From the DEPARTMENT OF BIOSCIENCES AND NUTRITION
Karolinska Institutet, Stockholm, Sweden

DISSECTING THE EPIGENETIC LANDSCAPES OF HEMATOPOIESIS AND FISSION YEAST

MICHELLE RÖNNERBLAD

Karolinska Institutet

Stockholm 2014

Published by Karolinska Institutet. Printed by Åtta.45 Tryckeri AB.

ISBN 978-91-7549-465-4

"*The important thing is not to stop questioning"*

- Albert Einstein (1879-1955)

ABSTRACT

The genome of eukaryotic cells is stored in the nucleus as chromatin, a DNA-protein complex that serves to compact and protect the DNA molecules. The basic unit of chromatin is the nucleosome composed of DNA wrapped around a histone protein core. In addition to condensing and protecting the genome, chromatin confers a number of regulatory properties employed for example in control of gene expression and stabilization of repetitive sequences. Chromatin also constitutes an obstacle that needs to be negotiated in processes such as transcription elongation, DNA replication and DNA repair. A wide range of chromatin modifying factors and mechanisms are involved in regulating the state of chromatin and affect all DNA related processes. These mechanisms, often referred to as epigenetic, include methylation of DNA, regulation by non-coding RNAs, remodeling of nucleosomes, posttranslational modifications of histones and incorporation of variant histones. The resulting chromatin state is called the epigenome and can, in contrast to the underlying DNA sequence, differ between cells in the same organism.

This thesis describes characterization of aspects of the egipenomes of hematopoietic cells and fission yeast. We show that in fission yeast, genes with related functions share common patterns of histone modifications in the promoter regions. We also demonstrate crosstalk between different histone modifications, including interdependence of histone H4 acetylation sites and regulatory roles of histone methylation for histone acetylation.

To better understand how chromatin factors influence human blood development we analysed expression of genes encoding chromatin modifying proteins in the hematopoietic system, including the hematopoietic stem cells and a wide range of mature blood cells. In doing so we could identify epigenetic factors that were expressed in cell type, cell lineage or cancer specific patterns, implicating them in regulation of blood development. We also found that several genes display differential use of alternative transcription start sites between cell types.

Finally we constructed an in-depth map of how DNA methylation and gene expression changes during human granulocyte development. Our experiments show that DNA methylation changes are linked to points of lineage restriction, implicating DNA methylation in control of cell fate. DNA methylation changes, most of which were decreases, were primarily located outside of CpG islands, which have been the focus of most DNA methylation studies historically. Interestingly, DNA methylation was especially dynamic in enhancer elements, and sites with decreasing DNA methylation overlapped with differentiation induced enhancers and increased expression of target genes. This result suggests a role of DNA methylation in regulating enhancer activity in granulopoiesis.

LIST OF PUBLICATIONS

I. Sinha I*, Buchanan L*, **Rönnerblad M**, Bonilla C, Durand-Dubief M, Shevchenko A, Grunstein M, Stewart AF, Ekwall K (2010). Genome-wide mapping of histone modifications and mass spectrometry reveal H4 acetylation bias and H3K36 methylation at gene promoters in fission yeast. *Epigenomics* **2**: 377-393

II. Prasad P*, **Rönnerblad M***, Arner E, Itoh M, Kawaji H, Lassman T, Daub C, Forrest AR, the FANTOM consortium, Andreas Lennartsson, Karl Ekwall. High-throughput transcription profiling identifies putative epigenetic regulators of hematopoiesis. *Blood, manuscript accepted for publication.*

III. **Rönnerblad M**, Andersson R, Olofsson T, Douagi I, Karimi M, Lehmann S, Hoof I, de Hoon M, Itoh M, Nagao-Sato S, Kawaji H, Lassman T, Carnici P, Hayashizaki Y, Forrest AR, Sandelin A, the FANTOM consortium, Ekwall K, Arner E, Lennartsson A. Analysis of the DNA methylome and transcriptome in granulopoieis reveal timed changes and dynamic enhancer methylation. *Blood, manuscript accepted for publication.*

**Equal contribution*

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

2-HG	2-hydroxyglutarate
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
Ac	Acetylation
ADD	ATRX-Dnmt3-Dnmt3L
ALL	Acute lymphoid leukemia
AML	Acute myeloid leukemia
ATP	Adenosine triphosphate
BAF	BRG1 associated factor
BER	Base excision repair
Bp	Basepair
CAGE	Cap-analysis gene expression
ChIP	Chromatin immunoprecipitation
CLL	Chronic lymphoid leukemia
CMP	Common myeloid progenitor
CRC	Chromatin remodeling complex
CTD	C-terminal domain
DMR	Differentially methylated region
DMS	Differentially methylated site
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
ESC	Embryonal stem cell
FACS	Flourescence-activated cell sorting
FAD	Flavin adenine dinucleotide
GFT	General transcription factor
GMP	Granulocyte-Macrophage progenitor
HDAC	Histone deacetylase
HP1	Heterochromatin protein 1
HSC	Hematopoietic stem cell
KAT	Lysine acetyltransferase
KDM	Lysine deacetylase
KMT	Lysine methyltransferase

MBD	Methyl-CpG-binding domain
MDS	Myelodysplastic syndrome
Me	Methylation
MLL	Mixed lineage leukemia
mRNA	Messenger RNA
MS	Mass spectrometry
NER	Nucleotide excision repair
Nt	Nucleotide
ORF	Open reading frame
PcG	Polycomb-group proteins
PIC	Preinitiation complex
Pol II	RNA polymerase II
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2
RLE	Relative log expression
RNA	Ribonucleic acid
SAM	S-adenosyl methionine
TET	Ten-eleven translocation
TFII	Transcription factor of pol II
TSS	Transcription start site
WGBS	Whole genome bisulfite sequencing
α-KG	α -ketoglutarate

1 INTRODUCTION

1.1 CHROMATIN

A human cell contains approximately two meters of DNA stored in a nucleus with a diameter of around six micrometres. For this to be possible, the genome must be efficiently compacted and organized to avoid physical damage. At the same time, DNA sequences such as genes and other functional loci must remain accessible upon demand. These requirements are fulfilled through packing the genome of eukaryotic cells in the form of chromatin, a complex of equal amounts of DNA and proteins. Chromatin organizes the genome into structures of varying compaction, the most extreme of which is the condensed metaphase chromosome. The first level of compaction consists of DNA wound around a core of histone proteins creating the nucleosome.

There are several forms of specialized chromatin. Heterochromatin is highly compacted and generally silent and includes for example certain repetitive regions and the silent X-chromosome. Euchromatin, on the other hand, is less compact and general more transcriptionally active while the specialized centromeric heterochromatin is crucial for chromosome segregation.

In addition to condensing and organizing the DNA in the nucleus, chromatin also contributes to regulation of DNA-related processes. Indeed, basically all DNA-related processes, including replication, repair and transcription, must take place in the context of chromatin, putting high demands on regulated accessibility.

Epigenetics has been defined as heritable changes in gene expression without changes to the underlying DNA sequence (Russo *et al*, 1996), but is commonly used to describe changes in chromatin state (Bird, 2007). Epigenetic mechanisms include addition and removal of posttranslational modifications to histones, incorporation of histone variants, rearrangement of nucleosomes by chromatin remodeling enzymes, methylation of DNA and regulation by non-coding RNA. These mechanisms of modifying chromatin play a central role in regulating chromatin states and functions. Epigenetics is recognized as an important determinant in normal development and differentiation, and epigenetic abnormalities are relevant in many diseases, including various malignancies.

1.2 THE NUCLEOSOME AND HISTONES

The basic unit of chromatin is the nucleosome, consisting of 146 bp wrapped 1.65 turns around a histone octamer core (Luger *et al*, 1997) (Figure 1A). Individual nucleosomes are connected by short stretches of DNA, called linker DNA, into nucleosomal arrays resembling beads on a string when viewed by electron microscopy. Nucleosomal arrays are further organized into chromatin fibres of increasing compaction, the most extreme of which is the metaphase chromosome.

The histone octamer core of the nucleosome consists of two copies each of histones H2A, H2B, H3 and H4. These small basic proteins possess a conserved "histone fold" structure with three alpha helices connected by two loops. Histones assemble into H2A-H2B and H3-H4 dimers. In the nucleosome two H3-H4 dimers form a central tetramer flanked by two H2A-H2B dimers (Luger *et al*, 1997). The unstructured N- and C-terminal tails of the histones extend from the nucleosome and are involved in interactions with neighbouring nucleosomes and other proteins. Linker histones, exemplified by histone H1, bind to the linker DNA and facilitate formation of higher order chromatin structures. The stability of the nucleosome is affected by several factors such as the underlying DNA sequence and the specific composition and modifications of the histone core.

1.2.1 Histone variants

The canonical histones are expressed and incorporated in a replication dependent manner. In addition to these, there are also a number of histone variants whose expression is replication independent. Most variants are for H2A and H3 and many are conserved between species (Talbert & Henikoff, 2010). The histone variants differ from their major-type counterpart in the amino acid sequence giving them unique properties and several have been associated with specific functions and locations in chromatin. For example, centromere specific H3 variants (CenH3) occupy centromeric chromatin and are required for kinetochore assembly (Talbert & Henikoff, 2010). H2A.Z has a conserved localization at 5´-ends of genes (Mavrich *et al*, 2008; Li *et al*, 2005; Zilberman *et al*, 2008; Barski *et al*, 2007), and is believed to be involved in, among other processes, transcription regulation. However, the precise effect seems to depend of the species/cell type and on posttranslational modifications (Li *et al*, 2005; Barski *et al*, 2007; Talbert & Henikoff, 2010; Millar, 2013). The histone variant H2A.X is highly similar to canonical H2A, but has a C-terminal serine that is phosphorylated at the site of double strand DNA lesions and is thought to recruit and/or retain repair machinery (Talbert & Henikoff, 2010).

1.3 HISTONE MODIFICATIONS

Histones are subject to a multitude of posttranslational modifications, most famously acetylation, methylation, phosphorylation, sumoylation and ubiquitination (Figure 1B). Recently several new histone modifications and modification sites have been discovered (Tan *et al*, 2011b), but in many cases the function of these novel modifications remains to be tested. Histones are preferentially modified on the protruding N-terminal tails, although some modifications localize to the globular domains. Histone modifications affect many aspects of chromatin biology including DNA repair, chromatin compaction, transcription initiation and elongation.

Considering the high number of different histone modifications and modification sites, the potential complexity is enormous. Histone modifications have been proposed to constitute a histone code where specific combinations of modifications give rise to specific effects in chromatin (Strahl & Allis, 2000). The histone code theory has been debated and several objections have been raised against it. Most importantly the limited complexity of observed modification patterns is argued to be incompatible with a true

code (Rando, 2012). Nevertheless, sets of modification patterns preferentially associated with genes of specific functional classes have been reported (Kurdistani *et al*, 2004).

This section will focus on acetylation and methylation of lysine residues, two of the best characterized types of histone modifications, and the enzymes involved in catalyzing their addition to and removal from histone substrates.

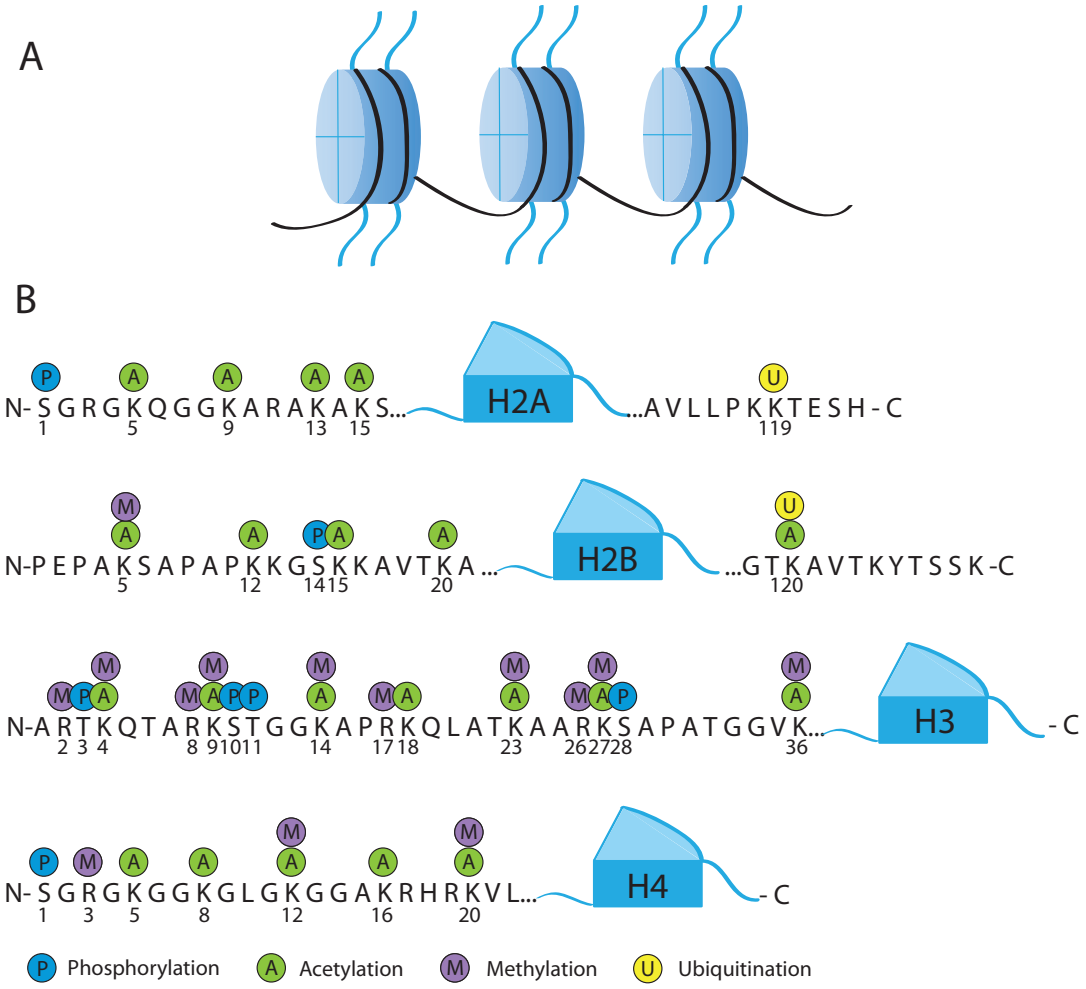

Figure 1 Histone modifications

A) Schematic picture of nucleosomes. DNA (black) is wound 1.65 turns around a histone protein core (blue) composed of two copies of histone H2A, H2B, H3 and H4. The unstructured histone tails protrude from the nucleosomes.

B) Histone tails, in particular the N-terminal tails, but also the globular domains, are subject to a plethora of posttranslational modifications. This picture shows some of the better known sites for acetylation (A), methylation (M), phosphorylation (P) and ubiquitination (U). In addition, histones may be modified by sumoylation, ADP-ribosylation, glycosylation, crotonylation, formylation, glycosylation, succinylation, oxidation and propionylation.

1.3.1 Histone acetylation, KATs and HDACs

The basic N-terminal tails of histones contain multiple lysine residues and lysine acetylation is one of the most studied histone modifications. Histone acetylation and methylation was described already in 1964 (Allfrey & Mirsky, 1964). Since then acetylation of non-histone proteins has also gained interest and is believed to be as extensive and important to cellular biology as protein phosphorylation (Kouzarides, 2000).

Despite the early description of histone acetylation, the enzymes adding and removing this mark were not discovered until 1995 (Kleff *et al*, 1995; Brownell & Allis, 1995; Taunton *et al*, 1996). Histone, or lysine, acetyltransferases (KATs) catalyze the transfer of an acetyl group from acetyl-CoA to the ε-amino group of lysines. Type A KATs constitute a diverse family that reside within the nucleus and can be classified into three groups: GNAT, MYST and CBP/p300 (Bannister & Kouzarides, 2011). The type B KATs are highly conserved and related to budding yeast Kat1p. Type B KATs are mainly cytoplasmic and acetylate free/newly synthesized histones on primarily H4K5 and H4K12 (Parthun, 2007). This modification is important for histone incorporation into chromatin, after which the mark is removed. Removal of acetyl groups is catalyzed by histone deacetylases (HDACs). There are four classes of HDACs. Class I, II and VI require zinc for catalytic activity while the class III HDACs, called sirtuins based on homology to yeast Sir2p, are NADH-dependent.

Acetylation is dynamic and the precise acetylation levels result from the opposing activities of KATs and HDACs, which often localize to the same sites simultaneously or even interact physically to maintain acetylation balance. Although there is some variation depending on the precise lysine residue involved, acetylation is generally considered to be an active mark present at open chromatin and at promoters of transcribed genes (Bannister & Kouzarides, 2011; Wang *et al*, 2008). Similarly, many KATs are considered to be coactivators and many HDACs corepressors (Bannister & Kouzarides, 2011). The effect of acetylation on chromatin is likely mediated by both the neutralization of the positive charge of lysines decreasing the interaction strength with DNA (Zentner & Henikoff, 2013; Hong *et al*, 1993) and by recruitment of proteins and complexes containing a bromodomain that specifically recognize and bind acetyllysines (Peserico & Simone, 2011). Bromodomains are present in many proteins and protein complexes including chromatin modifiers such as KATs, lysine methyl transferases (KMTs), chromatin remodeling complexes (CRCs), and components of the general transcription factor TFIID (Filippakopoulos & Knapp, 2012). Interstingly, H4K16ac appears to be directly involved in regulating chromatin compaction, as it disrupts formation of higher order chromatin structures (Shogren-Knaak *et al*, 2006).

1.3.2 Histone methylation, KMTs and KDMs

Histones can be methylated at both lysine and arginine residues, although lysine methylation is more extensively studied. In contrast to acetylation, lysines can be mono-, di- or trimethylated. The state of methylation impacts on the biological outcome, adding an additional level of complexity and regulation (Bannister & Kouzarides, 2011).

Lysine methyltransferases (KMTs) transfer a methyl group from S-adenosyl methionine (SAM) to the lysine ε-amino group (Bannister & Kouzarides, 2011). While all KMTs that act on the N-terminal histone tails have a common SET-domain, DOT1, which catalyses H3K79 methylation on the globular domain, does not and is structurally distinct from other KMTs (Greer & Shi, 2012; Feng *et al*, 2002). Most KMTs function in multisubunit complexes that target the H3 N-terminal tail and display a high level of specificity, both regarding the lysine site and the level of methylation (Butler & Dent, 2013; Bannister & Kouzarides, 2011). For example, humans have several H3K4 KMTs. SETD7 (KMT7) can only monomethylate H3K4 (Xiao *et al*, 2003; Wang *et al*, 2001). Other KMTs such as the MLL1-4 (KMT2A-D), SETD1A (KMT2F) and SETD1B (KMT2G) are able to catalyse mono-, di- and trimethylation of the same site (Greer & Shi, 2012).

