The Epigenome of Multiple Myeloma

From genome-wide analysis to pharmacological manipulation

ALBA ATIENZA PÁRRAGA
Nowadays epigenetic dysregulation is known to play a crucial role in virtually all cancers. In multiple myeloma (MM), an extensively heterogeneous malignancy, the key common feature among patients is the gene silencing imposed by the PRC2 complex through the addition of H3K27me3. This thesis focuses on the exploration of the MM epigenomic landscape, with an emphasis on both the interplay between H3K27me3 and other epigenetic tags, and on the effects of a series of inhibitors altering this profile.

In paper I we provide the genome-wide H3K27me3 distribution unique to MM and demonstrate that the silencing of genes in the profile correlates with an advanced and poor-outcome disease. Reduction of H3K27me3 using the EZH2 inhibitor UNC1999 reactivates genes with anti-tumor activity and induces apoptosis in vitro. EZH2 inhibition also leads to downregulation of the MM oncogenes IRF-4, BLIMP-1, XBP-1 and c-MYC. Paper II identifies miR-125a-3p and miR-320c, predicted to target these oncogenes, as part of the PRC2 targets induced upon treatment.

In addition, H3K27me3 can be recognized and bound by the PRC1 complex. In paper III we show that inhibition of PRC1 using PTC-209 induces apoptosis and this is further enhanced when PTC-209 is combined with UNC1999. Moreover, PTC-209 has been previously shown to reduce the expression of c-MYC. Combined treatment using PTC-209 and JQ1, demonstrated to downregulate c-MYC, results in additive and synergistic effects in reducing MM cell viability.

In paper IV we present the first catalogue of genomic regulatory regions in normal plasma cells, as predicted by their combinations of histone marks. Using this, we demonstrate that in MM a subset of TSSs and enhancers become targeted by H3K27me3 and display high DNA methylation, pointing towards a possible silencing. Conversely, poised TSSs lose H3K27me3 and seemingly become de novo activated. Furthermore, we show that EZH2 physically interacts with the DNA methyltransferase DNMT1 and that combined inhibition using UNC1999 and the DNA hypomethylating agent AZA blocks the G2/M arrest triggered by AZA and induces apoptosis.

In summary, this thesis highlights the complex interconnectivity of epigenetic mechanisms in MM and provides proof-of-principle of the anti-MM effects derived from inhibiting epigenetic components in single or combinatorial regimens.

Keywords: Multiple myeloma, epigenetics, Polycomb, EZH2, miRNA, DNA methylation, inhibitor.

Alba Atienza Párraga, Department of Immunology, Genetics and Pathology, Experimental and Clinical Oncology, Rudbecklaboratoriet, Uppsala University, SE-751 85 Uppsala, Sweden. Science for Life Laboratory, SciLifeLab, Box 256, Uppsala University, SE-75105 Uppsala, Sweden.

© Alba Atienza Párraga 2019

ISSN 1651-6206
ISBN 978-91-513-0650-6
urn:nbn:se:uu:diva-377438 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-377438)
“Dear, dear! How queer everything is today! And yesterday things went on just as usual. I wonder if I’ve been changed in the night? Let me think: was I the same when I got up this morning? I almost think I can remember feeling a little different. But if I’m not the same, the next question is, Who in the world am I? Ah, that’s the great puzzle!”

- Alice in “Alice’s adventures in Wonderland” by Lewis Carroll

To my family, who made me (this).
To my dad (the problem-solver);
to my mom (the curious);
to my brother (the perseverant).
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


* Indicates equal contribution
† Indicates shared senior and corresponding authorship

Reprints were made with permission from the respective publishers.
Related work by the author

## Contents

Introduction ................................................................................................... 13

Epigenetics .................................................................................................... 14
  Mechanisms of epigenetic regulation ....................................................... 16
  The epigenome – the genome’s blueprint ............................................. 21
  Reverting epigenetic tags ....................................................................... 24

Multiple myeloma ......................................................................................... 28
  Molecular pathogenesis .......................................................................... 29
  Genetic architecture .............................................................................. 30

Epigenetics of multiple myeloma ................................................................. 33
  MicroRNAs .......................................................................................... 34
  DNA methylation .................................................................................. 34
  Histone modifications ......................................................................... 36

Present investigations .................................................................................... 38
  Paper I. ............................................................................................... 38
  Paper II. ............................................................................................ 39
  Paper III. ........................................................................................... 40
  Paper IV. ........................................................................................... 41
  A bird’s-eye view ................................................................................. 43

Acknowledgements ....................................................................................... 45

References ..................................................................................................... 48
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGO2</td>
<td>argonaute 2</td>
</tr>
<tr>
<td>AID</td>
<td>activation-induced cytidine deaminase</td>
</tr>
<tr>
<td>AML</td>
<td>acute myeloid leukemia</td>
</tr>
<tr>
<td>AZA</td>
<td>5-azacytidine</td>
</tr>
<tr>
<td>BET</td>
<td>bromodomain and extra-terminal motif</td>
</tr>
<tr>
<td>BETi</td>
<td>BET inhibitor</td>
</tr>
<tr>
<td>BIM (BCL2L11)</td>
<td>Bcl-2-like protein 11</td>
</tr>
<tr>
<td>BLIMP-1 (PRDM1)</td>
<td>PR/SET domain 1</td>
</tr>
<tr>
<td>BMI-1 (PCGF4)</td>
<td>B lymphoma Mo-MLV insertion region 1</td>
</tr>
<tr>
<td>BMP6</td>
<td>bone morphogenetic protein 6</td>
</tr>
<tr>
<td>BMSC</td>
<td>bone marrow stromal cell</td>
</tr>
<tr>
<td>BRD2/3/4</td>
<td>bromodomain containing 2/3/4</td>
</tr>
<tr>
<td>BRG1 (SMARCA4)</td>
<td>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB Binding Protein</td>
</tr>
<tr>
<td>CBX2/4/6/7/8</td>
<td>chromobox2/4/6/7/8</td>
</tr>
<tr>
<td>CCND1/2/3</td>
<td>Cyclin D1/D2/D3</td>
</tr>
<tr>
<td>CD81</td>
<td>CD81 Molecule</td>
</tr>
<tr>
<td>CENP-A</td>
<td>centromere protein A</td>
</tr>
<tr>
<td>CGIs</td>
<td>CpG islands</td>
</tr>
<tr>
<td>ChIP-seq</td>
<td>chromatin immunoprecipitation followed by next-generation sequencing</td>
</tr>
<tr>
<td>CSR</td>
<td>immunoglobulin class-switch recombination</td>
</tr>
<tr>
<td>CTCF</td>
<td>CCCTC-binding factor</td>
</tr>
<tr>
<td>DAC</td>
<td>5-aza-2'-deoxycytidine</td>
</tr>
<tr>
<td>DICER</td>
<td>Dicer 1, ribonuclease III</td>
</tr>
<tr>
<td>DNMT</td>
<td>DNA methyltransferase</td>
</tr>
<tr>
<td>DNMTi</td>
<td>DNMT inhibitor</td>
</tr>
<tr>
<td>DOT1L</td>
<td>DOT1 like histone lysine methyltransferase</td>
</tr>
<tr>
<td>DROSHA</td>
<td>Drosha ribonuclease III</td>
</tr>
<tr>
<td>EED</td>
<td>embryonic ectoderm development</td>
</tr>
</tbody>
</table>
ENCODE

eRNA
EZH1/2
FDA
FGFR3
G9a (EHMT2)
GC
GWAS
H2AK119ub
H3K4me1
H3K4me3
H3K9me3
H3K27ac
H3K27me3
H3K36me2/3
H3K79me1/2/3
HAT
HDAC
hESCs
HGP
HP-1
IDH1/2
IGH
IL-6
IRF-4
KDM
KDMi
KLF2
KMT
L3MBTL2
LINE-1
lncRNA
MAFB
MDS
MEL18 (PCGF2)
MGUS
miRNA
MLL (KMT2A)
MM

ENCyclopedia of DNA Elements
enhancer RNA
enhancer of zeste homolog 1/2
food and drug administration
fibroblast growth factor receptor 3
euchromatic histone lysine methyltransferase 2
germinal center

gene-wide association study
histone H2A lysine 119 monoubiquitylation
histone H3 lysine 4 monomethylation
histone H3 lysine 4 trimethylation
histone H3 lysine 9 trimethylation
histone H3 lysine 27 acetylation
histone H3 lysine 27 trimethylation
histone H3 lysine 36 di-/trimethylation
histone H3 lysine 79 mono-/di-/trimethylation
histone acetyl transferase
histone deacetylase
human embryonic stem cells
human genome project
heterochromatin protein 1
isocitrate dehydrogenase 1/2
immunoglobulin heavy chain locus
interleukin-6
interferon regulatory factor 4
histone lysine demethylase
KDM inhibitor
kruppel like factor 2
histone lysine methyltransferase
lethal(3)malignant brain tumor-like protein 2
long interspersed nuclear element 1
long non-coding RNA
MAF BZIP transcription factor B
myelodysplastic syndrome
Polycomb group ring finger 2
monoclonal gammopathy of undetermined significance
microRNA
mixed lineage leukemia 1
multiple myeloma
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSET (NSD2)</td>
<td>multiple myeloma SET domain containing protein</td>
</tr>
<tr>
<td>MYC</td>
<td>avian myelocytomatosis viral oncogene homolog</td>
</tr>
<tr>
<td>NPC</td>
<td>normal plasma cell</td>
</tr>
<tr>
<td>PC</td>
<td>plasma cell</td>
</tr>
<tr>
<td>PCL</td>
<td>plasma cell leukemia</td>
</tr>
<tr>
<td>PRC1/2</td>
<td>Polycomb repressor complex 1/2</td>
</tr>
<tr>
<td>PROTAC</td>
<td>proteolysis-targeted chimera</td>
</tr>
<tr>
<td>PYGO2</td>
<td>Pygopus family PHD finger 2</td>
</tr>
<tr>
<td>RING1A/B</td>
<td>ring finger protein 2</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosine methionine</td>
</tr>
<tr>
<td>SAT-a</td>
<td>satellite alpha</td>
</tr>
<tr>
<td>SHM</td>
<td>somatic hypermutation</td>
</tr>
<tr>
<td>SMM</td>
<td>smoldering myeloma</td>
</tr>
<tr>
<td>snoRNA</td>
<td>small nucleolar RNA</td>
</tr>
<tr>
<td>snRNA</td>
<td>small nuclear RNA</td>
</tr>
<tr>
<td>SOCS-1</td>
<td>suppressor of cytokine signaling 1</td>
</tr>
<tr>
<td>SUZ12</td>
<td>suppressor of zeste 12 protein homolog</td>
</tr>
<tr>
<td>SWI/SNF</td>
<td>SWItch/sucrose non-fermentable</td>
</tr>
<tr>
<td>TET1/2</td>
<td>ten-eleven-translocation 1/2</td>
</tr>
<tr>
<td>TF</td>
<td>transcription factor</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>TSS</td>
<td>transcription start site</td>
</tr>
<tr>
<td>UTX (KDM6A)</td>
<td>ubiquitously transcribed tetratricopeptide repeat, X chromosome</td>
</tr>
<tr>
<td>XBP-1</td>
<td>X-box binding protein 1</td>
</tr>
</tbody>
</table>
Introduction

An average adult is composed of a total of 37 trillion human cells that can be classified in more than 200 different types, assembling all the different structures and tissues in the body. Yet they all carry the exact same DNA sequences.

