universität wien

MASTERARBEIT

Titel der Masterarbeit

„Methylome and Transcriptome Analyses in MO7e and MO7e$^{p210/BCR\text{-}ABL}$ cells“

verfasst von

Thais Topakian, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2015

Studienkennzahl lt. Studienblatt	A 066 877
Studienrichtung lt. Studienblatt	Masterstudium Genetik und Entwicklungsbiologie
Betreut von:	Ao. Univ.-Prof. Dr. Wolfgang Mikulits

Acknowledgement / Danksagung

Am Abschluss meines Studiums angelangt, möchte ich hiermit allen Personen, die mich unterstützt und gefördert haben, meinen herzlichsten Dank aussprechen. Allen voran gilt mein besonderer Dank Univ. Prof. Dr. Sabine Zöchbauer-Müller sowie Dr. Gerwin Heller, die mir ermöglicht haben in einem hervorragenden Labor mit einem überaus netten und hilfsbereiten Team zusammenzuarbeiten. Ganz speziell bedanke ich mich bei Dr. Heller für seine fortwährende Hilfestellung und großzügige Unterstützung während der gesamten Zeit. Für die äußerst gute und entspannte Atmosphäre möchte ich mich außerdem besonders bei meinen lieben Arbeitskolleginnen Barbara Ziegler und Mag. Corinna Altenberger bedanken, die mich stets gut beraten, mir in jeder erdenklichen Situation geholfen haben und mit welchen es auch außerhalb des Arbeitsplatzes immer Spaß gemacht hat, Zeit zu verbringen. In diesem Zuge danke ich auch herzlich den weiteren Arbeitskollegen unseres Großraumlabors Mag. Erwin Tomasich und Dr. Maximilian Marhold für ihre Hilfsbereitschaft und die lustige Zeit.

Ich bedanke mich außerdem bei Dr. Wolfgang Mikulits für seine Betreuung und Benotung dieser Masterarbeit.

Desweiteren gilt mein besonderer Dank meiner Familie und meinen Freunden, insbesondere meinen lieben Eltern für ihre Unterstützung in allen Bereichen und Lebenslagen.

Table of content

Ich habe mich bemüht, sämtliche Inhaber der Bildrechte ausfindig zu machen und ihre Zustimmung zur Verwendung der Bilder in dieser Arbeit eingeholt. Sollte dennoch eine Urheberrechtsverletzung bekannt werden, ersuche ich um Meldung bei mir.

1 Scientific Background

1.1 Epigenetics

1.1.1 History

In the 1950ies, Conrad Waddington introduced the term epigenetics, which originates from the greek word “epigenesis” and had been used to describe all developmental events that give rise to a mature organism (1). Waddington published his model of “the epigenetic landscape” which illustrates how different gene expression is connected to cell type differentiation (2). In spite of his pioneering idea, Waddington didn’t understand the underlying mechanisms yet. Thus, epigenetics was not a focus of research for the next 30 years but regained importance in the mid-1970ies and early 1980ies with the detection of DNA methylation and histone modifications as mechanisms for regulating transcriptional gene expression. At that time, Arthur Riggs proposed that DNA methylases have a gene-regulatory function. More substantial evidence supported this theory in 1980 when Jones and Taylor used the DNA methyltransferase inhibitor 5-azacytidine to link methylation and gene expression. Almost contemporary, Michael Grunstein and colleagues demonstrated the essential role of histone-tails in the regulation of gene expression (5). With the detection of nucleosome remodelling complexes and the role of non-coding RNAs as their possible recruiters, epigenetics has evolved to a generally accepted subdiscipline of molecular biology (6-13). Nowadays epigenetics is defined as a “stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence” (14).