Lysine demethylases (KDM) were discovered relatively recently. The earliest one to be described was lysine-specific demethylase 1 (LSD1 or KDM1A), which was first found in 2004 (Shi *et al*, 2004). This enzyme uses flavin adenosine dinucleotide (FAD) as a cofactor (Shi *et al*, 2004) and has different specificities depending on the complex it associates with. In the context of the CoREST corepressor complex LSD1 demethylates H3K4, while acting as a H3K9 demethylase when interacting with the androgen receptor (Klose & Zhang, 2007). As methylation of these sites have opposite effects on gene expression, LSD1 can have dual functions as both an activator and a repressor.

While the discovery of LSD1 was a major advance, this enzyme is only capable of removing mono- and dimethylation (Shi *et al*, 2004). Enzymes active on trimethylated lysines were not found until 2006, when the Jumonji demethylase family was discovered (Tsukada *et al*, 2005; Whetstine *et al*, 2006). These enzymes have a jumonji domain that is able to remove trimethylation using Fe(II) and α-ketoglutarate (α-KG) as cofactors (Whetstine *et al*, 2006). As KMTs, demethylases have high substrate specificity (Bannister & Kouzarides, 2011).

Unlike acetylation the addition of a methyl group does not change the charge of the histone protein. Futhermore, in contrast to acetylation, methylation has been strongly associated with both transcriptional activation and repression depending on the precise lysine site modified. For example, promoter H3K4me3 is strongly correlated with active gene expression (Barski *et al*, 2007; Justin *et al*, 2010) while H3K9me2/3 is a repressive mark that recruits heterochromatin protein 1 (HP1) and is pivotal in heterochromatin formation (Lachner *et al*, 2001; Bannister & Kouzarides, 2011). H3K27me3 is another repressive methylation mark that is added and read by polycomb-group proteins (PcG). H3K27 is methylated by the polycomb repressive complex 2 (PRC2), which includes the KMT EZH2 (KMT6). H3K27me3 is subsequently recognized and bound by PRC1 leading to transcriptional silencing and chromatin compaction (Simon & Kingston, 2013).

Some promoters show patterns of overlapping H3K4me3 and H3K27me3, commonly referred to as bivalent domains (Bernstein *et al*, 2006). Originally described in embryonic stem cells (ESC), bivalent promoters are often associated with genes involved in cell fate determination and differentiation (Bernstein *et al*, 2006). The

bivalent domains are believed to keep genes in a state that is transcriptionally silent, but poised for activation.

1.3.3 Crosstalk between histone modifications

There is significant crosstalk between different histone modifications, either *in situ* (the same site), in *cis* (of the same histone) or in *trans* (between different histone molecules). Both lysine and arginine can be modified in more than one way. Lysines for example may be acetylated, methylated, sumoylated or ubiquitinated, depending on the site (figure 1B). Since these modifications are mutually exclusive on the same lysine residue, it represents the most direct version of crosstalk. Importantly, various modifications of a particular site are commonly involved in different or even opposite processes. While H3K9me3 is associated with heterochromatin, H3K9ac is found at transcribed genes (Lachner *et al*, 2001; Bannister & Kouzarides, 2011; Wang *et al*, 2008). H3K36ac is common in active promoters while H3K36 is methylated in transcribed coding regions (Barski *et al*, 2007; Carrozza *et al*, 2005; Wang *et al*, 2008).

There are several examples of histone modifications at separate sites on the same histone affecting each other in *cis*. H3S10 phosphorylation is believed to cause acetylation of H3K14 (Edmondson, 2002) but blocks acetylation of H3K9 (Latham & Dent, 2007). Acetylation of histone H4 follows a pattern dubbed the "acetylation zip" for K16, K12, K8 and K5 (Turner *et al*, 1989; Zhang *et al*, 2002; Tweedie-Cullen *et al*, 2012). These lysines are acetylated from the globular domain and outward (i.e. first K16, then K12 etc.) suggesting some mode of crosstalk in *cis*. Crosstalk in *trans* tends to be somewhat more complicated. One well-know example is the requirement of H2BK120 ubiquitination (H2BK123 in budding yeast), for methylation of H3K4 and H3K79 in gene bodies during transcription (Latham & Dent, 2007). Another potential crosstalk in *trans* may be mediated through the physical interaction of the H3K4 KMT MLL (MLL1/KMT2A) and the KAT MOF (KAT8), possibly linking H3K4me and H4 acetylation (Dou *et al*, 2005).

1.4 CHROMATIN REMODELING COMPLEXES

Chromatin remodeling complexes (CRCs) utilize the energy of adenosine triphosphate (ATP) hydrolysis to regulate the structure of nucleosomal chromatin (Hargreaves & Crabtree, 2011; Clapier & Cairns, 2009). CRCs perform remodeling by evicting, rearranging or sliding nucleosomes along DNA and some are involved in replacing canonical histones with histone variants (Clapier & Cairns, 2009). These complexes participate in spacing nucleosomes after replication and moving them to allow passage of polymerases during transcription and replication, as well as ensuring access of the repair machinery upon DNA damage (Narlikar *et al*, 2013; Clapier & Cairns, 2009; Hargreaves & Crabtree, 2011). Remodelers also have regulatory functions by adjusting nucleosome positions to hide or expose DNA elements functioning as recognition sites (Clapier & Cairns, 2009; Hargreaves & Crabtree, 2011; Narlikar *et al*, 2013).

The catalytic subunits of CRCs share a DNA-dependent ATPase domain related to the yeast Snf2-helicase (Ryan & Owen-Hughes, 2011) and CRCs can be divided into four

main families based on the sequence of the ATPase subunit (Flaus, 2006; Clapier & Cairns, 2009). Although the INO80, CHD, SWI/SNF and ISWI families all contain the Snf2-related ATPase domain, they are distinguishable by the flanking domains (Flaus, 2006; Clapier & Cairns, 2009). The SWI/SNF ATPases, for example, have a C-terminal bromodomain, allowing for recognition of acetylated lysines, while the CHD remodelers contain N-terminal tandem chromodomains for methyllysine recognition (Clapier & Cairns, 2009).

As with other chromatin modifiers most CRC ATPases are incorporated into large multimeric complexes, the precise functions of which often depend on subunit composition. Accessory subunits are involved in regulating ATPase activity, interacting with other chromatin modifying factors or transcription factors and targeting for example to specifically modified histones (Narlikar *et al*, 2013; Clapier & Cairns, 2009). CRCs have both overlapping and specific functions in chromatin biology, and the precise division of labor has not been completely elucidated.

Many CRCs have been shown to be required for normal development. One example is the SWI/SNF BAF (BRG1 associated factor) complex. Mammalian BAF complexes include one of the ATPase subunits SMARCA4 (BRGI1) or SMARCA2 (BRM) (Wang *et al*, 1996b). *Smarca4* is an essential gene in mice (Bultman *et al*, 2000), while *Smarca2* mutation causes growth abnormalities (Reyes *et al*, 1998). Interestingly, BAF subunit composition has in many cases been shown to be specific for cell type or developmental stage, and subunit replacement may be involved in driving differentiation. For example, mouse embryonic stem cells (ESCs) have a specific BAF complex that is required for pluripotency (Ho *et al*, 2009). Similarly, one study showed that proliferation of neuronal progenitors requires BAF subunits BAF45a and BAF53a (Lessard *et al*, 2007). Transition into postmitotic neuronal cells was accompanied by replacement by these subunits by BAF53b, BAF45b and BAF45c and this switch was important for normal differentiation.

1.5 DNA METHYLATION

DNA can be methylated on the 5-carbon of cytosines, creating 5-methyl cytosine (5mC). The most common form of methylation is on cytosines in CpG dinucleotides, although significant non-CpG methylation has been reported in ESC (Ramsahoye *et al*, 2000; Lister *et al*, 2009) and in murine frontal cortex (Xie *et al*, 2012). CpG dinucleotides are significantly underrepresented in the genome, possibly because of the vulnerability of 5mC to deamination transforming it to thymine and leading to a possible C to T mutation (Jones, 2012).

In addition, CpGs are unevenly distributed in the genome and are concentrated in regions called CpG islands (CGI) (Jones, 2012; Illingworth & Bird, 2009). Although the criteria for CGIs have been subject to some debate, one common definition is a region of at least 200 bp, with at least 50% GC content and 60% of the expected CpG frequency (Gardiner-Garden & Frommer, 1987).

Whereas the bulk of CpGs are methylated in vertebrates, corresponding to approximately 1% methylation of the genome (Ehrlich *et al*, 1982; Bird & Taggart,

1980), CpG islands are generally unmethylated (Illingworth & Bird, 2009). Hypomethylation of CGI regions in germline cells explains why these regions have been protected against CpG depletion by deamination mutations (Jones, 2012).

60-70% of mammalian promoters are associated with CGIs, accounting for half of the CGIs in the genome (Illingworth & Bird, 2009; Illingworth *et al*, 2010). Genes with CGI promoters are predominantly housekeeping genes although some are tissue-specific or development regulatory genes (Deaton & Bird, 2011). Promoter or transcription start site (TSS) methylation is strongly associated with transcriptional repression but CGI promoters can be silent without being methylated (Weber *et al*, 2007; Deaton & Bird, 2011). Methylation of CGI promoters generally reflects long term and stable repression, for example of pluripotency genes in somatic cells (Mohn *et al*, 2008). While inhibiting transcription initiation, DNA methylation does not appear to block elongation as gene bodies are often significantly methylated (Laurent *et al*, 2010; Jones, 2012). Interestingly, gene body methylation is not uniform, but higher in exons than introns, which has inspired theories concerning regulation of splicing (Laurent *et al*, 2010; Jones, 2012).

In addition to transcriptional regulation, DNA methylation is also highly important for genome stability by silencing transposable elements, stabilizing repetitive sequences, X-chromosome inactivation and parental gene imprinting (Jones, 2012).

1.5.1 DNA methylation writers, DNMTs

DNA methylation is performed by DNA methyltransferases, of which there are three catalytically active members in mammals: DNMT3A, DNMT3B and DNMT1 (Moore *et al*, 2012). These enzymes transfer a methyl group from SAM to the C5 position carbon of cytosine (Moore *et al*, 2012). DNMT3A and DNMT3B are *de novo* enzymes active on previously unmethylated DNA (Okano *et al*, 1998; 1999), whereas DNMT1 is a maintenance enzyme, preferentially acting on hemimethylated DNA after replication to conserve methylation patterns (Pradhan *et al*, 1999). All three enzymes are required for proper development. *Dnmt1* and *Dnmt3b* knockouts are embryonically lethal in mice, whereas *Dnmt3a* knockout mice die a few weeks after birth (Okano *et al*, 1999; Li *et al*, 1992). Both DNMT3A and B are required for establishing methylation patterns during development (Jones, 2012; Okano *et al*, 1999), and although highly similar they have separate functions and expression patterns (Xie *et al*, 1999). While DNMT3B is preferentially expressed in stem cells and low in most differentiated tissues, DNMT3A and DNMT1 are relatively ubiquitously expressed (Xie *et al*, 1999; Yen *et al*, 1992).

DNMT1 localizes to the replication fork during S-phase and targets hemimethylated DNA through its accessory factor UHRF1 (Leonhardt *et al*, 1992; Bostick *et al*, 2007). DNMT3A and DNMT3B interact with DNMT3L, a non-catalytic DNMT family member that stimulates DNMT3A/B activity (Hata *et al*, 2002; Chen *et al*, 2005). DNMT3L is necessary for establishing methylation patterns in early development (Hata *et al*, 2002; Chen *et al*, 2005; Moore *et al*, 2012), but is not expressed in most adult tissues, except for in germ cells and thymus (Hata *et al*, 2002; Aapola *et al*, 2000).

1.5.2 Effects of DNA methylation, 5mC readers

The mechanisms behind DNA methylation effects on chromatin and transcription have not been completely elucidated. Broadly speaking, however, they can be classified into two categories; attraction of specific 5mC binding proteins and blocking binding of 5mC sensitive proteins (Klose & Bird, 2006).

Members of three protein families have been shown to bind 5mC: the methyl-CpG-binding domain (MBD) family, the zinc-finger (ZF) family and the SRA-family, including the DNMT1 associated factor UHRF1 and the related UHRF2 (Buck-Koehntop & Defossez, 2013). Of the MBD family, MECP2, MBD1, MBD2 and MBD4 have been shown to bind methylated DNA in different sequence contexts. MBD1, MBD2 and MECP2 have described functions in gene repression (Hendrich & Bird, 1998; Lewis *et al*, 1992; Cross *et al*, 1997; Buck-Koehntop & Defossez, 2013). Notably, this effect appears to be, at least in part, mediated through recruitment of other chromatin modifiers (Klose & Bird, 2006). For example both MBD2 and MECP2 are involved in targeting corepressors such as HDAC-containing complexes to methylated DNA (Ng *et al*, 1999; Jones *et al*, 1998) and MBD1 interacts with the H3K9 methyltransferae SETDB1 (KMT1E) (Sarraf & Stancheva, 2004).

DNA methylation also can block or decrease binding of proteins to DNA, including transcription factors such as MYC (Jones, 2012; Klose & Bird, 2006). Comparably, proteins harboring a ZF-CXXC domain have been shown to preferentially bind unmethylated DNA (Long *et al*, 2013). ZF-CXXC-containing proteins include the H3K36me demethylases KDM2A and KDM2B, CFP1 (which interacts with the SETD1 H3K4 methylase complex), and the H3K4 KMTs MLL and MLL2 (Long *et al*, 2013).

1.5.3 DNA methylation erasers

There are two conceivable mechanisms by which DNA can be demethylated: passive and active demethylation. Passive demethylation simply entails lack of remethylation of the daughter strand of newly synthesised DNA. Active demethylation is more controversial, but several putative demethylases have been proposed (Bhutani *et al*, 2011; Schomacher, 2013). Despite the on-going discussion concerning putative demethylases it is clear that active DNA demethylation does occur as postmitotic or non-dividing cells have been shown to loose DNA methylation at specific loci upon differentiation or gene induction (Bruniquel & Schwartz, 2003; Klug *et al*, 2010; Miller & Sweatt, 2008). In addition, active demethylation on a global scale has been observed in the paternal genome after fertilization, but before cell division, as well as in germ line progenitors (Mayer *et al*, 2000; Schomacher, 2013).

To date, several pathways of DNA demethylation have been proposed with varying amounts of evidence and counter-evidence. In plants, a 5mC-specific glycosylase has been shown to mediate DNA demethylation together with the base excision repair (BER) machinery, but to date no such mechanism has been confirmed in mammals (Schomacher, 2013). However, most of the proposed pathways involve DNA repair

mechanisms, epecially BER or nucleotide excision repair (NER), excising the 5mC or modified 5mC base (Franchini *et al*, 2012).

In one proposed DNA demethylation pathway, the DNA-deaminase AID has been suggested to deaminate modified cytosines, leading to their removal by BER pathways (Franchini *et al*, 2012). The GADD45 proteins are other examples of proposed mediators of DNA demethylation. These proteins have been suggested to direct the DNA repair machinery for removal of 5mC at specific loci, although the theory has been questioned due to inconclusive results (Kohli & Zhang, 2013; Schomacher, 2013).

The Ten-eleven translocation (TET) enzymes, first identified by the occurrence of the *TET1-MLL* fusion gene in acute myeloid leukemia (AML) (Ono *et al*, 2002), have been strongly implicated in DNA demethylation (Pastor *et al*, 2013), and are the most accepted candidates for mediators of active DNA demethylation. There are three members of this family, TET1, TET2 and TET3, with different expression patterns in different tissues (Guibert & Weber, 2013). TET enzymes oxidize 5mC in a series of steps to hydroxymethyl- (5hmC), formyl- (5fC) and finally carboxylcytosine (5caC) (Tahiliani *et al*, 2009; Ito *et al*, 2010; 2011). The modified 5mC may be diluted by replication, as DNMT1 is not active on hemi-hydroxymethylated DNA (Hashimoto *et al*, 2012). Alternatively, it could be removed by the BER machinery where the DNA glycosylase TDG excises 5fC or 5caC thereby causing demethylation (Kohli & Zhang, 2013). There is ample evidence supporting the function of TET enymes in DNA demethylation. For example, TET3 is required for the demethylation of the male pronucleus and the decrease of 5mC is associated with an increase of 5hmC, suggesting demethylation through hydroxylation (Wossidlo *et al*, 2011).

1.5.4 DNA hydroxymethylation

The discovery of the TET enzymes and their activity raised the possibility that the products of 5mC oxidation may be more than demethylation intermediates, and have distinct epigenetic roles. 5hmC in particular, has received attention in this capacity (Guibert & Weber, 2013). 5hmC has been detected in numerous tissues and is especially high in ESCs and brain (Tahiliani *et al*, 2009; Kriaucionis & Heintz, 2009; Guibert & Weber, 2013). Importantly it is far more abundant than 5fC and 5caC (Pastor *et al*, 2013). Studies show slightly different genome-wide distributions of 5hmC depending of the cell type examined. Studies in murine and human ESC and brain cells show that this modification is found at promoters, in gene bodies and cis-regulatory elements such as enhancers (see section 1.7.3), whereas other cell types appear to have 5hmC depletion in promoters (Shen & Zhang, 2013; Pastor *et al*, 2013). In ESC, differentiation induced enhancer activation is associated with increases in 5hmC levels (Serandour *et al*, 2012). However, more work must be perfomed to completely resolve whether 5hmC has a function of its own, rather than as a DNA demethylation intermediate. Interestingly, although not recognized by most MBD family members, one study indicates that 5hmC might be specifically bound by MBD3, which has a low affinity for 5mC (Yildirim *et al*, 2011).

1.6 CROSSTALK BETWEEN CHROMATIN MODIFIERS

As already touched upon, there is extensive crosstalk between different chromatin modifications. This is best illustrated by the presence of multiple chromatin modification activities in the same complexes, and by the role of chromatin modifications in recruiting these complexes. One excellent example is the NuRD complex that contains HDAC1 and 2 as well as the chromatin remodeler CHD3 or CHD4 (Allen *et al*, 2013). The complex also contains MBD2 or MBD3, enabling targeting to methylated DNA regions. In addition, methylation of H3K9 enhances binding to chromatin by the PHD (plant homeo domain) domains of CHD4, whereas H3K4 methylation reduces it (Musselman *et al*, 2012). Thus, a single chromatin-modifying complex is affected by histone modifications and DNA methylation, and will itself modify histones and rearrange nucleosomes.