Very early on researchers realized that the heritability of complex traits – e.g. height, predisposition to type 2 diabetes or Cronh’s disease – estimated by classical family correlation studies could not be fully explained by differences in the DNA sequence – i.e. polymorphisms – only. This led to the manufacturing of the term “missing heritability” [1], accounting for all those additional factors, not buried within the DNA sequence, that contribute to passing those traits on through generations.

So, if DNA cannot explain cell identity and complex trait heritability, what then? Start your music player of choice and listen to your favorite song. Now imagine it being played by a gospel choir or a punk-rock band. The song’s score is the same but the net result is very different. Each cell has its own set of instruments and vocalists, playing their music in a very unique and defining manner. Epigenetics is that set of instruments; some will add a pause – we call these silencers – and some will raise the beat – these are activators. And they all regulate the expression of that music’s score, the DNA.

And just like the instruments in the analogy, whose contribution can be adapted to satisfy the audience in a live concert, so can epigenetic mechanisms reversibly readjust in response to a changing environment. Due to its inherent capability of regulating cell plasticity, epigenetics is involved in a plethora of physiological processes. However, when out of tune, the same guardians that maintain proper gene expression can be behind the development of a variety of conditions.

The contents of this thesis will focus on how epigenetics contributes to the malignancy of multiple myeloma, a cancer that occurs when normal plasma cells suddenly acquire an unlimited proliferation potential. Key to understanding how these mechanisms are disrupted is the fact that they can be pharmacologically reverted, therefore enabling direct observation of the effects of their presence or absence.

Let us first talk about epigenetics and myeloma and then the epigenetics of myeloma.
Epigenetics

Back in the late 90s, the urge for understanding the underlying determinants of complex human traits led to the birth of a milestone project whose main goal was to provide a consensus human genome sequence that could be used as a reference for identifying trait-causing genetic variants: the Human Genome Project (HGP). Simultaneously, the privately-funded Celera Genomics Corporation announced a more cost-effective and faster sequencing method, starting an arms race with the HGP. Pioneering DNA scientists and world leading bioinformaticians united in one of the most well media-documented scientific efforts of all times, using the most cutting-edge technologies and consuming both private and public budgets. Finally, 2001 witnessed the publication of the first draft of the complete human genome as a joint attainment by the HGP and Celera [2, 3]. Among the discoveries enabled by this was the very shocking – and disappointing – realization that a vast majority of our genome does not contain information translatable into the functional units of life, proteins. In fact, it estimated that only roughly a 2% of the genome is composed of protein-coding sequences. The actual inferred number of genes is in the order of 20,000, equal to that of the simpler nematode C. elegans [4]. If the number of genes does not follow a direct relationship with organism complexity, what then? In 2003, the ENCyclopedia of DNA Elements (ENCODE) Project was launched with the aim of cataloguing all the elements of the genome with a defined function. Whereas the definition of functionality by the HGP was based on the potential for giving rise to a final protein product, the ENCODE project expanded this to include non-protein coding genes, transcription regulating units and elements influencing chromatin structure [5]. The underlying assumption was that complexity was achieved by the well-calibrated regulation of gene expression at a given time-point by non-protein-coding genomic elements. Nowadays we know that the genome is populated by a myriad of defined units – such as promoters, enhancers or insulators – whose interaction with a target gene determines the extent to which that particular gene is expressed. Understanding a transcription factor (TF) as a protein that directly binds to DNA and regulates the binding and/or processivity of the transcriptional machinery, a promoter is a proximal-to-gene TF-binding site while an enhancer is located distal to the target gene. Generally, interac-
tions between proximal promoters and distal enhancers increase the expression rate of a given gene. An insulator, however, can interfere with these interactions and therefore limit gene activity. In other words, physical contacts between a TF and its binding site and between different regulatory elements are at the core of gene expression regulation. It is not surprising to say then that the control of gene activity heavily relies on DNA accessibility.

Figure 1. Epigenetics is a determinant of the tolerance to phenotypic switches. Under physiological conditions (A), a given cell can be driven into only a certain number of phenotypic fates, determined by the tolerance to the action of transcription factors (TF) imposed by the epigenome. However, abnormal epigenetic regulation can lead to an excessively loose genome that allows for illegitimate transitions (B) or to a too rigid structure that hinders the cell’s ability to undergo phenotypic changes (C).

Any given human autosomal cell carries 6 billion base pairs (bp) packaged into 46 chromosomes within their nucleus. If stretched out, this translates into nearly 2 meters of DNA contained in a nucleus of average 3 micrometers (0.000003 meters) in radius [6]. This remarkable packaging level is achieved by several layers of high order folding that allows compartmentalization of the nucleus into more or less active territories. In genetics, a linear double-stranded helix of deoxyribonucleotides is termed DNA. In order to achieve a folded structure, a stretch of ~147bp negatively-charged DNA is tightly wrapped around a core nucleosome composed of positively-charged histone proteins. The resulting DNA-protein composite is referred to as chromatin. Nonetheless, as mentioned earlier, coordinated gene expression is only made possible by matched fluctuations in chromatin accessibility and, therefore, tolerance to TF binding. The probability of a TF binding its cognate locus depends on nucleosomal density but it is also influenced by the chemical modifications experienced by the DNA and the protruding histone tails as well as
the action of non-coding RNAs. These mechanisms of TF binding – and subsequently, gene expression – regulation that do not alter the underlying genetic sequence were baptized as epigenetics and are known to be stable over cell division, propagating the expression patterns to the daughter cells.

Epigenetics plays an essential role in determining cell fate. After fertilization, the parental epigenetic profiles are reset in the zygote in order to provide totipotent cells with the capability of assuming different epigenetic conformations and committing into a specific cell lineage. In this scenario, highly proliferative totipotent cells gradually decrease their replication rate while progressing through a certain differentiation path. In principle, the alternative paths that a given totipotent cell can follow are, at least, as many as the number of different cell types composing the entire organism. However, the likelihood of entering any particular differentiation route depends on the combination of TFs present and their recruitment to target loci, as well as the chromatin permissibility to those TFs. Therefore, while TFs drive the gene expression changes necessary for phenotype transitions, the chromatin environment determines the resistance imposed to those changes (Fig. 1A). Once the cell commits into a particular lineage, the epigenetic apparatus will lock the genome and stabilize the phenotype across cell divisions, preventing spurious reversal to the undifferentiated state. However, such a determining phenomenon is prone to severe consequences when proper regulation is lost; an overly loose chromatin may enable multidirectional cell transitions (Fig. 1B), whereas excessive chromatin compaction may lead to an inability of the cell to progress into differentiation (Fig. 1C). Hence, illegitimate chromatin regulation can have dreadful consequences. Indeed, recent comprehensive analysis of the driver mutations across 33 cancer types identified 43 different chromatin regulator proteins as bearing point mutations or small insertions/deletions [7]; the number of dysregulated epigenetic mechanisms, therefore, could be substantially higher if other alterations were to be considered. Nonetheless, the genome is nothing of a rigid structure but rather a very dynamic fluid that needs to adjust to the specific cellular needs. For instance, physiological processes like transcription, recombination, replication and mitotic condensation demand very different modes of chromatin compaction within the same cell. In addition, the cell must be capable of accommodating the potential changes in the environment by triggering specific expression programs. Therefore, epigenetic regulation of TF activity ought to be a very dynamic system of tightly, yet adjustably, controlling gene expression.

Mechanisms of epigenetic regulation

Since the first mention of the term epigenetics in an attempt to describe how gene interactions – or epistasis – drive pluripotent cells into a determined fate
[8], considerable progress has been gained in identifying the mechanisms that actively modify chromatin structure and TF binding to their target loci. Nowadays it is broadly accepted that epigenetics is not confined to developmental processes. Given its role in regulating all type of gene activity, it is an essential part of most physiological cellular processes, including senescence and integration of stimuli, but also the key to gene dysregulation in many diseases. As to date, most of these are categorized as to affecting 1/ the nucleosome density and position; 2/ the effects of non-coding RNAs; 3/ the covalent addition of methyl groups to the DNA sequence; and 4/ the post-translational modification of histone tails.

**Nucleosome remodeling.** Intuitively, if chromatin is bundled around the protein body of the nucleosome, the density and the turnover rate of these bodies can have a direct impact on chromatin compaction and accessibility. It is well established that regions of open chromatin have a tendency to display low nucleosomal density and high turnover, while the opposite is true for compact genomic regions [9]. Various remodeling proteins use the energy released by ATP hydrolysis to mobilize, exchange or alter the nucleosome structure – thereby temporarily exposing the underlying DNA sequence. Nucleosome remodeling is not only essential in the binding of TFs but it is also necessary during transcription elongation where, in concert with histone chaperones, nucleosomes need to be dislodged in order for RNA polymerase II to proceed along the gene [10]. In addition, the heavily chemically-modified histones contained in the nucleosome can be exchanged by unmodified proteins to reset the epigenetic landscape; this is, for instance, the case of nucleosome exchange by protamines during spermatogenesis [11, 12]. The core of the nucleosome is composed of two units of each the histones H2A, H2B, H3, and H4; however, nucleosome remodelers may also direct the switch of these by non-canonical histone variants that index the genome for specialized functions. The exchange of histone H2A for H2A.X is involved in the sensing of DNA damage and the recruitment of the repair machinery [13], whereas the replacement of histone H3 for CENP-A marks centromeric regions and participates in kinetochoore assembly during chromosome segregation [14]. Due to their multiple roles in regulating chromatin structure for adequate transcription, replication and DNA repair, mutations in nucleosome remodelers have been implicated in a variety of human syndromes [15-17], including several types of cancer [18, 19].

