites with H3K27me3, which leads to ubiquitination on H2AK119 [43]. It has also suggested that transcription factors or lncRNA could participate in Polycomb recruitment [43,45]. The noncanonical PRC1s (PRC1.1, 2, 5, and 6) have the RING1- and YY1-binding protein (RYBP) subunits that are involved with H3K27me3-independent recruitment [46]. For the independent pathway, the Kdm2b suggested to recognize CpG islands and help with PRC1 recruitment [47,48].

As mentioned before (Figure 4), in mammalian systems, there are multiple PRC1 and PRC2 complexes that are encoded by multicopy PcG genes that give them a diversity of function [49]. In addition, PcG machinery also linked to X inactivation [50], parent-of-origin imprinting [51], and cancer epigenetics [52].

The PcG complexes can lead to transcription repression by removing HATs from their target genes [53]. PRC2 depletion can lead to increasing the H3K27 acetylation globally, which is catalyzed by p300 and CBP. PRC1 are very dynamic complexes. They can evolve between different cell stat developments [41]. Chromatin locations that enriched with PRC1 and PRC2 complexes are dispersed throughout euchromatin and can overlap with each other's but rarely find in heterochromatin areas and silenced domains [41,54]. The Polycomb complexes silence essential target genes for the cell maintenance identity during cell states transitions [41].

1.1.2.5 Trithorax Group (TrxG)

The PcGs are not the only large multiprotein complexes that can change chromatin with catalyzing covalent modifications on histones and leading to chromatin structural changes. Silencing and activation need to be in dynamic balance [40].

Trithorax protein group (TrxG) is responsible for activation, which acts antagonistically from PcGs [40]. For instance, histone lysine methylation at Lys4 (H3K4) and Lys36 of histone H3 (H3K36), which catalyzed by Trx and Absent, small and homeotic discs 1 (Ash1) respectively, inhibits PRC2-mediated trimethylation at histone H3 lysine 27 (H3K27me3) [55,56]. In table 3, the main subunits for each complex described.

Another layer of antagonism operation for PcG and TrxG is through RNA polymerase II (Pol II) [40]. Histone lysine ubiquitylation of histone H2A mediated by the PRC1 complex, and it colocalizes with Pol II. This ubiquitylation is necessary for the existence of an unproductive, 'poised' Pol II with Ser5-phosphorylation, at bivalent genes in embryonic stem cells, that might prevent elongation step [57]. Although, phosphorylation at Ser2 in Pol II can occur via the TrxG protein bromodomain-containing 4 (BRD4) and may promote elongation [58].

PcG and Trx have opposite effects on transcription [40]. PRC1 promotes chromatin compaction [59], whereas Trx promotes an open configuration. The H3K27 acetylation facilitated by Trx can neutralize the positive charge of Lys and disrupt histone-DNA contacts [60] [40].

1.1.2.6 Non-coding RNA

Lots of studies have shown that non-coding RNAs (ncRNA) play an important role in epigenetic regulation [61]. There are different categories of ncRNA. They can categorize based on their size. Long non-coding RNAs (lncRNA) are longer than 200 nucleotides, and small non-coding RNAs are shorter than 200 nucleotides [62].

Long non-coding RNAs (lncRNAs) are associated with different mechanisms. They can regulate gene expression, coordinate chromatin structure, and involve in mRNA stability [63]. The lncRNAs are involved in cellular proliferation and differentiation and can act as oncogenes in different cancers [64].

MicroRNA is another group of ncRNAs that containing approximately 22 nucleotides, which can cleave from 70-100 nucleotide hairpin precursors and can hybridize with complementary mRNA target genes and inhibit their functions [13]. MicroRNAs have regulatory roles and are responsible for post-transcriptional gene silencing [65]. MicroRNAs dysregulation has found in several solid tumors and hematologic malignancies [66].

1.1.2.7 Chromatin Remodelers

ATP-dependent chromatin remodelers contain different assembled complexes that each of them has an ATPase subunit, which belongs to the SNF2 protein superfamily [67]. These enzymes depend on the existence of other conserved domains that are categorized into the mating type switching/sucrose non-fermenting (SWI/SNF), inositol (INO80), chromodomain helicase DNA-binding (CHD), and imitation switch (ISWI) families [67]. These remodelers are specific for cell-type and developmental-stage [68]. *In vitro* studies showed that they all increase nucleosome mobility [68].

1.1.2.8 Chromodomain Helicase DNA-binding (CHD) Family

CHD family has structural and functional domains that play roles in potential physical interactions with nucleosomes: (*a*) tandem chromodomains (chromatin organization modifier) located in the N-terminal region, which is in common with other chromatin-associated proteins such as Polycomb and heterochromatin protein 1 (HP1). (*b*) In the central region of the protein structure, the helicase/ATPase domain, which has a high similarity with the SWI2/SNF2 ATPase. (*c*) The C-terminal DNA-binding domain, with a preferential to bind DNA regions with A+T-rich sequences [67,69].

The CHD protein family has nine members and three subfamilies [69]. The subfamily I consists of CHD1 and CHD2, which contain all the common domains for the CHD family. The CHD3 and CHD4 belong to subfamily II which they do not have the DNA-binding domain, but they have a double plant homeodomain (PHD) zinc-finger domains at their N-terminal. The rest of the CHD members belong to subfamily III. The majority of this subfamily members contain a conserved terminal hairpin (TCH) motif or a DNA binding domain named SANT domain [70]. These remodelers considered as either transcriptional activators or repressors [70]. These families with more structure domains showed in figure 5.

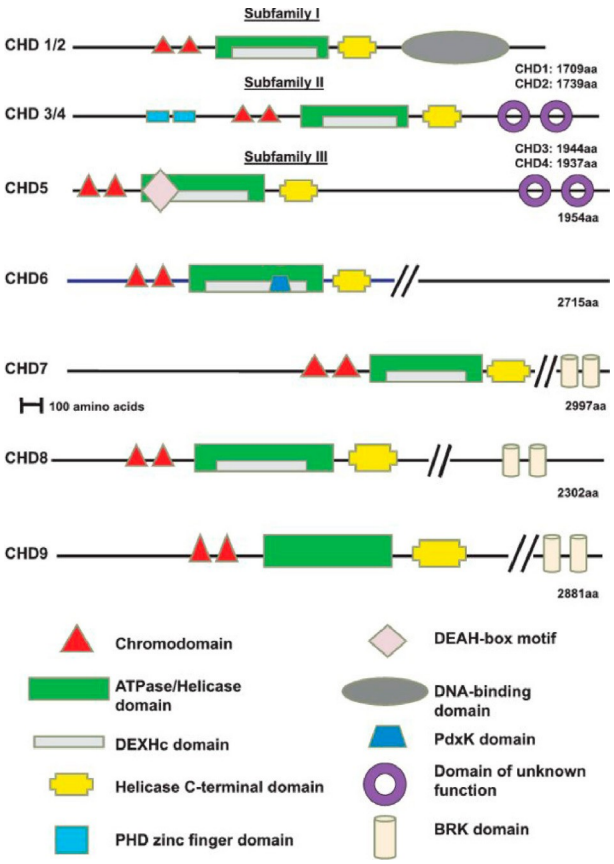

Figure 5. Schematic representation of all known CHD 1-9 proteins in human and their structural domains. The two truncated chromodomains in N-terminal, the SNF2-like helicase-ATPase domain at the center, and DNA-binding domain in C-terminal are the common CHD domains pictured in the figure. Figure reprinted with permission from the publisher [69].

It has shown that the CHD family can change nucleosome composition or its location [69]. CHD1 and CHD2 have a role in transcription activation and elongation as well as they have interactions with different elongation factors, transcription factors, activators, and co-activators [69].

1.1.3 3D Genome Organization

In higher eukaryotes, the genome organized non-randomly in the three-dimensional (3D) space of the nucleus. Instead, in the interphasic nucleus, individual chromosomes occupy specific delimited regions called “chromosome territories” (CT)

[70-73], which constitute a significant feature of nuclear architecture. In order to achieve the necessary degree of compaction to fit within these areas, the chromatin fiber needs to condense by looping into itself [74]. This organization within the nuclear space has functional implications in the regulation of gene expression and in other nuclear processes. Thus, the radial position of chromosomes and genes in the nucleus is cell-type and tissue-specific and is often altered in cancer and disease cells [74].

Chromosome territories further organized into sub-chromosomal domains. Chromosomes first organized into two different types of compartments: A ("active") and B ("inactive), in accordance with their transcriptional status and the degree of chromatin compaction. Within those compartments, chromatin further organized into topologically associating domains (TADs), self-interacting domains from several hundred kb up to 1-2 MB in size, with an average of around 800 kb [73] in mammals. TADs formed with the help of specific architectural proteins like CCCTC-binding factor (CTCF) and cohesin, which often found at their boundaries [75]. Likewise, chromosome territories, TADs can differ in structure following gene activation in different cell types or conditions [76].

1.1.3.1 Active and Inactive Domains of Genome Organization

The non-random organization of the genome has a direct correlation with gene density and transcriptional activity. In general, genes placed in gene-dense regions tend to be more active in comparison to genes located in gene-poor regions. This placement results in megabase sized domains, which switch between high and low transcriptional activity [77].

A clear example of that is the nuclear envelope, which has a regulatory role in transcription and gene regulation [74]. In most higher eukaryotes, the nuclear periphery is enriched in condensed heterochromatin and has a connection with transcriptional repression [73]. Those domains directly associated with the nuclear lamina are called lamina-associated domains (LADs) [78]. LADs contain gene-poor regions and contain developmentally repressed genes that emerge during differentiation. LADs can vary in their size, from approximately 10 kb to a few megabases [74]. Besides, LADs boundaries enriched for binding sites for the insulator protein CTCF [74].

LADs highly overlap with other heterochromatin domains enriched for histone H3 lysine 9 di-methylation (H3K9me2), named large organized chromatin lysine modifications (LOCKs). LOCKs and LADs can change in size during development, present cell-type specificity, and lose in cancer cells [78,79].

1.1.4 Transcription and Promoters

The number of different transcripts that one gene can have is unclear. There are less than 20000 genes in human that are encoding more than 80000 protein-coding transcripts, and this suggests there are extensive regulation mechanisms at the transcriptional, translational, and post-translational levels [80]. There are four different regulatory mechanisms for alternative transcription: a) alternative transcription initiation, b) alternative translation initiation, c) alternative splicing, and d) alternative polyadenylation [80,81]. The alternative transcription initiation is the outcome of using alternative promoters and transcription start sites (TSSs) in protein-coding transcripts [80]. In mammalian genomes, more than 70% of genes have multiple polyadenylation sites, more than half of genes have alternative TSSs, and approximately all genes have alternative splicing [81].

There are two classes of mammalian promoters: conserved TATA box-enriched promoters and CpG-rich promoters [82]. The TATA-box promoters are usually associated with tissue-specific genes and highly conserved across species, and they are a minority in both mouse and human [82]. The board distribution of CpG islands represents the majority of promoters in mammalian [82].

1.1.4.1 Alternative Transcription Start Site (TSS)

The regulation for using alternative AUG and translation isoform depends on the availability of translation initiation factor complex [83]. The binding of different transcription factor complexes at the regulatory elements in promoter sequences can result in more than one RNA transcript, which can lead to different transcription isoforms [83] (Figure 6).

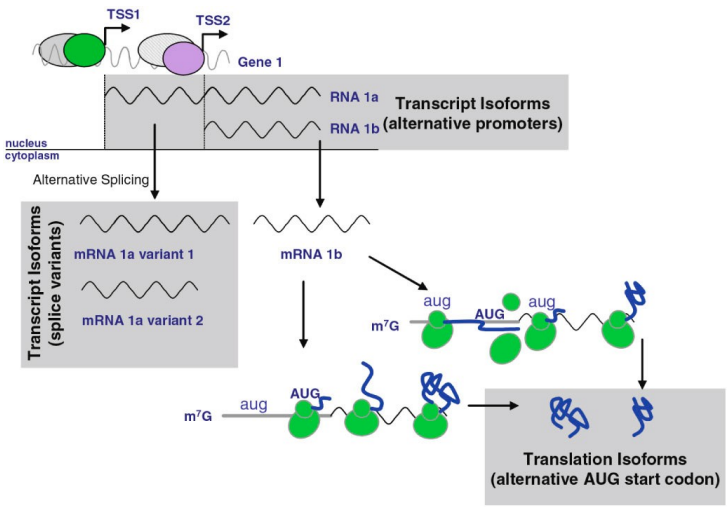

Figure 6. Mechanisms of isoform formation. Different transcription start sites (TSS), and alternative promoters, alternative splicing, and alternative translational start sites can result in different isoforms. Figure reprinted with permission from the publisher [83]. This research originally published in the International Journal of Hematology. Grech G, et al., Expression of different functional isoforms in haematopoiesis. Int J Hematol. 2014, 99(1), pp 4–11. The original publication is available at https://link.springer.com/journal/12185

Alternative transcript isoforms are essential for biological regulation, and their misexpression linked with different diseases, including cancer [81]. It has shown that alternative transcript isoform choice has tissue-specific regulation in the human genome, which is affecting approximately half of multi-exonic genes [81].