The best characterized crosstalk between DNA methylation and histone modifications concerns methylation of H3K4 and H3K9. DNMT3A and DNMT3B, as well as DNMT3L, have ADD-domains (ATRX-Dnmt3-Dnmt3L) that bind to histone H3 and this interaction is abolished by H3K4 methylation (Ooi *et al*, 2007; Otani *et al*, 2009; Zhang *et al*, 2010). In addition, several histone methyltransferases, including the H3K4 KMTs MLL and MLL2, have a CXXC domain specific for unmethylated DNA (Long *et al*, 2013). Similarly, several H3K9 KMTs interact with and recruit DNMTs, whereas MBD1 interacts with the two H3K9 KMTs SUV39H1 (KMT1A) and SETDB1 (KMT1E), demonstrating another two-way communication between histone modifications and DNA methylation (Hashimoto *et al*, 2010). It is clear that chromatin modifications exist and act in networks of interdependent mechanisms, the complexity of which we still have a lot to learn about.

1.7 TRANSCRIPTION

1.7.1 Basic transcription machinery

RNA polymerase II (pol II) transcribes protein coding genes as well as many functional RNA genes. Transcription regulation is required to ensure expression of the correct genes for a given cell type. Pol II transcription can be divided into three phases: initiation, elongation and termination. The minimal machinery required for transcription initiation consists of the pol II complex and five general transcription factors (GTFs) (TFIIB, - D, -E, -F and –H) (Liu *et al*, 2013). The GTFs and pol II bind sequentially to the core promoter of genes, starting with TFIID, to form the preinitiation complex (PIC). TFIIH unwinds the DNA helix at the transcription start site (TSS) to allow access of pol II to single stranded DNA in order to start RNA synthesis (Smolle & Workman, 2013). TFIIH phosphorylates ser5 of the C-terminal domain (CTD) of RBP1, the largest subunit of pol II, as pol II escapes the promoter. At this point pol II dissociates from the GTFs and enters into early elongation after which it acquires CTD ser2 phosphorylation (Smolle & Workman, 2013; Liu *et al*, 2013). The phosphorylated CTD recruits factors for efficient elongation, termination and mRNA processing. The RNA receives a 7-methylguanosyl cap in early elongation, and is polyadenylated at the 3’ end as transcription is terminated.

The core promoter is a minimal set of regulatory DNA-elements required for directing pol II transcription, often containing a TATA-box, and is located immediately upstream of the TSS (Butler & Kadonaga, 2002). However, pol II transcribed genes are associated with multiple regulatory DNA elements in addition to the core promoter. These cis-regulatory DNA sequences include the proximal promoter and enhancers, which contain binding sites for specific transcription factors (Butler & Kadonaga, 2002). While only the PIC is required for basal transcription, initiation can be greatly enhanced or repressed by the binding of specific transcription factors. Normally transcriptional activators bind to regulatory DNA-sequences and recruit co-activators and the transcriptional machinery, leading to formation of the PIC (Weake & Workman, 2010).

1.7.2 Epigenetics and transcription

Chromatin has a major influence on transcription by for example affecting regulation of initiation, posing as an obstacle to elongation and preventing cryptic transcription.
As mentioned in the previous section, transcriptional activators initiate gene induction by binding to regulatory sites and recruiting coactivator complexes that act on chromatin, including histone-modifying enzymes such as KATs and CRCs. These enzymes contribute to an accessible chromatin format facilitating initiation, but also to recruitment of further effector proteins recognizing histone modifications (Weake & Workman, 2010; Smolle & Workman, 2013). Correspondingly, repressors recruit chromatin modifiers such as HDACs and CRCs, leading to a closed and less permissive chromatin conformation, negatively affecting transcription.

Elongation requires the polymerase to negotiate nucleosomal DNA. This is accomplished through the activities of CRCs and histone chaperones proteins. CRCs evict and remodel nucleosomes in front of the elongating polymerase, and reassemble nucleosomes in its wake with the help of histone chaperones that accept the evicted histones and escort them to reassembly (Clapier & Cairns, 2009). This is not only important for efficient elongation, but also to prevent transcription from cryptic start sites behind pol II (Li *et al*, 2007; Pointner *et al*, 2012).

Histone modification patterns are dramatically different in promoters and in bodies of transcribed genes. For example, while histone acetylation is highest in the promoter regions of active genes, H3K36me3 is enriched in the body of transcribed genes (Smolle & Workman, 2013; Barski *et al*, 2007). Studies in yeast have shown that the H3K36 KMT Set2p is recruited to transcribed genes by association with the phosphorylated CTD of pol II and H3K36 methylation prevents cryptic transcription by recruitment of CRCs and HDACs (Carrozza *et al*, 2005; Smolle & Workman, 2013).

1.7.3 Enhancers

Enhancers are cis-regulatory elements that can positively regulate transcription from a cognate promoter over substantial genomic distances of up to thousands of kilobases (Calo & Wysocka, 2013). Enhancers are approximately 200-500 bp long and contain clusters of recognition sequences for DNA-binding proteins and thereby act as binding

platforms for transcription factors (Calo & Wysocka, 2013). They are believed to favor transcription through DNA-looping, bringing enhancer and target promoter into close proximity and allowing contact between the proximal promoter and enhancer-bound transcription factors as well as delivery of accessory factors needed for transcription (Calo & Wysocka, 2013; Smallwood & Ren, 2013). Studies and estimates suggest that genes are often regulated by multiple enhancers. As different combinations of transcription factors may bind different enhancers this would allow for complex transcription patterns. Correspondingly genes that need to be coordinately activated under certain circumstances may share similar enhancers. This confers an additional layer of control and flexibility to transcriptional regulation (Cho, 2012).

A major breakthrough in studying enhancer biology was the realization that these regions are associated with certain chromatin features. Among these, the most well known are the presence of DNA hypersensitive sites, binding of the KAT p300 (KAT3B) and the CRC ATPase SMARCA4 as well as enrichment of H3K4me1 but lack of H3K4me3 (Thurman *et al*, 2012; Heintzman *et al*, 2007; Rada-Iglesias *et al*, 2011). In addition, epigenetic marks can be used to separate active from inactive enhancers. Active enhancers are associated with H3K27ac while enhancers that are in a poised state may have H3K27me3 (Creyghton *et al*, 2010; Rada-Iglesias *et al*, 2011).

The discovery of enhancer chromatin signatures inspired whole genome experiments aiming to identify enhancers based on these characteristics. The ENCODE project identified 400.000 putative enhancers, with current estimates predicting the real number to be around one million (The ENCODE Project Consortium, 2012; Smallwood & Ren, 2013). Although the functionality of most of these sites remains to be tested, experiments support the predictive power of epigenetic features to identify enhancers (May *et al*, 2012; Visel *et al*, 2009; Calo & Wysocka, 2013).

In 2013 several studies reported the finding of "super-enhancers"; clusters of multiple enhancers bound by high levels of mediator and master transcription factors (Whyte *et al*, 2013; Lovén *et al*, 2013; Hnisz *et al*, 2013). Super-enhancers are generally associated with genes for cell type specific transcription factors or other factors with cell type specific functions and are therefore likely involved in regulation of cell fate.

1.8 HEMATOPOIESIS: LINEAGES AND CELLS

Hematopoiesis is the process by which all blood cells develop from a common hematopoietic stem cell (HSC) pool (figure 2). The descendants of the multipotent HSC pass through increasing degrees of restriction, finally giving rise to thrombocytes, erythrocytes and white blood cells of both the innate and the acquired immunesystem (Orkin, 2000). Postnatal hematopoiesis is located in the bone marrow. HSCs are rare and divide infrequently, but because of the high proliferation rates in later stages of blood development a HSC is capable of producing $1*10^6$ mature cells after only 20 rounds of proliferation (Hoffbrand & Moss, 2011). Approximately $1*10^{10}$ blood cells are formed each day, with potential for increased production if needed (Hoffbrand & Moss, 2011) Differentiation proceeds from the HSC to the lineage committed common myeloid progenitor (CMP) or common lymphoid progenitor (CLP) (Orkin, 2000). Further lymphoid development gives rise to T cells, B cells and NK-cells, whereas the

myeloid lineage includes the mast cells, monocyte/macrophages and granulocytes such as neutrophils, basophils and eosinophils. The megakaryocyte-erythrocyte progenitor (MEP), from which erythrocytes and thrombocytes develop, also stem from CMP (Iwasaki & Akashi, 2007). Dendritic cells, on the other hand, can have either myeloid or lymphoid origin.

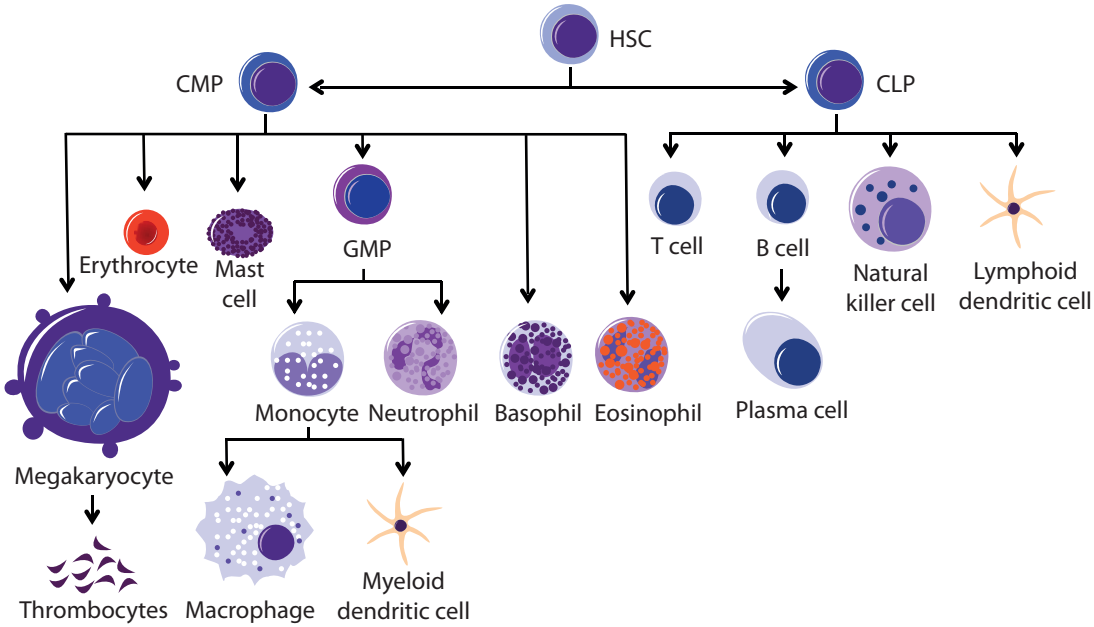

Figure 2 Hematopoeisis

Schematic illustration of hematopoietic development and mature blood cell types. All blood cells descend from a common hematopoietic stem cell (HSC) and differentiation progresses through stages of increasing restriction. The common myeloid progenitor (CMP) gives rise to cells of the myeloid lineage, whereas the lymphoid cells develop from the common lymphoid progenitor (CLP). The granulocyte/macrophage progenitor (GMP) is bipotent and develops into monocytes or neutrophils. Several intermediary cell stages and progenitors have been omitted for clarity.

Lineage choice may depend on chance or external signals in the form of growth factors. These are mainly produced by bone marrow stromal cells and stimulate self renewal and multipotency of stem cells, as well as proliferation and differentiation. The signal is transmitted into the nucleus and the transcriptional program by transcription factors, the combination and levels of which control the differentiation process. For example, the transcription factor GATA-2 regulates HSC survival (Tsai & Orkin, 1997), PU.1 and the C/EBP family are involved in myeloid commitment (Iwasaki & Akashi, 2007) whereas GATA-1 is required for erythroid and megakaryocytic development (Orkin *et al*, 1998; Fujiwara *et al*, 1996) and IKAROS (IKF1) is essential for lymphoid differentiation (Wang *et al*, 1996a) although many of these are important for multiple lineages.

Hematopoietic cell populations can be distinguished and purified based on the expression of surface proteins. A wide range of well characterized surface markers have been described allowing isolation of quite specific cell types. For example, HSCs

and early progenitors, but not more mature cell types, express CD34. As the HSC develop into CMP or CLP progenitor cells, they start expressing CD38 in addition to CD34, whereas HSCs are CD38 negative.

1.8.1 Granulopoiesis and neutrophils

The white blood cells of the myeloid lineage, i.e. monocytes and the granulocytes (neutrophils, eosinophils and basophils) are phagocytes. As parts of the innate immune system these cells constitute our first line of defence against pathogens. Of the granulocytes, the neutrophils are by far the most abundant with blood counts reaching $1*10^{10}$ cells/L (60-70% of leukocytes in blood), and a production rate of $1\text{-}2*10^{11}$ per day in an adult human (Hoffbrand & Moss, 2011; Borregaard, 2010). The cytoplasm of the mature neutrophil contains granules and the nucleus is characteristically polymorphic with two or more lobes.

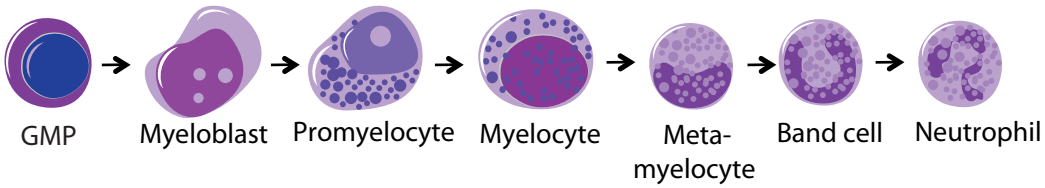

Figure 3 Overview of granulopoiesis

Among the first differentiation stages after granulocytic commitment is the highly proliferative myeloblast. The characteristic granules start forming in the promyelocyte stage and cells seize proliferation as myelocytes. Late granulopoiesis is accompanied by changes in nuclear morphology to the trade-mark multilobed shape in mature bone marrow neutrophils.

Neutrophil differentiation is regulated by several transcription factors at different stages. For example, granulocytes and monocytes develop from a common progenitor (GMP). The differentiation choice after this stage involves the balance of transcription factors C/EBPα and PU.1, both required for granulopoiesis (Iwasaki & Akashi, 2007). High PU.1 expression pushes cells toward monocytic differentiation whereas C/EBPα and somewhat lower PU.1 levels promote granulocyte development (Mak *et al*, 2011). The repressor GFI-1 is also required for granulopoiesis and is involved in the same transcriptional network (Karsunky *et al*, 2002).

Cellular maturation progresses through a series of cell stages, where the early and intermediate stages are highly proliferative until the myelocyte stage, after which the precursor cells are postmitotic (figure 3). There is a large cache of neutrophils in the bone marrow, but after release into the blood stream the cells circulate for only 6-10 hours before migration out into the surrounding tissues, where they normally survive 4-5 days (Hoffbrand & Moss, 2011). Granule synthesis begins at the promyelocyte stage and continues throughout differentiation (Borregaard, 2010). The granules act as stores of adhesion molecules, antimicrobial peptides, proteolytic enzymes and other defensive factors. There are several different types of granules with specific function, formed sequentially during differentiation and distinguished by the presence of distinct components (Gullberg *et al*, 1997).

Neutrophils converge on sites of inflammation by chemotaxis and internalize microbes by phagocytosis. The microbes are subsequently killed by fusion of the phagosome with proteolytic granules and production of reactive oxygen species. Granules can also be released into the extracellular environment to combat pathogens. In addition, neutrophils produce chemokines that aid in attracting more immune cells to the site of inflammation (Borregaard, 2010; Mocsai, 2013). The important role of neutrophils in inflammation and defence against microorganisms is illustrated by the severe effects of neutrophil deficiencies. Neutropenia is associated with recurrent infections that may become life threatening without treatment (Lakshman & Finn, 2001).

1.9 EPIGENETIC MECHANISMS IN HEMATOPOIESIS

Epigenetic mechanisms and chromatin modifications have proven to be highly important in development and control of differentiation. In fact, epigenetic states have been proposed to both stabilize cellular identity as well as drive differentiation. To emphasize the importance of this regulation in normal development, abnormal epigenetic patterns have been associated with a number of malignancies. This section will provide an overview of the role of chromatin modifications in normal hematopoiesis, and also discuss how epigenetic deregulation contributes to hematopoietic malignancies.

1.9.1 DNA methylation in hematopoiesis

1.9.1.1 DNMTs in hematopoiesis

While DNA methylation was long considered to be a relatively fixed and static chromatin mark after embryonic development, it has in recent years been found to be surprisingly plastic even in adult differentiation (Meissner *et al*, 2008). Both DNMT3A and DNMT1 have been shown to have vital, but very different, roles in hematopoiesis, while the role of DNMT3B has been less investigated (Bröske *et al*, 2009; Challen *et al*, 2012; Trowbridge *et al*, 2009). Experiments using a mouse model with a tissue specific cre-lox knockout system of *Dnmt3a* in hematopoietic cells, demonstrated that DNMT3A is necessary for differentiation of HSC, but not for lineage choice (Challen *et al*, 2012). Upon reimplantation of DNMT3A-/- HSC into irradiated recipients, there was an accumulation of HSC. The effect was caused by reduced differentiation potential, supported by the observation that stem cell maintenance genes were upregulated while lineage differentiation genes were downregulted. However, when differentiation proceeded, no significant effect was observed on lineage choice, except for a slight skewing to B cell differentiation.

In contrast, DNMT1 is required for maintenance of HSC and highly important for lineage specification (Bröske *et al*, 2009; Trowbridge *et al*, 2009). One study observed rapid death of mice due to bone marrow failure after conditional knockout of *Dnmt1* in hematopoietic cells (Bröske *et al*, 2009). By instead using an inducible hypomorphic *Dnmt1* system in a repopulation assay, it could be concluded that HSC with hypomorphic *Dnmt1* had reduced self-renewal capacity likely due to the downregulation of stem cell maintenance genes. There was also a marked skewing of differentiation toward the myeloid lineage, whereas lymphoid commitment and B cell

development was blocked. Similarly, myeloerythroid transcription factors were upregulated in *Dnmt1* hypomorphic mice, whereas lymphoid regulators were downregulated. These results clearly indicate that DNMT1, and maintenance of methylation, is important for lineage choice.