**Non-coding RNAs.** One of the main outcomes from the ENCODE project was the annotation of more than 19,000 transcribed, non-protein-coding – yet functional, as proven by a plethora of forthcoming studies – RNAs [5]. The inventory of such RNAs encompasses tRNAs – that bring the right amino acid during translation –, rRNAs – that associate with a set of proteins to form the
RNA molecules, including functional ribosomes, snRNAs involved in intron splicing, snoRNAs that participate in the processing of pre-rRNA, eRNAs, miRNAs, and lncRNAs. Non-coding RNAs, and in particular miRNA, provide unique examples of true epigenetic inheritance across generations via transfer from the sperm or the oocyte to the zygote [20, 21]. While the role of enhancer RNA (eRNA) is only beginning to be understood, microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) have had a richer portrayal in research. MiRNAs are transcribed as longer pri-miRNAs that are cleaved by DROSHA in the nucleus into a hairpin-shaped precursor that is then exported to the cytoplasm. There, DICER performs a second cleavage in order to generate a mature, ~22 nt long, miRNA that assembles in a RISC complex with the Argonaught protein (AGO2) [22]. The miRNA seed sequence – the 5’-most 8 nucleotides – defines target specificity for the given mRNA. However, the short length of the recognition site entails certain promiscuity; thus, every miRNA is believed to have multiple mRNA targets and, in turn, most mRNA are thought to be regulated by multiple miRNA. The most canonical effects of this targeting are mRNA degradation and translation inhibition, when complementarity is complete or partial, respectively [23]. Therefore, miRNA targeting is inexorably associated with functional mRNA silencing in the cytoplasm. In contrast, long non-coding RNAs (lncRNAs) have been assigned a number of additional mechanisms of expression control. Although it is unknown whether all of them are ultimately functional, most of the transcribed non-protein-coding sequences belong to the category of lncRNA [24]. These comprise a diverse group a long regulatory RNAs – more than 200 nt in length – that can localize to both the cytosol and the nucleus, are spliced, and that may or may not be polyadenylated [25] and still are not capable of producing a protein product. Due to their nature, lncRNAs have the inherent capability of adopting different structures and mediate varied interactions, partly accounting for the tremendous number of functions they have been ascribed. One of this is the interaction with histone modifying enzymes that, as further described below, play a key role in determining chromatin structure and TF binding. For instance, a variety of lncRNAs have been documented to interact with the histone modifying complex PRC2 [26]. LncRNAs may also act as scaffold units, bringing two separate factors together as in the case of the androgen receptor and the protein PYGO2, promoting target expression [27]. In other scenarios, they can also act as decoys – or sponges – and prevent protein function as in the case of the lncRNA GAS5 and the glucocorticoid receptor [28].

DNA methylation. DNA methylation is the historical icon of epigenetic regulation since the early times of its identification as the first mechanism of propagating expression programs across cell divisions. Mechanistically, it involves the covalent addition of a methyl group to the fifth carbon of cytosine in a context of CpG dinucleotides and it is mediated by a group of enzymes generally termed DNA methyltransferases (DNMTs). While DNMT3A and
DNMT3B are regarded as de novo methyltransferases, DNMT1 binds hemimethylated DNA and perpetuates the methylation profile after replication. The methyl groups are provided by the S-adenosine methionine (SAM) donor, the intracellular levels of which are highly dependent on the availability of diet-derived precursor molecules [29], exemplifying the connection between the environment – i.e. nutritional status – and the effects on epigenetic regulation. The inverse process of demethylation can result from a passive dilution of the mark after consecutive rounds of replication or by an active process of demethylation via oxidation by the TET enzymes [30]. Over the years it has become increasingly clear that the consequences of DNA methylation on gene expression are largely dependent on the genomic context. Overall, the mammalian genome has a lower-than-expected CpG content due to their intrinsic tendency to deaminate. However, certain gene promoters carry stretches of dense CpG clusters – named CpG islands or CGIs – that, when unmethylated, mark active transcription [31]. The gene bodies of active genes are, in contrast, highly methylated and this has been suggested to regulate elongation [32] and splicing [33] as well as to prevent activation of cryptic intragenic transcription [34]. Beyond these, the genome is globally CpG-poor and methylation of distal regulatory elements limits their functionality, partly by interfering with the binding of the corresponding regulatory factor. For instance, the presence of DNA methylation at insulators disrupts the binding of CTCF, a boundary-defining element that prevents promoter-enhancer interaction, thereby increasing the expression of the target gene [35]. In addition, DNA methylation strongly anti-correlates with chromatin accessibility, preventing TF binding to their cognate binding site, including proximal promoters and distal enhancers [36]. However, whether DNA methylation is instructive or a consequence of TF binding is unclear. Because of it being a veteran in epigenetic research, DNA methylation has been studied in an interminable catalog of settings and has been found to be altered in a multitude of conditions. Early life trauma [37], exposure to endocrine disruptors [38] or to alcohol [39] among a variety of other environmental insults have been documented to associate with changes in DNA methylation patterns. Interestingly, global loss of DNA methylation is a common feature of a majority of cancers and it has been suggested as a mechanism of increasing chromatin permissibility, thereby promoting stochastic gene activation and allowing cancer cells to sample between different transcriptional states [40].

**Histone modifications.** As briefly pointed out earlier, eukaryotic DNA is condensed in the nucleus into chromatin fibers consisting of 147 bp of DNA wrapped around histone proteins, forming the basic subunit of a nucleosome. Each nucleosome contains two units of the 4 core histone proteins H2A, H2B, H3 and H4, while the nucleosome-linker histone H1 stabilizes the assembly and folding of nucleosomes into high-order structures [41, 42]. The histone’s N-terminal domain protrudes from the nucleosome and is subject to a number
of post-translational modifications. The most thoroughly-investigated modifications are histone acetylation and methylation. Nonetheless, a number of other histone tags have been reported, including phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, and deamination [43], and the list continues to expand, now comprising histone tail clipping as well [44, 45]. Adding to this, recent evidence indicates that the lateral surface of histone proteins, in intimate contact with DNA, can suffer an array of chemical modifications too [46]. Initially thought to control DNA-templated processes – including transcription, replication, recombination and repair – by directly regulating chromatin compaction, we know now that histone barcoding also mediates its effects in an indirect manner by directing or prohibiting the binding of effector proteins. Histone acetylation is deposited by a group of enzymes called histone acetyltransferases (HATs) and erased by histone deacetylases (HDACs). The addition of acetyl groups – deriving from the cofactor acetyl-CoA – neutralizes the histone’s positive charge that maintains its tight bond with DNA, thus relaxing the interaction. In addition, acetylated histones can act as landing surfaces for proteins containing an acetyl-binding bromodomain. As a general rule, histone acetylation associates with open chromatin and active transcription. On the contrary, histone methylation does not alter the net histone charge and therefore its consequences are not as clear-cut. Likewise, the fact that sequential moieties (mono-, di- and tri- methyl) can be added to a single histone residue makes the understanding of methylation substantially more troublesome. Although methylation has been observed on various residues in both histone H3 and H4, lysine methylation on histone H3 is by far the best characterized. Proteins involved in depositing the mark are generally termed histone lysine methyltransferases (KMTs) whereas the erasers are histone lysine demethylases (KDMs). Illustrating the complex nature of histone methylation is the association of tri-methylation on histone H3 lysine 4 and 36 – i.e. H3K4me3 and H3K36me3, respectively – with actively transcribed genomic regions and H3K27me3 and H3K9me3 with tightly sealed, silenced chromatin. The output in gene expression triggered by these histone tags fundamentally relies on the recognition by specific factors containing a methyl-lysine binding domain, such as PHD fingers, chromodomains or Tudor, PWWP and MBT domains. For the sake of the good understanding of the coming chapters, let us now focus on a particular histone-modifying group of proteins with a key role in cellular differentiation and lineage fidelity – two mechanisms profoundly compromised in cancer.

The Polycomb group proteins (PcG) were first discovered in *Drosophila melanogaster* as critical components in the regulation of development and body formation and they are now known to play an essential role in mammalian embryonic cells by controlling key differentiation pathways. Moreover, the PcG proteins have also been involved in stem cell differentiation, X-chromosome inactivation and cellular senescence. Two major multimeric com-
plexes are now well-defined in mammals; the PRC1 and the PRC2. The inherent complexity of the PRC1 complex is being widely recognized although the impact on gene regulation has not been fully delineated. The canonical PRC1 complex consists of RING1A/B, a chromodomain-containing CBX subunit and either MEL18 (PCGF2) or BMI-1 (PCGF4), constituting the PRC1.2 and PRC1.4, respectively. The chromodomain subunits recognize and bind to the PRC2-mediated H3K27me3, while RING1A/B bears the catalytic activity and deposits the H2AK119ub histone mark. A number of CBX subunits (CBX2, CBX4, CBX6, CBX7 and CBX8) incorporate in a mutually exclusive manner. In hematopoietic stem cells, the expression of CBX7 enhances self-renewal, while CBX2, CBX4 and CBX8 are involved in differentiation [47]. Further increasing the complexity, alternative assembly of the components gives rise to non-canonical PRC1 complexes, such as those lacking CBX or those including KDM2B (PRC1.1) or L3MBTL2 (PRC1.6). The incorporation of KDM2B targets the complex to unmethylated CpG islands independently of H3K27me3 [47]. In turn, PRC2 is comprised of the core components EED, SUZ12 and one of the enzymatic subunits EZH2 or EZH1. The latter contain a SET domain responsible for the transfer of methyl groups from the SAM donor to histone H3K27. In mammalian cells, PRC2 recruitment to target genes is believed to be mediated by protein partners and non-coding RNAs [48]. The direct consequences of this epigenetic modification are not yet totally understood but hypotheses point to the recruitment of silencing factors (e.g.: PRC1 and nucleosome remodelers), interference with deposition of activation marks or transcription factor eviction. Moreover, PRC2 has been suggested to associate with several other repressive proteins with epigenetic modification abilities, such as HDACs, KDMs and DNMTs [49]. In vertebrate embryonic and somatic stem cells, PRC2 sets particular genes in a “poised” state, ready for activation or repression depending on lineage-specific signals. These loci are regarded as “bivalent” genes due to the concomitant presence of the H3K27me3 repressive mark and the H3K4me3 activation mark and are involved in stem cell identity maintenance [50].