1.2 HEMATOPOIETIC SYSTEM

1.2.1 Hematopoiesis

The most regenerative tissue in human body is the blood which, can produce up to 10^{12} cells per day [84]. Hematopoietic stem cells (HSCs) have self-renewal capacity. As a result of this capacity, both the common myeloid progenitor (CMP) and the common lymphoid progenitor (CLP) are generated from HSCs [85,86].

The adult hematopoietic system consists of two separate lineages: myeloid and lymphoid. The lymphoid lineage includes the B, T, and natural killer (NK) cells. The myeloid lineage is more diverse and includes monocytes, macrophages, erythrocytes, megakaryocytes, granulocytes (neutrophils, eosinophils, and basophils), and mast cells. [86,87]. (Figure 7)

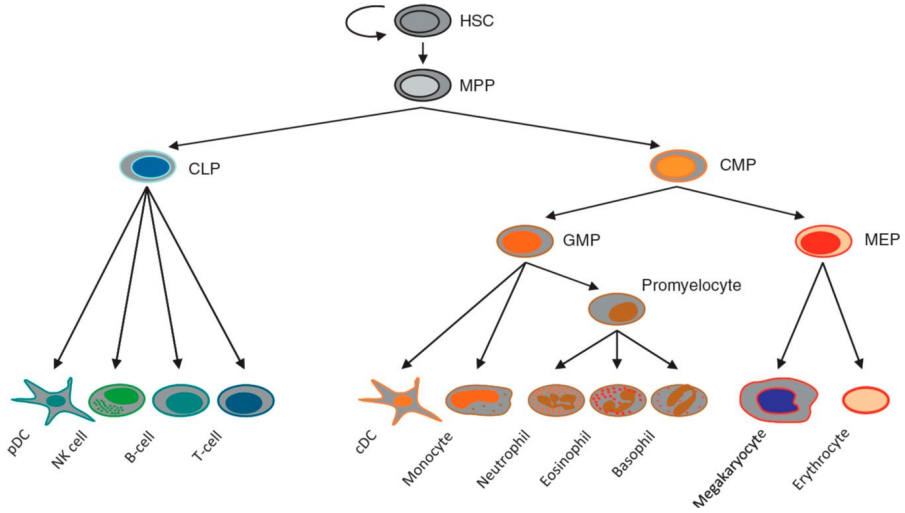

Figure 7. Hematopoiesis. Hematopoietic stem cells give rise to two major cell lineages, the myeloid and lymphoid. HSC, hematopoietic stem cell; MPP, multipotent progenitor; CMP, common myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-monocyte progenitor; MEP, megakaryocyte-monocyte progenitor. Figure reprinted with permission from the publisher [87].

The formation of blood cells or hematopoiesis occurs in the bone marrow niche [88]. A hematopoietic stem cell has two specific functions: self-renewal capacity and multilineage differentiation potential. Asymmetric division can provide an identical stem cell and a more mature cell [87,89].

1.2.2 Epigenetics Regulation in Hematopoiesis

Hematopoiesis is a good model to study epigenetic mechanisms. For example, DNA methylation and different gene expression regulation are so critical for cell-fate and HSC differentiation into different blood lineages [90]. DNA methyltransferases enzymes play a critical role in hematopoiesis. DNA methylation increases during lymphoid differentiation and decreases during myeloid lineage development [90]. DNMT3A and DNMT3B are involved in HSC renewal and differentiation. DNMT3B has more specific expression patterns than DNMT3A and is expressed only in hematopoietic stem cell and hematopoietic progenitor cells (HPCs) [91,92].

As mentioned before, histone modifications are epigenetic regulatory factors. Their role in chromatin status and their enzyme modifiers are essential in the regulation of hematopoiesis differentiation; their dysregulation has reported in different types of leukemia [92]. As it comes before, both PRC1 and PRC2 are important for HSC self-renewal and hematopoiesis regulation [93].

Another mechanism that is important for regulation is transcription factors and cytokine receptors. The levels of the transcription factors differ in different cell lineages. For example, GATA-2 expression is in all intermediate myeloid progenitors, or stoichiometry between GATA-1 and PU.1 is important for megakaryocytic-erythroid (MegE) and granulocytic-macrophage (GM) lineages commitment [94].

Different types of leukemia can occur depending on the level of hematopoietic cell differentiation when first neoplastic transformation happens, and based on which PcG gene is involved [37]. The transcriptional activation of PcG-targeted genes has shown to correlate with methylation- to-acetylation change in MLL-AF9-transduced HSC cells [29].

1.2.3 Epigenetics Regulation in Acute Myeloid Leukemia

Acute myeloid leukemia (AML) is an aggressive clonal malignancy characterized by the accumulation of abnormally differentiated or poorly differentiated cells in bone marrow due to somatic genetic mutations in hematopoietic progenitor cells that change standard mechanisms of proliferation, self-renewal, and differentiation [66,95]. The most common type of acute leukemia in adults is AML, which is the leading cause of death among leukemias in the United States [96].

The data from the Cancer Genome Atlas AML sub-study revealed that mutations involved in AML can classify into one of these nine classifications: DNA methylation-related genes, transcription factor fusions, myeloid transcription fac-

tor genes, the NPM1 gene, chromatin-modifying genes, tumor suppressor genes, cohesin complex genes, signaling genes, and spliceosome complex genes [97].

The epigenetic aberrations have a significant role in AML occurrence [92]. Several mutations in epigenetic regulators have detected, for example, DNMT3A with 26% to 16%, Isocitrate dehydrogenase 1 and 2 (IDH1/2) enzymes with 33% to 15% and Ten-eleven translocation 2 (TET2) enzyme with 23% to 7% are the most common mutations in epigenetic regulators in AML [92].

Another layer of dysregulation in AML occurs in histone modifier enzymes. Mutations in EZH2, ASXL1 (additional sex-comb like-1), and MLL (mixed-lineage leukemia) have reported in different patients studies [98,99]. MLL has reported as the most dysregulated histone modifier in AML [100].

As mentioned before, the nuclear organization is important for gene expression regulation [101]. Different epigenetic regulators are involved in 3D nuclear structures, and their dysregulation reported in different cancers and especially AML. Cohesin is one of these regulators. All the members of the cohesin complex reported being mutated in AML patients [102,103]. In AML, mutations in cohesin can associate with mutations in TET2, DNMT3A, RUNX1, or NPM1 [104].

Besides, hypermethylation that can cause silenced genes is involved in myeloid malignancies as both prognostic markers and therapeutic targets [66]. Therefore genome-wide epigenetic profiling is critical for understanding AML to have a more accurate molecular therapy [66].

1.3 CIRCADIAN RHYTHM MACHINERY

The term "*circadian*" comes from the Latin "*circa diem*" which means "about a day". Circadian rhythms defined as the physiological and biochemical properties of the human body that recur with approximately 24-hour cycles [105]. Sleep/wake cycles, for instance, represent a manifest of this internal timing [106]. Circadian clocks exist in most of the life forms, giving the organism the ability to predict daily variations in the environment and have appropriate physiological responses to adapt to it [107].

In mammals, the master circadian clock situated at the suprachiasmatic nucleus (SCN), located at the anterior part of the hypothalamus, and it controls oscillating circadian rhythms of many physiological and behavioral responses [108]. Circadian clocks need to readjust daily by external time cues or Zeitgebers [106]. At the SCN, light is the predominant zeitgeber, while at the peripheral organs, feeding-fasting rhythms are more dominant [109].

1.3.1 Circadian Clock Regulation

The circadian clock is under the control of negative transcriptional and translational feedback loops [110]. Two basic helix-loop-helix (bHLH) transcription factors in mammals, Circadian Locomotor Output Cycles Kaput (CLOCK), and Brain and Muscle ARNT-Like 1 (BMAL1), constitute the positive limb of the feedback loop. Upon activation, these transcription factors heterodimerize and bind to conserved E-box regulatory sequences in their target promoters [111] to promote transcriptional activation of clock-controlled genes (CCGs) such as the Cryptochrome-encoding genes (Cry1-2), and Period-encoding genes (Per1-3) [112]. CRY and PER proteins make then a complex in the cytoplasm that translocate back to the nucleus and can inhibit CLOCK: BMAL1-mediated gene expression [111,112], conferring the negative limb of the feedback loop. This regulation by CLOCK: BMAL1 heterodimers affects a broad range of physiological functions [113].

The circadian machinery controls cellular transcription to provide proper adaptation to the environment regarding the diurnal cycle. Between 2-30% of all mammalian transcripts undergo circadian oscillation depending on cell or tissue type [113,114].

1.3.2 Circadian Clocks and Epigenetics Regulations

There have been studies showing that not only DNA methylation but also histone post-translational modification (PTMs) can associate with the circadian machinery [115]. Specific epigenetic remodelers are under the coordination of the molecular clocks. For instance, histone H3K4-specific methyltransferase mixed-lineage leukemia 1 (MLL1) interacts with CLOCK: BML1 complex to promote oscillating circadian transcription [116,117]. Furthermore, the circadian machinery can affect chromatin architecture and DNA topology [115]. Chromatin conformation capture (3C-based) techniques showed that circadian chromatin loops occur to control specific promoter-enhancer interactions that regulate circadian transcription [118,119].

CLOCK has intrinsic histone acetyltransferases (HAT) activity which, is necessary for its gene expression and circadian function [113,120,121] and can help it to act as a chromatin modifier [122]. This function can be enhanced by BMAL1, its heterodimer partner [120].

2 AIM OF THE THESIS

The overall aims for this thesis were to study the epigenetic and transcriptional regulation of cellular development and differentiation.

Study I:
Investigate novel roles for epigenetic factors during differentiation of hematopoietic cells

Study II:
Identify the role of Polyhomeotic homolog proteins (PHC) subunits of Polycomb repressive complex 1 during myeloid differentiation

Study III:
To study the potential roles for alternative Transcription Start Sites (TSS) on protein domains exclusion

Study IV:
To understand the role of 3D chromatin structure on transcription

3 MATERIALS AND METHODS

This section provides a brief description of some of the specific methods used in studies I-IV. For more details and remaining methods please see the materials and methods for each study.

3.1 Cell Culture

The human blood cell line K-562 (ATCC® CCL-243™), established from a 53-year-old female with chronic myelogenous leukemia in terminal blast crisis, was cultured in Iscove's Modified Dulbecco's Medium (IMDM) (12440061, ThermoFisher Scientific) supplemented with 10% fetal bovine serum (10270106, ThermoFisher Scientific). K-562 cell line was used for transfection and differentiation study in paper I. Since we were not able to establish stable Cas9 expression in HL-60 or U-937 cell lines.

The KG-1 cell line (ATCC® CCL-246), established from a 59-year-old Caucasian male with erythroleukemia that evolved into acute myelogenous leukemia, was cultured in RPMI 1640 medium (21875034, ThermoFisher Scientific) supplemented with 10% fetal bovine serum in paper II.

Jurkat, Clone E6-1 (ATCC® TIB-152), established from the peripheral blood of a 14-year-old boy, was cultured in RPMI 1640 medium (21875034, ThermoFisher Scientific) supplemented with 10% fetal bovine serum in paper III.

Human colon cancer cell line HCT116 cell line (ATCC® CCL-247™), established from a male with colorectal carcinoma, was cultured in McCoy's 5A medium (ThermoFisher Scientific, 26600023) supplemented with 10 % fetal bovine serum and 1% penicillin-streptomycin (15140122, ThermoFisher Scientific). Serum shock treatments were performed [123]. HCT116 cells were cultured with serum-rich medium with 50% horse serum (16050122, ThermoFisher Scientific, 16050122) for 2 hours. Cells were cultured with serum free McCoy's 5A medium for indicated periods in paper IV.

Human female embryonic stem cells (HS181) (HESCs) were cultured on irradiated male feeder fibroblasts [124], and human embryoid bodies (HEBs) were differentiated *in vitro* from HS181 cells in paper IV.

3.2 Colony Forming Unit Assay

In vitro colony Forming Unit (CFU) Assay can be used to measure and quantify the ability of proliferation, differentiation and colony forming capacity [125].