1.1.2 Chromatin structure

Considering the size of the human genome which is comprised of 6 billion base pairs per cell, the DNA necessarily is packaged in three levels organizing and establishing the chromatin structure. Mainly, DNA is wrapped around a positively charged group of histone proteins (15). An octamer of these histones including 2 copies of H2A, H2B, H3, and H4, respectively, forms one nucleosome (16). Each nucleosome contains a stretch of approximately 166bp DNA and displays the basic unit that is repeated throughout the structure (17). To build higher-order chromatin, the nucleosomes coil into a short fiber termed “30-nanometer fiber”. They are stabilized by histone H1 binding to the “linker DNA” which represents the DNA in between adjacent nucleosomes (18). During cell cycle these fibers are further packaged to form coils and loops in order to build an even more densely packaged metaphase-chromosome (19).

Two major forms of chromatin are distinguished. Euchromatin depicts the less condensed chromatin structure that is accessible for factors of the transcription machinery (transcription factors, RNA polymerase), while heterochromatin is densely packed and transcriptionally inactive (Fig.1) (20). Moreover, heterochromatin is further split in a constitutive and a facultative heterochromatin (21). Although both are associated with transcriptional inhibition, facultative heterochromatin can regain transcriptional activity in a temporal context for processes such as X-inactivation or tissue specification while constitutive heterochromatin ascertains genomic integrity by stably silencing repetitive elements, centromeres and telomeres (22-24).

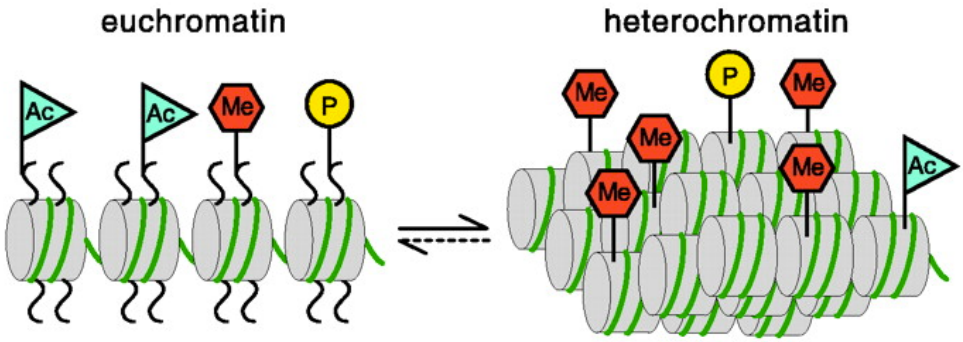


Figure 1. The difference between euchromatic and heterochromatic domains. Euchromatin is accessible for transcription marks and constitutes the open chromatin structure. Heterochromatin carries repressive marks inhibiting transcription. Adapted from (25).

1.1.3 Epigenetic mechanisms and their role in the regulation of DNA accessibility

Unlike the largely static feature of the DNA sequence, the epigenome can be altered by environmental signals and these changes can be passed to the next cell generation (26). The networks controlling the conformational state of the DNA include different mechanisms of DNA methylation, post-translational modifications of histones (PTMs), histone variants, chromatin remodelling and non-coding RNAs (Fig. 2) (5, 6, 27-29). These modifications contribute to maintain genomic stability and to shape a cells gene expression pattern.

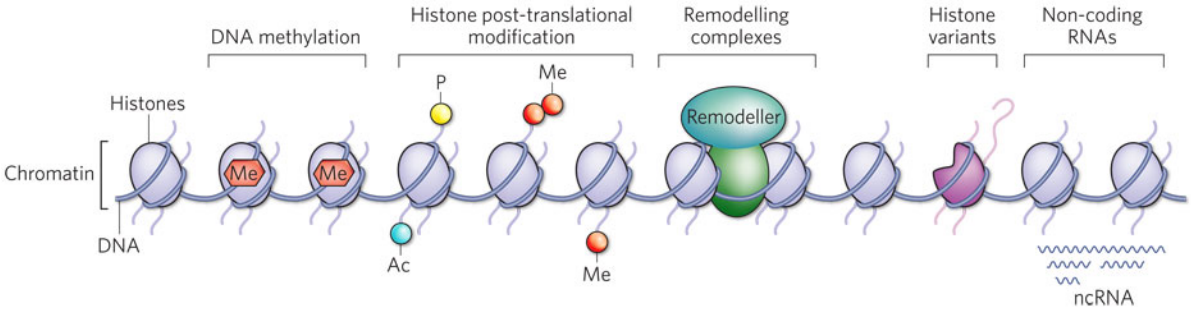


Figure 2. The implication of epigenetic mechanisms in chromatin regulation**.** Five interrelated mechanisms are known to regulate chromatin structure and affect cellular processes like development and differentiation: DNA methylation, post-translational modifications of histones, remodelling complexes, insertion of histone variants and non-coding RNAs (30).