The epigenome – the genome’s blueprint

Codes are ubiquitously present in our lives – let it be language codes, religious guidelines, laws, the value of money or any other sort of code. This had for very long kept culture aside from nature, being the former just a set on randomly-assigned rules and the latter a direct link to physical and chemical properties. However, the discovery of the genetic code in the 60s threatened this widely-extended belief. Long thought to be constraint by the rules of physics and chemistry, this assumption has, however, proven difficult to validate [51]. The emerging field of Code Biology describes itself as “the study of all codes of life, from the genetic code to the codes of culture, with the standard methods
of science” [52]. Codes are defined as a set of rules that translate a sign into a meaning and have to fulfill the following criteria: 1/ the code must bridge two independent worlds – e.g. the highway code connects street signals to the driver’s behavior; 2 – there has to be a bridging unit between the two worlds, i.e. an adaptor; 3 – the set of rules must be arbitrary in such a way that, if manipulated, they still maintain a code, perhaps with a different meaning. In the genetic code, triplets of nucleotides transcribed from DNA to mRNA – codons – are recognized by complementary nucleotides – anticodons – on tRNA molecules loaded with a specific amino acid. In this scenario, a sequence of codons (sign) determines a polypeptide (meaning), one being an array of nucleotides and the other of amino acids (two independent worlds), brought together by a ribosome (adaptor) through an association that is not dependent on physical or chemical laws (it is arbitrary). This provides the most eminent example of an organic code. Nonetheless, complex organisms seem to carry a number of other organic codes. And so is the case of the epigenetic code (Fig. 2).

Epigenetics indexes the genome into regulatory elements that are more or less accessible to TFs and orchestrates chromatin folding and in so doing ultimately controls gene expression. The epigenetic tags (signs) are recognized by a DNA-binding element (adaptor) that directly or indirectly regulates the activity of the transcriptional machinery (meaning). Nonetheless, epigenetic tags are not isolated events; instead they come in a variety of combinations. Generally, similar epigenetic marks – e.g. associated with active chromatin – cluster together, possibly preventing the strong consequences of accidental loss of one of the marks. For instance, active promoters are very frequently painted with H3K27ac and H3K4me3 and devoid of DNA methylation. Nonetheless, antonymous combinations are also relatively common in situations of a need for versatility. It is the case of bivalent genes, bearing both the positive H3K4me3 and the negative H3K27me3 marks in embryonic stem cells, where upon differentiation, these will resolve in either a fully active or fully silenced locus. Consequently, an atlas of all the genome-wide epigenetic combinations – the epigenome – can provide tremendous information of the underlying patterns of gene expression; just like a house’s blueprint can provide the instructions for the construction, the epigenome can lay down the basis for coordinated genomic activity [53].
Figure 2. The combination of epigenetic tags composes a code that indexes the genome’s regulatory units. Epigenetic tags are rarely isolated events; instead, their regulatory effect results from the combination of multiple elements. For instance, enhancers and TSSs are marked by H3K4me1 and H3K4me3, respectively. The activation of this, however, is marked by the presence of H3K27ac and the lack of DNA methylation (grey balloons). In turn, actively transcribed genes carry H3K36me2/3 or H3K79me2/3 and high levels of DNA methylation (black balloons) in their intergenic regions. These loci display a largely relaxed chromatin conformation. In contrast, the TSS of silenced genes is heavily DNA methylated and these genes accumulate H3K27me3 and/or H3K9me3. Inactive genes are located at tightly packed chromatin regions. The table at the bottom depicts the enzymes responsible for the establishment and erasure of these epigenetic tags.

The beauty of such a concerted calibration prompted the merging of multiple layers of epigenomic information in big genomic charts. Following this approach, Ernst and Kellis were able to connect the regulatory elements in the CD4+ T cell genome to the presence of certain histone mark combinations and the patterns of CTCF and RNA pol II binding [54]. Promoter regions were unambiguously marked by H3K4me3 and a majority also carried numerous histone acetylations. Transcribed regions were enriched for H3K4 methylation (mono- and di-), H3K36me3 and H3K79 methylations (mono-, di- and tri-). Large-scale repressed and heterochromatic regions were found associated with H3K27me3 and H3K9me3. More recently, the NIH Roadmap Epigenomics Program generated a catalog of 111 reference epigenomes not limited to histone patterns, but also involving DNA methylation, chromatin accessi-
bility and expression profiling [55]. Indeed, these are all interrelated mechanisms as illustrated by the recruitment of DNMT1 to H3K9me3-marked regions by the HP-1 protein [56] or by the pre-marking of genes for de novo DNA methylation by H3K27me3 in cancer [57-59]. In addition, not only epigenetic marks maintain an intense chat with each other, chromatin looping and 3D nuclear structure also play a key part in the conversation. During the priming of embryonic stem cells into differentiation, as an example, H3K27me3 has been shown to be involved in the formation of long-range chromatin interactions [60].

This illustrates how chromatin tagging generates a lexicon in which each individual mark is only a component of the alphabet, arranged in a particular grammar that generates a vocabulary for the DNA-binding proteins to read and interpret.

**Reverting epigenetic tags**

As conductors of the genomic orchestra, epigenetic mechanisms are behind an extensive number of physiological processes. DNA methylation and histone acetylation have a very well established role in synaptic plasticity, enabling learning and memory [61]. Moreover, DNA methylation patterns have been seen to change in parallel with ageing, leading to the formulation of an epigenetic clock theory that establishes that profiles of DNA methylation can accurately predict individuals’ age [62]. As a refinement to the system, this genome’s conductor is also able to respond to external cues. Diet – e.g. folate intake during pregnancy [63] –, stress – e.g. early-life trauma [64] –, chemical insults – e.g. arsenic exposure [65] – and physical exercise [66], among others, all have an impact on the epigenome. It is, hence, a predictable consequence that alteration of such an exquisite structure is linked to many pathological conditions, and to date these range from autoimmune diseases [67] to neurodegenerative and psychological disorders [68], addiction [69] and cancer [70].

Two basic features make epigenetics a very attractive target for drug development. On the one hand, owing to its fluctuating nature, epigenetics is a reversible system. On the other hand, it is genome-wide, meaning that each single epigenetic mechanism regulates a multitude of loci and thus, disabling one unleashes a cascade of effects and reduces the likelihood of developing resistance. Although advanced molecular techniques have allowed the on-demand manipulation of the genome, we will now focus on drug-based strategies, with a more attainable translation into clinical practice.

Epigenetic mechanisms have been time and again classified in three categories: writers – depositing epigenetic tags –, readers – interpreting and conveying the information in the tags – and erasers – removing the tags.
Targeting the writers. The classical example of drugs against an epigenetic writer are DNA hypomethylation agents that inhibit the function of DNMTs (DNMTi). The first-generation DNMTi, widely used and currently approved for the treatment of myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML), were 5-azacytidine (azacytidine, AZA) and 5-aza-2′-deoxycytidine (decitabine, DAC). These are cytidine analogues that intercalate onto RNA and DNA, where they trap DNMTs and lead to their degradation. Nonetheless, due to their mechanism of action these are potentially toxic and have a limited pharmacokinetic profile. Second-generation agents include prodrugs as well as non-intercalating DNMT inhibitors [71, 72]. Marking of histone proteins has also been widely investigated. Lysine methyltransferases – e.g. EZH2, DOT1L, G9a –, lysine acetyltransferases – p300 and CBP – and arginine methyltransferase can all be specifically inhibited by the use of one or another drug. Due to the high prevalence of alterations in EZH2 – the histone methyltransferase depositing H3K27me3 – in both solid and hematological tumors [73-82], numerous inhibitors with increasing selectivity and potency and minimized toxicity have arisen in the last years. The majority of these compounds rely on the dependence of EZH2 on the SAM donor and compete for its binding pocket being, thus, very specific against the protein. However, EZH1 can substitute EZH2 in the formation of the PRC2 complex; therefore, subsequent compounds were designed to target both EZH2 and EZH1 simultaneously. In line with this, Konze et al. described in 2013 the development of the first orally bioavailable bi-specific EZH2/EZH1 inhibitor, UNC1999 [83]. Other alternatives have involved the allosteric inhibition of the EZH2 partner EED [84] and the disassembly of the entire PRC2 complex by the use of stapled peptides [85]. Inhibitors of the Polycomb group proteins have also extended to PRC1, with PTC-209 directed towards the BMI-1 subunit of the complex [86, 87]. Similarly to EH2, DOT1L – methylating H3K79 – has been found to be a promising target in MLL-rearranged leukemias in clinical studies (NCT02141828 and NCT01684150 trials) while G9a and MMSET (or NSD2) inhibitors – H3K9 and H3K36 methyltransferases, respectively – are still in the preclinical stage [88-90]. The patterns of histone lysine acetylation and arginine methylation are often times found to be altered in a variety of malignancies and selective small molecule inhibitors have been developed in order to revert this as well [91, 92].

Targeting the readers. Epigenetic readers are those with a recognition pocket for a certain epigenetic tag that, upon binding, can directly or indirectly alter the activity of the transcriptional machinery. Members of this category include proteins with a bromodomain – for acetylated histones – or a chromodomain, Tudor or MBT domain – recognizing methylated histones. The bromodomain-containing BET family (BRD2, BRD3, BRD4, and BRDT) has been implicated in a number of malignancies, favoring the initial development of first-
generation pan-BET inhibitors. These competitively bind the protein’s bromodomain thereby preventing its binding to chromatin and its function in TF recruitment. One distinguished member of this family is JQ1, which has been widely used as an oncogene downregulator of, for instance, MYC. This effect has been proposed to be partially explained by the disruption of super-enhancers – i.e. clusters of physically proximal enhancers [93]. Second-generation BETi’s in current investigation encompass member-specific BETi [94], bivalent BETi [95] and proteolysis-targeted chimera (PROTAC)-based small molecules [96] aimed at degrading the target BET protein.

Targeting the erasers. Drug-development against epigenetic erasers has mainly focused on histone deacetylases (HDACs) – removing acetyl groups on histones – and lysine demethylases – removing the methyl marks on histones’ lysines–. The HDAC family contains a multitude of members whose action is not limited to histones but can also deacetylate a variety of other cellular proteins. Therefore, first-generation pan-HDACi come at the price of a partial toxicity. Nonetheless, these have been shown to be of clinical benefit and are in current use for the treatment of peripheral T-cell lymphoma, cutaneous T-cell lymphoma and multiple myeloma [97-99]. One single lysine demethylase inhibitor (KDMi) against the H3K4me1 demethylase KDM1A has reached clinical evaluation and is now in use for the treatment of psychological disorders [100], while others directed towards KDM4B – H3K9me3/H3K36me3 – are still in the pre-clinical setting [101].