Interestingly, the *Dnmt1* hypomorphic mice were protected against induction of myeloid leukemia using the oncogenic *Mll-AF9* fusion gene, possibly due to the lack of self-renewal capability and forced myeloid differentiation. This could also provide the reason behind the successful treatment of myeloid malignancies with the DNA methylation inhibitor 5-aza-cytidine (Kaminskas, 2005). Indeed, 5-aza-cytidine treatment of HSC gave similar results as *Dnmt1* hypomorphic mutation (Bröske *et al*, 2009), and the related inhibitor 5-aza-2'deoxycytidine causes increased numbers of myeloid progenitors at the expense of lymphoid progenitors (Ji *et al*, 2010).

1.9.1.2 Genome-wide methylation changes in hematopoiesis

In agreement with the observation that maintained methylation is required for lymphoid, but not for myeloid commitment, several studies have described loss of methylation in myeloid and erythroid cells during differentiation (Hogart *et al*, 2012; Shearstone *et al*, 2011; Bock *et al*, 2012; Ji *et al*, 2010; Hodges *et al*, 2011; Bocker *et al*, 2011). By contrast, lymphoid cells appear to show a net gain in methylation. For example, out of neutrophils, B cells and HSCs, neutrophils have the highest number of hypermethylated regions, while B cells have the fewest (Hodges *et al*, 2011). Importantly, this difference is already evident in the committed progenitors, as CLP has more sites of increased methylation than CMP (Bock *et al*, 2012).

In line with the suggested regulatory role of DNA methylation in lineage specification, methylation changes are often associated with genes important for hematopoietic control, such as transcription factors, and genes involved in functions of the mature cells (Ji *et al*, 2010; Hodges *et al*, 2011; Bock *et al*, 2012; Bocker *et al*, 2011). The cell specific genes become unmethylated in the appropriate lineage and methylated in other cells. Methylation changes have also been described in binding sites for hematopoietic transcription factors and putative enhancers (Hodges *et al*, 2011; Lee *et al*, 2012; Bock *et al*, 2012; Schmidl *et al*, 2009; Deaton *et al*, 2011). In addition, myeloid transcription factors and their binding sites are specifically methylated in lymphoid cells. These observations support a role for cell type specific DNA methylation both in lineage specification and in safeguarding against activation of a myeloid transcription program in the lymphoid lineage. Interestingly, one study reported that differentially methylated regions (DMRs) between mature myeloid (neutrophils) and lymphoid (B cells) cells displayed intermediate methylation levels in earlier stem/progenitor cells, possibly in preparation for either outcome (Hodges *et al*, 2011).

It should be noted that two studies, one performed on material from human fetal bone marrow on B cell development and one comparing CD34+ progenitor cells with T cells, reported that lymphoid commitment is associated with general demethylation in contrast to the increased methylation reported in other studies (Lee *et al*, 2012; Schmidl *et al*, 2009). The reason for this discrepancy is unclear, but could be caused by methodological differences.

In addition to early lineage choices, DNA methylation also appears to be involved in later hematopoietic development as even closely related cell types, such as human conventional and regulatory T cells, display regions of differential methylation (Schmidl *et al*, 2009). Again, DMRs were associated with cell type specific genes and commonly located in promoter distal sites with methylation sensitive enhancer activity. Methylation changes are generally smaller between more differentiated cell types than earlier in development, and differ more between distant tissues (Deaton *et al*, 2011; Lee *et al*, 2012; Bock *et al*, 2012; Bocker *et al*, 2011).

Studies comparing samples from individuals of different ages indicate a predominant DNA hypomethylation effect of ageing in both human HSC (Bocker *et al*, 2011) and T cells (Heyn *et al*, 2012). The T cell study found that the hypomethylated genes were enriched for genes that are differentially expressed in ageing HSC. The genes found to be hypomethylated as a consequence of age in HSC, on the other hand, overlapped significantly with myeloid specific genes that are demethylated during myelopoiesis. Intriguingly, age is also associated with skewing of hematopoiesis toward the myeloid lineage (Pang *et al*, 2011).

Although DNA methylation studies have historically been focused on CGIs in promoters, several of the studies mentioned above, as well as studies in other tissues, report that the majority of methylation changes in normal development occur outside of CGIs (Ji *et al*, 2010; Lee *et al*, 2012; Irizarry *et al*, 2009; Hogart *et al*, 2012). Data suggests that changes are particularly common in CGI-adjacent regions, called shores, and that these changes may be better correlated with changes in expression than changes in the CGIs themselves (Ji *et al*, 2010; Irizarry *et al*, 2009). Interestingly many cell type specific differentially methylated sites (DMSs), including differences in CGIs, are not in immediate proximity to the promoter or TSS, but instead located in promoter distal elements, as already mentioned, or in gene bodies (Lee *et al*, 2012; Hodges *et al*, 2011; Schmidl *et al*, 2009; Deaton *et al*, 2011).

1.9.2 Histone modifications in hematopoiesis

Several studies have shown that, like DNA methylation, histone modifications may be involved in regulating plasticity and differentiation. Bivalent domains (see section 1.3.2) appear to participate in regulating genes related to hematopoietic control. Many HSC and progenitor-specific genes are associated with bivalent domains already in ESCs but lose the repressive H3K27me3 in early blood cells (Adli *et al*, 2010; Abraham *et al*, 2013). Later in differentiation they become associated with K27me3 and silenced again (Abraham *et al*, 2013). Notably, many of these genes have important roles in differentiation including critical transcription factors (Abraham *et al*, 2013; Adli *et al*, 2010). In HSCs bivalent domains are present at lineage specifying genes. Strikingly, the level of H3K4me3 at HSC bivalent promoters reflects the number of differentiated cells the gene will be transcribed in and is accordingly high at master regulators of blood lineages (Adli *et al*, 2010).

In general HSC and multipotent progenitors have a more permissive chromatin state, allowing access to genes of multiple lineages. In line with this, hematopoietic

commitment is preceded by low or background transcription of genes belonging to several lineages (Hu *et al*, 1997; Miyamoto *et al*, 2002).

Comparison of global H3 and H4 acetylation levels in CD34+ cells *vs.* CD34- cells indicates that HSC and progenitor cells have higher levels of acetylation than committed cells (Chung *et al*, 2009). In addition, acetylation turnover was higher in more immature cells indicting more dynamic and open chromatin. In general, more promoters are associated with bivalent or active marks in HSCs than in committed or differentiated blood cells (Cui *et al*, 2009; Weishaupt *et al*, 2010; Abraham *et al*, 2013). In accordance with bivalent genes being poised for transcription in later development, active and bivalent chromatin marks in HSC have been detected at promoters of lineage specific genes, despite the genes being silent or lowly expressed (Maes *et al*, 2008; Orford *et al*, 2008; Weishaupt *et al*, 2010; Abraham *et al*, 2013; Hattangadi *et al*, 2011). Interestingly, bivalent marks of cell specific genes may persist in closely related cell types, although they are lost in distant cell types, possibly indicating some residual plasticity (Abraham *et al*, 2013).

A study on umbilical cord blood focusing on a panel of 20 lymphoid and myeloerythroid genes reported that HSCs had a broader acetylation profile with genes of multiple lineages acetylated, whereas lineage committed cells only displayed acetylation at genes appropriate for that lineage (Maes *et al*, 2008). Interestingly several lymphoid genes showed H3K4me2 in HSCs, and during B cell development H3K4me2 was replaced by H3K4me3 at B cell specific genes while methylation was lost at these genes in T cells. The authors suggested that H3K4me2 may be an additional indicator of a poised state. A similar observation was made in murine erythroid differentiation where a group of silent promoters in HSC were characterized by H3K4me2, but absence of H3K4me3 (Orford *et al*, 2008). This subset was enriched for lineage specific genes and upon erythroid differentiation, non-erythroid genes lost H3K4me2, while genes that gained H3K4me3 showed increased expression. The authors concluded that the wider H3K4me2 profile of HSCs reflected the differentiation potential.

Collectively, available data demonstrates that genes relevant to hematopoietic differentiation or function of the mature blood cell types are associated with bivalent or permissive chromatin states in HSCs and progenitor cells, priming them for expression. Therefore, the chromatin state of hematopoietic genes can be said to define the plasticity of HSCs.

1.9.3 Chromatin modifying factors in hematopoiesis

As expected from the global changes of epigenetic marks in blood cell differentiation, many individual chromatin regulators have been shown to have important roles in hematopoiesis. In addition to the requirement for DNMT3A and DNMT1 in hematopoiesis (discussed in section 1.9.1.1), the putative DNA demethylase TET2 is also required for normal blood development. TET2 and TET3 are both expressed in peripheral blood, whereas TET2 is the only TET enzyme in bone marrow (Lorsbach *et al*, 2003). TET2 deficient mice show an expansion of the HSC population with decreased differentiation potential that favor monocytic development (Butler & Dent,

2013). The TET2 deficient mice also develop leukemia at an early age (Butler & Dent, 2013).

Several histone modifying enzymes, or subunits of histone modifying complexes, have been demonstrated to be important for hematopoiesis. For example, the KAT MOZ (KAT6A) is required for maintenance of the HSC and progenitor pool, but not for differentiation (Thomas, 2006), and overexpression of HDAC1 inhibits myeloid differentiation (Wada *et al*, 2009).

The H3K4 methyltransferase MLL is highly relevant in the hematopoietic system as indicated by its frequent involvement in hematological malignancies (discussed in section 1.9.4.1) (Chung *et al*, 2012). MLL is required for HSC self-renewal, and mice with induced *Mll* deletion in the bone marrow die after three weeks from bone marrow failure and HSC depletion (Jude *et al*, 2007). In contrast, committed progenitors do not require MLL (Jude *et al*, 2007).

Several PcG factors of both PRC1 and PRC2 (see section 1.3.2) are known to have roles in normal hematopoiesis. BMI1, a subunit of PRC1, is important at several stages of blood development. BMI1 deficiency in mice causes increased B cell development due to premature derepression of regulators of B cell development, the promoters of which are normally marked with bivalent domains in HSCs (Oguro *et al*, 2010). *Bmi1* mutation also causes HSC depletion in mice (Park *et al*, 2003). Correspondingly, overexpression of BMI1 or the PRC2 KMT subunit EZH2 increases HSC self-renewal (Iwama *et al*, 2004; Kamminga, 2006). In addition, EZH2 overexpression also leads to abnormal myeloid expansion in mice (Herrera-Merchan *et al*, 2012), while *Ezh2* mutation causes deficiencies in early B cell development (Su *et al*, 2002; Mochizuki-Kashio *et al*, 2011).

LSD1 regulates lineage specific hematopoietic genes through its interaction with the repressor GFI1 (Saleque *et al*, 2007). *Lsd1* knockdown in mice causes expansion of myeloerythroid progenitors by enhancing proliferation, while inhibiting terminal differentiation of granulocytes, megakaryocytes and erythrocytes but stimulating monocyte development (Sprüssel *et al*, 2012).

Several CRCs have critical roles in blood development. Deletion of the gene encoding the ATPase CHD4 of the NuRD complex, generally associated with gene repressive functions, in bone marrow of mice causes expansion of the HSC pool and accumulation of erythroid progenitors (Yoshida *et al*, 2008). In contrast, lymphoid and other myeloid progenitors are reduced. CHD4 is also required for several steps of T cell development (Williams *et al*, 2004). Likewise, the BAF complex has been demonstrated to be essential in hematopoiesis. A recent study revealed that the BAF53a (ACTL6A) subunit of the BAF complex is required for HSC self-renewal and proliferation of myeloid progenitors. In line with this, BAF53A depletion in mice resulted in bone marrow failure (Krasteva *et al*, 2012). In addition, expression of a dominant negative version of SMARCA4, one of the two possible ATPase cores of BAF, causes a block in myeloid maturation in culture (Vradii *et al*, 2006). SMARCA4 is also required for normal T cell development (Chi *et al*, 2003).

1.9.4 Epigenetics in hematopoietic malignancies

Approximately 7% of all cancer cases are hematopoietic malignancies. As in all malignancies, these are the result of accumulation of mutations that may be influenced by inherited and environmental factors, causing a clonal expansion of transformed cells. There are numerous kinds on hematological malignancies. The uncontrolled proliferation can occur either in the bone marrow, as with leukemias, which are characterized by expansion of white blood cells, or in the peripheral tissue as in lymphomas where the malignancy starts in the lymphoid tissue. Subdivisions further group malignancies according to for example the cell/lineage of origin and if it has acute or chronic progression. In addition, there are other proliferative diseases of the blood that may in time develop into malignancies. One such disease is myelodysplastic syndrome (MDS) which progresses to acute myeloid leukemia (AML), the most common acute leukemia in adults.

Perhaps the earliest evidences of the pivotal role of epigenetics in hematopoietic regulation came from the involvement of chromatin modifiers in blood malignancies. In this section some examples of the involvement of epigenetic mechanisms in hematopoietic malignancies will be presented, with emphasis on AML and DNA methylation.

1.9.4.1 DNA methylation in hematopoietic malignancies

Aberrant DNA methylation is one of the hallmarks of cancer, and typically involves general hypomethylation causing destabilization of the genome and local hypermethylation of CpG islands, causing silencing of tumor suppressor genes (Sandoval & Esteller, 2012). Changes in DNA methylation are recognized as important contributors to many hematopoietic malignancies (Cahill & Rosenquist, 2013; Schoofs & Müller-Tidow, 2011; Issa, 2013). *DNMT3A* is recurrently mutated in AML and MDS causing reduced or abolished enzymatic activity and *DNMT3A* mutations are associated with poor prognosis (Ley *et al*, 2010; Walter *et al*, 2011; Yan *et al*, 2011).

TET2 mutations are common in various myeloid malignancies (Chan & Majeti, 2013). Also, mice without TET2 are prone to develop hematopoietic malignancies, especially of the myeloid lineage (Butler & Dent, 2013). One study showed that *TET2* mutations in patients with myeloid malignancies correlated with decreased 5hmC levels and increased DNA methylation (Ko *et al*, 2010). Gain-of-function mutations in the IDH1 and IDH2 enzymes are also common in AML, but mutually exclusive with TET2 mutations, and are associated with hypermethylation of leukemic cells (Figueroa *et al*, 2010; Chan & Majeti, 2013). These mutations cause the enzymes to convert their normal product α-KG to 2-hydroxyglutarate (2-HG), which is structurally similar to α-KG (Figueroa *et al*, 2010). TET2 requires α -KG as a cofactor and is inhibited by 2-HG, accounting for the similarity between TET2 and IDH caused AMLs (Figueroa *et al*, 2010).

The role of DNA demethylation related factors in myeloid malignancies is striking considering the decreases of DNA methylation associated with normal myeloid development, and how *Dnmt1* mutation hampers induction of AML. It is tempting to

speculate that the possible DNA demethylation deficiency caused by *TET2*/*IDH* mutations in AML blocks differentiation and is the mechanism behind leukemic transformation. However, it is more difficult to reconcile the apparent oncogenic effect of the *TET2/IDH* associated demethylation deficiency with the oncogenic effect of *DNMT3A* mutations, as they should intuitively have opposite effects. Still, *TET2*/*IDH* and *DNMT3A* mutations, are not only common in the same type of leukemias, but have also been found in the same clones (Welch *et al*, 2012; Patel *et al*, 2012). One possible explanation may be some level of locus specificity that influences the final outcome. Indeed, although *DNMT3A* mutation causes no global change in DNA methylation in HSCs, locus specific hypomethylation was observed at genes frequently overexpressed in leukemias (Challen *et al*, 2012). Intriguingly, many of the detected DMRs showed increased methylation, especially in CGIs. Also, a recent study describes expansive regions of hypomethylation, called canyons, the borders of which are maintained by DNMT3A (Jeong *et al*, 2014). In HSC these regions were enriched for TF binding sites and genes deregulated in leukemia.

Many studies have reported changes in DNA methylation in AML, both associated with and independent of the cytogenetic subtype of the patient (Schoofs & Müller-Tidow, 2011). Several studies also find correlations between methylation profiles and clinical outcome (Schoofs & Müller-Tidow, 2011). However, the precise methylation changes described sometimes differ, possibly due to the different methodologies or bioinformatics analysis employed. For example, MeDIP-seq (see section 2.6) on leukemic cells from 12 AML patients, including different cytogenetic subtypes, found no net change in DNA methylation compared to normal bone marrow (Saied *et al*, 2012). However, several site specific alterations were located, including at known leukemia-associated genes. A larger study of 344 AML patients using a microarray-based approach showed that cytogenetic subtype was the strongest factor for determining the DNA methylation profile (Figueroa *et al*, 2010). Still, many genes were consistently altered in AML irrespective of cytogenetic subtype. Notably, the study also identified a set of 15 genes, the methylation status of which could predict overall survival. Similarly, a study on normal karyotype AML showed that some mutations, including *IDH* mutations, were predictors of the DNA methylation profile of AML cells (Deneberg *et al*, 2011). Interestingly this study found a correlation between positive clinical outcome and the degree of methylation of PcG target genes. PcG targets in general were hypermethylated in AML and higher methylation was associated with disease-free survival.

An extensive study produced by The Cancer Genome Atlas carefully examined samples from 200 AML patients using the Infinium 450k array (see section 2.6.1) as well as whole genome sequencing, expression profiling by microarrays and RNA-seq (The Cancer Genome Atlas Research Network, 2013). A large proportion of patients had mutations in DNA methylation related genes or chromatin modifiers, 44% and 30% respectively, making epigenetic genes one of the most common functional groups to be mutated in AML. In accordance with the results presented by Figueroa *et al* (2010), the DNA methylation patterns largely reflected the underlying genotype, and depending on the mutations present, both hyper- and hypomethylation phenotypes were observed. Importantly, changes were most common in CpG poor regions where many methods of DNA methylation analysis used in earlier studies lack coverage.

Finally, the importance of DNA methylation in myeloproliferative disorders is highlighted by the efficacy of treatment with drugs targeting the methylation machinery. In fact, the DNMT inhibitor 5-aza-cytidine has been approved by the FDA for treatment of MDS and some types of AML (Kaminskas, 2005).

1.9.4.2 Mutations in Chromatin regulators

In line with their proposed roles in control of normal hematopoiesis, many of the factors discussed in section 1.9.3 have been found to be mutated in blood malignancies. For example, mutations of PRC2 components have been described in several myeloid malignancies. Interestingly, while loss-of-function of EZH2 has been reported for myeloid leukemias, gain-of-function mutations have been described in lymphoid malignancies (Butler & Dent, 2013). This distinction is in line with the effect of *Ezh2* mutation on myeloid and lymphoid development (See section 1.9.3).