Epigenetic mechanisms regulate transcriptional gene expression. DNA methylation in promoter regions is associated with transcriptional gene silencing (31). Modifications of histone proteins may result either in gene activation or gene inactivation depending on the specific modification and the affected histone residue (32). Considered as the main histone modifications, methylation, phosphorylation and acetylation play a key role in chromatin organization (33-35). Acetylation exclusively takes place at lysine residues of histones 3 and 4 and is associated with transcriptional gene activity (36). Major targets for histone methylation are lysines and arginines. Depending on the methylated histone residue and the number of methyl-groups added, histone methylation results either in gene activation or silencing (33, 37). Similar to acetylation, phosphorylation of serines, threonines and tyrosines induces a conformational change by introducing negative charges and frequently results in a less condensed, transcriptionally

active chromatin. However, phosphorylations at specific residues also appear to be involved in chromosome condensation during cell division and through crosstalk with other PTMs in further biological processes such as DNA repair, apoptosis or activation of signalling transduction (38, 39).

Histone variants are described as the non-canonical histones and are mainly found in single gene copies with constitutive expression. Compared to canonical histones, histone variants slightly differ in their primary amino acid sequence and show strong diversity within eukaryotes. Exchange of histones with histone variants can be matched to gene expression, DNA repair, chromosomal segregation or kinetochore assembly (40). Figure 3 summarizes histone-tail PTMs and histone variants associated with a specific transcriptional state.

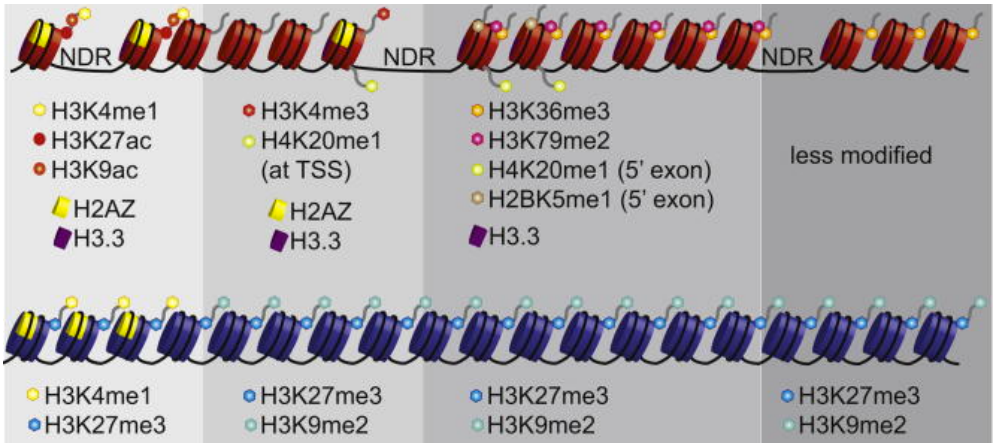


Figure 3. Combinations of histone PTMs and variants regulating transcription. DNA wrapped around red histones is accessible for transcription factors and the transcriptional machinery, while DNA densely packed with blue histones is inaccessible (41).

Incorporation of new histones or histone variants requires ATP-dependent chromatin remodellers which can remove, exchange or slide histones along the DNA (42). Just recently, focus is put on non-coding RNAs by demonstrating that they are able to recruit chromatin-modifying enzymes (43).