Other epigenetic targets. Crucial to the control of gene expression is nucleosome positioning and density and members of the SWI/SNF chromatin remodeling complex are amongst the most frequently mutated in cancer. Most of these mutations result, however, in the loss-of-function of the complex. Therefore, the strategy of protein inhibition has been replaced by synthetic lethality, which exploits the dependency of the cell on the remaining wild-type allele for survival. For example, inhibition of BRG1 (or SMARCA4) has demonstrated potential in a BRG1-mutant lung tumor xenograft model [102]. Another hotspot for mutations in glioma and acute myeloid leukemia (AML) is in the IDH1 and IDH2 proteins [103, 104]. These participate in the tricarboxylic acid cycle and convert isocitrate into α-ketoglutarate, a cofactor of certain dioxygenases including the DNA demethylation enzymes TET1/2. Mutations in IDH1/2, however, produce 2-hydroxyglutarate instead of α-ketoglutarate, leading to a blockage in TET1/2 activity and subsequent DNA hypermethylation. Several inhibitors of the mutant IDH1/2 are now being evaluated [105]. Non-coding RNAs, in turn, are powerful players in the game of genome regulation too. Strategies to manipulate these consist in posing a steric hindrance to their target recognition, directing them for RNase-mediated degradation or, if abnormally missing, reconstituting their function. The molecular devices employed for these purposes span from miRNA mimics – that
attempt to restore lost miRNA function – to antagomiRs – which inhibit the function of the corresponding miRNA – and interference RNA (RNAi) strategies towards malfunctioning lncRNAs. Yet their introduction into clinical practice has been hampered by their low stability and by the need for adequate delivery vessels.

Drugs targeting epigenetic mechanisms – or epidrugs – are endlessly expanding and have proven to have a tremendous therapeutic potential. Since the primary mechanism of epigenetic regulation is to alter the tolerance to transcriptional changes, epidrugs have been indicated as sensitizers of many agents. In addition, epidrugs have been shown to prevent the development of resistance to conventional treatment by blocking the associated transcriptional reprogramming. Furthermore, these molecules provide invaluable tools for the investigation of the effects mediated by their targets.
Multiple myeloma

Multiple myeloma (MM) is the second most common hematological malignancy, with an estimated annual incidence of 114,000 new cases and constituting a 1% of all cancer-related deaths worldwide [106]. Phenotypically, the MM cell resembles a post-germinal center, long-lived, antibody-producing plasma cell (PC) that has, however, retained its proliferative potential. During normal B cell development, naïve B cells originating in the bone marrow (BM) migrate to secondary lymphoid tissues – e.g. lymph nodes – to undergo a process of affinity maturation involving somatic hypermutation (SHM) and immunoglobulin class-switch recombination (CSR) (Fig. 3). These mechanisms, however, require an extensively accessible chromatin structure and favor the introduction of genetic lesions that can eventually lead to an increased proliferation rate.

The MM cell typically resides and expands in the bone marrow, actively secreting monoclonal antibodies and compromising the normal niche for the development of both the myeloid and lymphoid lineages. The defects in normal hematopoiesis result in the anemia, fatigue and elevated risk for infections commonly displayed by MM patients. In addition, the myeloma cells establish extensive interactions with the BM that support growth and increase drug resistance while inducing an osteoclastic phenotype. This phenotypic switch from osteoblasts to osteoclasts leads to bone destruction and hypercalcemia and, in combination with the high levels of secreted immunoglobulins in plasma – i.e. serum paraprotein or M-protein –, to impaired renal function [107].

Current treatment protocols include proteasome inhibitors – e.g. bortezomib –, that take advantage of the high protein load in these cells, immunomodulatory agents – e.g. lenalidomide and thalidomide –, corticosteroids – e.g. dexamethasone and prednisone –, histone deacetylase inhibitors – e.g. Panobinostat – and chemotherapeutics – e.g. doxorubicin, melphalan, cyclophosphamide. In spite of the development of second-generation compounds with increased activity and reduced side-effects, patients often develop a refractory form of the disease and MM is, therefore, still considered incurable [108].
Molecular pathogenesis

The mechanisms underlying the maintenance of the long-lived PC pool in the BM are of particular relevance in the context of MM, being cells capable of self-perpetuating while retaining a terminally-differentiated phenotype. Examination of the specific transcriptional programs regulating the generation of PC has identified BLIMP-1, XBP-1 and IRF-4 as essential mediators involved in the PC maturation process (Fig. 3).

Figure 3. B-cell malignancies arise from cells at different stages of B-cell development. The normal B-cell development program is a process that involves rapid cell proliferation and genomic restructuring during somatic hypermutation (SHM) and class-switch recombination (CSR), essential for antibody’s affinity maturation, in the germinal center (GC) of secondary lymph nodes. However, during these processes cells are vulnerable to the accumulation of genetic abnormalities. Therefore, a variety of malignancies have been associated with different stages of B-cell development. At the plasmablast to plasma cell transition, cells exchange BCL-6 and MYC expression for BLIMP-1, XBP-1 and IRF-4. These factors are not only crucial for plasma cell survival but also play a key role in maintaining the MM cell pool.
BLIMP-1 is a transcriptional repressor that, through interaction with histone deacetylases and histone methyltransferases, regulates the expression of key factors for B cell identity [109-111]. Mice in which BLIMP-1 has been depleted display deficient plasma cell formation and maintenance [112]. In MM, BLIMP-1 is highly expressed and its reduction decreases tumor growth in a plasmacytoma-prone transgenic mouse model [113]. In turn, XBP-1 regulates the immunoglobulin mRNA processing and secretion and participates in the unfolded protein response (UPR), triggered by an increased protein load in the endoplasmic reticulum [114]. The induced expression of XBP-1 leads to the development of an MGUS/MM phenotype in mice [115]. XBP-1 is highly expressed in MM primary cells [116] and is a direct target of the proteasome inhibitors [117], partly accounting for the response to this type of therapeutic agents. In addition, murine models of Irf4 deletion exhibit defective differentiation of memory B cells into plasma cells [118]. In 2008, Shaffer et al. demonstrated that interference with IRF-4 abolishes viability of multiple MM cell lines with different genetic abnormalities [119]. Among direct targets of this transcriptional activator are BLIMP-1 and MYC, the latter implicated in the progression of MM [119, 120].

Additionally, the bone marrow niche and the interactions established between MM and the BM stromal cells (BMSCs) are well known to have an important role in MM pathogenesis. Cell-cell interactions induce the secretion of cytokines involved in growth, survival and drug resistance [121]. Antibodies and kinase inhibitors are currently under clinical trials in MM, exemplifying the therapeutic potential of blocking the signaling pathways activated by the MM-BM interactions.

Genetic architecture

Multiple myeloma is an extensively genetically heterogeneous disease. However, the identification of pre-malignant stages preceding the development of overt myeloma allowed the characterization of genetic events involved in initiation and progression. Nearly all cases of MM develop from an asymptomatic phase of abnormal accumulation of BM antibody-producing plasma cells denoted as monoclonal gammopathy of undetermined significance (MGUS) and further smoldering myeloma (SMM). Moreover, progression of MM can lead to migration at extramedullary sites such as peripheral blood, giving rise to plasma cell leukemia (PCL).

Primary genetic events associated to disease initiation include karyotype aneuploidy and are sub-classified into hyperdiploid and non-hyperdiploid cases (Fig. 4). The hyperdiploid subtype features trisomies of odd-numbered chromosomes while the non-hyperdiploid group is characterized by translocations involving the immunoglobulin heavy chain (IGH) genes. Traditionally, the hyperdiploid group has been regarded as having a more favorable outcome; however recent evidence indicates that only patients harboring trisomy
3 or trisomy 5 display a better overall survival [122]. The translocations found in the non-hyperdiploid subset mainly affect three groups of genes: 1/ cyclin D1 and D3 (CCND1, CCND3), 2/ c-MAF and MAFB and 3/ FGFR3/MMSET. The t(11;14) and t(6;14) translocations place CCND1 and CCND3, respectively, under the control of the IGH enhancer, upregulating their expression and directly affecting cell cycle regulation. C-MAF is the splice variant overexpressed as a consequence of t(14;16) while MAFB overexpression originates from t(14;20) and both act as transcriptional regulators of a number of genes, including CCND2. The t(4;14) translocation involves FGFR3, with a role in the activation of the MAPK/ERK pathway, and MMSET, a transcriptional activator mediating H3K36me3 [123].

Figure 4. The spectrum of MM genetic abnormalities. All MM cases develop from a pre-symptomatic stage called MGUS, characterized by a low level of monoclonal proteins in blood (M-protein) and a low accumulation of plasma cells in the bone marrow. The progression to MM involves a higher M-protein secretion, higher number of malignant cells and the development of symptomatic features (CRAB: Calcium elevation, Renal dysfunction, Anemia, Bone disease). The figure displays the most common genetic abnormalities, with early events categorized as hyperdiploidy or translocations affecting the IGH locus. Secondary events include deletions, translocations and point mutations in a variety of genes, as well as the APOBEC mutational signature, that entails the conversion of C bases to U which are further substituted with T or G.
Strikingly, all these abnormalities converge in the activation of cyclin D genes either by cis mechanisms – i.e. juxtaposition of cyclin D genes to IGH enhancers – or by trans activation – i.e. overexpression of genes regulating cyclin D transcription or activity. This provides a model in which deregulation of cyclin D genes would be an early initiating event in the development of MGUS. Notably, however, induced expression of these genes in murine models fails to recapitulate the MM phenotype and only c-MYC activation in AID-expressing B cells has proven to promote the development of plasma cell tumors [124]. Therefore, primary translocations are regarded as necessary, but not sufficient to drive transformation into MM.

A plethora of mutations at different frequencies have been detected when aiming to identify secondary lesions present in MM but not in MGUS. These include whole chromosome arm gains and losses – i.e. gain of 1q or deletion of 17p, 12p and 1p –, translocations resulting in overexpression of MYC and point mutations in members of important signaling pathways for MM, such as the MAPK and the NF-κB pathways, as well as in genes with a role in DNA repair and epigenetic modifiers [123, 125, 126].

The existence of subclones bearing mutations in different components provides the niche for adaptation that underlies the development of refractory tumors. This was initially highlighted by results showing that, in high-risk MM patients, subclones alternate dominance in response to therapy [127]. In addition to this, two independent studies have recently demonstrated the existence of subclones within the tumor at different stages of differentiation and with distinct drug resistance patterns. The authors suggest that epigenetic gene regulation may play a role in determining the drug-resistant phenotype [128, 129].
Epigenetics of multiple myeloma

In addition to the well-defined genetic aberrations in MM, accumulating data now emphasizes the role of epigenetic dysregulation for disease initiation, progression and response to therapy. As previously discussed, disturbance of epigenetic mechanisms has profound implications for correct cellular functioning and hence, it is commonly observed in a variety of disorders. In MM, both mutations and altered expression profiles have been documented for a number of epigenetic components. Although abnormalities in other epigenetic regulators have been documented in MM, the spotlight will now be on the mechanisms that will come to relevance in the work included in this thesis.