HDACs are other examples of chromatin modifiers that are highly relevant in blood malignancies as they commonly associate with oncogenic fusion proteins to cause aberrant repression of genes promoting differentiation (Glozak & Seto, 2007). HDAC inhibitors show promising results in treatment of myeloproliferative disorders, especially in combination therapy with e.g. DNA methylation inhibitors (Quintás-Cardama *et al*, 2011). Furthermore, early experiments with inhibitors of bromodomains also show encouraging results against AML and mixed lineage leukemia (MLL) cells, further emphasizing the role of histone acetylation in blood malignancies (Zuber *et al*, 2011).

1.9.4.3 Leukemic fusion proteins

In comparison to other cancers, hematological malignancies have a high incidence of chromosomal translocations producing oncogenic fusion proteins. Strikingly, one partner of these chimeras is often an epigenetic factor. Acute leukemias with translocations involving the *MLL* gene have been called the prototypical epigenetically driven cancer (Neff & Armstrong, 2013). *MLL* translocations are found in 5-10% of AML cases and more than 70% of childhood acute lymphoid leukemia (ALL) cases (Neff & Armstrong, 2013). The precise mechanism of transformation is not fully understood. MLL-fusion proteins appear to bind and upregulate inappropriate genes by causing aberrant H3K4me and in some cases H3K79 methylation through recruitment of the KMT DOT1 (Okada *et al*, 2005; Neff & Armstrong, 2013; Krivtsov *et al*, 2008). In fact, one study indicated that DOT1 may be a promising target for future therapy development against MLL leukemias (Daigle *et al*, 2011).

Many other translocations involving epigenetic regulators have been reported in blood malignancies, such as for example the KATs MOZ and MORF (KAT6B) fused to the KAT CBP (KAT3A) (Rozman *et al*, 2004; Panagopoulos *et al*, 2001) and the KMT NSD3 fused to the nuclear pore protein NUP98 in AML (Rosati, 2002). Even non-epigenetic fusion proteins often employ chromatin regulators in their mode of transformation. Fusion proteins involving a transcription factor as one partner, such as PML-RARα and AML-ETO, recruit chromatin-modifying enzymes to inappropriate targets. For example PML-RARα, found in the majority of acute promyelocytic leukemia (a subtype of AML) cases, recruits HDACs to genes important for

differentiation, causing transcription silencing and differentiation arrest (Rice *et al*, 2007).

In summary, the transforming abilities and high incidence of epigenetic abnormalities in hematopoietic malignancies, as well as the efficacy of drugs targeting these pathways in the clinic, clearly state the importance of chromatin regulators in blood development.

2 METHODS

2.1 INTRODUCTION TO MICROARRAYS

Microarray technology represents a relatively cheap and efficient method of assessing the levels of thousands of specific nucleic acid sequences in a sample of DNA or RNA. This methodology has been adapted to a wide range of applications and makes use of complementary sequences, usually of DNA, immobilized on a solid surface in a location that is predetermined (spotted arrays) or can be assessed (bead arrays). On bead arrays, each bead is covered in multiple copies of the same probe, which also contains a sequence for identification. After application of the beads to a surface, the location of each bead-probe can be determined.

For most applications and array platforms the query sample is labelled, usually by flourophores or biotin for subsequent flourophore binding, and hybridized to the array. Scanning produces intensity values for each spot/bead corresponding to the amount of the complementary sequence in the sample. Microarray analysis requires careful consideration of data processing, including normalization methods and quality control assessment, but provides large amounts of data.

2.2 EXPRESSION PROFILING

In the work included in this thesis, two genome-wide methods for expression profiling have been used: expression microarrays and cap analysis gene expression (CAGE). As discussed below these methods are highly informative and have complementary uses.

2.2.1 Expression microarrays

Expression profiling using microarrays is a widely used application of microarray technology. Most commonly, the purified RNA sample is reverse transcribed into cDNA that is labelled and hybridized to a microarray. While there are many platforms for this application, they usually contain one or a few probes per gene, thereby yielding quantitative information for full-length transcripts, but usually not about TSS. Still, expression arrays provide transcription rate data for most known genes at a relatively low cost and require less computational analysis than CAGE or RNA sequencing.

2.2.2 CAGE

In paper II and III transcriptome data generated by CAGE as part of the FANTOM 5 consortium was used. This method is based on the purification and 5'-end sequencing of capped transcripts yielding information on both transcription levels and the precise TSS and is highly quantitative and reproducible (Kanamori-Katayama *et al*, 2011; Shiraki *et al*, 2003). The protocol has been adapted to work with the Helicos single molecule sequencer, which circumvents the need for the PCR amplification and cloning step of the original protocol, thus avoiding the potential sequence biases these methods may bring (Kanamori-Katayama *et al*, 2011).

Briefly, RNA is purified and first strand synthesis is performed with random primers. The 5' cap is biotinylated by initial oxidation of the diol-group of the cap structure, followed by addition of biotin hydrazide causing covalent linkage of biotin to the cap (Carninci *et al*, 1996). The RNA-cDNA duplexes are treated with RNaseI to remove RNAs with incomplete cDNAs. The RNA-cDNA hybrids are purified excluding oligonucleotides shorter than 100 bp before isolation of capped RNAs with streptavidin coated beads. The cDNA is released and receives a 3' poly-A tail by terminal deoxynucleotidyl transferase. The poly-A tail binds to poly-dT primers in the Heliscope flow cell surface and individual cDNA molecules are sequenced.

Previous cloning based CAGE protocols had a read length of 20 nt but with the Helicos adapted protocol the median read length has been increased to 33 nt facilitating accurate mapping to the genome. Reads are filtered for size (20-70 nt) and alignment score and normalized to tags per million (TPM). Normally, the cutoff for detection is set at 10 TPM corresponding to approximately 5 RNA molecules per cell. Expression of a gene is determined either by summing up all TPM normalized tags over the refseq gene or, more recently, by using relative log expression (RLE) normalized data and merging overlapping tags into tag clusters representing TSSs. Clusters are TPM transformed, assigned to genes and added up to describe that gene's expression.

For RLE normalization we use the edgeR Bioconductor package for R. The RLE normalization method creates a median library by taking the geometric mean for each gene across all libraries. A scaling factor is calculated through the relationship of each gene of each library to the median library. The advantage of the RLE normalization is that it renders libraries of different sizes comparable. Only performing a TPM normalization allows for example very highly expressed genes, such as hemoglobin in erythrocytes, to have too much influence on total tag count and skew the data set.

For some genes CAGE signal can be detected within exons, a phenomenon called exon painting. This is likely caused by recapping of longer transcripts (Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project, 2009) or by second strand synthesis at the reverse transcription step. However, using robust tag clusters (containing at least 10 detected molecules with the same TSS in a library) minimizes the effect of exon painting on data sets.

2.3 CHIP-ON-CHIP

In paper I histone modifications of the fission yeast *Schizosaccharomyces pombe* were mapped genome-wide using chromatin immunoprecipitation coupled to microarray analysis (ChIP-on-chip). In ChIP, antibodies specific for the protein in question are used to pull down the DNA fragments the protein is binding, after which the DNA is isolated and analyzed to determine where in the genome the protein binds.

In the protocol used for paper I, cells are fixed using formaldehyde, crosslinking DNA and proteins. After fixation, cells are lysed and the chromatin sheared by sonication, before incubation with antibodies specific for example for histone modifications. The antibody-chromatin complexes are isolated by binding to protein A-covered beads

before reversal of the crosslink and protein removal. The purified DNA can then be analyZed by a number of techniques, including, as in paper I, using microarrays.

2.4 SCHIZOSACHAROMYCES POMBE AS A MODEL SYSTEM

The fission yeast *S. pombe* is an excellent model system for epigenetic studies. It is easily and cheaply maintained and can with little effort be cultured in large quantities. *S. pombe* is also easy to genetically modify through homologous recombination, for example introducing point mutations and allowing retention of endogenous promoters when expressing tagged versions of proteins. *S. pombe* has a small, well-annotated genome, which facilitated the early development of genome-wide microarrays for chromatin and expression studies. In addition, it has human-like epigenetic features, such as the structure of the centromeres and heterochromatin. However it lacks for example DNA methylation and H3K27 methylation.

2.5 BLOOD AS A MODEL FOR DIFFERENTIATION

Understanding the mechanisms governing hematopoiesis is highly relevant to understanding leukaemia and other blood disorders, and for the future development of treatments, including regenerative medicine and cancer therapy. In addition blood constitutes a convenient system for studying differentiation since it is composed of several cell types stemming from a common HSC. Unlike most other tissues the different cell types can be readily purified with FACS (fluorescence-activated cell sorting) from bone marrow, avoiding culturing steps known to induce abnormal changes in for example methylation (Meissner *et al*, 2008; Shen, 2006).

2.6 METHODS OF ANALYZING DNA METHYLATION

Recent years have brought great advances in analyses of DNA methylation, introducing genome-wide technologies and shedding the limitations of locus-specific methods. DNA methylation analysis can be divided into three main classes: restriction enzyme digestion-based methods, affinity-based methods and bisulfite treatment-based methods. The out-put DNA from these three approaches can be analysed by microarray techniques or sequencing.

Restriction enzyme-based techniques take advantage of methylation-sensitive enzymes, causing methylated sequences to be protected against digestion. Affinity-based methods use antibodies specific for 5mC (MeDIP) or methyl-binding domains to pull down methylated regions. In bisulfite-based methods, the DNA is initially treated with bisulfite, converting unmethylated cytosines to uracil and thereby to thymine in subsequent PCR amplification steps. Some methods combine these approaches, such as reduced representation bisulfite sequencing (RRBS). This method uses DNA digestion, in this case with a methylation-insensitive enzyme with a CpG-containing recognition sequence, to enrich the sample for high CpG areas (Meissner, 2005).

As with all methods, these approaches have strengths and weaknesses. Notably, most methods only accurately assess high/moderate density CpG areas. Both restriction enzyme- and affinity-based methods lack coverage in low CpG areas. In addition,

affinity-based methods give the relative methylation degree of a region, but do not identify methylation status of specific CpGs. The best coverage is offered by whole genome bisulfite sequencing (WGBS). However, this method requires extensive sequencing, leading to high costs. This problem has been circumvented in RRSB by using initial endonuclease digestion, to enrich for high-CpG areas, offering excellent coverage of these regions, but as mentioned failing to analyze low-CpG regions.

2.6.1 The 450k array

The Infinium HumanMethylation450 beadchip, or 450k array, from Illumina offers an affordable alternative to WGBS. It covers more than 480000 cytosines in the human genome with base pair resolution. Although probing fewer CpGs than methods such as RRBS, it is not dependent on CpG density and includes sites in and near CGIs as well as in low-CpG areas (Bibikova *et al*, 2011).

The 450k platform is a bead array that requires bisulfite treated PCR-amplified DNA as input. The sample is hybridized to the array and a single base extension causes incorporation of a fluorophore. There are two types of probes on the array. The Infinium I probes, used for earlier methylation arrays, use two separate probes for each site, bridging the target CpG; one corresponding to the methylated sequence (ending with CG) and one corresponding to the unmethylated sequence (ending with CA to be complementary to CpT resulting from bisulfite conversion). The probe is used as a primer to perform incorporation of a fluorophore marked base complementary to the next base in the target DNA. For Infinium II probes, the probe sequence ends between the C and G of the target CpG, and the methylated or unmethylated state is determined by which base is incorporated. For both probe types the ratio of methylation to total signal (methylated/(methylated+unmethylated)) can be expressed as a β-value, representing the ratio of cells in which the site is methylated.

A weakness of bisulfite-based methods, such as WGBS and the 450k array, is the inability to distinguish between 5mC and 5hmC, which are both protected against conversion to uracil by bisulfite treatment. However, several methods have already been developed to circumvent this problem (Rivera & Ren, 2013).

3 RESULTS AND DISCUSSION

3.1 PAPER I: GENOME-WIDE MAPPING OF HISTONE MODIFICATIONS AND MASS SPECTROMETRY REVEAL H4 ACETYLATION BIAS AND H3K36 METHYLATION AT GENE PROMOTERS IN FISSION YEAST

In paper I we investigated patterns and coexistence of histone modifications in the fission yeast *S. pombe* to gain insight into effects of modification combinations and their relationship to gene expression. We employed two powerful and complementary techniques. Using ChIP-on-chip with antibodies specific for 12 histone modifications, we could assess the genomic distribution of marks with regard to each other and to genes. However, ChIP does not address the coexistence of histone modifications on the same histone molecule. To this end we used quantitative mass spectrometry (MS) and detected 10 acetylation sites as well as H3K36, H3K4, H3K9 and H4K20 mono-, di-, and tri-methylation. To elucidate the relationship between histone modifications and gene expression we performed transcriptome profiling using expression microarrays.

Comparing the occupancy of histone modifications in promoter regions of genes, we found that H3 acetylation marks had the best correlation with each other, as did H4 acetylation marks. In general, histone acetylation in the promoter region correlated with transcription, as expected. The correlation was stronger for H3 acetylation than for H4 acetylation, in particular for H3K18ac, H3K9ac and H3K27ac. Similar results were obtained for open reading frames (ORFs).

Clustering genes based on the pattern of histone modifications in their promoters produced 31 clusters. Gene ontology analysis revealed that 29 of these clusters were enriched for at least one functional term. Further grouping could divide the clusters into 4 groups based on similarity, where each group showed unique characteristics. For example, group I had high H3K36me2 and low H3ac, while group III had high H3K27ac and low H3K36me2. Interestingly, clusters of the same group were often linked to similar functional categories. These results indicate that, in *S. pombe,* genes with similar functions share common patterns of histone modifications in the promoter.

Clustering ORFs based on H4 acetylation indicated a preference for core-proximal lysine acetylation, where genes that had the distal H4K5ac, usually had the other sites acetylated as well, and genes that had H4K8ac had high levels of K12ac and K16ac. MS results confirmed that the acetylation bias was true also on single histones peptides. Similar observations have been made in human cells as well as mouse brain and called a H4 acetylation zip (Turner *et al*, 1989; Zhang *et al*, 2002; Tweedie-Cullen *et al*, 2012). Comparing H4 acetylation on genes with different length, we observed that shorter genes tended to have several of the H4 lysines acetylated, whereas longer genes often had only K16ac. Furthermore, genes with 3-4 H4 lysines acetylated had higher average expression than genes with only K16ac. Our results demonstrate a H4 acetylation zip in fission yeast as well as a connection between H4 acetylation in ORFs and transcription levels.

Previous studies have demonstrated that methylation of H3K4 and H3K36 affects histone acetylation in budding yeast. H3K4 methylation by the KMT Set1p recruits KATs causing H3K4 trimethylated histones to be highly acetylated (Jiang *et al*, 2007; Taverna *et al*, 2006). In contrast, Set2p associates with the phosphorylated CTD of elongating pol II and methylates H3K36, which recruits the HDAC Rpd3p causing deacetylation of the ORF (Carrozza *et al*, 2005). To test if the same crosstalk exists in fission yeast, we analysed histone acetylation levels with MS in *set1* and *set2* mutant strains. As predicted, acetylation levels on both H3 and H4 were decreased in the *set1Δ* background and increased in the *set2Δ* background, confirming that H3 methylation is involved in regulation of histone acetylation.

When comparing histone modification levels in promoter regions, the lowest correlation was observed for H3K27ac and H3K36me2. This prompted us to investigate the relationship between these two modifications on the same histone peptide with MS. The results showed a clear antagonistic relationship between H3K36me2/3 and H3K27ac. Furthermore, H3K27ac levels were elevated in the *set2Δ* background, consistent with H3K36me recruitment of HDAC activity.

Because H3K36me recruits the HDAC Rpd3p in budding yeast, we next looked at H3K27ac levels in a strain with mutated *clr6*, the fission yeast ortholog of *rpd3*. Indeed, global levels of H3K27ac were elevated in the *clr6-1* strain. To address the potential significance of the observed K27ac/K36me crosstalk at promoters, we looked at expression of the previously identified gene clusters in the *set2Δ* strain relative to wild type cells. Interestingly, whereas most clusters showed some degree of increased expression, the two most upregulated clusters belonged to group 1, which had high H3K36me2 in wild type cells. In contrast, *set1Δ* caused a general downregulation of expression. Collectively, this demonstrates crosstalk between H3K27ac and H3K36me with a potentially regulatory role at gene promoters in *S. pombe*.

In summary, in paper I we showed that H3 acetylation correlates with transcription, in particular acetylation of lysines K9, K18 and K27, and that genes with similar function share common histone modification patterns in the promoter region. H4 acetylation in *S. pombe* follows a zip model, with bias for core-proximal sites. Genes with several H4 sites acetylated in the ORF tend to have higher expression than monoacetylated genes. We also describe crosstalk between H3 methylation and histone acetylation. H3K36 methylation was found at promoters of genes and showed an antagonistic relationship with K27ac, possibly through recruitment of the HDAC Clr6.

3.2 PAPER II: HIGH-THROUGHPUT TRANSCRIPTION PROFILING IDENTIFIES PUTATIVE EPIGENETIC REGULATORS OF HEMATOPOIESIS

Many chromatin-modifying factors have lineage or cell specific functions in hematopoiesis, as discussed in section 1.9. Despite this, there are no studies with comprehensive analysis detailing expression patterns of epigenetic factors in blood cells. For dependable and comparable results across cell types, expression using the same methodology needs to be co-analysed. In paper II we took advantage of the vast number of cell types analysed by CAGE in the FANTOM 5 consortium and set out to

create a map of the expression of an extensive list of chromatin modifiers and accessory subunits.

RLE normalized CAGE data was obtained for mature blood cells as well as for stem/progenitor cells, and expression data for 199 chromatin modifiers was extracted for analysis. Clustering of the cell types and replicates, as well as PCA analysis, largely separated the myeloid from the lymphoid lineage, and also separated progenitor cells. The two dendritic subtypes, myeloid and lymphoid, clustered with their lineage of origin, indicating some level of epigenetic memory of descent.

We divided the genes into six functional groups based on their involvement in DNA methylation, histone acetylation/deacetylation, histone methylation/demethylation or chromatin remodeling. Heatmap analysis revealed several interesting patterns. Some genes are relatively evenly expressed, such as HDAC1 and HDAC2, indicating a ubiquitous role in hematopoiesis. However, many genes show differential expression between cell types. DNMT3B, for example, is only expressed in progenitor cells, in line with previous studies (Mizuno, 2001; Xie *et al*, 1999), whereas DNMT1 levels are generally higher in most lymphoid cells than in myeloid cells.