In order to understand the role of the chromatin structure and its regulation of transcription, replication and damage repair, the interaction of all these mechanisms, their coordination in time and their reversibility should be considered. So far, studies particularly highlighted the relationship of DNA methylation and histone modification. Fuks et al. reported the binding of DNA methyltransferases (DNMTs) to histone deacetylases (HDACs) and demonstrated the interaction of

DNMTs with the histone methyltransferase SUV39H1 leading to an expression-repressive state (44, 45). Additionally, other studies demonstrated how the transcriptional repressive histone mark H3K36me3 enables the binding of DNMT3a to the chromatin and that histone deacetylation facilitates DNA demethylation (46, 47). Besides others, Robertson and Wolffe illustrated a basic model of epigenetic transcription repression demonstrating that DNA methylation can trigger histone deacetylation (48) (Fig. 4).

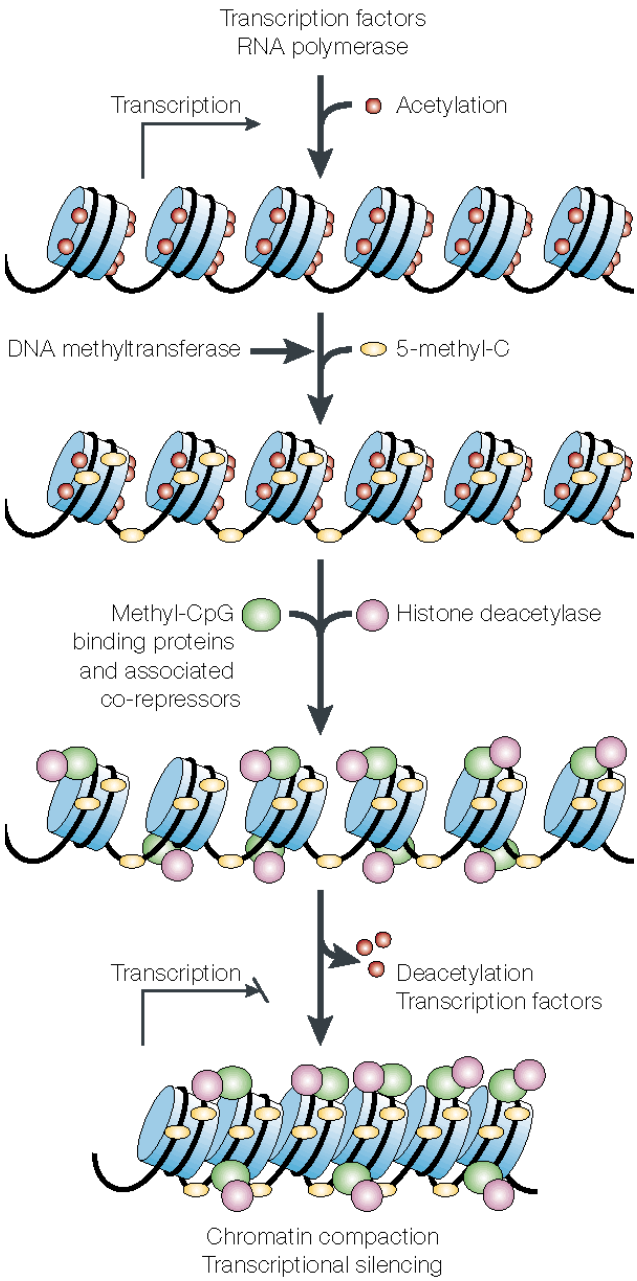


Figure 4. Cooperation of DNA methylation and histone deacetylation for transcriptional inhibition. To transcriptionally repress a genomic sequence, DNA gets methylated and Methyl-CpG binding proteins and co-repressors such as histone deacetylases are recruited. Adapted from (48).

1.1.4 DNA methylation

1.1.4.1 The methyltransfer reaction

In mammalian cells, DNA methylation predominantly occurs within CpG dinucleotides and is catalysed by DNA methyltransferases (DNMTs) which transfer a methyl-group from S-Adenosyl-methionine (SAM) to the 5'-carbon of cytosines. As proposed by Zangi et al. the DNA acts as substrate but also as a co-factor of the biochemical reaction (Fig. 5) (49). To enable the methylation reaction, DNMT flips the cytosine out of the DNA helix and modifies it in its active centre (50). Through a nucleophilic attack of the enzyme on C6 a covalent bond is formed favouring a further nucleophilic attack of C5 on SAM and the transfer of the methyl-group. To release the DNMT, a subsequent deprotonation of C5 and elimination of the enzyme nucleophile takes place (49).