Figure 5. The epigenetic alterations in MM affect oncogenes and tumor suppressor genes. The mechanisms for abnormal regulation of oncogenes and tumor suppressor (TS) genes include altered miRNA expression and DNA methylation profiles, as well as overactivation of several histone modifying enzymes.
MicroRNAs

Abnormal miRNA expression and function in MM has been linked to the presence of certain genetic lesion – e.g. chromosomal translocations and copy number variation – but also to anomalies in the miRNA processing machinery – e.g. DICER and AGO2 – and in the DNA methylation patterns [130-133]. By their mechanism of action, miRNAs can either have an oncogenic role – i.e. when targeting tumor suppressor genes – or have themselves a tumor suppressor activity – i.e. when targeting oncogenes (Fig. 5).

For example, oncogenic miRNAs – oncomiRs – have been shown to be involved in the IL-6 signaling pathway – essential for MM survival – and to govern the activity of cell cycle regulators, pro-apoptotic genes and other well-known tumor suppressor genes in MM [134, 135]. Furthermore, analysis of miRNA profiles have identified miR-20a-5p and miR-148a to have an elevated expression in MM as compared to normal PCs [136]. Among the overexpressed miRNAs, the miR-17-92 cluster, some Let-7 family members, miR-106b-25 cluster and miR-125b have been ascribed a role in controlling the expression of the IL-6 negative regulator SOCS-1 and the pro-apoptotic BIM as well as having a role in promoting angiogenesis and regulating p53 activity [137-141].

In turn, a large collection of miRNAs are found underexpressed in MM as compared to normal PCs as a result of excessive DNA methylation at their promoters [142, 143]. Illustrating this is the case of the miR-125a-3p, miR-198, miR-601, miR-765, miR-877-5p, miR-1290, miR-223-3p, miR-320c and miR-630. All of these have been documented as silenced in MM and as responsive to the DNA hypomethylating action of AZA [137, 143-146]. In addition, other miRNAs with a tumor suppressor profile seem to regulate the activity of the positive cell cycle regulator CCND2 [144].

The multiple roles of miRNAs in MM pathogenesis have stimulated the use of antagomiRs or miRNA mimics in an attempt to elucidate their mechanisms of action and their potential as therapeutic targets in MM. The silencing of miRNAs with oncogenic potential, for instance miR-106b-25 and miR-17-92, by the use of antagomiRs has proven successful in reducing tumor growth in vivo [137]. An additional strategy is the overexpression of tumor suppressor miRNAs, as exemplified by the reduction of tumor volume in vivo after injection of Let-7 mimics [147].

DNA methylation

Recently, whole-exome sequencing of MM patients at diagnosis and targeted sequencing of previously treated MM cases revealed the mutation of numerous epigenetic modifiers, including the TET2 and DNMT3A enzymes. Furthermore, mutations in any of the DNA methylation modifiers correlated with
a shorter overall survival in these patients, indicating the clinical importance of mutations in epigenetic regulators [148].

In MM, as in many other malignancies, the pattern of DNA methylation is characterized by regional hypermethylation embedded in a globally hypomethylated genome [149]. Hypomethylation of repetitive elements such as long interspersed nuclear element 1 (LINE-1), Alu and satellite alpha (SAT-a) correlates with disease progression [150] and poor prognosis [151]. Silencing of repetitive elements is a constraint imposed by selective pressure and it is well conserved across species; aberrant reactivation of these repeats irreversibly increases genomic instability (Fig. 5). In addition, promoter hypermethylation gradually increases in the transformation process from normal plasma cells, reaching its peak at the PCL stage and in cell lines [152]. Among the genes aberrantly hypermethylated in MM are cell cycle regulators – e.g. p16, p73 – and genes involved in apoptosis – e.g. DAPK –, DNA repair – e.g. MGMT and hMLH1 –, and critical signaling pathways – e.g. SOCS-1, RASSF1A, SHP1 and SYK [153-155].

Importantly, the analysis of DNA methylation has traditionally focused on the examination of promoter CGI, but mounting evidence now suggests that the effects of DNA methylation are largely dependent on the genomic context, emphasizing the need for unbiased and genome-wide studies. Following this line, Agirre et al. [156] demonstrated that the progression from normal plasma cells to MGUS and MM is characterized by a gradual increase in hypomethylation, whereas hypermethylation occurs at the MGUS stage and does not increase in MM. Hypermethylation was not found in proximity to genes but mainly occurred outside CGI promoters and corresponded to enhancers in normal immature B cells. These sites were found to be hypermethylated in stem cells and to lose methylation upon normal B cell differentiation, indicating that re-methylation of enhancer regions may play a role in a hypothetical scenario of dedifferentiation of MM cells during transformation. These findings encourage the evaluation of the MM DNA methylome in a comprehensive and genome-wide manner that allows the examination of the interactions with other regulatory regions and/or machineries.

The DNA methylation machinery is, in addition, highly connected to other epigenetic regulatory networks. DNA methylation has been reported to silence numerous miRNAs with oncogenic properties and their reactivation by induced expression or the use of hypomethylating agents has been shown antmyeloma effects [143, 146]. Also, the relationship between DNA methylation and histone marks is particularly intriguing in the context of H3K27me3. This histone tag has been found to preferentially localize at CpG-rich regions and several reports suggest that perturbation of the DNA methylome by interference with the DNMTs may alter the global distribution of H3K27me3 [157, 158]. Moreover, genes marked by H3K27me3 in human embryonic stem cells (hESCs) become DNA hypermethylated in cancer and aged tissues [159]. In line with this notion, a subset of H3K27me3 targets as defined in hESCs were
found to be hypermethylated and to correlate with poor prognosis in MM [160].

Owing to the well-reported changes in DNA methylation in MM, hypomethylating agents have been tested in vitro and in vivo and demonstrate that this strategy shows anti-MM potential either as a single agent or as a sensitizer to other drugs. Currently, several clinical trials using hypomethylating agents are ongoing, many of them combining these with classical MM therapies – e.g. lenalidomide, dexamethasone or HDACi.

Histone modifications

The t(4;14) translocation, present in about 15% of MM cases and associated with an adverse prognosis, results in the overexpression of MMSET (also known as NSD2), a histone methyltransferase responsible for the dimethylation of histone H3 lysine 36 (H3K36me2), a mark associated to active transcription [161]. Induced expression of MMSET has been shown to activate transcriptional programs that promote myelomagenesis [162, 163]. Detailed examination of the effects triggered by t(4;14) on chromatin determined that accumulation of H3K36me2 throughout the genome displaces the silencing mark H3K27me3 to ectopic sites and leads to inactivation of genes expressed in normal germinal center B cells and of known MYC targets in B cells [164].

Due to its role in gene silencing of important cellular programs, PRC1 – in charge of H2AK119ub – has been implicated in several malignancies. For instance, CBX7 has been linked to impaired senescence in several solid tumors and to leukemic transformation [47, 165]. In MM, genome-wide association studies (GWAS) identified the CBX7 as a risk locus for the development of the disease [166]. In addition, the BMI-1 subunit is overexpressed and correlates with poor prognosis in a variety of solid and hematological malignancies. Furthermore, BMI-1 collaborates with c-MYC in driving the development of B-cell lymphomas [167]. In MM, BMI-1 overexpression promotes cell growth and is associated with poor prognosis [168]. BMI-1 knock-down, however, impairs MM cell growth by upregulating BIM and its pharmacological inhibition leads to the re-activation of negative cell cycle regulators, thereby promoting apoptosis and blocking cell cycle progression [168, 169].

The suggested role of Polycomb proteins and the H3K27me3 chromatin mark in maintenance of stemness, along with reports showing its correlation with cell proliferation and resistance to cytostatic drugs, underpins the attention that these proteins have gained in the context of cancer initiation and progression in the recent years. Importantly, EZH2 is a major regulator of germinal center B cells and, upon differentiation towards plasma cells, its expression becomes downregulated [78]. Nonetheless, EZH2 and other components of the PRC2 complex are highly expressed in MM as compared to normal plasma cells [170]. The cause of this induced expression has been suggested
to involve IL-6 signaling, activation by c-MYC or activated EZH2 expression as an effect of the downregulation of its targeting miRNA, miR-26a. The role of EZH2 as an oncogene in MM has been demonstrated by the ectopic expression leading to growth factor-independent proliferation and by growth arrest resulting from its interference [82]. In addition, the H3K27me3-mediated gene silencing increases in advanced stages of MM [170]. Opposing EZH2 activity, the histone demethylase UTX (KDM6A) acts in a complex with KMT2C and KMT2D to remove H3K27me3. Mutations and deletion of UTX correlate with an adverse prognosis and its loss promotes MM proliferation *in vitro*, further emphasizing the role of H3K27me3 in MM pathogenesis [148, 171].

In a recent study, Ishiguro et al. [172] detected a high expression of the active transcription-associated H3K79 methyltransferase DOT1L in MGUS, SMM and MM. The DOT1L inhibitors SGC0946 and EPZ-5676 triggered a reduction in cell proliferation *in vitro* and reduced tumor formation in SCID mice after *ex vivo* treatment. Among the DOT1L targets, the authors identified the downregulation of IRF-4, MYC and BLIMP-1 as partially accounting for the anti-MM effects. In addition, the dysregulation in H3K9 methylation marks has also become apparent by reports identifying G9a – H3K9 methyltransferase – as a potential target in MM mediating the silencing of the pro-apoptotic gene BIM [173, 174] and by the participation of KDM3A – H3K9 demethylase – in the regulation of KLF2 and IRF-4 [175] as well as glycolytic and apoptotic pathways [176].

The multiple abnormalities in histone modification profiles reported in MM have encouraged the design and pre-clinical evaluation of inhibitors directed towards the underlying epigenetic machineries. Although the HDAC inhibitor Panobinostat is currently the only FDA-approved epigenetic treatment for MM, several other epigenetic inhibitors are in different stages of clinical trials. These include BET bromodomain inhibitors, known to impair the expression of MYC, a number of histone deacetylase inhibitors and a single EZH2 inhibitor.
Present investigations

MM is a malignancy responsible for about 80,000 annual deaths worldwide. The relative inefficacy of current therapeutic strategies is rooted in the extensive intra- and inter-patient variability that hinders the design of precision-based treatment. A variety of genetic alterations have been proposed to underlie MM initiation and progression; nonetheless, these are seldom recurrent.

In 2010, however, a major leap in MM research resulted from the identification in our lab of a gene silencing profile as the one common denominator among patients [170]. The genes in this profile were significantly enriched among PRC2 targets and their progressive silencing correlated with advanced stages of the disease. This set the stage for the evaluation of the abnormal epigenetic patterns in MM, with a focus on the consequences of their erasure.