Interestingly, expression of specific BAF complex subunits was both variable and cell type specific. Out of the two possible ATPase subunits, SMARCA2 was significantly more highly expressed compared to SMARCA4 in most cell types. BAF53A (ACTL6A) was low in most cell types, but very abundant in progenitor cells, in agreement with requirement of this protein in HSCs (Krasteva *et al*, 2012). In contrast, BAF45C (DPF3) was exclusively expressed in B cells, while BAF60C (SMARCD3) was present in a subset of myeloid cells. These results point to the existence of BAF-complexes with different subunit composition in hematopoietic cells, which may be involved in control of cell type specific genes and differentiation, as observed in other tissues (see section 1.4).

Next, we used edgeR analysis to search our data set for chromatin modifiers that were differentially expressed between the myeloid and the lymphoid lineages. We found that the two histone chaperones ASF1B and JDP2 are only expressed in myeloid cells. Several other factors, such as the CRC components CHD3 and INO80D, as well as MLL, KDM2B and the bromodomain protein ATAD2 have significantly higher transcription rates in lymphoid than in myeloid cells. These factors may therefore be involved in establishing lineage choice and identity.

We also included CAGE data from 21 leukemic cell lines in our analysis. PCA results demonstrated that the leukemic cells cluster together with progenitor cells, separate from mature blood cells. However, we identified several genes that showed differential expression between malignant cells and normal blood cells, including progenitors. SMARCA4 and the PRC1 component CBX8 are both overexpressed in leukemic cells. In contrast, the chromatin remodeler CHD2 and another PRC1 factor, EPC1, are significantly more lowly expressed in leukemic cells compared to normal blood cells. CBX8 and EPC1 have previously been implicated in AML and ALL respectively (Tan *et al*, 2011a; Nakahata *et al*, 2009). CHD2 on the other had, has been found to be recurrently mutated in chronic lymphoid leukemia (Quesada *et al*, 2012).

Finally, we utilized the ability of CAGE to identify precise TSSs to search for genes with multiple TSSs that show differential use between cell types. We found that several genes had multiple active TSSs in blood cells, and in some cases the relative contribution between TSSs to total transcription differed between cells. For RBBP7, HDAC5 and CHRAC1, the dominant TSS was not the same in all cell types. For RBBP7, for example, we identified two TSSs, of which TSS1 was the preferred site in B cells, TSS2 the preferred site in myeloid dendritic cells, and some cells, such as CD8+ T cells, used both sites equally. Determining the biological consequences of alternative TSS use requires experimental validation, but it could possibly reflect differences in transcriptional regulation.

In conclusion, the results from paper II include identification of differential expression of multiple chromatin modifiers in the hematopoietic system. The results from our analysis reveal expression of epigenetic factors that is specific for cell types, lineages and leukemic cells, indicating possible roles for numerous factors in control of differentiation and lineage commitment. In addition, we identified differential use of alternative TSSs between cell types. This comprehensive expression map of chromatin modifiers in blood cells can hopefully serve as a foundation for functional characterization of the role of specific factors in hematopoiesis.

3.3 PAPER III: ANALYSIS OF THE DNA METHYLOME AND TRANSCRIPTOME IN GRANULOPOIEIS REVEAL TIMED CHANGES AND DYNAMIC ENHANCER METHYLATION

As discussed in section 1.9.1, DNA methylation is clearly important for making the choice between myeloid and lymphoid lineage commitment. In paper III we aimed to gain further insight into changes of DNA methylation patterns in granulopoiesis. We used bone marrow from healthy human donors and FACS to isolate granulopoietic cell populations at four different differentiation stages, namely CMP, GMP, PMC (promyelocytes and myelocytes) and PMN (including metamyelocytes, band cells and mature neutrophils). The methylome of these cell populations was analysed using the Infinium 450k array. We also performed transcriptome profiling using both expression arrays and CAGE. This analysis produced a genome-wide map of methylation and expression changes at high cellular resolution, ranging from the multipotent CMP to mature bone marrow neutrophil, allowing us to pinpoint the stage at which changes occur in granulopoietic differentiation.

By comparing β-values (defined as the ratio of methylation for a specific CpG site) between the four cell types, we identified 10156 differentially methylated sites (DMSs) during granulopoiesis. A majority of these sites showed decreased methylation (8973 DMSs), while a smaller set gained methylation (930 DMSs). Interestingly, a small set of sites (253 DMSs) showed changes in both directions between different stages. Plotting of DMSs between consecutive cell stages made it abundantly clear that changes were not uniform over time, but occurred at specific stages. Most of the increases were between CMP and GMP, whereas the major decrease happened between GMP and PMC. Therefore, the major methylation changes occur at points of lineage restriction, suggesting a role in control of cell fate.

Examining the genomic distribution of DMSs we found that, in agreement with previous work, changes were greatly underrepresented in CGIs (Ji *et al*, 2010). However, we did not observe any overrepresentation of DMSs in CpG island shores (0-2 kb from CGI). Instead, DMSs were enriched in the shelf (2-4 kb from CGI) and the open sea regions (>4kb from CGI).

Gene ontology analysis was performed on differentially expressed genes. As could be expected, the list of genes upregulated during granulopoietic differentiation was enriched for functional terms relating to neutrophil function and defence responses. Correspondingly, for downregulated genes the terms included functions in cell division, biosynthesis and DNA replication reflecting terminal differentiation. Comparing the expression and DNA methylation data, we found that there was a significant overlap between upregulated genes and genes with decreased DNA methylation, and *vice versa*. This result agrees with the generally repressive nature of DNA methylation. Upon closer examination of several known granule genes, it was clear that the change in DNA methylation was concomitant with, and opposite to, the change in gene expression. Importantly, we confirmed this trend at several key hematopoietic transcription factor genes, including *PU.1*, *GATA2* and *GFI1*. Furthermore, the changes in DNA methylation and expression of these transcriptional regulators was accompanied by a corresponding change in motif activity. The motif activity reflects the transcription levels of genes with consensus binding sites for a particular transcription factor. This indicates that DNA methylation is involved in controlling the transcriptional network that governs hematopoiesis.

The 450k array probes are annotated for location in chromatin-defined enhancers. Comparing β-value variation across samples for all probes on the array showed that there was greater DNA methylation variability in enhancers than in the rest of the genome. This was supported by the overrepresentation of enhancer-annotated probes in the DMS list.

To further explore the role of DNA methylation in enhancers during granulopoiesis, we used a new CAGE-based method for enhancer annotation. This method, developed by Andersson *et al* (2014), is founded on the observation that enhancers are flanked by divergent transcription, and the expression of enhancers is proportional to enhancer activity. When we examined the overlap between DMSs and enhancers active in CD34+ cells compared to neutrophils and *vice versa*, we found that DMSs are significantly enriched in neutrophil specific enhancers, but not in progenitor specific enhancers. As most of the DMSs showed decreased methylation, this indicated that enhancer demethylation correlates with activation. Indeed, enhancers with decreasing DMSs showed increased transcription. Importantly, expression of downstream genes also increased, suggesting that DNA methylation is involved in regulating differentiation induced enhancer activation, and thereby transcription of downstream target genes.

To summarize, we characterized the methylome and transcriptome in granulopoiesis using high cellular resolution, which allowed us to closely define when in differentiation changes occur. We found that there is a general demethylation of

granulopoietic cells, and that changes in methylation occur at specific time points in conjunction with lineage restriction. Changes were most common in regions distant to CGIs and reciprocal to changes in transcription. Importantly DNA methylation may control expression of key hematopoietic transcription factors. Finally DNA methylation is dynamic in enhancer elements and may be involved in controlling the activity of these cis-regulatory elements during differentiation.

4 CONCLUDING REMARKS

This thesis has aimed to clarify the functions and consequences of epigenetic mechanisms on a large scale, with focus of their involvement in hematopoiesis. In paper I we used the fission yeast *S. pombe* as a model system to investigate combinatorial patterns of histone modifications. We demonstrated that there are specific modification patterns associated with genes of particular functions. We also demonstrated several cases of histone modification crosstalk. These results highlight the importance of taking a larger, more comprehensive picture into account when studying chromatin biology. As discussed in the introduction, epigenetics is rarely about a single factor or mark, but more accurately about interdependent networks.

This is the reason why studies such as paper II are both interesting and important. By examining the expression of a wide range of chromatin modifiers, we could identify cell, lineage and cancer specific epigenetic factors, suggesting a role for these proteins in blood development. Although experimental validation of their precise function in hematopoiesis is required, it provides the groundwork for identifying combinations of chromatin modifiers essential and characteristic for differentiation of different blood lineages, which would potentially reflect the epigenomes of these cells.

In paper III we zoomed in on granulocyte development and analyzed DNA methylation and gene expression of closely related cell populations representing different stages of granulopoiesis. We found that DNA methylation changes predominantly consist of decreases in granulpoiesis, and that changes occur at specific time points. Whereas decreased DNA methylation in neutrophil development has been described elsewhere (Ji *et al*, 2010; Hodges *et al*, 2011; Bocker *et al*, 2011), previous studies lacked the high cellular resolution to ascertain exactly when changes occur.

An interesting question is whether the observed demethylation is a result of active or passive mechanisms. Cells around the GMP stage, where the decrease is most obvious, are highly proliferative and CAGE data from paper II indicated that DNMT1 is somewhat more weakly expressed in a subset of myeloid cells, including neutrophils, possibly arguing for passive demethylation.

Still, considering the high incidence of *TET2* and *IDH1/2* mutations in AML with accompanying hypermethylation (discussed in section 1.9.4.1), it would be of interest to evaluate whether the DNA demethylation in granulopoiesis is mediated by TET2 and hydroxymethylation. Unfortunately, the TET enzymes were not included in our expression analysis in paper II and one of the weaknesses of the 450k array is the inability to distinguish methylation from hydroxymethylation. If demethylation was preceded by hydroxymethylation it would argue strongly for active demethylation.

450k array analysis of granulopoiesis revealed that DNA methylation is especially dynamic in enhancer elements. Importantly, decreased enhancer methylation overlapped with increased enhancer activity and expression of downstream genes. Since enhancers are preferably located in CpG poor areas (Andersson *et al*, 2014) and the 450k array lacks complete coverage in the genome, many enhancer regions are only

represented by a single probe. It may be relevant to investigate the methylation status of the surrounding region to address whether the methylation change has site-specific effects or reflects the overall chromatin state of that region. The recent discovery of super-enhancers, extensive enhancer clusters involved in regulation of cell fate (Whyte *et al*, 2013; Lovén *et al*, 2013; Hnisz *et al*, 2013), creates new questions. It would be interesting to see if any of our demethylated enhancers reside in such clusters. Super-enhancers are commonly associated with cell type specific genes. Indeed, among the genes controlled by the demethylated enhancers were genes relevant for neutrophil development and function such as cytokines and receptors for cytokines and growth factors.

5 ACKNOWLEDGEMENTS

The best part of my PhD is without question the people I had the honor of working with. To all the KEK-group members, past and present, I can only say thank you for making these year to something I will remember with happiness. I'm grateful to all the people who contributed to the work in this thesis, professionally, personally or both. I would especially like to thank:

Karl, my co-supervisor, for giving me the chance to do a PhD in your lab. I appreciate all your help and advice.
Andreas, Oh, brave leader and dauntless supervisor! How can I ever thank you for all your help? You are something incredibly rare, a genuinely caring person with a will to understand other people. As much as I appreciate your guidance and all the effort you put in to make my thesis happen, I appreciate still more our non-scientific conversations that gave me a lot of valuable insights. I shall certainly miss disagreeing with you and contradicting you!
Babett, for many of my favorite reading experiences, not to mention everything I know about R, which basically gave me my papers. You are incredibly talented and such a good friend. May your future be full of good wool!
Punit, I really enjoyed working with you. I have the greatest respect for you scientific mind as well as your warm personality, patience and last, but not least, cooking skills! Your hard work will pay off…
Laia, you are so sweet, so beautiful and so funny! I'm looking forward to the Iranian food….
Agata, for arguing convincingly against giving up when I felt tempted to. We've missed you!
Annelie, for putting my feet back on earth countless times when I needed it. You have no idea how often I wished I was more like you, so determined and pragmatic.
Ulrika, rarely has any officemates shared so many tears as you and I have. If I could only have half of your energy! Never forget you're brilliant…
Jenna, I owe you many, many laughs and cookie recipes. You are always there, ready to listen and give honest opinions. I cannot thank you enough for your patience with me while writing this thesis, or for making me realize my talent for cat imitations.
Lee, you always take time to listen to problems and offer good advice.
Victoria, for practically radiating energy! I swear, the room becomes a little bit brighter when you step in. You inspire me!
Christos, one of the warmest and funniest people I've met! You brought so much life and good humor to the lab.
Lina, you were really a breath of fresh air when the lab needed it!
Wenbo, your smile always makes me feel better!
Julian and Mickaël, for giving me a great start in the lab. Everyone should have teachers like you when they start!
Marianna, I enjoyed working with you and wish it could have been longer.
Ingela, I will always admire your strength.
Jhang, good luck with your projects!
Olga, at one point you asked me not to thank you for the chocolate. So I won't (no matter how good it was. And it was GOOD). Instead I will thank you for the many,

many hours you have spent encouraging me, giving me advice (both on science and on life in general) and for being such a great friend. What will I do without my morning hug?
Carolina, just imagine how much you could have made if you charged me for our therapy sessions! I miss our talks, your (not so common) common sense and warmth. You are such an amazing person, caring and honest. I miss you!
Indranil, for always answering questions so pedagogically. And for the lessons in indian cooking!
Birgitta, for always having time to chat and always making me laugh. Also, thank you for being so bad at aiming…
Peter, number one Reiseleiter! What else needs to be said?

Also many thanks to my collaborators outside the KEK-group, especially: **Eric Arnér**, for guiding us through the CAGE-jungle and answering my endless questions on bioinformatics. **Robin Andersson,** who did most of the enhancer analysis. **Mohsen Karimi**, most excellent teacher and Pyro-expert without whom my paper still wouldn't have been finished. **Fredrik, Jessica, Patrik & David** at the BEA core facility for all your help.

I'm also grateful to my friends and family for all their support. In particular:

Ying, for all the discussions, scientific and otherwise. I love your sense of humor, appreciate you thoughtfulness and envy you your powerful mind.
Kristina Volkova, for the energy you have given me the past few months, by just being who you are. Your courage and drive inspires me and makes me want to be as brave as you are (yes you are, there's no point in arguing). You're next!
Mille, my mother, for always having my back, fighting with me when I need it and knowing when I just need a hug. I love you!
Oliver, the world's greatest baby brother! You impress me constantly with your talent, your intellect and most of all your huge warm heart and sense of justice. The future is yours!
Bo, my grandfather, who so often led by example and inspired me to be brave enough to do something new. And gave me a kick-ass cover for my thesis!
Kjell, my father. You have always been my rock, my one point of solid ground no matter what. I love you more than I can say, and I owe you everything.
Annika, for listening, many, many hours when I needed it, even though you didn't have to.
Java & Caramel, my sweet, furry babies! Without your belly-fluff and fuzz-therapy I would never have gotten this far.
Bella, for that ever-wagging tail, all the walks in the woods when I needed to clear my head. Also, for not eating my cats…
Yvonne, Gunnar, Emelie & Joel, for welcoming me into your family with open arms and for being so understanding when I've been stressed.
Thomas, Words cannot describe what you mean to me. I'll have to settle for "I love you", but it is so much more. Just knowing there are people like you in this world gives me hope. I promise I will be a nicer person to live with, now that this is over…

6 REFERENCES

Aapola U, Shibuya K, Scott HS, Ollila J, Vihinen M, Heino M, Shintani A, Kawasaki K, Minoshima S, Krohn K, Antonarakis SE, Shimizu N, Kudoh J & Peterson P (2000) Isolation and Initial Characterization of a Novel Zinc Finger Gene, DNMT3L, on 21q22.3, Related to the Cytosine-5- Methyltransferase 3 Gene Family. *Genomics* **65:** 293–298

Abraham BJ, Cui K, Tang Q & Zhao K (2013) Dynamic regulation of epigenomic landscapes during hematopoiesis. *BMC Genomics* **14:** 1–1

Adli M, Zhu J & Bernstein BE (2010) Genome-wide chromatinmaps derived fromlimited numbers ofhematopoietic progenitors. *Nature Publishing Group* **7:** 615–618

Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project (2009) Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature* **457:** 1028–1032

Allen HF, Wade PA & Kutateladze TG (2013) The NuRD architecture. *Cell. Mol. Life Sci.* **70:** 3513–3524

Allfrey VG & Mirsky AE (1964) Structural Modifications of Histones and their Possible Role in the Regulation of RNA Synthesis. *Science* **144:** 559–559

Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, Ntini E, Arner E, Valen E, Li K, Schwarzfisher L, Glatz D, Raithel J, Lilje B, Rapin N, Bagger FO, et al (2014) An atlas of active enhancers across human cell types and tissues. *Nature*

Bannister AJ & Kouzarides T (2011) Regulation of chromatin by histone modifications. *Nature Publishing Group* **21:** 381–395

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones DE, Wang Z, Wei G, Chepelev I & Zhao K (2007) High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129:** 823–837

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL & Lander ES (2006) A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125:** 315–326

Bhutani N, Burns DM & Blau HM (2011) Perspective. *Cell* **146:** 866–872

Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, Delano D, Zhang L, Schroth GP, Gunderson KL, Fan J-B & Shen R (2011) High density DNA methylation array with single CpG site resolution. *Genomics* **98:** 288–295

Bird A (2007) Perceptions of epigenetics. *Nature* **447:** 396–398

Bird AP & Taggart MH (1980) Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Research* **8:** 1485–1497

Bock C, Beerman I, Lien W-H, Smith ZD, Gu H, Boyle P, Gnirke A, Fuchs E, Rossi DJ & Meissner A (2012) DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. *Molecular Cell* **47:** 633–647

Bocker MT, Hellwig I, Breiling A, Eckstein V, Ho AD & Lyko F (2011) Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *Blood* **117:** e182–e189

Borregaard N (2010) Neutrophils, from Marrow to Microbes. *Immunity* **33:** 657–670

Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S & Jacobsen SE (2007) UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science* **317:** 1760–1764

Brownell JE & Allis CD (1995) An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proc. Natl. Acad. Sci. U.S.A.* **92:** 6364–6368