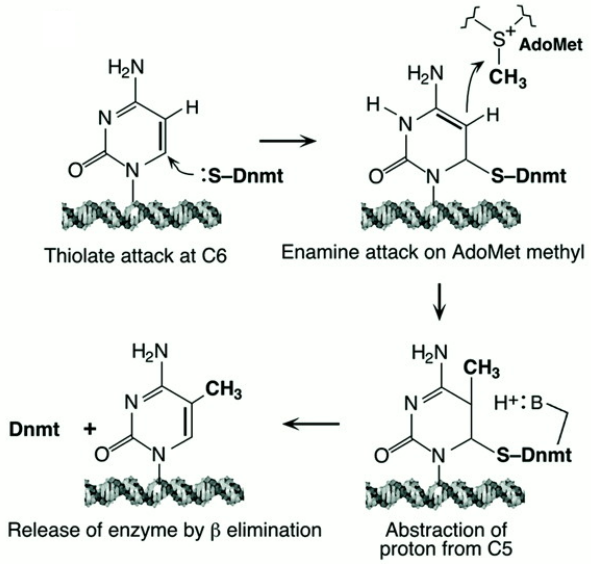


Figure 5. Chemical reaction of DNA methylation. By the nucleophilic attack of the DNMT's sulphur atom on cytosine C6, C5 attacks the methyl-group donor SAM (or AdoMet) and enables the transfer of the methyl-group ($-CH_3$). Subsequent β-elimination releases the enzyme-cytosine interaction. Adapted from (51).

1.1.4.2 The DNA methylation machinery

To set up cellular DNA methylation patterns, two classes of DNA methylation are distinguished: maintenance and de novo methylation.

DNMT1 was found to have an increased preference for hemi-methylated DNA (52). In comparison to other DNMT enzymes, DNMT1 is the most abundant in somatic cells and occurs relative to other cell cycle phases in an especially high level during S-phase (53, 54). DNMT1 co-localizes to the proliferating cell nuclear antigen (PCNA), slides along with the replication fork and copies the methylation marks from the original to the newly synthesized strand (55, 56). As a result, DNMT1 maintains the methylation patterns throughout cell division and restores the original methylation pattern in case of DNA damage (57).

More recently, the strongly conserved family of de novo DNMTs (DNMT3a, 3b) has been characterized. In contrast to DNMT1, de novo enzymes are highly expressed in germ cells, embryonic stem cells and embryos but in a lower extent in somatic cells. Thus, their major function is the initial establishment of developmental- and tissue-specific methylation patterns and the immobilization of newly incorporated transposons (58, 59). In addition, they are thought to harbour a proofreading function in case of methylation errors left by DNMT1 (60). Figure 6 summarizes the operating mechanisms of de novo and maintenance DNMTs.