To better understand the effect of the knock out of our target gene (*CHD2*) in K-562 cell line, we performed CFU assay. We used semi-solid methylcellulose medium (MethoCult™ H4034 Optimum, StemCell Technologies) in the presence of all cytokines including recombinant human stem cell factor (SCF), recombinant human erythropoietin (EPO), recombinant human granulocyte colony-stimulating factor (G-CSF), recombinant human interleukin 3 (IL-3), and recombinant human granulocyte-macrophage colony-stimulating factor (GM-CSF). In each sample, approximately 1000 cells were resuspended to 1mL of Iscove's Modified Dulbecco's Medium (IMDM) with 2% Fetal Bovine Serum (07700, StemCell Technologies). The cell mixture was vortexed vigorously and seeded on a 35 mm dish (27100, StemCell Technologies). The dishes were kept in humidity at 37 °C and colonies were counted after 11 days using an inverted microscope.

3.3 CRISPR-Cas9 Screen

Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) technique is a new and powerful method for genome manipulation. CRISPR-Cas9 is an evolved defense system in bacteria and archaea against viruses and plasmids [126]. It depends on small RNAs for sequence-specific detection and silencing of foreign nucleic acids [126]. CRISPR/Cas9 consists of two components: single guide RNA (sgRNA) and Cas9 endonuclease [127].

The sgRNA has two parts: a constant part which forms a stem-loop scaffold for binding to Cas9, and a 20-nt part at 5'-end for complementary binding at different target DNA sites [127].

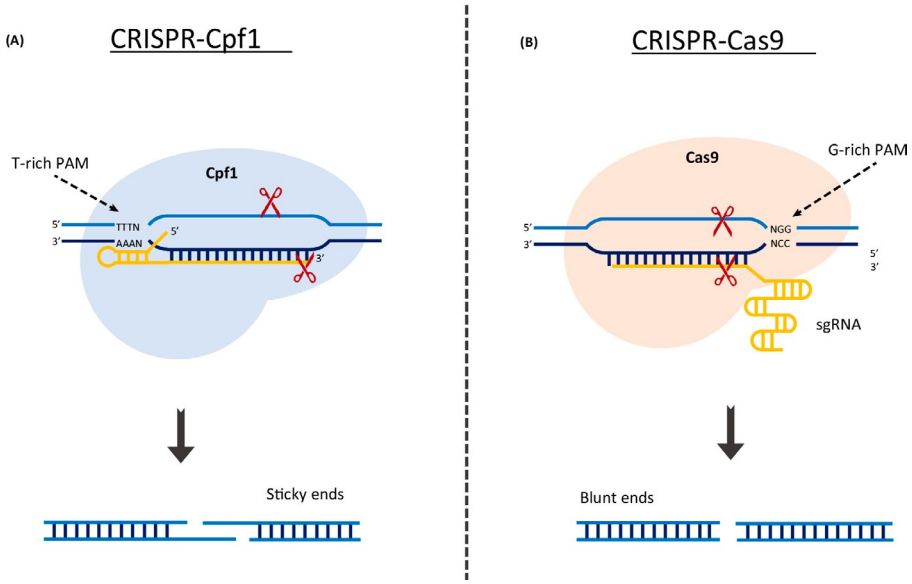

Figure 8. Schematic view for Clustered Regularly Interspaced Palindromic Repeats -Associated Proteins9 (CRISPR-Cas9) editing system and compare it with CRISPR-from *Prevotella* and *Francisella* 1 (Cpf1). In Cpf1, protospacer adjacent motif (PAM) is a T-rich region (5'-TTTN-3') in comparison with a G-rich region (5'-NGG-3') for Cas9. In the Cpf1 editing system, cohesive overhangs create after double-strand breaks (DSBs) compare with blunt ends in the Cas9 system. In both systems, the DSBs repair through Homology-directed repair (HDR) and nonhomologous end-joining (NHEJ). Figure reprinted with permission from the publisher [128].

CRISPR/Cas9 can be used for loss of function, repressing or activating the expression of a specific gene [127,129]. CRISPR/Cas9 method is easier and more efficient compare to other gene editing technologies, e.g. zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). Unlike other methods that bind to specific DNA sequence with protein-DNA recognition, CRISPR/Cas9 binds to the target by the help from the 20-nt sequence at the 5' end of sgRNA as it has shown in the figure 8. The CRISPR/Cas9 is also cheaper than other techniques but there are some drawbacks as well, for example on/off-target efficiency [127]. Different algorithms are used to overcome the efficacy and specificity in designing a good sgRNA. Another way to have lower off-target is to use CRISPR from *Prevotella* and *Francisella* 1 (Cpf1) protein instead of Cas9 because it only needs mature crRNAs (CRISPR RNA) for targeting, while Cas9 system requires both tracrRNA (transactivating crRNA) and crRNA [127]. And as it pictured in figure

8, Cpf1 generating cohesive ends in compare with blunt ends in Cas9 which also helps to increase the efficiency of its insertion [128].

In the first part of paper I, K-562 cell line with a stable Cas9 expression was used to study a library targeting 1092 epigenetic factors. The CRISPR_Cpf1 was used for gene-specific study in paper I.

3.4 siRNA Transfection

The Neon™ Transfection System 100 µL Kit (MPK10096, ThermoFisher Scientific) was used for siRNA transfection in paper II and III.

ON-TARGETplus Human PHC1 (1911) siRNA SMARTPOOL (L-011850-00-0005, Horizon Discovery); ON-TARGETplus Human PHC2 (1912) siRNA SMARTPOOL (L-021410-00-0005, Horizon Discovery); ON-TARGETplus Human PHC3 (80012) siRNA SMARTPOOL (L-015805-01-0005, Horizon Discovery) were used in paper II.

In paper III, we used pre-designed siRNA: ON-TARGETplus Human KDM2B siRNA (J-014930-07, J-014930-08-0, Horizon Discovery) and also two siRNA which designed and ordered from ThermoFisher Scientific as below:

KDM2B k1: GGCAGAAAGACTCTGGAAGAAGA (target on exon1)

KDM2B k3: CAACTATGAGTACAGAGAGAA (target on exon3)

Lipofectamine RNAiMAX Transfection Reagent (13778150, Thermo Fisher Scientific) was used to transfect CTCF siRNA (h) (sc-35124, Santa Cruz Biotechnology), GFP siRNA (sc-45924, Santa Cruz Biotechnology) or PARP1 siRNA (h) (sc-29437, Santa Cruz Biotechnology) in HCT116 cell line in paper IV.

3.5 RNA/DNA-FISH Analysis

Bacterial artificial chromosome/clone (BAC) was used to generate probes for *H19/IGF2, TLK1, VAT1L, PARD3, TARDBP*, LADs and 4C interactors in Paper IV. The BACs probes were sonicated to 500-2000 bps range followed by labelling using Bioprime Array CGH kit (18095-011, Invitrogen). Equal amounts of each labelled products were used as FISH probe.

In paper IV cells cultured on 8 wells chamber slides (154534, ThermoFisher Scientific) were crosslinked with 1 or 3 % formaldehyde for 15 minutes at room

temperature. The cells were permeabilized with 2X sodium salt citrate (SSC)/0.5% Triton for 10 minutes. The crosslinked slides kept in 70% Ethanol for storage at -20°C until further use.

In DNA-FISH the crosslinked cells were denatured in 2X SSC/ 50% formamide (F9037, Sigma-Aldrich) for 40 minutes at 80°C, cells were kept in ice-cold 2X SSC (93017, Sigma-Aldrich) for 5 minutes. The FISH probe was mixed with human Cot-1 DNA (15279011, ThermoFisher Scientific) and hybridized to the slides in a buffer containing 10 % dextran sulfate sodium (D8906, Sigma-Aldrich), 2X SSC, and 50% formamide, overnight at 37°C. Cells were washed twice with 2X SSC/ 50% formamide for 15 minutes at 42°C and with 2X SSC for 15 minutes at 42°C, and followed by mounting with Vectashield mounting medium containing 4,6-diamidino-2-phenylindole (DAPI) (H-1200, Vector Labs).

For RNA-FISH, without denaturation step at 80°C following hybridization and washing steps were performed as described before.

4 RESULTS

4.1 Study I

Epigenetic regulators are essential for normal hematopoiesis, and they are involved in different well-known pathways such as self-renewal, proliferation, and differentiation. In this study, we aimed to investigate for new roles for epigenetic regulators in hematopoietic differentiation. We used a lentivirus construct Cas9-sgHPRT with blasticidin resistance to transduce K-562 cells and create stable K-562-Cas9 cells. K-562-Cas9 cells transduced with lentivirus library with 5048 sgRNAs targeting 1092 epigenetic regulators plus 320 controls in duplicate. The transfected cells treated with Phorbol 12-myristate 13-acetate (PMA) for megakaryocytic differentiation, and cells sorted after 72 hours of treatment for megakaryocytic cell surface markers CD61/CD41 [130]. The morphological changes after PMA treatment studied by microscopy and phenotypic analysis carried out by flow cytometry. Three different gating settings used to collect different cell populations after PMA induced differentiation. We collected undifferentiated cells, which were negative for both CD61/CD41 markers (P1). The second population was double positive for both CD61/CD41 markers (P2). The last population was only positive for CD61 (P3). Two biological replicas sorted, and for each sample group, the genomic DNA sequenced for guides and UMI sequences. Each library analyzed in comparison with unsorted cells. The overlap for the top 10% of sgRNAs selected if they were positive for all four sgRNAs and had log fold change more than 0.2 for each sgRNA. These criteria helped us to narrow our top gene list to 14 candidate genes in P1, 13 genes for P2, and 30 genes for the P3 population. One of the genes in the P3 population, called *CHD2*, belongs to the chromatin remodeler family, in which different members have been shown to be important for pluripotency in myeloid cells and also differentiation of muscle cells.

For further validation of top candidates, in this study, *CHD2* were knocked out in the same cell model (K-562) and differentiation treatment. For this part, we decided to use the CRISPR-Cpf1 system since it has higher efficacy and less off-target effects compare with CRISPR-Cas9. We designed four different sgRNA located at different exons in *CHD2* (Exons 3, 7, 14, and 28) with the help of the Benchling web tool and cloned them into the vector pY095. The cloned vectors confirmed with Sanger sequencing. A mixture of all four sgRNA transfected to K-562 cells and in parallel, the original pY095 transfected as the control. After 72 hours, cells sorted for GFP signal and collected for single cells in 96 well plates. These single cells expanded and analyzed for *CHD2* KO with western blot, and the knock out cells confirmed with Sanger sequencing for each specific exon.

The mono clones induced to differentiation treatment with the same protocol as the screen, but we used 24 hours treatment instead of 72 hours, to be able to detect earlier effects on differentiation. We validated that the *CHD2* KO induced megakaryocytic differentiation by analysis differentiation after *CHD2* KO, without PMA induction by comparing with the control samples. This data was in agreement with the results from the screen. The sgRNA targeting *CHD2* enriched in the P3 population. Hence *CHD2* may have the potential to inhibit differentiation, which we confirmed in our single KO studies, demonstrating that *CHD2* KO cells were more differentiated than control cells. Also, the *CHD2* KO cells had a stronger differentiation response to PMA treatment in comparison with control cells, as the cell population for CD61/CD41 positive (P2) were significantly larger in comparison to controls. To analyze whether the induced differentiation coupled with cell proliferation, the cells seeded in low cell density, and their logarithmic growth was followed every 24 hours for four days. Our comparisons showed that *CHD2* KO cells have a lower proliferation rate in low cell density conditions. Next, we decided to analyze the ability to form new colonies in our *CHD2* knocked out cells in a colony-forming assay. Indeed *CHD2* KO cells also have a lower ability to form colonies in CFU assays.

Since RNA polymerase II is necessary for CHD2 recruitment to the active transcription start sites [131], we wanted to analyze the effect on transcription in *CHD2* KO cells. For this purpose, we used CHD2 CHIP-sequencing data for the K-562 cell line from the ENCODE project to find CHD2-binding genes. We identified 8872 CHD2 target genes. The G-ontology analysis showed that CHD2-target genes are involved in different cell functions, such as chromatin organization, histone modification, and cell cycle. In addition, we analyzed the K-562 CAGE data from the FANTOM 5 consortium, which showed that the expression level for CHD2-target genes is significantly higher in comparison with CHD2 non-target genes. To further determine the role of CHD2 on transcription, we performed RNA-sequencing on our *CHD2* KO cells and controls. The RNA-sequencing data showed the importance of CHD2 in active transcription. CHD2 target genes were significantly repressed in *CHD2* KO compare to the control cells. We also analyzed the role of *CHD2* in AML patients in the Cancer Genome Atlas (TCGA) cohort for 162 *de novo* AML patients. Our analysis demonstrated a significant overlap between *CHD2* co-expressed genes in AML patients and CHD2-target genes in K-562 cells. These data revealed that *CHD2* might be involved to promote genes transcription in AML patients.