Bröske A-M, Vockentanz L, Kharazi S, Huska MR, Mancini E, Scheller M, Kuhl C, Enns A, Prinz M, Jaenisch R, Nerlov C, Leutz A, Andrade-Navarro MA, Jacobsen SEW & Rosenbauer F (2009) DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nature Genetics* **41:** 1207–1215

Bruniquel D & Schwartz RH (2003) Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat Immunol* **4:** 235–240

Buck-Koehntop BA & Defossez P-A (2013) On how mammalian transcription factors recognize methylated DNA. *epigenetics* **8:** 131–137

Bultman S, Gebuhr T, Yee D, La Mantia C, Nicholson J, Gilliam A, Randazzo F, Metzger D, Chambon P, Crabtree G & Magnuson T (2000) A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Molecular Cell* **6:** 1287–1295

Butler JEF & Kadonaga JT (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & Development* **16:** 2583–2592

Butler JS & Dent SYR (2013) The role of chromatin modifiers in normal and malignant hematopoiesis. *Blood* **121:** 3076–3084

Cahill N & Rosenquist R (2013) Uncovering the DNA methylome in chronic lymphocytic leukemia. *epigenetics* **8:** 138–148

Calo E & Wysocka J (2013) Modification of Enhancer Chromatin: What, How, and Why? *Molecular Cell* **49:** 825–837

Carninci P, Kvam C, Kitamura A, Ohsumi T, Okazaki Y, Itoh M, Kamiya M, Shibata K, Sasaki N, Izawa M, Muramatsu M, Hayashizaki Y & Schneider C (1996) High-efficiency full-length cDNA cloning by biotinylated CAP trapper. *Genomics* **37:** 327–336

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia W-J, Anderson S, Yates J, Washburn MP & Workman JL (2005) Histone H3 Methylation by Set2 Directs Deacetylation of Coding Regions by Rpd3S to Suppress Spurious Intragenic Transcription. *Cell* **123:** 581–592

Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, Bock C, Vasanthakumar A, Gu H, Xi Y, Liang S, Lu Y, Darlington GJ, Meissner A, Issa J-PJ, Godley LA, Li W & Goodell MA (2012) Dnmt3a is essential for hematopoietic stem cell differentiation. *Nature Genetics* **44:** 23–31

Chan SM & Majeti R (2013) Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia. *Int J Hematol* **98:** 648–657

Chen Z-X, Mann JR, Hsieh C-L, Riggs AD & Chédin F (2005) Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. *J. Cell. Biochem.* **95:** 902–917

Chi TH, Wan M, Lee PP, Akashi K, Metzger D, Chambon P, Wilson CB & Crabtree GR (2003) Sequential roles of Brg, the ATPase subunit of BAF chromatin remodeling complexes, in thymocyte development. *Immunity* **19:** 169–182

Cho KWY (2012) Enhancers. *WIREs Dev Biol* **1:** 469–478

Chung YR, Schatoff E & Abdel-Wahab O (2012) Epigenetic alterations in hematopoietic malignancies. *Int J Hematol* **96:** 413–427

Chung YS, Kim HJ, Kim TM, Hong SH, Kwon KR, An S, Park JH, Lee S & Oh IH (2009) Undifferentiated hematopoietic cells are characterized by a genome-wide undermethylation dip around the transcription start site and a hierarchical

epigenetic plasticity. *Blood* **114:** 4968–4978
Clapier CR & Cairns BR (2009) The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* **78:** 273–304
Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA & Jaenisch R (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107:** 21931–21936
Cross SH, Meehan RR, Nan X & Bird A (1997) A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nature Genetics* **16:** 256–259
Cui K, Zang C, Roh T-Y, Schones DE, Childs RW, Peng W & Zhao K (2009) Chromatin Signatures in Multipotent Human Hematopoietic Stem Cells Indicatethe Fate of Bivalent Genes during Differentiation. *Stem Cell* **4:** 80–93
Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, Johnston LD, Scott MP, Smith JJ, Xiao Y, Jin L, Kuntz KW, Chesworth R, Moyer MP, Bernt KM, Tseng J-C, Kung AL, Armstrong SA, Copeland RA, Richon VM, et al (2011) Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. *Cancer Cell* **20:** 53–65
Deaton AM & Bird A (2011) CpG islands and the regulation of transcription. *Genes & Development* **25:** 1010–1022
Deaton AM, Webb S, Kerr ARW, Illingworth RS, Guy J, Andrews R & Bird A (2011) Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Research* **21:** 1074–1086
Deneberg S, Guardiola P, Lennartsson A, Qu Y, Gaidzik V, Blanchet O, Karimi M, Bengtzen S, Nahi H, Uggla B, Tidefelt U, Hoglund M, Paul C, Ekwall K, Dohner K & Lehmann S (2011) Prognostic DNA methylation patterns in cytogenetically normal acute myeloid leukemia are predefined by stem cell chromatin marks. *Blood* **118:** 5573–5582
Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, Allis CD, Chait BT, Hess JL & Roeder RG (2005) Physical Association and Coordinate Function of the H3 K4 Methyltransferase MLL1 and the H4 K16 Acetyltransferase MOF. *Cell* **121:** 873–885
Edmondson DG (2002) Site-specific Loss of Acetylation upon Phosphorylation of Histone H3. *Journal of Biological Chemistry* **277:** 29496–29502
Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA & Gehrke C (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Research* **10:** 2709–2721
Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K & Zhang Y (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.* **12:** 1052–1058
Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, Schifano E, Booth J, van Putten W, Skrabanek L, Campagne F, Mazumdar M, Greally JM, Valk PJM, LOwenberg B, Delwel R & Melnick A (2010) DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia. *Cancer Cell* **17:** 13–27
Filippakopoulos P & Knapp S (2012) The bromodomain interaction module. *FEBS Letters* **586:** 2692–2704
Flaus A (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Research* **34:** 2887–2905
Franchini D-M, Schmitz K-M & Petersen-Mahrt SK (2012) 5-Methylcytosine DNA Demethylation: More Than Losing a Methyl Group. *Annu. Rev. Genet.* **46:** 419–441

Fujiwara Y, Browne CP, Cunniff K, Goff SC & Orkin SH (1996) Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci. U.S.A.* **93:** 12355–12358

Gardiner-Garden M & Frommer M (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.* **196:** 261–282

Glozak MA & Seto E (2007) Histone deacetylases and cancer. *Oncogene* **26:** 5420–5432

Greer EL & Shi Y (2012) Histone methylation: a dynamic mark in health, disease and inheritance. **:** 1–15

Guibert S & Weber M (2013) Functions of DNA Methylation and Hydroxymethylation in Mammalian Development 1st ed. Copyright © 2013 Elsevier Inc. All rights reserved.

Gullberg U, Andersson E, Garwicz D, Lindmark A & Olsson I (1997) Biosynthesis, processing and sorting of neutrophil proteins: insight into neutrophil granule development. *Eur. J. Haematol.* **58:** 137–153

Hargreaves DC & Crabtree GR (2011) ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Nature Publishing Group***:** 1–25

Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X & Cheng X (2012) Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Research* **40:** 4841–4849

Hashimoto H, Vertino PM & Cheng X (2010) Molecular coupling of DNA methylation and histone methylation. *Epigenomics* **2:** 657–669

Hata K, Okano M, Lei H & Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129:** 1983–1993

Hattangadi SM, Wong P, Zhang L, Flygare J & Lodish HF (2011) From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood* **118:** 6258–6268

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE & Ren B (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics* **39:** 311–318

Hendrich B & Bird A (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol. Cell. Biol.* **18:** 6538–6547

Herrera-Merchan A, Arranz L, Ligos JM, de Molina A, Dominguez O & Gonzalez S (2012) Ectopic expression of the histone methyltransferase Ezh2 in haematopoietic stem cells causes myeloproliferative disease. *Nature Communications* **3:** 623

Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, Diez J, Sanchez-Mut JV, Setien F, Carmona FJ, Puca AA, Sayols S, Pujana MA, Serra-Musach J, Iglesias-Platas I, Formiga F, Fernandez AF, Fraga MF, Heath SC, Valencia A, et al (2012) Distinct DNA methylomes of newborns and centenarians. *Proc. Natl. Acad. Sci. U.S.A.* **109:** 10522–10527

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke HA & Young RA (2013) Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **155:** 934–947

Ho L, Ronan JL, Wu J, Staahl BT, Chen L, Kuo A, Lessard J, Nesvizhskii AI, Ranish J & Crabtree GR (2009) An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci. U.S.A.* **106:** 5181–5186

Hodges E, Molaro A, Santos Dos CO, Thekkat P, Song Q, Uren PJ, Park J, Butler J, Rafii S, McCombie WR, Smith AD & Hannon GJ (2011) Directional DNA

Methylation Changes and Complex Intermediate States Accompany Lineage Specificity in the Adult Hematopoietic Compartment. *Molecular Cell* **44:** 17–28

Hoffbrand V & Moss P (2011) Essential Haematology 6 ed. John Wiley & Sons

Hogart A, Lichtenberg J, Ajay SS, Anderson S, NIH Intramural Sequencing Center, Margulies EH & Bodine DM (2012) Genome-wide DNA methylation profiles in hematopoietic stem and progenitor cells reveal overrepresentation of ETS transcription factor binding sites. *Genome Research* **22:** 1407–1418

Hong L, Schroth GP, Matthews HR, Yau P & Bradbury EM (1993) Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 'tail' to DNA. *J. Biol. Chem.* **268:** 305–314

Hu M, Krause D, Greaves M, Sharkis S, Dexter M, Heyworth C & Enver T (1997) Multilineage gene expression precedes commitment in the hemopoietic system. *Genes & Development* **11:** 774–785

Illingworth RS & Bird AP (2009) CpG islands – 'A rough guide'. *FEBS Letters* **583:** 1713–1720

Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr ARW, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R & Bird AP (2010) Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. *PLoS Genet* **6:** e1001134

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S & Feinberg AP (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics* **41:** 178–186

Issa JPJ (2013) The myelodysplastic syndrome as a prototypical epigenetic disease. *Blood* **121:** 3811–3817

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC & Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466:** 1129–1133

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C & Zhang Y (2011) Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* **333:** 1300–1303

Iwama A, Oguro H, Negishi M, Kato Y, Morita Y, Tsukui H, Ema H, Kamijo T, Katoh-Fukui Y, Koseki H, van Lohuizen M & Nakauchi H (2004) Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1. *Immunity* **21:** 843–851

Iwasaki H & Akashi K (2007) Myeloid Lineage Commitment from the Hematopoietic Stem Cell. *Immunity* **26:** 726–740

Jeong M, Sun D, Luo M, Huang Y, Challen GA, Rodriguez B, Zhang X, Chavez L, Wang H, Hannah R, Kim S-B, Yang L, Ko M, Chen R, Göttgens B, Lee J-S, Gunaratne P, Godley LA, Darlington GJ, Rao A, et al (2014) Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nature Genetics* **46:** 17–23

Ji H, Ehrlich LIR, Seita J, Murakami P, Doi A, Lindau P, Lee H, Aryee MJ, Irizarry RA, Kim K, Rossi DJ, Inlay MA, Serwold T, Karsunky H, Ho L, Daley GQ, Weissman IL & Feinberg AP (2010) Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467:** 338–342

Jiang L, Smith JN, Anderson SL, Ma P, Mizzen CA & Kelleher NL (2007) Global Assessment of Combinatorial Post-translational Modification of Core Histones in Yeast Using Contemporary Mass Spectrometry: LYS4 TRIMETHYLATION CORRELATES WITH DEGREE OF ACETYLATION ON THE SAME H3 TAIL. *Journal of Biological Chemistry* **282:** 27923–27934

Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics* **13:** 484–492
Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J & Wolffe AP (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics* **19:** 187–191
Jude CD, Climer L, Xu D, Artinger E, Fisher JK & Ernst P (2007) Unique and Independent Roles for MLL in Adult Hematopoietic Stem Cells and Progenitors. *Cell Stem Cell* **1:** 324–337
Justin N, De Marco V, Aasland R & Gamblin SJ (2010) Reading, writing and editing methylated lysines on histone tails: new insights from recent structural studies. *Current Opinion in Structural Biology* **20:** 730–738
Kaminskas E (2005) FDA Drug Approval Summary: Azacitidine (5-azacytidine, VidazaTM) for Injectable Suspension. *The Oncologist* **10:** 176–182
Kamminga LM (2006) The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107:** 2170–2179
Kanamori-Katayama M, Itoh M, Kawaji H, Lassmann T, Katayama S, Kojima M, Bertin N, Kaiho A, Ninomiya N, Daub CO, Carninci P, Forrest ARR & Hayashizaki Y (2011) Unamplified cap analysis of gene expression on a single-molecule sequencer. *Genome Research* **21:** 1150–1159
Karsunky H, Zeng H, Schmidt T, Zevnik B, Kluge R, Schmid KW, Dührsen U & Möröy T (2002) Inflammatory reactions and severe neutropenia in mice lacking the transcriptional repressor Gfi1. *Nature Genetics* **30:** 295–300
Kleff S, Andrulis ED, Anderson CW & Sternglanz R (1995) Identification of a gene encoding a yeast histone H4 acetyltransferase. *J. Biol. Chem.* **270:** 24674–24677
Klose RJ & Bird AP (2006) Genomic DNA methylation: the mark and its mediators. *Trends in Biochemical Sciences* **31:** 89–97
Klose RJ & Zhang Y (2007) Regulation of histone methylation by demethylimination and demethylation. *Nat Rev Mol Cell Biol* **8:** 307–318
Klug M, Heinz S, Gebhard C, Schwarzfischer L, Krause SW, Andreesen R & Rehli M (2010) Active DNA demethylation in human postmitotic cells correlates with activating histone modifications, but not transcription levels. *Genome Biol* **11:** R63
Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R, Liu XS, Aravind L, Agarwal S, Maciejewski JP & Rao A (2010) Impaired hydroxylation of 5-methylcytosine inmyeloid cancers with mutant TET2. *Nature* **468:** 839–843
Kohli RM & Zhang Y (2013) TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502:** 472–479
Kouzarides T (2000) Acetylation: a regulatory modification to rival phosphorylation? *EMBO J.* **19:** 1176–1179
Krasteva V, Buscarlet M, Diaz-Tellez A, Bernard MA, Crabtree GR & Lessard JA (2012) The BAF53a subunit of SWI/SNF-like BAF complexes is essential for hemopoietic stem cell function. *Blood* **120:** 4720–4732
Kriaucionis S & Heintz N (2009) The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* **324:** 929–930
Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, Xia X, Jesneck J, Bracken AP, Silverman LB, Kutok JL, Kung AL & Armstrong SA (2008) H3K79 Methylation Profiles Define Murine and Human MLL-AF4 Leukemias. *Cancer Cell* **14:** 355–368
Kurdistani SK, Tavazoie S & Grunstein M (2004) Mapping global histone acetylation patterns to gene expression. *Cell* **117:** 721–733
Lachner M, O'Carroll D, Rea S, Mechtler K & Jenuwein T (2001) Methylation of

histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410:** 116–120
Lakshman R & Finn A (2001) Neutrophil disorders and their management. *J. Clin. Pathol.* **54:** 7–19
Latham JA & Dent SYR (2007) Cross-regulation of histone modifications. *Nat Struct Mol Biol* **14:** 1017–1024
Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J & Wei CL (2010) Dynamic changes in the human methylome during differentiation. *Genome Research* **20:** 320–331
Lee ST, Xiao Y, Muench MO, Xiao J, Fomin ME, Wiencke JK, Zheng S, Dou X, de Smith A, Chokkalingam A, Buffler P, Ma X & Wiemels JL (2012) A global DNA methylation and gene expression analysis of early human B-cell development reveals a demethylation signature and transcription factor network. *Nucleic Acids Research* **40:** 11339–11351
Leonhardt H, Page AW, Weier HU & Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865–873
Lessard J, Wu JI, Ranish JA, Wan M, Winslow MM, Staahl BT, Wu H, Aebersold R, Graef IA & Crabtree GR (2007) An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development. *Neuron* **55:** 201–215
Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F & Bird A (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69:** 905–914
Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, Harris CC, Lichti CF, Townsend RR, Fulton RS, Dooling DJ, Koboldt DC, Schmidt H, Zhang Q, Osborne JR, Lin L, et al (2010) DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* **363:** 2424–2433
Li B, Carey M & Workman JL (2007) The role of chromatin during transcription. *Cell* **128:** 707–719
Li B, Pattenden SG, Lee D, Gutiérrez J, Chen J, Seidel C, Gerton J & Workman JL (2005) Preferential occupancy of histone variant H2AZ at inactive promoters influences local histone modifications and chromatin remodeling. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 18385–18390
Li E, Bestor TH & Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915–926
Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo Q-M, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B & Ecker JR (2009) Human DNA methylomes at baseresolution show widespread epigenomicdifferences. *Nature* **462:** 315–322
Liu X, Bushnell DA & Kornberg RD (2013) Biochimica et Biophysica Acta. *BBA - Gene Regulatory Mechanisms* **1829:** 2–8
Long HK, Blackledge NP & Klose RJ (2013) ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. *Biochem. Soc. Trans.* **41:** 727–740
Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST & Downing JR (2003) TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* **17:** 637–641
Lovén J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI & Young RA (2013) Selective Inhibition ofTumor Oncogenes by Disruption of Super-Enhancers. *Cell* **153:** 320–334
Luger K, Mäder AW, Richmond RK, Sargent DF & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389:** 251–260

Maes J, Maleszewska M, Guillemin C, Pflumio F, Six E, Andre-Schmutz I, Cavazzana-Calvo M, Charron D, Francastel C & Goodhardt M (2008) Lymphoid-affiliated genes are associated with active histone modifications in human hematopoietic stem cells. *Blood* **112:** 2722–2729

Mak KS, Funnell APW, Pearson RCM & Crossley M (2011) PU.1 and Haematopoietic Cell Fate: Dosage Matters. *International Journal of Cell Biology* **2011:** 1–6

Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, Gilmour DS, Albert I & Pugh BF (2008) Nucleosome organization in the Drosophila genome. *Nature* **453:** 358–362

May D, Blow MJ, Kaplan T, McCulley DJ, Jensen BC, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Afzal V, Simpson PC, Rubin EM, Black BL, Bristow J, Pennacchio LA & Visel A (2012) Large-scale discovery of enhancers from human heart tissue. *Nature Genetics* **44:** 89–93

Mayer W, Niveleau A, Walter J, Fundele R & Haaf T (2000) Demethylation of the zygotic paternal genome. *Nature* **403:** 501–502

Meissner A (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Research* **33:** 5868–5877

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R & Lander ES (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*