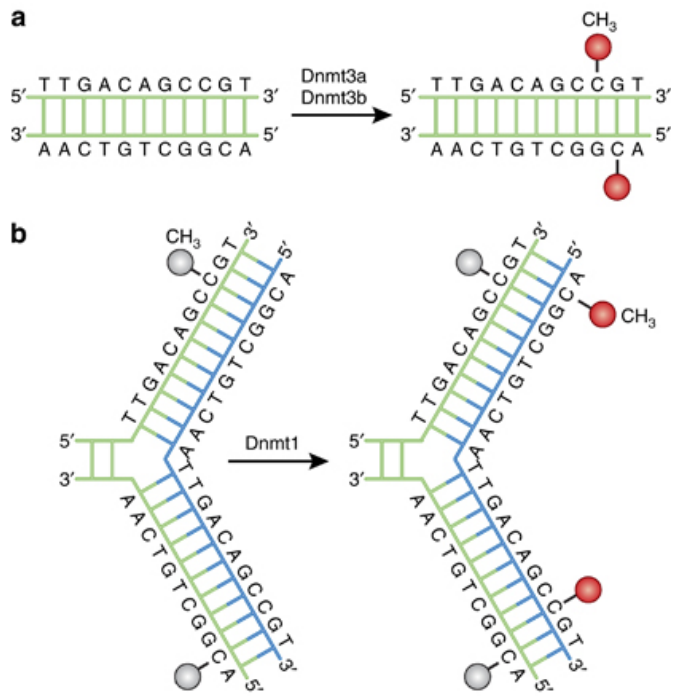


Figure 6. Pathways of DNA methylation. **(A)** Methyltransferases DNMT3a and 3b operate de novo on unmethlyated DNA. **(B)** DNMT1 maintains methylation marks by copying them to the newly synthesized strand during DNA replication (61).

The importance of DNMTs was highlighted in murine DNMT1 knockout and knockdown studies demonstrating the severe effects on X-inactivation, genomic imprinting and embryonic developmental processes (62-64). Mouse embryos lacking DNMT1 showed a genome-wide loss of DNA methylation resulting in reactivation of genomic imprints of both parental genomes and of the X chromosome that has undergone inactivation during early development (65, 66). Similar to DNMT1 deficient embryos which die early within embryonic development, DNMT3b and 3a deficiency results in death in the embryonic stage or shortly after birth (67).

1.1.4.3 DNA methylation is reversible - mechanisms of DNA demethylation

While somatic cells mainly keep their DNA methylome, global DNA demethylation plays an important role in the development of embryos and primordial germ cells (PGCs). Before PGCs reset their methylation pattern during gametogenesis according to the sex of the developing embryo, they primarily undergo a complete eradication of 5-methylcytosines (5mCs) (68). A similar process can be observed during embryogenesis, where a constitutive loss of 5mCs can be monitored in both parental genomes until the early blastocyste stage (69, 70). In addition, DNA demethylation occurs during differentiation according to the tissue's identity as it was shown for hematopoietic cells, myocytes, neurons and hepatocytes (71-74). Several studies suggested active and passive underlying mechanisms and highlighted that their impairment might strongly contribute to tumorigenesis (75-79).

1.1.4.3.1 Passive demethylation

Passive demethylation relies on the lack of DNMT1 activity during several rounds of replication resulting in a proceeding dilution of 5mCs. This mechanism was linked to the occurrence of 5-hydroxymethylcytosines (5-hmCs) which are established through oxidation of 5mCs by a group of proteins termed ten-eleven translocation proteins (TET) (80). TET proteins 1-3 are highly conserved in metazoans and particularly show abundance in oocytes, zygotes, hematopoietic cells and embryonic stem cells (80-85). In addition, TET proteins are capable of

inducing further modifications of 5-hmCs such as 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (86, 87).

Several scenarios have been suggested linking the active modification of 5mCs to a passive dilution during replication. A plausible model proposes that the methylation machinery does not recognize the oxidative derivative given the fact that DNMT1 activity is reduced on hemi-hydroxymethylated DNA *in vitro* (88). This is further supported as the signal for TET-mediated oxidation of 5mCs increases in the paternal genome after fertilization which is accompanied by elevated levels of 5fC and 5caC. Thus, due to impaired cognition of the modified and methylated bases, DNA methylation is reduced after several cell cycles (83, 89).

1.1.4.3.2 Active demethylation

Active demethylation describes the direct removal of 5mCs. Similar to passive dilution of 5mCs which has been linked to TET activity, active mechanisms can depend on the 5mC oxidation by TET proteins. Thus, TET proteins can catalyse enzymatic modifications of 5mCs to 5fCs and 5caCs which are deaminated, excised by glycosylases and repaired by the base excision repair pathway (BER) or the nucleotide excision pathway (NER) (90, 91).