Paper I.

The PRC2 complex is aberrantly active in multiple myeloma and its inhibition re-activates tumor suppressor genes.

As a first approach in understanding the role of epigenetic dysregulation, we aimed at defining the collection of genes marked by H3K27me3 – the mark deposited by PRC2 – in the MM genome and at evaluating the effects of its removal.

By means of chromatin immunoprecipitation coupled to next-generation sequencing (ChIP-seq) we categorized all the PRC2 targets as bearing only H3K27me3 or carrying this mark along with H3K4me3 – i.e. K27-only and bivalent, respectively. PRC2 targets were substantially lowly expressed as compared to H3K4me3-marked genes and were overrepresented among silenced genes in embryonic tissues, pointing to a poorly differentiated phenotype in MM. Illustrating the connectivity between epigenetic mechanisms, PRC2 targets overlapped with genes previously annotated as bound by CBX8 – a component of the PRC1 complex.

Since H3K27me3 marking is also present in normal plasma cells (NPCs), we filtered these genes out in order to obtain MM-unique profiles. By doing this, we found that the silencing of K27-only targets correlated with advanced disease stages and with poor survival, irrespectively of the genetic background. This suggests a possible use as a prognostic factor and highlights the potential of H3K27me3 loss for the treatment of high-risk MM patients.
In turn, bivalent marking was more profound in MM than in NPC, similar to the disease-specific increase in these reported in germinal center-derived B-cell lymphoma [50]. The fact that these genes carry opposing histone marks implies a potential for activity switch given the loss of one of them. Indeed, we found bivalent genes to correlate with genes reactivated by the knockdown of Polycomb proteins.

With all evidence highlighting the potential of H3K27me3 removal in anti-MM treatment, we decided to investigate the effects of two specific EZH2/1 inhibitors – i.e. UNC1999 and GSK343. Treatment with either inhibitor led to reduction in MM cell viability and this was true for both cell lines and samples obtained from newly-diagnosed patients. MM cell lines treated with EZH2 inhibitors displayed an apoptotic phenotype and were less capable of forming colonies in vitro. In fact, analysis of the effects in gene expression revealed a reactivation of apoptosis pathways. Others have linked the blockage in apoptosis activation to unresponsiveness to treatment in MM. EZH2/1 inhibition, thus, presents as an opportunity for treatment-resistant MM patients. Strikingly, reduction in H3K27me3 was also accompanied by a decreased expression in MM oncogenes.

Paper II.

**Inhibition of PRC2 not only induces re-activation of protein-coding genes but also tumor suppressor miRNAs.**

Intrigued by the finding that loss of a silencing pressure – H3K27me3 – triggered gene downregulation (Paper I), here we focused on identifying the subjects on this downregulation and exploring the mechanism behind this.

Envisaging a potential secondary mechanism as responsible for gene downregulating, we exposed MM cell lines to a longer EZH2/1 inhibition, thereby allowing intermediates to operate. This strategy revealed a reduction in 344 genes, including well-established MM oncogenes. XBP-1, BLIMP-1 and IRF-4 play a key role in PC differentiation and their manipulation severely impairs MM viability, while MYC overexpression has been proposed as an important factor in the progression from MGUS to MM [109-120]. This establishes the benefits of EZH2/1 inhibition as to be dual: through direct up-regulation of tumor suppressor genes and through indirect downregulation of oncogenes.

In order to evaluate whether these effects could be mediated by miRNA-directed silencing, we performed miRNA expression profiling in cells treated with the inhibitor. This discovered 118 and 88 up- and downregulated, respectively. This was the first report on H3K27me3 mediating miRNA dysregulation in MM.

Our analysis detected miRNAs with a reduced expression after treatment. These included miR-17-92 and miR-106b-25 clusters and members of the Let-
7 family. These are overexpressed in MM and are regarded as oncogenic due to their regulation of apoptosis, angiogenesis and p53 signaling [137-141].

The induced miRNAs included members found to be underexpressed in MM as compared to NPC due to promoter DNA hypermethylation and to be reactivated upon hypomethylating treatment with AZA [137, 143-146]. This provides yet another example of the interplay between different epigenetic machineries. Target prediction identified miR-125a-3p and miR-320c as directed towards XBP-1, BLIMP-1 and IRF-4, potentially explaining the down-regulation observed in these oncogenes. Interestingly miR-125a-3p is preferentially expressed in proliferating centroblasts and its levels are decreased upon differentiation into PC [177]. This further supports the notion that PRC2 silencing promotes maintenance of proliferation over differentiation. In addition, we found miR-494 expression to be induced upon treatment. This miRNA has been shown to regulate MYC in other B-cell malignancies [178]. This could possibly underlie the downregulation of MYC observed after EZH2/1 inhibition. Examination of the miR-125a-3p and miR-320c loci validated the accumulation of H3K27me3, which was severely lost upon EZH2/1 inhibition.

Paper III.

Combining PRC1 inhibition with PRC2 or MYC synergistically reduces multiple myeloma viability.

The newly-defined PRC2 targets in MM were enriched among targets of the PRC1 subunit CBX8 (Paper I). The CBX family contains proteins with a chromodomain that recognizes and binds H3K27me3, therefore bridging the two Polycomb repressor complexes. The PRC1 complex is, in addition, formed by the assembly of the RING1A/B ubiquitin ligase and its supporter protein BMI-1. The latter has been found overexpressed in MM and to promote cell growth in vitro and in vivo [169, 179, 180]. In addition, high levels of BMI-1 correlate with poor prognosis in MM patients treated with the conventional therapies [168]. BMI-1 has also been found to partner with MYC in lymphomagenesis and loss of BMI-1 downregulates MYC in MM [168, 181, 182]. With this in mind, we determined to examine the role of BMI-1 for MM survival and the effects of its combined inhibition with EZH2/1 or MYC inhibitors.

To this end, we used PTC-209, a compound initially developed to decrease BMI-1 expression. PRC1 inhibition in this manner profoundly altered the viability of MM cell lines and primary samples. Importantly, our collection of primary samples included both newly-diagnosed as well as relapsed patients and the effects were irrespective of refractoriness, highlighting the potential of the strategy in therapy-resistant MM patients. We showed that induction of apoptosis accounted for the reduced viability. The pro-apoptotic gene BIM was induced by PRC1 inhibition, partly explaining the consequences.
We also reported that PTC-209 reduces the BMI-1 protein and, consequently, the H2AK199ub mark, whereas the mRNA levels remain constant. In fact, other BMI-1 inhibitors – i.e. PTC596 – post-transcriptionally regulate the protein by targeting it for hyperphosphorylation and subsequent degradation [183]. We hypothesize that this mechanism may also be at play with PTC-209.

Combining PRC1 with either PRC2 or MYC inhibition – using JQ1 – led to synergistic and additive anti-proliferative effects. As others have pointed out in other scenarios, combinatorial treatment seems to be a valuable strategy in MM therapy.

Paper IV.

The multiple myeloma epigenome is profoundly re-wired, making it more susceptible to combined inhibition of PRC2 and DNMTs.

Our results inevitably pointed towards a complex regulatory network of epigenetic interactions in MM. With the aim of inspecting this, we set up to provide a map of the regulatory regions in NPC as defined by their histone modification patterns and to use this to investigate the genome-wide alterations in MM. Our previous data suggested a connection between PRC2 targeting and DNA methylation (Paper II), therefore we included DNA methylation profiles in our analysis. To further gain insights into the interplay between the two machineries, we tested the effect of combined inhibition.

By exploring the effects of EZH2/1 inhibition in a larger panel of MM cell lines, we noticed that only a fraction of this responds to the treatment. In addition, inhibition-resistant cell lines displayed very limited transcriptional changes. The sensitivity profiles, however, were independent of the genetic background or the basal level of EZH2/1 or H3K27me3. Moreover, H3K27me3 was efficiently depleted in both resistance groups.

Prompted by our findings of H3K27me3 targets overlapping with DNA methylated genes, we envisaged a potential cross-talk between the two machineries. The interplay between PRC2 and DNMT has, however, proven controversial. In osteosarcoma, PRC2 has been shown to directly interact with DNMTs, while other have suggested that the two mechanisms alternate during malignant transformation [57, 184]. Our data indicated that, in MM, EZH2 physically interacts with DNMT1. Therefore, we resolved to examine the effects of combined inhibition of these components.

To this end, we optimized a strategy of prolonged and low-dose treatment with the DNMT inhibitor AZA with the aim of minimizing potential toxicity. Our data indicated that this approach efficiently downregulates DNMTs protein levels and leads to a global loss of DNA methylation. This effect did not correlate with dependency upon external additional of IL-6, shown to control the expression and nuclear translocation of DNMT1 [185, 186]. In spite of this, treatment with AZA triggered G0/G1 and G2/M arrest, possibly arising
from a certain degree of DNA damage, as accounted by the increase in γ-H2AX phosphorylation. Nonetheless, combined treatment with EZH2/1 inhibition blocked cell cycle arrest and induced apoptosis. Novel approaches in cancer treatment include the prevention of G2/M arrest [187], emphasizing the potential of the strategy. At the protein level, we noticed a reduction in H3K27me3 and DNMT1 when using the EZH2/1 inhibitor or AZA, respectively. Nonetheless, combined inhibition showed a tendency towards partially mitigating the reduction in H3K27me3, suggesting a possible compensation from the global loss in DNA methylation.

As a means towards elucidating the interplay between different epigenetic tags, we used ChIP-seq data for the distribution of several of the most well-characterized histone marks as well as chromatin accessibility data – i.e. ATAC-seq. To this, we added expression – RNA-seq – as well as DNA methylation – high-density arrays – data. This generated the first epigenomic profile of the regulatory elements in the NPC genome.

Direct comparison with MM demonstrated that although the general patterns of DNA methylation are maintained, Polycomb and heterochromatic regions display a larger heterogeneity, suggesting the existence of different subclones bearing varying levels of DNA methylation at these sites. We also found that subsets of TSSs and enhancers had a higher median DNA methylation value. Previously, enhancer DNA methylation has been noticed when comparing MM to immature B cells. Here we demonstrate that these remain unmethylated in NPCs and only gain DNA methylation in the transition towards MM.