Millar CB (2013) Organizing the genome with H2A histone variants. *Biochem. J.* **449:** 567–579

Miller CA & Sweatt JD (2008) Covalent Modification of DNA Regulates Memory Formation. *Neuron* **59:** 1051–1051

Miyamoto T, Iwasaki H, Reizis B, Ye M, Graf T, Weissman IL & Akashi K (2002) Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Developmental Cell* **3:** 137–147

Mizuno SI (2001) Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. *Blood* **97:** 1172–1179

Mochizuki-Kashio M, Mishima Y, Miyagi S, Negishi M, Saraya A, Konuma T, Shinga J, Koseki H & Iwama A (2011) Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. *Blood* **118:** 6553–6561

Mocsai A (2013) Diverse novel functions of neutrophils in immunity, inflammation, and beyond. *Journal of Experimental Medicine* **210:** 1283–1299

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M & Schübeler D (2008) Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Molecular Cell* **30:** 755–766

Moore LD, Le T & Fan G (2012) DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38:** 23–38

Musselman CA, Ramírez J, Sims JK, Mansfield RE, Oliver SS, Denu JM, Mackay JP, Wade PA, Hagman J & Kutateladze TG (2012) Bivalent recognition of nucleosomes by the tandem PHD fingers of the CHD4 ATPase is required for CHD4-mediated repression. *Proc. Natl. Acad. Sci. U.S.A.* **109:** 787–792

Nakahata S, Saito Y, Hamasaki M, Hidaka T, Arai Y, Taki T, Taniwaki M & Morishita K (2009) Alteration of enhancer of polycomb 1 at 10p11.2 is one of the genetic events leading to development of adult T-cell leukemia/lymphoma. *Genes Chromosom. Cancer* **48:** 768–776

Narlikar GJ, Sundaramoorthy R & Owen-Hughes T (2013) Mechanisms and

Functions of ATP-Dependent Chromatin-Remodeling Enzymes. *Cell* **154:** 490–503

Neff T & Armstrong SA (2013) Recent progress toward epigenetic therapies: the example of mixed lineage leukemia. *Blood* **121:** 4847–4853

Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D & Bird A (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature Genetics* **23:** 58–61

Oguro H, Yuan J, Ichikawa H, Ikawa T, Yamazaki S, Kawamoto H, Nakauchi H & Iwama A (2010) Short Article. *Stem Cell* **6:** 279–286

Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G & Zhang Y (2005) hDOT1L Links Histone Methylation to Leukemogenesis. *Cell* **121:** 167–178

Okano M, Bell DW, Haber DA & Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247–257

Okano M, Xie S & Li E (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics* **19:** 219–220

Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H & Hayashi Y (2002) LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* **62:** 4075–4080

Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin S-P, Allis CD, Cheng X & Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448:** 714–717

Orford K, Kharchenko P, Lai W, Dao MC, Worhunsky DJ, Ferro A, Janzen V, Park PJ & Scadden DT (2008) Differential H3K4 Methylation Identifies Developmentally Poised Hematopoietic Genes. *Developmental Cell* **14:** 798–809

Orkin SH (2000) Diversification of haematopoietic stem cells to specific lineages. *Nature Reviews Genetics* **1:** 57–64

Orkin SH, Shivdasani RA, Fujiwara Y & McDevitt MA (1998) Transcription factor GATA-1 in megakaryocyte development. *Stem Cells* **16 Suppl 2:** 79–83

Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M & Shirakawa M (2009) scientific report. *EMBO reports* **10:** 1235–1241

Panagopoulos I, Fioretos T, Isaksson M, Samuelsson U, Billström R, Strömbeck B, Mitelman F & Johansson B (2001) Fusion of the MORF and CBP genes in acute myeloid leukemia with the t(10;16)(q22;p13). *Human Molecular Genetics* **10:** 395–404

Pang WW, Price EA, Sahoo D, Beerman I, Maloney WJ, Rossi DJ, Schrier SL & Weissman IL (2011) Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc. Natl. Acad. Sci. U.S.A.* **108:** 20012–20017

Park I-K, Qian D, Kiel M, Becker MW, Pihalja M, Weissman IL, Morrison SJ & Clarke MF (2003) Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423:** 302–305

Parthun MR (2007) Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene* **26:** 5319–5328

Pastor WA, Aravind L & Rao A (2013) TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol* **14:** 341–356

Patel JP, Gönen M, Figueroa ME, Fernandez H, Sun Z, Racevskis J, Van Vlierberghe P, Dolgalev I, Thomas S, Aminova O, Huberman K, Cheng J, Viale A, Socci ND,

Heguy A, Cherry A, Vance G, Higgins RR, Ketterling RP, Gallagher RE, et al (2012) Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med* **366:** 1079–1089

Peserico A & Simone C (2011) Physical and functional HAT/HDAC interplay regulates protein acetylation balance. *J. Biomed. Biotechnol.* **2011:** 371832–10

Pointner J, Persson J, Prasad P, Norman-Axelsson U, Strålfors A, Khorosjutina O, Krietenstein N, Svensson JP, Ekwall K & Korber P (2012) CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in S. pombe. *EMBO J.* **31:** 4388–4403

Pradhan S, Bacolla A, Wells RD & Roberts RJ (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J. Biol. Chem.* **274:** 33002–33010

Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, Ramsay AJ, Beà S, Pinyol M, Martínez-Trillos A, López-Guerra M, Colomer D, Navarro A, Baumann T, Aymerich M, Rozman M, Delgado J, Giné E, Hernández JM, González-Díaz M, et al (2012) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature Genetics* **44:** 47–52

Quintás-Cardama A, Santos FPS & Garcia-Manero G (2011) Histone deacetylase inhibitors for the treatment of myelodysplastic syndrome and acute myeloid leukemia. *Leukemia* **25:** 226–235

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA & Wysocka J (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470:** 279–283

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP & Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U.S.A.* **97:** 5237–5242

Rando OJ (2012) Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Current Opinion in Genetics & Development* **22:** 148–155

Reyes JC, Barra J, Muchardt C, Camus A, Babinet C & Yaniv M (1998) Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). *EMBO J.* **17:** 6979–6991

Rice KL, Hormaeche I & Licht JD (2007) Epigenetic regulation of normal and malignant hematopoiesis. *Oncogene* **26:** 6697–6714

Rivera CM & Ren B (2013) Mapping Human Epigenomes. *Cell* **155:** 39–55

Rosati R (2002) NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). *Blood* **99:** 3857–3860

Rozman MA, Cam s M, Colomer D, Villamor N, Esteve J, Costa D, Carri A, Aymerich M, Aguilar JL, Domingo AC, Sol F, Gomis F, Florensa L, Montserrat E & Campo E (2004) Type IMOZ/CBP (MYST3/CREBBP) is the most common chimeric transcript in acute myeloid leukemia with t(8;16)(p11;p13) translocation. *Genes Chromosom. Cancer* **40:** 140–145

Russo VEA, Martienssen RA & Riggs AD (1996) Epigenetic mechanisms of gene regulation Woodbury: Cold Spring Harbor Laboratory Pr

Ryan DP & Owen-Hughes T (2011) Snf2-family proteins: chromatin remodellers for any occasion. *Current Opinion in Chemical Biology* **15:** 649–656

Saied MH, Marzec J, Khalid S, Smith P, Down TA, Rakyan VK, Molloy G, Raghavan M, Debernardi S & Young BD (2012) Genome Wide Analysis of Acute Myeloid Leukemia Reveal Leukemia Specific Methylome and Subtype Specific Hypomethylation of Repeats. *PLoS ONE* **7:** e33213

Saleque S, Kim J, Rooke HM & Orkin SH (2007) Epigenetic Regulation of

Hematopoietic Differentiation by Gfi-1 and Gfi-1b Is Mediated by the Cofactors CoREST and LSD1. *Molecular Cell* **27:** 562–572
Sandoval J & Esteller M (2012) Cancer epigenomics: beyond genomics. *Current Opinion in Genetics & Development* **22:** 50–55
Sarraf SA & Stancheva I (2004) Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular Cell* **15:** 595–605
Schmidl C, Klug M, Boeld TJ, Andreesen R, Hoffmann P, Edinger M & Rehli M (2009) Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Research* **19:** 1165–1174
Schomacher L (2013) Mammalian DNA demethylation: Multiple faces and upstream regulation. *epigenetics* **8:** 679–684
Schoofs T & Müller-Tidow C (2011) Cancer Treatment Reviews. *Cancer Treatment Reviews* **37:** S13–S18
Serandour AA, Avner S, Oger F, Bizot M, Percevault F, Lucchetti-Miganeh C, Palierne G, Gheeraert C, Barloy-Hubler F, Peron CL, Madigou T, Durand E, Froguel P, Staels B, Lefebvre P, Metivier R, Eeckhoute J & Salbert G (2012) Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Research* **40:** 8255–8265
Shearstone JR, Pop R, Bock C, Boyle P, Meissner A & Socolovsky M (2011) Global DNA Demethylation During Mouse Erythropoiesis in Vivo. *Science* **334:** 799–802
Shen L & Zhang Y (2013) 5-Hydroxymethylcytosine: generation, fate, and genomic distribution. *Current Opinion in Cell Biology* **25:** 289–296
Shen Y (2006) Abnormal CpG island methylation occurs during in vitro differentiation of human embryonic stem cells. *Human Molecular Genetics* **15:** 2623–2635
Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA & Shi Y (2004) Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell* **119:** 941–953
Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, Fukuda S, Sasaki D, Podhajska A, Harbers M, Kawai J, Carninci P & Hayashizaki Y (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 15776–15781
Shogren-Knaak M, Ishii H, Sun J-M, Pazin MJ, Davie JR & Peterson CL (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311:** 844–847
Simon JA & Kingston RE (2013) Occupying Chromatin: Polycomb Mechanisms for Getting to Genomic Targets, Stopping Transcriptional Traffic, and Staying Put. *Molecular Cell* **49:** 808–824
Smallwood A & Ren B (2013) Genome organization and long-range regulation of gene expression by enhancers. *Current Opinion in Cell Biology* **25:** 387–394
Smolle M & Workman JL (2013) Biochimica et Biophysica Acta. *BBA - Gene Regulatory Mechanisms* **1829:** 84–97
Sprüssel A, Schulte JH, Weber S, Necke M, Händschke K, Thor T, Pajtler KW, Schramm A, König K, Diehl L, Mestdagh P, Vandesompele J, Speleman F, Jastrow H, Heukamp LC, Schüle R, Dührsen U, Buettner R, Eggert A & Göthert JR (2012) Lysine-specific demethylase 1 restricts hematopoietic progenitor proliferation and is essential for terminal differentiation. *Leukemia* **26:** 2039–2051
Strahl BD & Allis CD (2000) The language of covalent histone modifications. *Nature*

403: 41–45
Su I-H, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT & Tarakhovsky A (2002) Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol* **4:** 124–131
Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L & Rao A (2009) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **324:** 930–935
Talbert PB & Henikoff S (2010) Histone variants — ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11:** 264–275
Tan J, Jones M, Koseki H, Nakayama M, Muntean AG, Maillard I & Hess JL (2011a) CBX8, a Polycomb Group Protein, Is Essential for MLL-AF9-Induced Leukemogenesis. *Cancer Cell* **20:** 563–575
Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, Lu Z, Ye Z, Zhu Q, Wysocka J, Ye Y, Khochbin S, Ren B & Zhao Y (2011b) Identification of 67 Histone Marksand Histone Lysine Crotonylationas a New Type of Histone Modification. *Cell* **146:** 1016–1028
Taunton J, Hassig CA & Schreiber SL (1996) A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272:** 408–411
Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP & Chait BT (2006) Yng1 PHD Finger Binding to H3 Trimethylated at K4 Promotes NuA3 HAT Activity at K14 of H3 and Transcription at a Subset of Targeted ORFs. *Molecular Cell* **24:** 785–796
The Cancer Genome Atlas Research Network (2013) Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N Engl J Med* **368:** 2059–2074
The ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **488:** 57–74
Thomas T (2006) Monocytic leukemia zinc finger protein is essential for the development of long-term reconstituting hematopoietic stem cells. *Genes & Development* **20:** 1175–1186
Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, Garg K, John S, Sandstrom R, Bates D, Boatman L, Canfield TK, Diegel M, Dunn D, Ebersol AK, Frum T, et al (2012) The accessible chromatin landscape of the human genome. *Nature* **488:** 75–82
Trowbridge JJ, Snow JW, Kim J & Orkin SH (2009) Short Article. *Stem Cell* **5:** 442–449
Tsai FY & Orkin SH (1997) Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood* **89:** 3636–3643
Tsukada Y-I, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P & Zhang Y (2005) Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439:** 811–816
Turner BM, O'Neill LP & Allan IM (1989) Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS Letters* **253:** 141–145
Tweedie-Cullen RY, Brunner AM, Grossmann J, Mohanna S, Sichau D, Nanni P, Panse C & Mansuy IM (2012) Identification of Combinatorial Patterns of Post-Translational Modifications on Individual Histones in the Mouse Brain. *PLoS ONE* **7:** e36980
Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM & Pennacchio LA (2009) ChIP-

seq accurately predicts tissue-specific activity of enhancers. *Nature* **457:** 854–858

Vradii D, Wagner S, Doan DN, Nickerson JA, Montecino M, Lian JB, Stein JL, van Wijnen AJ, Imbalzano AN & Stein GS (2006) Brg1, the ATPase subunit of the SWI/SNF chromatin remodeling complex, is required for myeloid differentiation to granulocytes. *J. Cell. Physiol.* **206:** 112–118

Wada T, Kikuchi J, Nishimura N, Shimizu R, Kitamura T & Furukawa Y (2009) Expression Levels of Histone Deacetylases Determine the Cell Fate of Hematopoietic Progenitors. *Journal of Biological Chemistry* **284:** 30673–30683

Walter MJ, Ding L, Shen D, Shao J, Grillot M, McLellan M, Fulton R, Schmidt H, Kalicki-Veizer J, O'Laughlin M, Kandoth C, Baty J, Westervelt P, DiPersio JF, Mardis ER, Wilson RK, Ley TJ & Graubert TA (2011) Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia* **25:** 1153–1158

Wang H, Cao R, Xia L, Erdjument-Bromage H, Borchers C, Tempst P & Zhang Y (2001) Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Molecular Cell* **8:** 1207–1217

Wang JH, Nichogiannopoulou A, Wu L, Sun L, Sharpe AH, Bigby M & Georgopoulos K (1996a) Selective defects in the development of the fetal and adult lymphoid system in mice with an Ikaros null mutation. *Immunity* **5:** 537–549

Wang W, Côté J, Xue Y, Zhou S, Khavari PA, Biggar SR, Muchardt C, Kalpana GV, Goff SP, Yaniv M, Workman JL & Crabtree GR (1996b) Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15:** 5370–5382

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ & Zhao K (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics* **40:** 897–903

Weake VM & Workman JL (2010) Inducible gene expression: diverse regulatory mechanisms. *Nature Reviews Genetics* **11:** 426–437

Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M & Schübeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics* **39:** 457–466

Weishaupt H, Sigvardsson M & Attema JL (2010) Epigenetic chromatin states uniquely define the developmental plasticity of murine hematopoietic stem cells. *Blood* **115:** 247–256

Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, Wartman LD, Lamprecht TL, Liu F, Xia J, Kandoth C, Fulton RS, McLellan MD, Dooling DJ, Wallis JW, Chen K, Harris CC, Schmidt HK, Kalicki-Veizer JM, Lu C, et al (2012) The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150:** 264–278

Whetstine JR, Nottke A, Lan F, Huarte M, Smolikov S, Chen Z, Spooner E, Li E, Zhang G, Colaiacovo M & Shi Y (2006) Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases. *Cell* **125:** 467–481

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI & Young RA (2013) Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **153:** 307–319

Williams CJ, Naito T, Arco PG-D, Seavitt JR, Cashman SM, De Souza B, Qi X, Keables P, Andrian Von UH & Georgopoulos K (2004) The chromatin remodeler Mi-2beta is required for CD4 expression and T cell development. *Immunity* **20:** 719–733

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M,

Arand J, Nakano T, Reik W & Walter JOR (2011) 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Communications* **2:** 241–8

Xiao B, Jing C, Wilson JR, Walker PA, Vasisht N, Kelly G, Howell S, Taylor IA, Blackburn GM & Gamblin SJ (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* **421:** 652–656

Xie S, Wang Z, Okano M, Nogami M, Li Y, He WW, Okumura K & Li E (1999) Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene* **236:** 87–95

Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL & Ren B (2012) Base-Resolution Analyses of Sequence and Parent-of-Origin DependentDNA Methylation in the Mouse Genome. *Cell* **148:** 816–831

Yan X-J, Xu J, Gu Z-H, Pan C-M, Lu G, Shen Y, Shi J-Y, Zhu Y-M, Tang L, Zhang X-W, Liang W-X, Mi J-Q, Song H-D, Li K-Q, Chen Z & Chen S-J (2011) Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nature Genetics* **43:** 309–315

Yen RW, Vertino PM, Nelkin BD, Yu JJ, el-Deiry W, Cumaraswamy A, Lennon GG, Trask BJ, Celano P & Baylin SB (1992) Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Research* **20:** 2287–2291

Yildirim O, Li R, Hung J-H, Chen PB, Dong X, Ee L-S, Weng Z, Rando OJ & Fazzio TG (2011) Mbd3/NURD Complex Regulates Expression of 5-Hydroxymethylcytosine Marked Genes in Embryonic Stem Cells. *Cell* **147:** 1498–1510

Yoshida T, Hazan I, Zhang J, Ng SY, Naito T, Snippert HJ, Heller EJ, Qi X, Lawton LN, Williams CJ & Georgopoulos K (2008) The role of the chromatin remodeler Mi-2 in hematopoietic stem cell self-renewal and multilineage differentiation. *Genes & Development* **22:** 1174–1189

Zentner GE & Henikoff S (2013) Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol* **20:** 259–266

Zhang K, Williams KE, Huang L, Yau P, Siino JS, Bradbury EM, Jones PR, Minch MJ & Burlingame AL (2002) Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol. Cell Proteomics* **1:** 500–508

Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W & Jeltsch A (2010) Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Research* **38:** 4246–4253

Zilberman D, Coleman-Derr D, Ballinger T & Henikoff S (2008) Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456:** 125–129

Zuber J, Shi J, Wang E, Rappaport AR, Herrmann H, Sison EA, Magoon D, Qi J, Blatt K, Wunderlich M, Taylor MJ, Johns C, Chicas A, Mulloy JC, Kogan SC, Brown P, Valent P, Bradner JE, Lowe SW & Vakoc CR (2011) RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478:** 524–528