1.1.4.4 Methylation distribution throughout the genome

Approximately 1% of all nucleotides in DNA of human somatic cells are 5-methylcytosines (5mC). 70-80% of all CpG dinucleotides within the human genome are methylated (92). Depending on the genomic location, methylated cytosines are either associated with transcriptional gene activity (gene bodies) or gene silencing (CpG islands, CGI) (Fig.7) (93, 94).

Although DNA methylation in promoter regions is generally associated with gene silencing, the occurrence of DNA methylation within gene bodies which are considered as the regions downstream of exon 1 was demonstrated to positively affect gene expression (93, 95, 96). Several studies reported that gene body methylation enhances transcription and consider it as the main cause to prevent transcription initiation on intragenic promoters (97, 98).

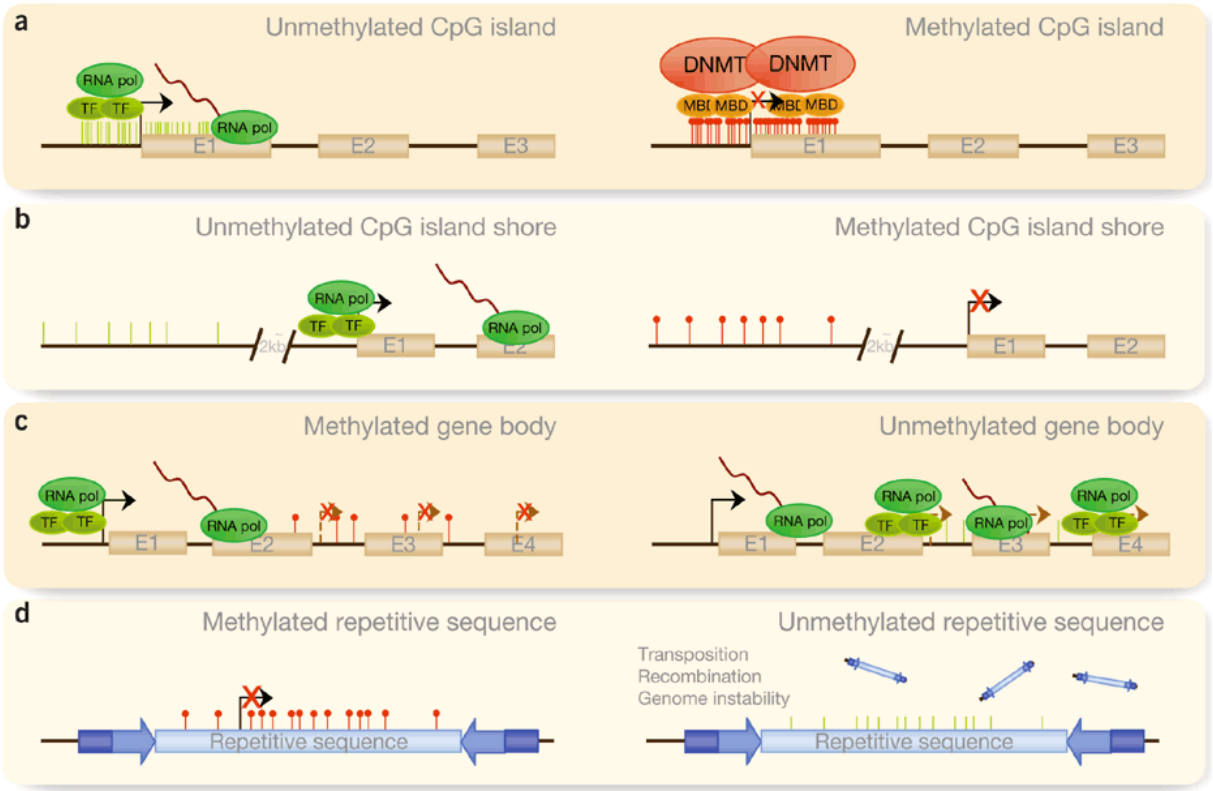


Figure 7. DNA methylation patterns of specific genomic regions. **(A)** An unmethylated CGI upstream of a gene (left) leads to transcriptional inhibition, if methylated (right). **(B)** CpG island shores, located 2kB upstream of CGIs, lead to same results as illustrated in (A) after methylation. **(C)** Methylated gene bodies (left) loosing their CH3-marks (right) lead to unregulated and incorrect transcription. **(D)** Initially methlyated repetitive sequences (left) result in chromosomal recombination, transposition and genomic instability due to methylation loss (right) (99).