4.2 Study II

Epigenetic modifiers and specifically the Polycomb repressive complex 1 and 2 are essentials for cell lineage differentiation, and also self-renewal capacity in stem cells [132]. A lot of studies have been done to understand the functional roles of each PRC1 and -2, core subunits in their complexes and also in cellular developments. The Ring1A and Ring1B carry E3 ubiquitin ligase activity in the PRC1 complex, and in PRC2 complex EZH1-2 are responsible for trimethylation on histone H3 lysine 27 [60]. In this study, we focused on less-studied canonical subunits of PRC1 complex named Polyhomeotic homolog proteins (PHC) 1, 2 and 3, and try to understand their potential roles in myeloid differentiation.

We used publicly available datasets to analyze the differences between the expression levels for all three PHC subunits during myeloid differentiation. The expression pattern differs between the subunits, with a high expression for PHC1 at the early stages of hematopoiesis, while expression levels for others two subunits are low. The expression levels for these subunits change during myeloid differentiation.

We used the KG-1 cell line as a model to study the role of PHC1-3 in myelopoiesis. It has been described that KG-1 can differentiate to monocyte/macrophage lineage at the presence of PMA [133]. So, the KG-1 cells underwent differentiation in the presence of 200 nM PMA for 48 hours, and differentiation confirmed both with morphology changes as well as increased level of the CD68 pan-macrophage surface marker expression. The only PHC subunits that showed changes at the mRNA level after treatment was PHC1. The PHC1 mRNA level was reduced by half approximately while there were no significant changes for PHC2 and PHC3. However, western blot data showed a reduction at the protein level for PHC1 and increased levels for PHC3 while we were unable to find a suitable antibody for PHC2.

In the next step, we could establish an efficient and very specific knocked down system for each PHC subunit using pre-designed pool siRNAs, which gave more than 90% knock down for both PHC1 and PHC2 and around 60% for PHC3. The knocking down for each subunit was still stable after PMA differentiation for 48 hours. The response to the PMA treatment was not strong enough and we only observed a trend in *PHC2* KD sample. Our analysis for the expression level of CD68 mRNA showed an increased level after *PHC2* KD.

To explore the molecular mechanism for each PHC subunit during PMA differentiation, RNA-sequencing performed on each specific *PHC* subunit KD samples in the

presence of PMA (200 nM) for 48 hours. Our analysis showed different clustering for each set of *PHC* KD, which suggested that each subunit has distinct regulatory effects during differentiation, as well as shared gene targets in their downstream pathways. In our analysis, PHC2 showed the unique set of differentially expressed genes (2071 genes for PHC2 in comparison to 197 genes for PHC1, and 529 genes for PHC3), indicating that the different PHCs regulate specific gene sets. This pattern can be because of different knocked down efficiency in comparison to other samples, especially for PHC3, which needs to be improved.

We performed the volcano plot analyzes for differential gene expression on RNA-sequencing data. The analysis confirmed the specificity of each *PHC* KD and showed the top differentially expressed genes for each sample against the control.

Gene set enrichment analyses on the RNA-sequencing data demonstrated that despite being part of the same complexes, PHC1-3 regulate different pathways in myeloid differentiation, such as changes in the expression pattern of interferon response, myeloid developmental genes, HOXA9 targets, and EZH2 targets. EZH2 is the catalytic subunit of PRC2 [134]. In our analysis, EZH2 target responses showed opposite regulation in *PHC1* KD in comparison with *PHC2* and *PHC3* KDs. Analyzing public data sets for both EZH1 and EZH2 expressions, we noticed these two subunits have almost the same level of expression in the hematopoietic stem cell in the bone marrow, but they switch their expression during differentiation. EZH2 expression level goes down with differentiation while, EZH1 has higher expressions in polymorphonuclear cells both in the bone marrow and the peripheral blood.

To better understand the regulating mechanism underlaid PHC1-3 KD, we studied the potential involvement of the canonical PRC1 complex. Both MEL-18 and BMI-1 are part of canonical subunits with PHC in the PRC1 complex. Taking advantage of publicly available data for CHIP-sequencing in K-562 cell line for both MEL-18 and BMI-1, and compared them with our gene list from each PHC 1-3 KD. We showed a considerable overlap of PHC regulated genes and MEL-18 BMI-18 target genes and pathways, which can indicate the common downstream pathways between these factors.

4.3 Study III

Mammalian genome uses different mechanisms to diversify its transcript pool. Alternative splicing sites or alternative promoters are used in mammalians to produce multiple protein isoforms. It has been suggested that approximately half of the protein-coding genes have alternative promoters [135]. One of the critical steps to understand the regulation of gene transcription and development is to identify where the start site for a specific mRNA is located and how the isoforms are involved in the regulation of different developmental steps.

In this study, we used the data from the FANTOM 5 database for transcription start sites (TSS) to investigate how the usage of alternative TSS can cause exclusion of coding sequence to regulate biological processes. We analyzed data from 890 human primary cells cap analysis of gene expression (CAGE) libraries data from 176 different cell types.

In the beginning, we decided on different controlling parameters to run our analysis. First, we overlapped different tags for each TSS and grouped them into tag clusters (TC) with an extra 500 bp from upstream; then we chose the TCs that have at least 1 or 10 tags per million (TPM) in any of the included cell types. Here we only focused on the TCs that have gene annotations. We had different hierarchical filters to dissect all different TSS subclasses and their cellular specificity. We noticed that known TSS were commonly used in different cell types, but TSS in intragenic regions or antisense strands were more specific to the cell type.

Then we re-run our analysis to find out if the TSS distribution were different across different cell types and identify outliners for each group. We notified that hematopoietic cells are among outliers in two groups with TSS in intragenic regions (10% instead of 6%) or protein-coding gene (20% instead of 39%). To dissect this finding more, we chose 11 primary hematopoietic libraries to characterize the usage percentage for each different TCs group. Our analysis showed that TCs within 5' UTRs and known TSS in coding genes are more frequent in progenitor cells but not in the myeloid lineage. On the other hand, TSS within the coding region is more in favour of myeloid cells than progenitors. Besides, lymphoid cells showed preference pattern to somewhat in between progenitors and myeloid cells.

These differences for TCs within protein-coding regions were interesting for us since this can cause truncated proteins with domain loss. In our analysis, 7.8% of our mapped TCs to known coding genes belonged to this group. Expression for some of

these TCs is highly cell-specific. Gene ontology analysis did not show any functional classification, although immune cells and blood cells showed a specific subgroup.

Then we decided to explore the definition of an alternative promoter, as it is commonly believed the most upstream TSS is the main TSS, but our analysis showed that this is the case for only 33% of the TCs in our libraries. Our analysis showed that the most upstream promoter used ubiquitously, but it is not the specific promoter.

Using the FANTOM 5 data, we could show that alternative TSS to transcribe Vinexin utilized in different cell types. Also, other studies have shown that different protein variants, of vinexin alpha and beta, have cell type specific functions [136-138].

To investigate the functional impact of the protein isoforms that generated from alternative TSS, we set some parameters to find TCs, which leads to domain loss in specific cell types. These analyses showed that 78 protein domains from 36 genes have alternative TSS that cause a protein domain loss. When we did our analysis only on the hematopoietic cells, we could show the domain loss happens in 60 different proteins based on lineage or cell-type-specific. Focusing on epigenetic and transcriptional regulators, we validated the alternative TSS in *KDM2B*, *PRDM1*, and *RERE*, with real-time PCR data, performed in different hematopoietic celltypes.

We continued to investigate the role of different TSS on *KDM2B* isoforms in Jurkat cells since the mice studies have confirmed there are two Kdm2b isoforms, and we speculated it might also be the case in human. We targeted the two most expressed TSSs (TSS1 and TSS3) for *KDM2B* with specific siRNA for knocking down and investigate the functional outcomes for knocking down different isoforms with RNA-sequencing and CHIP-sequencing. Our analysis revealed differences between two isoforms, with the importance of the long isoform in transcription regulation, while almost no transcriptional effects from knocking down the short isoform. The H2AK119ub CHIP-sequencing data showed the same results.

In the end, we followed the changes in domain usage during cell differentiation. In this part, we looked at 16 different time courses, and our analysis showed 76 different genes that have a change in their TSS activity for short and long isoforms, which some of them are even novel.

4.4 Study IV

In this study, our goal was to understand how chromatin organization in 3D in different sub-nuclear compartments impacts transcriptional regulation. To answer this question, we performed circular chromatin conformation capture (4C) coupled to sequencing [139] to the well-studied epigenetically regulated H19 imprinted control region (*ICR*) [140]. We used this region as a bait in human embryonic stem cells (hESCs) and derived human embryoid bodies (hEBs). We identified 518 different intra- and inter-chromosomal chromatin fiber interactions. These interactions within the network were later confirmed by 3D DNA-FISH analysis, showing that interactors with high reads counts in the 4C-seq were significantly closer to the main bait (i.e., *H19 ICR* and *VAT1L*) compared with regions with the lower number of reads.

Given the unique capacities of our 4C assay to capture more than two simultaneously interacting sequences, we could reconstruct the interactions between the interactors of the *H19 ICR* to define a network of interactions within our bait. We then reasoned that regulated encounters within the network might be facilitated by dynamic molecular ties. As we previously showed that PARylation of CTCF was essential for the long-range chromatin insulation in cis [141], we decided to test if it was also necessary to form chromatin network interactions in trans. Our data showed that the removal of PAR by PAR glycohydrolase (PARG) activity led to the disassembly of the majority of the chromatin networks. In parallel, and since CTCF can activate PARP-1 [142], we observed a reduction of PAR levels upon CTCF knock-down, suggesting that PARylation in chromatin complexes might be the result of CTCF and PARP-1 interactions.

We hypothesized that genomic loci occupied by PARP-1 formed dynamic complexes with other chromatin regions that carried factors binding to PAR with high affinity such as CTCF. To prove that we treated hESCs with the PARP-1 inhibitor, Olaparib, which not only inhibits PARP-1 activity but also disrupts the interaction between PARP1 and CTCF, for 24 hours. Our results showed that Olaparib treatment led to a significant reduction in the proximity between chromatin network hubs under these conditions. Our results also showed that the interaction between CTCF and PARP1 is crucial for the connection between *H19 ICR* and its interacting chromatin network.

Because it was previously shown that PARP-1 activity oscillates by feeding [143]. Our analysis in the 4C library showed interactions between LADs and circadian controlled genes. We decided to investigate whether our network represented

some fine-tune mechanism of regulation of circadian genes. We could not establish circadian synchronization in hESCs as it was reported before [144], nor in hEBs because of their production procedures. Instead, we used the human colon cancer cell line HCT116 cells as an appropriate model for circadian synchronization upon serum shock [145]. Data from *in situ* proximity ligation assay (isPLA) revealed that CTCF and PARP-1 proximity oscillates in a circadian manner, with peaks of interactions at 8 and 32 hours after serum shock, and mainly occurring at the nuclear periphery. Additionally, 3D DNA-FISH analysis for our 4C bait (*IGF/H19 ICR*) and circadian network nodes (*VAT1L*, *TARDB,P* and *PARD3*) showed a rhythmic pattern of recruitment of these loci to the repressive environment of the nuclear periphery which correlates with oscillating transcriptional attenuation (RNA-FISH analysis). A more detailed examination of the *PARD3* locus with a more detailed kinetics between 8 and 16 hours after serum shock revealed that following the arrival to the nuclear periphery at 10 hours, the transcription attenuation of this locus occurred later, between 10 and 12 hours, together with the acquisition of the repressive chromatin mark H3K9me2, investigated by assessing the proximity between this hPTM and the PARD3 locus by chromatin *in situ* proximity analysis (ChrISP) [146]. The importance of H3K9me2 acquisition in circadian transcription was further proved by showing that upon its depletion by inhibiting the methyl transferase G9a/Glp, *PARD3* recruitment to the nuclear periphery and its transcriptional circadian oscillation were abolished. Thus, we concluded that circadian recruitment of active alleles to the nuclear periphery preceded the acquisition of the H3K9me2 repressive mark and circadian transcriptional attenuation.

5 DISCUSSION

Although the genome of different cells in an organism is the same, different regulatory mechanisms are involved to allow different transcriptomes from the same genetic material at the different developmental stages and differentiation levels. The epigenetic machinery and its impact on the genome and transcription levels are the most critical parts of the regulatory mechanisms.

To study different parts of the epigenetic machinery, we took advantage of the hematopoietic system in studies I-III. One of the main reasons for that is the frequently mutated status for epigenetic regulators in hematological malignancies [147].