Moreover, we found an extensive epigenomic reconfiguration in MM. Heterochromatic regions, marked by H3K9me3 in NPCs, exchange this for H3K27me3 in MM. In prostate cancer, a similar switch in epigenetic marking has been documented in late-replicating domains and the authors hypothesized that this may sensitize these regions to DNA damage and repair errors [188]. If this would prove to be true in MM, it would translate into a higher sensitivity towards intercalating agents such as AZA. Regarding H3K27me3 we found an overall increase along the genome except for a focal, selective, enrichment in TSSs and enhancers. H3K27me3-mediated silencing has been proposed to arise as a secondary effect to inactive transcription [189]; however, our data implies a rather directed mechanism. Among the loci showing this pattern we found a region in the CD81 3’UTR. This locus gains H3K27me3 in connection with the loss of active marks and signs of expression in MM. CD81 silencing has been shown to correlate with better overall prognosis and with PC differentiation [129, 190]. This is an intriguing finding and could simply reflect an enrichment in our patient cohort in MM samples with a more differentiated phenotype. However, this remains to be confirmed. Despite the increase EZH2 expression in MM and the overall increase in H3K27me3 we also found loci losing the mark while gaining active histone patterns and expression. We exemplify this with the case of the BMP6 TSS. Bone morphogenetic proteins,
including BMP6, have been found overexpressed in MM [191]. Our analysis provides the mechanistic explanation for this *de novo* activation in MM.

In the next steps, the evaluation of the chromatin changes – by ChIP-seq – as well as in expression patterns – by RNA-seq – triggered by either treatment or the combination will provide further insights on how these two mechanisms cross-talk and how this relates to the phenotypic changes observed – i.e. induction of apoptosis and inhibition of cell cycle arrest. In order to accurately analyze the changes, a spike-in strategy will be followed. This relies on the external addition of material – chromatin or RNA, respectively – in known ratios used to calibrate the measurements. Although this is a well-established strategy in RNA-seq experiments, only recently the awareness of the need in ChIP-seq has started to appear [192].

**A bird’s-eye view**

The work here presented expands the knowledge on the molecular mechanisms underlying MM pathogenesis, portraying malignant cells as dependent on the rewiring of their epigenomic patterns.

The acquisition of a new gene silencing profile mediated by the PRC2 – through the addition of the H3K27me3 mark – is present as a common feature in MM and the progressive silencing of these corresponds to MM cases with an aggressive and poor-prognosis disease. Nonetheless, a key characteristic of epigenetics is its reversibility and doing so prompts MM cells into increasing the expression of tumor suppressor genes and drifting into apoptosis. Importantly, however, epigenetic machineries rarely act in isolation and, in MM, the H3K27me3-marked genes share similarities with previously defined PRC1 targets. Indeed, interference with the two complexes simultaneously triggers a combinatorial reduction in viability. PRC2 inhibition, moreover, is capable of reducing the expression of oncogenes. Among these oncogenes are classical examples of MM dependency for survival, such as IRF-4, XBP-1, BLIMP-1 and MYC. MM cells exploit the H3K27me3-mediated silencing for maintaining miRNAs targeting these factors low. However, loss of the mark leads to the reduction of these essential oncogenes and to cell death. As another example of the interplay between epigenetic regulators, these miRNAs are also silenced by DNA methylation in MM. In fact, DNMT1 appears to interact with the PRC2 component EZH2 and cells treated with the DNA hypomethylating agent AZA display a higher susceptibility to PRC2 inhibition. Genome-wide, the DNA methylation profiles are very similar between NPCs and MM, except for PRC2 targets and heterochromatic regions. These gain further H3K27me3 marks but display a largely heterogeneous enrichment in DNA methylation. This points to a cross-talk between one and another mechanism, however the direction of this remains to be characterized. Loss of H3K27me3 does not appear to affect the distribution of DNA methylation;
however, the inverted effect can only be elucidated by inspecting the epigenome after single or combined inhibition. Furthermore, the epigenomic alterations are not limited to DNA methylation. Certain loci, particularly at TSSs and enhancers become enriched for silencing marks and show patterns of low expression activity, whereas a fraction of poised TSSs display signs of de novo activation. The net result for MM malignancy arising from these changes is, at the moment, unexplored.

The results here presented evidence the complexity of epigenetic regulation, with multiple synchronized mechanism and bidirectional regulation of one another. This is a basic component of physiological epigenomes; however, it also reflects the exponential amplification of abnormalities when individual components acquire an aberrant behavior.
Acknowledgements

Wow, it’s been quite of a journey. Time to unfasten your seat belt. Let me undo mine too and tell you now about something important.

A bit of history: [12-16yo – High school] the one thing that catches my attention is the beauty of biological systems, with bits and pieces all perfectly aligned and working in synchrony. [16-18yo – Upper high school] among the living beings, I choose humans and I choose to understand how that biological perfection is shaken by the appearance of disease. [18-22yo – University of Barcelona] I am fascinated by genetics and, in particular, epigenetics. Between my notes, I find myself underlining (twice!!) the words Polycomb and Trithorax. [22-23yo – University of Bergen] speaking of damaged perfection, there’s no clearer example of things going south than cancer. That’s what I need to know more about: cancer epigenetics. I did, but not all by myself. My most sincere gratitude:

To Helena Jernberg-Wiklund, my main supervisor. Thank you for letting me get closer to my dreams. No one could have ever made me feel more welcome and valued. I remember something that you said once: nobody should ever need to be pat on the back. I didn’t understand or agree at that time. I do now. I don’t need you to pat my back because I feel your trust in me in many other ways. Thanks for that, as a PhD student the scariest part is that your job is no longer memorizing the right answer but finding yourself one and I would have never overcome that fear without feeling your support.

To Fredrik Öberg, my co-supervisor. Having you around has always brought some sort of aura of brilliance. I deeply admire you and I look up to you. You’re a wonderful combination of funny, kindness and knowledge. And your knowledge expands to so many grounds! Thanks for always bringing a new perspective when I got blind at my own data after looking too much into it.

To Antonia Kalushkova Nair, my second co-supervisor. Nothing I could write here would ever do justice to how much you’ve given me. You have nourished me, made me grow, taught me everything and, always, as a friend. With you I transitioned from a fearful student to an independently-thinking researcher. All I have learnt I owe it to you but, most importantly, I owe you my self-confidence. I will always remember all these years but the main thought coming afloat will be how you and I could always discuss any science as equals, from day one. Thank you.
To all my other friends in the HJW lab: Prasoon, Charlotte, Pernilla, Lotta, Mohammad and Patrick. Pernilla and Lotta, you two are awesome, nothing in this lab would have come to reality without your hard work and you are a reference encyclopedia when it comes to experiment trouble-shooting, thank you!! Mohammad, you were the first one to spot me and to open the door to your lab, thanks for giving me the opportunity! Patrick, you’re genuinely smart and funny, which makes you a fabulous lab-mate and friend. We’ve been through a lot of failed experiments and disappointments in the last months, a lot of swearing and hair-pulling, but it’s alright. You always find the right joke at the right time and that has made you irreplaceable! You still have plenty of time ahead, enjoy it and know that you can always count on me, just like I’ve known I could always count on you.

To the other people that, for one reason or another, I have happened to work very close with. To Klev, Sathish, Marco and Marcus. Klev, you and I very often disagree in many things however, that only makes our meetings a lot more fun and our conclusions a lot more flavorful, thanks for disagreeing (not always). Sathish, I can say with a 99% confidence that I now know ChIP inside-out thanks to you. Your project has been challenging but working with you has always been smooth like butter. Marco and Marcus, my two consecutive teaching-mates. The days spent in teaching have been tedious and very long. Yet, you two have managed to turn them into something I actually looked forward to. Thanks for bringing the fun and for allowing me to practice my “control-freakiness” in organizing things my way.

To all my friends, here and somewhere else. Here: Alberto, Khayrun, Milena, Juulia, Argyris, Iliana, Miguel, Sara, Max B., Yavor, Alessandra, Yang, Matko, Priya, Anne, Mahsa, My and a whole lot of other people at Rudbeck and BMC. Thanks for making me feel that I belong and for showing me that I can be whoever I want. To all the people who I did my masters with, especially to Kiki, Maria G., Maria V., Maria K., Ilektra (not in the same program but you were still part of the gang) Vanessa and Hanan, so many memories together! Thanks for not being only classmates but my surrogate family when I had just landed in this country. To Petar, for pushing my boundaries and making me grow as a pair of a character and for making me strive for better day by day. Can’t wait to see what’s coming. To all the people somewhere else (but all the time with me): Laura O., Laura R., Pol, Gerard and Carlos. Laurins, os quiero con locura. De todas las amigas que jamás he tenido o jamás tendré sois las más importantes y las amistades más bonitas. Me muero por volver a veros y por retomar las viejas costumbres que han tenido que esperar todos estos años. Pol, els anys a la uni van ser molt durs per a mi per molts motius però veure’t cada dia i compartir experiències amb tu va ser el meu únic flotador tots aquells anys, gràcies per llavorens i gràcies per continuar a dia d’avui sent un amic fantàstic. Gerard, la nostra historia és llarga i, durant molt de temps, molt dolorosa. Tot i així, estaré eternament orgullosa d’haver sigut capaços de deixar tot allò enrere i poder dir avui que
ets un amic. Sempre seràs part de mi, amb tu vaig aprendre les lliçons de vida més importants i és per tu que vaig decidir anar-me’n d’Erasmus i, en comptes de tornar-me’n, continuar el viatge cap a Suècia. És a dir, gran part del que sóc avui, incloent aquest doctorat, és gràcies a tu. Carlos, mi compañero durante tanto tiempo y mi mejor amigo. Mi balance, mi fuerza, mi consejero y mi apoyo. Hemos pasado por mucho, tú lo sabes, y sin ti nunca hubiese podido con todo esto. De todo lo que Escandinavia me ha dado, tú eres el mayor regalo. ¡Te debo tanto! Toda mi tesis, empezando por la frase de “Alicia en el país de las maravillas” gira alrededor de cambios. Y cambios veo por todas partes. Sin embargo, mi única constante todos estos años has sido tú y sé que eso nunca va a cambiar.

To my loving family. Mama, papa y tete: sois la mejor familia que nunca nadie puede tener. Os quiero y os admiro de aquí a la luna, ida y vuelta. Los retos han sido muchos estos años, pero cada vez que algo se interpone en mi camino pienso: “¡Recuerda quién eres, eres una Atienza-Párraga!” y eso es lo único que me da el valor para seguir adelante. Sois mi mayor fuente de orgullo y os lo debo todo. Gracias es poco porque yo soy por vosotros.

The work included in this thesis was supported by funding from the Swedish Cancer Society, the Swedish Research Council for Medicine, the Hans von Kantzow Foundation, the eSSence program (to Jan Komorowski), the Polish Academy of Sciences (to Jan Komorowski) and grant R01GM103893 (to Jian Jin) from the U.S. National Institutes of Health.
References


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)