5.1 Study I

In this study, we could show that our CRISPR-Cas9 library was an efficient method for study epigenetic regulators in the hematopoietic system. In our analysis on CRISPR-Cas9 library for 1092 epigenetic factors, we were not only able to identify new regulators but also reconfirm the regulatory role for some of the previously published epigenetic factors, such as *ARID4B*, *KMT2A/MLL*, and *ASH1L*. In total, we could show a potential regulatory role for approximately 5% of our library (57 epigenetic factors). With our analysis, we also revealed a potential new regulatory role for some of the epigenetic factors in human hematopoiesis differentiation pathways.

In our analysis, chromodomain helicase DNA-binding 2 (*CHD2*) was among our top genes on the list that we selected to validate with further analysis to show its role in megakaryocytic differentiation. Recently, other members of the chromodomain helicase DNA-binding family has shown to be important in the pluripotency of stem cells and also the progression of childhood AML [148,149]. It has already been shown that Chd2 can regulate muscle differentiation [150]. The genome-wide analysis showed that MyoD transcription factor is dependent on Chd2 for H3.3 deposition into specific gene promoters that induce muscle differentiation [150]. In our previous study, we noticed that the involvement of CHD2 in H3.3 deposition is not specific to the muscle cells, but also in myeloid cells [131]. Additionally, we showed in the K-562 cell model that both CHD1 and CHD2 participate in regulating the chromatin architecture in transcribed active loci. RNA-polymerase II is involved in their recruitment [131].

In this study, we demonstrated that *CHD2* knocked out in the K-562 cell line promotes megakaryocytic differentiation and inhibits cell proliferation. This data suggested that CHD2 is important to inhibit megakaryocytic differentiation. Our data for *CHD2* KO in K-562 is in concordant with data from Chd2 KO in mice. It has shown that Chd2 is an essential gene for development and survival. The Chd2 KO mice have a deficiency in hematopoietic stem cell differentiation, which leads to lymphomas [151].

Our analysis of the ENCODE data revealed that CHD2 could affect transcription via the Polymerase II machinery. RNA-sequencing data on our *CHD2* KO cells confirmed the role of CHD2 in transcription specifically for CHD2-target genes.

Our data clarify that CHD2 is important for proliferation, and its loss of function can promote megakaryocytic differentiation in K-562 cells.

5.2 Study II

Different studies have shown that PRC complexes are crucial for gene silencing and regulation of hematopoietic differentiation, and try to dissect the roles for different PRC complexes subunits in hematopoiesis [36]. This study aimed to investigate the role of the different Polyhomeotic homolog proteins (PHC) subunits in the PRC1 complex during hematopoietic differentiation. Our analysis of publicly available data suggested reverse regulatory roles for different PHC1-3. This pattern confirmed with our PHC knock down and RNA-sequencing experiments. Our data showed several potential genes and pathways, which are important for PRC1 complex functions in general and, more specifically, for PHC subunits. We believe that further experiments are needed to improve our data and clarify the potential regulatory roles for different PHC subunits. One way to do it, is to use a more efficient knock down system, and obtain stable knocked down cells. This issue is especially for *PHC3* that had only around 30 folds down regulation in comparison to 50 and 150 folds for *PHC1* and *PHC2*, respectively. We tried to use the CRISPR-Cpf1 system to knock out each PHC subunit in KG-1 cells, but the cells did not survive after transfection. We will investigate the PHC roles with shRNA (short-hairpin RNA) for each specific PHC subunit in our KG-1 model. The construction for specific insertion into the lentiviral vector is ongoing.

5.3 Study III

The technical development in RNA-sequencing such as cap analysis of gene expression (CAGE) allowed to profile gene expression with a focus on 5' end, and allowed a better understanding for transcription start sites (TSS) [152]. Intensive CAGE data analysis revealed lots of new TSSs and alternative promoters [153]. In this study, we demonstrated that different human cell types are using a similar fraction of TSSs except in immune and blood cells that have different TSS usage. As mentioned before, in the progenitors: 5' UTR and known TSS in coding genes, in the myeloid cells: TSS within coding genes are frequent, and the lymphoid pattern is in between the progenitors and myeloid cells.

In our analysis, we have noticed two occurring patterns for alternative TSS. As the case for NFXL1, the alternative TSSs can help to have different isoforms with different repeats of individual domains, which can help with the protein function and its efficiency. Another pattern can be like what we saw for MY010; the alternative TSSs can cause different isoforms with domain loss and have a regulatory effect.

KDM2B was among the genes with domain loss in hematopoietic cells. Published studies in mice also support this finding. However, our functional studies were unable to demonstrate specific roles for different KDM2B TSSs and its potential isoforms.

This study revealed a novel phenomenon of alternative TSS usage of genes according to hematopoietic differentiation. Although, as with all RNA-based studies, the mRNA isoforms need to prove also to become translated to protein isoforms. We were unable to confirm the existence of protein isoforms for KDM2B due to the technical problems with antibodies and further experiments for confirmation are needed. However, our results still inspire the further exploration of epigenetic regulation mediated by KDM2B isoforms during differentiation.

5.4 Study IV

Chromatin has active and inactive domains. Their transcriptional status is prone to segregate into distinct sub-nuclear compartments so they can maintain their stable expression patterns. In paper IV, we demonstrated that an inter-chromosomal network emerging from hESCs and derived hEBs connects active loci enriched in circadian genes to transcriptionally repressed lamina-associated domains (LADs). PARP1 and its co-factor CTCF regulate this interactome. The association between these two proteins does not only mediate chromatin fiber interactions but also

assist the recruitment of circadian genes to the nuclear lamina. Serum shock in HCT116 colon cancer cells induced synchronization. This synchronization promotes oscillations in PARP1-CTCF interactions and followed by oscillating recruitment of circadian loci to the lamina. The repressive H3K9me2 marks acquisition and transcriptional attenuation followed. Additionally, H3K9me2/3 depletion by inhibition of PARP activity with Olaparib, inhibiting G9a/Glp, or down regulation of PARP1 or CTCF expression not only counteracts with recruitment to the nuclear envelope but also circadian transcription. PARP1- and CTCF can regulate interactions between circadian loci and the repressive chromatin environment at the nuclear lamina, as a result, moderate circadian transcriptional plasticity.

6 ACKNOWLEDGEMENTS

Life is not easy as it is full of unknowns and when it mixes with moving to a new country and starting a Ph.D., it is going to be challenging. This period of pursuing my Ph.D. has been full of ups and downs with lots of lessons that I never could find in any textbook or article. These are the moments when life teaches us new lessons- for good or bad! Those that know me are familiar with this long Ph.D. chapter of my life. The reason I am able to finish it is because of those who contributed to my studies or gave me support when I needed it. To those of you who were simply there for me to laugh or cry throughout this long journey, I would like to extend my gratitude here. If your name isn't here it doesn't mean I am not grateful for you being in my life.

I would like to start by expressing my special appreciation to my main supervisor **Andreas Lennartsson**. I can't tell you how much you have helped me during my time in your lab. You were so patient and kind. I regained my confidence in scientific life and for that, I am always grateful to you. I am not sure I could ask for a better supervisor to help me to finish my Ph.D. journey. Thanks again for always being so understanding and supportive. Your faith in me gave me the confidence that I could do any task. You are balanced in giving me constructive criticism and much-needed advice while also giving me the freedom to be an independent scientist.

My co-supervisors who helped me during my study: **Peter Svensson**, thanks for being so critical and challenging during all our meetings. I have learned a lot from you. **Sören Lehmann**, thanks for all the input towards my projects during our many meetings and retreats. **Anna Palau De Miguel**, thanks for being so supportive and always available during the time that you become my co-supervisor in addition to being my colleague. Thanks for all your scientific help and for listening to me when I needed you.

Julian Walfridsson, I can't express how lucky I was to have you as my mentor and I will always be thankful for your help during the time that I really needed it the most. I appreciate that you always found time to listen to me and give me advice and encouragement.

Marianne Schultzberg, I can't express enough my gratitude toward you. It was a privilege to meet you. Without your support and help, I wouldn't have been able to get the chance to finish my Ph.D. study.

Rolf Ohlsson, I wouldn't be able to finish my Ph.D. in KI if it wasn't for the first chance to start it in your lab in the beginning. Thanks for the first invitation to work in KI. I have learned a lot during my time in your lab. **Anita Göndör**, being able to work with you has taught me a lot of lessons that I am going to have with me no matter where I go.

I would like to say thank to all collaborators for their support that made it possible for me to conduct my research as fast as possible:

To the supportive people at **High Throughput Genome Engineering Facility (HTGE)** and especially **Jenna Persson** and **Bernhard Schmierer**. To the kind assistance of **Aleksandra Krstic** during all sorting and FACS analysis. I would like to say a big thanks to **Fredrik Fagerström-Billai**, **Ashwini Gajulapuri**, **Tassos Damdimopoulos**, **Patrick Muller**, **Marika Rönnholm** and **Susann Fält** at **Bioinformatics and Expression Analysis core facility (BEA)** for their work at different projects especially during my paper revision. **Mickael Durand-Dubief**, thanks for all the analyses and your help with the paper.

Victoria Menendez Benito, thanks for accepting the role of chairperson of my dissertation. In addition, thanks for our many conversations during lunchtime and Christmas dinners.

Lars-Arne Haldosen, thanks for being available to give me advice and for your kind support. **Velmurugesan Arulampalam**, thanks for all your support during my time at the MTC department. **Andreas Nyström**, thanks for all the support during the times I needed it most. Without your help, I wouldn't have been able to fight for my rights. **Sara Elg**, our lives crossed paths because of some unfortunate circumstances, but I am so happy I got the chance to get to know you. It was a loss for us, Ph.D. students, in KI not having you as Ombuds, but I am happy to have you in my life as a friend.

I should say a big thanks to all the great people at the administrative in both Bionut and MTC departments. A special thanks to **Monica Alhberg** for her big smile and always being ready to help. To **Linda Strand** and **Maryam Sagafian**, thank you all for your help processing and tracking down our missing orders.

I would like to say thanks to my former and current colleagues in Lennartsson's group. Thank you all for all the time we shared, in the lab or offices, lunchtime, after work, and all the different activities we had over the last four years: **Sophia**, thanks for all your help, especially with running western blot. **Wenbo**, I hope it is

soon your time to defend your thesis, good luck with all the paperwork :) **Sylvain**, thanks for the constant advice to improve my laptop security! **Anna**, thanks for all the advice and help on the lab bench. **Elizabeth**, good luck with your future and hope you can start your Ph.D. soon. **Wlaa**, it was great to share an office with you and I hope you can defend your thesis soon. **Katarzyna**, it was fun to work with you for a short time and good luck with the rest of your Ph.D.! **Xingfu**, welcome to the lab. **Naochika**, good luck with your project. **Birgitta**, Tack så mycket for allt din hjalp, for being so cheerful and kind, for being so patient and practicing Swedish with me and also your help with ordering when it was urgent to do it.

I want to extend my thanks to all former and current members of Sören's group for their input on my projects during our meetings and retreats: **My**, for the best memories from Estoril. **Anne**, good luck with the rest of your Ph.D. :) **Anna**, **Albin**, and **Christer**, thanks for all the help to better understand all the clinical parts of the AML world. **Naomi**, welcome to the group. **Sofia**, and **Anna** thanks for all your help. A special thanks to **Huthayfa**, for all the advice during these a few last months while we were struggling to finish all the paperwork for defense. Good luck and congratulations in advance :)

Big thanks to my former colleagues at **MTC** who were there for me when I really needed you and always made me feel included even when I left. Thanks for tolerating me with my camera and accepting my passion for taking new pictures:

First of all, **Piri**, I want to give you a big hug for your kindness, you were always ready to help us with our ordering and different assignments and problems. I was so lucky to be able to learn how to work with human stem cells from you. **Anna**, I am always amazed by your strength and energy and will never forget those long nights and early mornings that we worked together in the lab. It was fun to work side by side with you and I hope you are the next to get the title! **Samer**, it was a H… but then I had the opportunity to know you and also still be friends with you so I can't say I regret it! Thanks for fixing all of my computer problems and teaching me how to use the microscope and not to see everything green! It was always fun to pay you by CHOCOLATEs and then help you eat them! We kept the currency in the lab even after you graduated and I even tried to upgrade it ;D **Xingqi**, I am amazed by your hard work and motivation. It was a great opportunity to work with you and congratulations on your new position. **Chengxi**, my best memories of you aren't from the lab, but they are from your wedding day in China! No wonder you were calling me a party girl! Hope to see you and **Yan** soon ;) **Moumita**, it was a great opportunity to know you and work with you. You were so kind and

helpful, especially when I joined the lab. **Noriyuki**, I learned a lot from you. It was a great time to work and discuss both science and non-scientific topics with you. Your energy for science is inspiring, and I hope to see you soon in Japan. **Alejandro**, thanks for all of your help, especially when you were trying to make our life easier with all your bioinformatics magic! I still have lots of Persian words to teach you! :) **Maria**, it was a great opportunity to share the office with you and I was always amazed by your depth of knowledge on different topics. **Jesus**, I still can't believe how we were lost in Gamla stan and couldn't find Noble museum for more than three hours! I blame it on being our first week in Stockholm. I hope to take you there next time you are here ;) **Sara**, thanks for your support when I just joined the lab and didn't let me feel homesick! **Li-Sophie**, thanks for all the time that you spent teaching me how to perform Thai Chi. I'm not sure that I can say I still remember them ;), but it was fun to work with you. **Honglei**, or should I now call you **Dr. LiLi**, Christmas always reminds me of you and our memories in the lab. After all, it was my first Christmas and we had the opportunity to be alone in the whole department! **Gözde**, I have changed this part so many times depending on your visa situation and different factors! But you know who you are and you have already said enough in your book, so I am not going to repeat them. You can check it in your thesis. **Manos**, we still need to discuss a bit more about history :) **Lluis**, it was a great time to work with you and thanks for all your help when we were working in the same lab. **Carolina**, although the first time I saw you were in the darkness of the microscope room, you were with me during my times of happiness and also sadness. Thank you for your true friendship and even more. **Barbara**, I am so happy to have the chance to work with you and I am even more happy to have you as a friend now. **Richelle**, it was fun to teach you lab works and thanks for helping me to learn more. **Aga**, it was fun to work with you and trying to solve all the troubles with the microscope or different protocols. Hope to see you and **Bartek** soon ;) **Vanessa**, thanks for speaking not very British and in a way that we could understand and let us think our English was improving! Good luck with your studies. **Alisa**, it was fun to be your co-supervisor for your master thesis and I learned a lot during that time. I am happy for you and your Ph.D. journey! **Mirco**, we didn't spend so much time in the lab but I am happy to know you and be your friend, despite disagreeing with each other 95% of the time :) **Ilias**, it was fun to teach you how to work with the microscope when you had just joined the lab. **Johana**, we never had the chance to work together but thanks all the unforgettable memories especially, from Gotland.

Andreas, **Huma**, and **Nancy** I had so much fun in Novum when we were sharing an office. Thanks for all the different topics and discussions during coffee time. **Fatemeh**, thanks for all the talks and encouragement while we were almost sharing the office with you but without seeing each other thanks to those thin walls ;)

My friends at **BioNut**: **Marjan**, when should we go to the gym? Although this question is more about me!!! You are my motivator! **Yasaman**, I hope the cake for my party for once will be the one that I asked for! I am going to miss you girls. Thanks for all the memories during dinners, coffee breaks, and birthday parties. You made my life full of fun. Thanks for being who you are so I can talk to you when I needed it. Thanks in advance for all the help with my dissertation and the party. **Marjan**, thanks for accepting to be the toastmaster, just don't forget to make the music playlist :) **Alex**, thanks for all the discussions during lunchtime and practicing Swedish on Wednesdays' lunchtime. **Lisa** and **Jenna,** thanks for all the shared memories during your time in the BioNut. **Parisa**, we always need your help to cut the cakes in the end to a good shape; after all, there are so many edges and we need to get rid of them! You are a lifesaver! **Mohammad**, thanks for all the discussion and your help with eating chocolates and cookies in the office ;) **Laia**, thanks for never forgetting us when you went home and always bringing the best gifts that we could ever ask for --the familiar taste of Iranian sweets ;) **Marzieh**, good luck with the new job! **Olga**, it was so funny how you learned one Persian phrase (Mage chie?!) and almost could use it in the best moments :) **Lina**, thanks for all your kind support, especially when I need something urgent in the lab. **Vlad**, thanks for being so patient when the girls and I were talking in secret and you somehow always found out what the plan was about ;D **Serena**, thanks for all the chats and encouragement. You are the best partner for carving a pumpkin for Halloween :) **Ani**, thanks for all the discussions and encouragement. **Jonathan**, thanks for all the discussions in the cell culture room ;) **Leticia**, you should come back and help me solve my biggest dilemma from our Swedish class, after all which one is correct: Aja Baya or Aja Paya! **Tales**, thanks for all the fun in the Base. **Agustin**, thanks for all encouragement and reminders to write my thesis ;) **Su**, good luck with the rest of your study.

My friends at **HERM: Shabnam**, you are no longer with us in Stockholm, but this doesn't change anything. I am always looking forward to seeing you and thankful for your friendship. **Esmat**, it was so great to meet you in Stockholm. I am always missing you and your big heart, but then I am happy to have you in my life. **Jelveh**, you are such a kind person that talking for a short time in a corridor always helps lift my spirit. Thanks for all your help. **Nadir**, thanks for all

your advice on not making mistakes in science and my life that I could not follow them most of the time ;) and also for not putting my name on the blocked list on your phone! **Teresa**, thanks for your help and for teaching me CFU protocol. **Thuy**, thanks for all dinners and game nights. **Monika**, good luck with the new job. **Lakshmi**, thanks for all the encouragement and chats especially close to the ice machine :) **Saeid**, thanks for your kind scientific advice. **Mahin**, thanks for all your kindness and help.

My CHaSE friends: **Zack**, **Azadeh**, **Kai**, **Natali**, **Kadir**, **Yan**, **Maryam**, **Josefin**, and **Vanja**, I am so happy to get to know all of you and learned a lot during all those meetings and discussions. I think we did great work at CHaSE 2019.

My old friends at MTC, especially **Habib**, **Mariam**, **Afrouz**, **Sunitha,** thanks for all the **MSA** board meetings. It was fun to skip the lab at that time even for a short time. It is even more joyful to see you now when it happens. **Tomek**, when should we start working on our space shuttle :D

My friends in **KI** and especially those who I know from my time in KI housing in Varberg, it is always fun to see you either in a party or international dinner or just to meet you in different corridors and buildings in KI: **Karen**, **Yogan**, **Tomek**, **Fadwa**, **Rita**, **Jorge**, **Laetitia**, **Lorena**, **Mellina**, **Naida**, **Mike**, **Shane**, **Avinash**, **Sunjay**, **Caitrin**, **Yildiz**.

It was impossible for me to work over the last eight years without having new friends in my life. Most of them are like a family to me and I love you all. **Sonia** and **Henric**, I am so lucky to have you in my life. Going to different Swedish classes didn't help me learn it, but it certainly helped me find one of my best friends who is more like a sister to me. **Henric**, now that I have a Ph.D. in Biology you should trust me more with my knowledge about animals. It is a panda's picture on the wall--not a koala! Let's not forget the little **Lionel-Nivan** and how much joy he brought to our life with so many sleepless nights! **Misia** and **Samer**, thanks for your true friendship during all these years. It is a journey to the middle of nowhere to come and visit you, but once there, it is always fun thanks to **Ada** and **Vincent** :) **Ania** and **Uffe**, I am always amazed every time I come to your place for a visit. There are always so many changes and progress with the house renovations, how you are managing? Let's not forget how much I enjoyed playing with **Daniel** and **David** and how much I tried to improve my Swedish with them. Not that they ever approved :) **Moumita** and **Punit**, thanks for all those spicy foods and I hope to see you soon and hopefully it will be in India or somewhere else. The world is so

small after all! **Fatima** and **Jake**, it is sad that you live in Bergen, but it makes it really fun for me to travel there and spend time with you especially when **Jake** is cooking ;) Thanks for all the video calls and your support during all these years. As always, I look forward to seeing you soon ;) **Caro** and **Frans**, I am so lucky to have you in my life and it is always great to be able to join your family gatherings and spend time with you all. Let's not forget the little **Olivia** and fun time with her :) **Andres**, thanks for being a good friend and also a good listener that I can tell everything and nag about anything. You know how much I love you! **Anu**, thanks for being a good friend and always late, so giving me a reason to nag about something to **Andres**: why is **Anu** never on time!!! :)

Iranian friends in Sweden: **Hanna** and **Amir**, thanks a lot for not only being my friends but also my family here in Sweden. Big thanks to **Hanna** for all those special Tahdig and foods that never let me feel homesick. Thanks for teaching me the basic lessons for Persian cooking and the difference between different spices and how to use them! Amir, thanks for not pretending to eat whatever I cooked and always asking me not to cook, although I never listened, but only tried to cook better! A big thanks to **Abtin** for teaching me dinosaurs' names, but I always needed a reminder. After all, I can't know and remember everything in biology! **Suzan**, I am so lucky to have you in my life and for letting me be part of your family. For all the Yalda's night, shopping for Norouz, movie nights, phone calls, and your endless kindness and smile. You know how much I love you. **Sahar**, thanks for your friendship and your big smile. **Mehrdad**, thanks for all those times that we either watched movies or discussed them. We can't watch them together now that you live in Canada, but I still enjoy catching up in our next discussion. I just wish you had more skills to play a game just like me :) **Sareh**, thanks for your friendship and all the encouragement. You are the next to defend! **Sadna** and **Reza**, thanks for all the fun memories, especially during all the 13bedars and BBQ ;) **Khalil** and **Santa**, thanks for your friendships and all the fun times with your BBQ skills. **Atefeh**, **Daniel**, and **Hamid**, thanks for all the fun and game nights. **Hadis** and **Mehdi**, thanks for all your kindness and friendship. **Marzieh**, and **Shahla** thanks for all the fun memories and dinner times. **Manizhe** and **Yaser**, thanks for your true friendship. Meeting you again in Stockholm was a big gift for me especially when we were living in the same building and you were helping me with all of my cooking and baking ;) **Yasi Golab** and **Behrooz**, your friendship are an inspiration to me with all of your traveling and adventures tips. I never forgot our memories from Gran Canaria :) **Maryam**, it was fun to be in Swedish class with you. At least I could practice Persian with you! Thanks for being so understanding during these last few months with all the applications and forms. Soon it will be your

time, so good luck! **Amir**, thanks for all the time that we have spent on the road, either on a bike or a car! **Anahita** and **Peyman** thanks for your friendship and enjoyable times, especially with your cute boy. **Azi** and **Maryam**, it was my luck to meet you, thanks for all the laughs and fun time. **Azi**, keep doing what you do and soon it is going to be your time to graduate.

I can't finish this part if I don't mention some of my friends' names that although we couldn't see each other as much as we would have liked, you all will always be a part of my life no matter where I live. You all helped me have a better life and gave me energy and motivation with your endless video calls, emails, messages and meeting you whenever we could be in the same place. **Mahshid**, thanks for all the encouragement during all the time that I've known you and for the best memories from Italy. I hope to see you soon. **Fereshteh**, you have taught me a big lesson: to follow my happiness and leave a place when it doesn't make me happy. Also, thanks for showing me to take risks and also letting me risk my life with your driving ;) **Saeide**, **Faezeh**, **Nina**, **Simin**, **Shahla**, **Zahra, Maryam**, **Elahe**, **Roya**, **Sohila**, **Fahimeh**, and **Zahra** thanks for your true friendship during all these years. For checking on me with your messages and emails, especially thanks for when I come back to visit and treating me as a tourist ;) **Mahboobeh**, and **Hoda** thanks for all those memories and thanks for still keeping in touch with me. **Mahboubeh**, thanks for all of your support in teaching me the basics of molecular biology when I was a Master's student, for helping me to find my Ph.D. position and for all your encouragement during these years.

Maryam, I don't have a word to describe how much I missed you during these years, but thanks for always being there for me when I need someone to listen to me nag or complain. After all, you know me very well. **Mohammad** and **Davood**, thanks for all the fun times and especially dancing time when I am back home and not on a road! **Behnaz**, you are my oldest friend and it is too bad that we can't see each other as often, but thanks for always sending me pictures so old that even I have trouble identifying myself. You are always there for me with some nostalgia brought on by old memories. Love you and both your kids. **Naiereh**, you are as dear to me as **Behnaz**, thanks for accepting me into your family. **Ghazal** and **Peter**, thanks for your hospitality during my visit to Montreal and I hope to see you soon :) **Mohadeseh**, I will never forget how we met and how I didn't like you in the beginning--or it was you who didn't like me! Either way, you are one of my dearest friends now. We missed so many events over these years due to the distance, but nothing can change our friendship --even your constant problems with your phone or your internet! I want to see **Haana**. Thanks to her we almost

have the same birthdate and now you can't say you forgot my birthday anymore! **Najmeh**, what can I say about you? Sometimes you are the only way that I can get some news from **Mohadeseh** when she has one of her problems with her phone, so thanks for being a good messenger ;) **Neda**, thanks for not only being a great teacher to me but also a good friend. I will never forget that you gave me my very first book in genetics. All my troubles started from that day! Just kidding ;D

My dearest **Setareh** and **Mohsen**. I have learned a lot by being your student. You were great mentors when I was an undergraduate student and those times in the Genetic Consulting Center were some of the best times in my life. I have learned so many life lessons in that center. Thanks for all your support and advice during all these years. I was so lucky to meet you in Toronto and I hope to see you soon. Hopefully, this time we will reunite in Stockholm :)

Last, but not least, I would like to thank my family. No matter how far I go they are always there for me. **Sohila**, from the time that you joined our family you brought lots of joy and happiness with you. After all, you are the sister that I never wanted, but now are forced to have and I have to love you more than my brothers since it is the sister-code! Let's not forget the newest family member, my little nephew, **Kourosh**. Unfortunately, I haven't been able to witness firsthand how much you are growing, but I am happy you have **Sohila** as your mom so she can teach you how to live your life while the rest of the family, including me, tries to spoil you ;) **Bahram**, you are more than a big brother for me and I cannot express how much I love you! **Bahman**, what can I do, you are the last kid so I have to love you maybe a little more than **Bahram**. I love you both from the bottom of my heart. I thank both of you for making my childhood full of joy, playfulness, and fights until death! Thanks for all of your support and for being my wings until I learned to fly. Thanks for being such good brothers that I never wanted to have a sister. I can't tell you how lucky I am to have you both in my life even if you are far away from me :) **Shekoofeh** and **Khalil**, thanks for being my parents and teaching me how to live without any fears or barriers. Thanks for forcing me to have big dreams and letting me fight for them! You are the reason that I am always going far to try to be a better version of myself.

7 REFERENCES

1. Waddington CH. The epigenotype. 1942. Int J Epidemiol. 2012 Feb;41(1):10-3.
2. Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. Cell. 2007 Feb 23;128(4):635-8.
3. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003 Mar;33 Suppl:245-54.
4. Anders H. Lund MvL. Epigenetics and cancer. Genes & development. 2004 Oct 1;18(19):2315-2335.
5. Karolin Luger AWMd, Robin K. Richmond, David F. Sargent & Timothy J. Richmond. Crystal structure of the nucleosome core particle at 2.8 A ˚ resolution. Nature. 1997 Sep 18;389:251-60.
6. Roger D. Kornberg JTa. Chromatin structure; oligomers of the histones. Science (New York, NY). 1974 May 24;184(4139):865-8.
7. Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. Nat Rev Cancer. 2013 Jul;13(7):497-510.
8. Felsenfeld G GM. Controlling the double helix. Nature. 2003 Jan 23;421(6921):448-53.
9. Bhalla KN. Epigenetic and chromatin modifiers as targeted therapy of hematologic malignancies. J Clin Oncol. 2005 Jun 10;23(17):3971-93.
10. Bird A. DNA methylation patterns and epigenetic memory. Genes & development. 2002 Jan 1;16:6-21.
11. Maria Tsompana MJB. Chromatin accessibility- a window into the genome. Epigenetics & Chromatin. 2014 Nov 20;7(33).
12. Roberts CW, Orkin SH. The SWI/SNF complex--chromatin and cancer. Nat Rev Cancer. 2004 Feb;4(2):133-42.
13. Bartel DP. MicroRNAs: Genomics,Biogenesis, Mechanism, and Function. Cell. 2004 January 23;116:281–29.
14. Inagaki T, Iwasaki S, Matsumura Y, et al. The FBXL10/KDM2B scaffolding protein associates with novel polycomb repressive complex-1 to regulate adipogenesis. J Biol Chem. 2015 Feb 13;290(7):4163-77.
15. Holliday R PJ. DNA Modification Mechanisms and Gene Activity during Development. Science (New York, NY). 1975 Jan 24;187(4173):226-32.

16. Riggs AD. X inactivation, differentiation, and DNA methylation. Cytogenetics and cell genetics. 1975;14(1):9-25.

17. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009 Nov 19;462(7271):315-22.

18. Blackledge NP, Klose R. CpG island chromatin: a platform for gene regulation. Epigenetics. 2011 Feb;6(2):147-52.

19. Li E, Zhang Y. DNA methylation in mammals. Cold Spring Harb Perspect Biol. 2014 May 1;6(5):a019133.

20. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet. 2009 Feb;41(2):178-186.

21. Berdasco M, Esteller M. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. Dev Cell. 2010 Nov 16;19(5):698-711.

22. Han JA, An J, Ko M. Functions of TET Proteins in Hematopoietic Transformation. Mol Cells. 2015 Nov;38(11):925-35.

23. Shinsuke Ito LS, Qing Dai, Susan C. Wu, Leonard B. Collins, James A. Swenberg, Chuan He, Yi Zhang. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science (New York, NY). 2011 2 SEPTEMBER;333(6047):1300-3.

24. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol. 2013 Jun;14(6):341-56.

25. Bhaumik SR, Smith E, Shilatifard A. Covalent modifications of histones during development and disease pathogenesis. Nat Struct Mol Biol. 2007 Nov;14(11):1008-16.

26. Kouzarides T. Chromatin modifications and their function. Cell. 2007 Feb 23;128(4):693-705.

27. Vakoc CR, Mandat SA, Olenchock BA, et al. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell. 2005 Aug 05;19(3):381-91.

28. Berger SL. Histone modifications in transcriptional regulation. Curr Opin Genet Dev. 2002 Apr;12(2):142-8.

29. Sun XJ, Man N, Tan Y, et al. The Role of Histone Acetyltransferases in Normal and Malignant Hematopoiesis. Front Oncol. 2015;5:108.

30. Sharon Y. Roth JMD, and C. David Allis. Histone acetyltransferases. Annual Review of Biochemistry. 2001;70:81-120.

31. Lachner M, O'Sullivan RJ, Jenuwein T. An epigenetic road map for histone lysine methylation. J Cell Sci. 2003 Jun 01;116(Pt 11):2117-24.

32. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000 Jan 06;403(6765):41-5.

33. Turner BM. Histone acetylation and an epigenetic code. BioEssays : news and reviews in molecular, cellular and developmental biology. 2000 Sep;22(9):836-45.

34. Lee DY, Teyssier C, Strahl BD, et al. Role of protein methylation in regulation of transcription. Endocr Rev. 2005 Apr;26(2):147-70.

35. Y. S. Histone lysine demethylases: emerging roles in development, physiology and disease. Nat Rev Genet. 2007;8(11):829-33.

36. Di Carlo V, Mocavini I, Di Croce L. Polycomb complexes in normal and malignant hematopoiesis. J Cell Biol. 2019 Jan 7;218(1):55-69.

37. Radulovic V, de Haan G, Klauke K. Polycomb-group proteins in hematopoietic stem cell regulation and hematopoietic neoplasms. Leukemia. 2013 Mar;27(3):523-33.

38. Morey L, Helin K. Polycomb group protein-mediated repression of transcription. Trends Biochem Sci. 2010 Jun;35(6):323-32.

39. Xu C, Bian C, Yang W, et al. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc Natl Acad Sci U S A. 2010 Nov9;107(45):19266-71.

40. Steffen PA, Ringrose L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. Nat Rev Mol Cell Biol. 2014 May;15(5):340-56.

41. Vidal M, Starowicz K. Polycomb complexes PRC1 and their function in hematopoiesis. Exp Hematol. 2017 Apr;48:12-31.

42. Cao R, Tsukada Y, Zhang Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell. 2005 Dec 22;20(6):845-54.

43. Blackledge NP, Rose NR, Klose RJ. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. Nat Rev Mol Cell Biol. 2015 Nov;16(11):643-649.

44. Ru Cao LW, Hengbin Wang, Li Xia, Hediye Erdjument-Bromage, Paul Tempst, Richard S. Jones, Yi Zhang. Role of histone H3 lysine 27 meth-

ylation in Polycomb-group silencing. Science (New York, NY). 2002 Nov 1;298(5595):1039-1043.

45. Brockdorff N. Noncoding RNA and Polycomb recruitment. RNA (New York, NY). 2013 Apr;19(4):429-42.

46. Tavares L, Dimitrova E, Oxley D, et al. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell. 2012 Feb 17;148(4):664-78.

47. Wu X, Johansen Jens V, Helin K. Fbxl10/Kdm2b Recruits Polycomb Repressive Complex 1 to CpG Islands and Regulates H2A Ubiquitylation. Molecular Cell. 2013;49(6):1134-1146.

48. Farcas AM, Blackledge NP, Sudbery I, et al. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. Elife. 2012 Dec 18;1:e00205.

49. Simon JA, Kingston RE. Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol. 2009 Oct;10(10):697-708.

50. Payer B, Lee JT. X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet. 2008;42:733-72.

51. Wu HA, Bernstein E. Partners in imprinting: noncoding RNA and polycomb group proteins. Dev Cell. 2008 Nov;15(5):637-8.

52. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer. 2006 Nov;6(11):846-56.

53. Pasini D, Malatesta M, Jung HR, et al. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic Acids Res. 2010 Aug;38(15):4958-69.

54. Ho JWK, Jung YL, Liu T, et al. Comparative analysis of metazoan chromatin organization. Nature. 2014;512(7515):449-452.

55. Yuan W, Xu M, Huang C, et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J Biol Chem. 2011 Mar 11;286(10):7983-9.

56. Schmitges FW, Prusty AB, Faty M, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. Mol Cell. 2011 May 6;42(3):330-41.

57. Brookes E, de Santiago I, Hebenstreit D, et al. Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. Cell Stem Cell. 2012 Feb 3;10(2):157-70.

58. Devaiah BN, Lewis BA, Cherman N, et al. BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain. Proc Natl Acad Sci U S A. 2012 May 1;109(18):6927-32.

59. Nicole J. Francis REK, Christopher L. Woodcock. Chromatin Compaction by a Polycomb Group Protein Complex. Science (New York, NY). 2004 Nov 26;306(5701):1574-7.

60. Endoh M, Endo TA, Endoh T, et al. Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. PLoS Genet. 2012;8(7):e1002774.

61. Bernstein E AC. RNA meets chromatin. Genes Dev. 2005 Jul 15;19(14):1635-55.

62. Zampetaki A, Albrecht A, Steinhofel K. Long Non-coding RNA Structure and Function: Is There a Link? Front Physiol. 2018;9:1201.

63. Forrest ME, Khalil AM. Review: Regulation of the cancer epigenome by long non-coding RNAs. Cancer Lett. 2017 Oct 28;407:106-112.

64. Kondo Y, Shinjo K, Katsushima K. Long non-coding RNAs as an epigenetic regulator in human cancers. Cancer Sci. 2017 Oct;108(10):1927-1933.

65. Fabian MR, Sonenberg N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nat Struct Mol Biol. 2012 Jun 5;19(6):586-93.

66. Marcucci G, Haferlach T, Dohner H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. J Clin Oncol. 2011 Feb 10;29(5):475-86.

67. Marfella CG IA. The Chd family of chromatin remodelers. Mutat Res. 2007 May 1;618(1-2):30–40.

68. Ho L, Crabtree GR. Chromatin remodelling during development. Nature. 2010 Jan 28;463(7280):474-84.

69. Hall JA, Georgel PT. CHD proteins: a diverse family with strong ties. Biochem Cell Biol. 2007 Aug;85(4):463-76.

70. Bickmore WA. The spatial organization of the human genome. Annu Rev Genomics Hum Genet. 2013;14:67-84.

71. Dostie J, Bickmore WA. Chromosome organization in the nucleus - charting new territory across the Hi-Cs. Curr Opin Genet Dev. 2012 Apr;22(2):125-31.

72. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet. 2001 Apr;2(4):292-301.

73. Dekker J, Mirny L. The 3D Genome as Moderator of Chromosomal Communication. Cell. 2016 Mar 10;164(6):1110-21.

74. Dekker J, Misteli T. Long-Range Chromatin Interactions. Cold Spring Harb Perspect Biol. 2015 Oct 01;7(10):a019356.

75. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012 Apr 11;485(7398):376-80.

76. Dekker J, Heard E. Structural and functional diversity of Topologically Associating Domains. FEBS Lett. 2015 Oct 07;589(20 Pt A):2877-84.

77. Bickmore WA, van Steensel B. Genome architecture: domain organization of interphase chromosomes. Cell. 2013 Mar 14;152(6):1270-84.

78. Zullo JM, Demarco IA, Pique-Regi R, et al. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. Cell. 2012 Jun 22;149(7):1474-87.

79. Wen B, Wu H, Shinkai Y, et al. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat Genet. 2009 Feb;41(2):246-50.

80. de Klerk E, t Hoen PA. Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. Trends Genet. 2015 Mar;31(3):128-39.

81. Reyes A, Huber W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. Nucleic Acids Res. 2018 Jan 25;46(2):582-592.

82. Carninci P, Sandelin A, Lenhard B, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet. 2006 Jun;38(6):626-35.

83. Grech G, Pollacco J, Portelli M, et al. Expression of different functional isoforms in haematopoiesis. Int J Hematol. 2014 Jan;99(1):4-11.

84. Doulatov S, Notta F, Laurenti E, et al. Hematopoiesis: a human perspective. Cell Stem Cell. 2012 Feb 3;10(2):120-36.

85. Cui K, Zang C, Roh TY, et al. Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell. 2009 Jan 09;4(1):80-93.

86. Iwasaki H, Akashi K. Myeloid lineage commitment from the hematopoietic stem cell. Immunity. 2007 Jun;26(6):726-40.

87. Luis TC, Killmann NM, Staal FJ. Signal transduction pathways regulating hematopoietic stem cell biology: introduction to a series of Spotlight Reviews. Leukemia. 2012 Jan;26(1):86-90.

88. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 2008 Feb 22;132(4):631-44.

89. Attar EC, Scadden DT. Regulation of hematopoietic stem cell growth. Leukemia. 2004 Nov;18(11):1760-8.

90. Ji H, Ehrlich LI, Seita J, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature. 2010 Sep 16;467(7313):338-42.

91. Prasad P, Ronnerblad M, Arner E, et al. High-throughput transcription profiling identifies putative epigenetic regulators of hematopoiesis. Blood. 2014 Apr 24;123(17):e46-57.

92. Eriksson A, Lennartsson A, Lehmann S. Epigenetic aberrations in acute myeloid leukemia: Early key events during leukemogenesis. Exp Hematol. 2015 Aug;43(8):609-24.

93. Goyama S, Kitamura T. Epigenetics in normal and malignant hematopoiesis: An overview and update 2017. Cancer Sci. 2017 Apr;108(4):553-562.

94. Laiosa CV, Stadtfeld M, Graf T. Determinants of lymphoid-myeloid lineage diversification. Annu Rev Immunol. 2006;24:705-38.

95. Dohner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. N Engl J Med. 2015 Sep 17;373(12):1136-52.

96. O'Donnell MR, Tallman MS, Abboud CN, et al. Acute Myeloid Leukemia, Version 3.2017, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2017 Jul;15(7):926-957.

97. Döhner H EE, Grimwade D, Amadori S, Appelbaum FR, Büchner T, Dombret H, Ebert BL, Fenaux P, Larson RA, Levine RL, Lo-Coco F, Naoe T, Niederwieser D, Ossenkoppele GJ, Sanz M, Sierra J, Tallman MS, Tien HF, Wei AH, Löwenberg B, Bloomfield CD. Diagnosis and management of AML in adults- 2017 ELN recommendations from an international expert panel. Blood. 2017 Jan 26;129(4):424-447.

98. Conway O'Brien E, Prideaux S, Chevassut T. The epigenetic landscape of acute myeloid leukemia. Adv Hematol. 2014;2014:103175.

99. Larsson CA, Cote G, Quintas-Cardama A. The changing mutational landscape of acute myeloid leukemia and myelodysplastic syndrome. Mol Cancer Res. 2013 Aug;11(8):815-27.

100. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat Rev Cancer. 2007 Nov;7(11):823-33.

101. Sexton T, Schober H, Fraser P, et al. Gene regulation through nuclear organization. Nat Struct Mol Biol. 2007 Nov;14(11):1049-55.

102. Thol F, Bollin R, Gehlhaar M, et al. Mutations in the cohesin complex in acute myeloid leukemia: clinical and prognostic implications. Blood. 2014 Feb 6;123(6):914-20.

103. Thota S, Viny AD, Makishima H, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. Blood. 2014 Sep 11;124(11):1790-8.

104. Leeke B, Marsman J, O'Sullivan JM, et al. Cohesin mutations in myeloid malignancies: underlying mechanisms. Experimental hematology & oncology. 2014;3:13.

105. KL. T. Basic science review on circadian rhythm biology and circadian sleep disorders. Ann Acad Med Singapore. 2008 Agu;37(8):662-8.

106. Reppert SM, Weaver DR. Coordination of circadian timing in mammals. Nature. 2002 Aug 29;418(6901):935-41.

107. Eckel-Mahan K, Sassone-Corsi P. Metabolism and the circadian clock converge. Physiol Rev. 2013 Jan;93(1):107-35.

108. Schibler U S-CP. A web of circadian pacemakers. Cell. 2002 Dec 17;111(7):919-22.

109. Stokkan KA, Yamazaki S, Tei H, et al. Entrainment of the circadian clock in the liver by feeding. Science (New York, NY). 2001 Jan 19;291(5503):490-3.

110. Brown SA, Kowalska E, Dallmann R. (Re)inventing the circadian feedback loop. Dev Cell. 2012 Mar 13;22(3):477-87.

111. Gekakis N, Staknis D, Nguyen HB, et al. Role of the CLOCK protein in the mammalian circadian mechanism. Science (New York, NY). 1998 Jun 05;280(5369):1564-9.

112. Takahashi JS. Transcriptional architecture of the mammalian circadian clock. Nat Rev Genet. 2017 Mar;18(3):164-179.

113. Nakahata Y, Grimaldi B, Sahar S, et al. Signaling to the circadian clock: plasticity by chromatin remodeling. Curr Opin Cell Biol. 2007 Apr;19(2):230-7.

114. Duffield GE BJ, Meurers BH, Bittner A, Loros JJ, Dunlap JC. Circadian programs of transcriptional activation, signaling, and protein turnover revealed by microarray analysis of mammalian cells. Curr Biol. 2002 Apr 2;12(7):551-7.

115. Pacheco-Bernal I, Becerril-Perez F, Aguilar-Arnal L. Circadian rhythms in the three-dimensional genome: implications of chromatin interactions for cyclic transcription. Clin Epigenetics. 2019 May 15;11(1):79.

116. Katada S, Sassone-Corsi P. The histone methyltransferase MLL1 permits the oscillation of circadian gene expression. Nature Structural & Molecular Biology. 2010;17(12):1414-1421.

117. Aguilar-Arnal L, Katada S, Orozco-Solis R, et al. NAD(+)-SIRT1 control of H3K4 trimethylation through circadian deacetylation of MLL1. Nat Struct Mol Biol. 2015 Apr;22(4):312-8.

118. Mermet J, Yeung J, Hurni C, et al. Clock-dependent chromatin topology modulates circadian transcription and behavior. Genes & development. 2018 Mar 1;32(5-6):347-358.

119. Kim YH MS, Zhang Y, Steger DJ, Won KJ, Lazar MA. Rev-erbalpha dynamically modulates chromatin looping to control circadian gene transcription. Science. 2018 Mar 16;359(6381):1274-1277.

120. Doi M, Hirayama J, Sassone-Corsi P. Circadian regulator CLOCK is a histone acetyltransferase. Cell. 2006 May 05;125(3):497-508.

121. Altmeyer M, Hottiger MO. Poly(ADP-ribose) polymerase 1 at the crossroad of metabolic stress and inflammation in aging. Aging. 2009 May 20;1(5):458-69.

122. Dardente H, Cermakian N. Molecular circadian rhythms in central and peripheral clocks in mammals. Chronobiol Int. 2007;24(2):195-213.

123. Balsalobre A DF, Schibler U. A serum shock induces circadian gene expression in mammalian tissue culture cells. Cell. 2998 Jun 12;93(6):929-37.

124. Imreh MP WS, Unger C, Gertow K, Aints A, Szeles A, Imreh S, Hovatta O, Fried G, Dilber S, Ahrlund-Richter L. Culture and expansion of the human embryonic stem cell line HS181, evaluated in a double-color system. Stem Cells Dev. 2004 Aug;13(4):337-43.

125. Pereira C, Clarke E, Damen J. Hematopoietic colony-forming cell assays. Methods in molecular biology (Clifton, NJ). 2007;407:177-208.

126. Jinek M CK, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (New York, NY). 2012 Aug 17;337(6096):816-21.

127. Cui Y, Xu J, Cheng M, et al. Review of CRISPR/Cas9 sgRNA Design Tools. Interdiscip Sci. 2018 Jun;10(2):455-465.

128. Zaidi SS, Mahfouz MM, Mansoor S. CRISPR-Cpf1: A New Tool for Plant Genome Editing. Trends Plant Sci. 2017 Jul;22(7):550-553.

129. Gilbert LA, Horlbeck MA, Adamson B, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell. 2014 Oct 23;159(3):647-61.

130. Huang R, Zhao L, Chen H, et al. Megakaryocytic differentiation of K562 cells induced by PMA reduced the activity of respiratory chain complex IV. PLoS One. 2014;9(5):e96246.

131. Siggens L, Cordeddu L, Ronnerblad M, et al. Transcription-coupled recruitment of human CHD1 and CHD2 influences chromatin accessibility and histone H3 and H3.3 occupancy at active chromatin regions. Epigenetics Chromatin. 2015;8(1):4.

132. van den Boom V, Maat H, Geugien M, et al. Non-canonical PRC1.1 Targets Active Genes Independent of H3K27me3 and Is Essential for Leukemogenesis. Cell Rep. 2016 Jan 12;14(2):332-46.

133. St Louis DC, Woodcock JB, Franzoso G, et al. Evidence for distinct intracellular signaling pathways in CD34+ progenitor to dendritic cell differentiation from a human cell line model. Journal of immunology (Baltimore, Md : 1950). 1999 Mar 15;162(6):3237-48.

134. Shen X, Liu Y, Hsu YJ, et al. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell. 2008 Nov 21;32(4):491-502.

135. Davuluri RV, Suzuki Y, Sugano S, et al. The functional consequences of alternative promoter use in mammalian genomes. Trends Genet. 2008 Apr;24(4):167-77.

136. Kioka N, Ito T, Yamashita H, et al. Crucial role of vinexin for keratinocyte migration in vitro and epidermal wound healing in vivo. Exp Cell Res. 2010 Jun 10;316(10):1728-38.

137. Kioka N, Sakata S, Kawauchi T, et al. Vinexin: a novel vinculin-binding protein with multiple SH3 domains enhances actin cytoskeletal organization. J Cell Biol. 1999 Jan 11;144(1):59-69.

138. Kimura A, Baumann CA, Chiang SH, et al. The sorbin homology domain: a motif for the targeting of proteins to lipid rafts. Proc Natl Acad Sci U S A. 2001 Jul 31;98(16):9098-103.

139. Göndör A, Rougier C, Ohlsson R. High-resolution circular chromosome conformation capture assay. Nature Protocols. 2008;3(2):303-313.

140. Sandhu KS, Shi C, Sjolinder M, et al. Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. Genes & development. 2009 Nov 15;23(22):2598-603.

141. Yu W, Ginjala V, Pant V, et al. Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. Nat Genet. 2004 Oct;36(10):1105-10.

142. Guastafierro T, Cecchinelli B, Zampieri M, et al. CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. J Biol Chem. 2008 Aug 8;283(32):21873-80.

143. Asher G, Reinke H, Altmeyer M, et al. Poly(ADP-ribose) polymerase 1 participates in the phase entrainment of circadian clocks to feeding. Cell. 2010 Sep 17;142(6):943-53.

144. Yagita K, Horie K, Koinuma S, et al. Development of the circadian oscillator during differentiation of mouse embryonic stem cells in vitro. Proc Natl Acad Sci U S A. 2010 Feb 23;107(8):3846-51.

145. Relogio A, Thomas P, Medina-Perez P, et al. Ras-mediated deregulation of the circadian clock in cancer. PLoS Genet. 2014;10(5):e1004338.

146. Chen X, Shi C, Yammine S, et al. Chromatin in situ proximity (ChrISP): single-cell analysis of chromatin proximities at a high resolution. BioTechniques. 2014;56(3):117-8, 120-4.

147. Plass C, Pfister SM, Lindroth AM, et al. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. Nat Rev Genet. 2013 Nov;14(11):765-80.

148. Heshmati Y, Turkoz G, Harisankar A, et al. The chromatin-remodeling factor CHD4 is required for maintenance of childhood acute myeloid leukemia. Haematologica. 2018 Jul;103(7):1169-1181.

149. Gaspar-Maia A, Alajem A, Polesso F, et al. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature. 2009 Aug 13;460(7257):863-8.

150. Harada A, Okada S, Konno D, et al. Chd2 interacts with H3.3 to determine myogenic cell fate. EMBO J. 2012 Jun 29;31(13):2994-3007.

151. Nagarajan P, Onami TM, Rajagopalan S, et al. Role of chromodomain helicase DNA-binding protein 2 in DNA damage response signaling and tumorigenesis. Oncogene. 2009 Feb 26;28(8):1053-62.

152. Kanamori-Katayama M, Itoh M, Kawaji H, et al. Unamplified cap analysis of gene expression on a single-molecule sequencer. Genome Res. 2011 Jul;21(7):1150-9.

153. Takahashi H, Lassmann T, Murata M, et al. 5′ end–centered expression profiling using cap-analysis gene expression and next-generation sequencing. Nature Protocols. 2012;7(3):542-561.