CGIs are defined as stretches of DNA with a length of at least 200bp and a GC content higher than 50% (100). CpG dinucleotides are underrepresented in genomes of species that extensively use DNA methylation for gene regulation, since methylated cytosines are frequently converted to thymines. Repair of the resulting TpG mismatches restores the CpG dinucleotides or results in TpA dinucleotides. Due to the two repair possibilities, CpG sites are found in lower than expected frequencies. CGIs are therefore defined with a CpG observed versus expected ratio (CpG O/E ratio) of less than 0.65 (101-103). Moreover, CGIs are usually located next to transcription start sites (TSS) and overlap with the promoters of 60-70% of human genes. Although the majority of CpG sites within the genome is methylated, CpG dinucleotides within CGIs are usually not methylated in normal cells with exception of CGIs associated with X-linked, imprinted or tissue-specific expressed genes (104-106). Thus, genes associated with non-methylated CGIs are frequently connected to a transcriptionally permissive chromatin structure (107). Although the responsible mechanisms for

maintaining the majority of CGIs unmethylated remain to be clarified, specific CGI sequence properties, static binding of TFs preventing the association of DNMTs or CpG-binding proteins targeting CGIs for demethylation are considered (108-111). Figure 7 illustrates how the methylation state of a specific genomic region affects gene expression.

1.1.4.5 Mechanism of DNA methylation mediated gene silencing

DNA methylation suppresses gene transcription by directly or indirectly interfering with the binding of transcription factors (TFs). Thus, the presence of CpG methylation can negatively affect their affinity towards TF recognition sequences (112). In addition, transcriptional silencing can be co-accomplished by methylcytosine binding proteins (MBP) which lead to condensed chromatin after recruiting co-repressors and histone deacetylases (113-115). Figure 8 summarizes different models for transcriptional regulation by DNA methylation.

Based on comparative studies investigating MeCP2 domain conservation, several human proteins were currently found to share a methyl-CpG binding domain (MBD) including 2 histone methyltransferases (SETDB1, SETDB2), 2 histone acetyltransferases (BAZ2A, BAZ2B), MeCP2/MBD proteins (MBD1, MBD2, MBD3, MBD4, MeCP2), proteins of the Kaiso family (Kaiso, ZBTB4, ZBTB38) and 2 proteins containing a SRA-domain (UHRF1, UHRF2) (116, 117). Despite MBD conservation, not all of the MBPs selectively bind methylated DNA but collaborate with each other or other proteins to form heterochromatic structures. Thus, members such as MBD3 inhibit gene expression by association with histone deacetylases and nucleosome remodellers, while MBD1, MBD2 and MeCP2 posses a transcriptional repression domain (TRD) (114, 115, 118). Additionally, based on the number of cysteine-rich CXXC domains, MBD1 isoforms with 3 motifs (MBD1v1, MBD1v2) were shown to bind DNA regardless of CpG methylation, while isoforms with 2 CXXC motifs (MBD1v3, MBD1v4) were found to silence methylated promoters only (119).

The determinants for MBP distribution and association throughout the genome are not clear yet. One scenario hypothesizes a random occupation of methylated CpG sites that is just dependent on the abundance of MBP family members and the genome-wide methylation density, while others propose

sequence-specificities beyond methylated CpGs of MBPs or MBP associated factors. Indeed, evidence indicates the presence of consensus sequences for MBD family members that allow for discrimination between different methylated genomic regions (120, 121).

Detailed underlying mechanisms how MBPs establish and maintain DNA methylation remain unclear. However, growing molecular evidence demonstrates that cancer cells exploit these proteins in order to exert transcriptional control. Thus, MeCP2, MBD1, MBD2 were found to be involved in gene silencing in a variety of cancers including lung, colorectal and haematological malignancies (122-128).

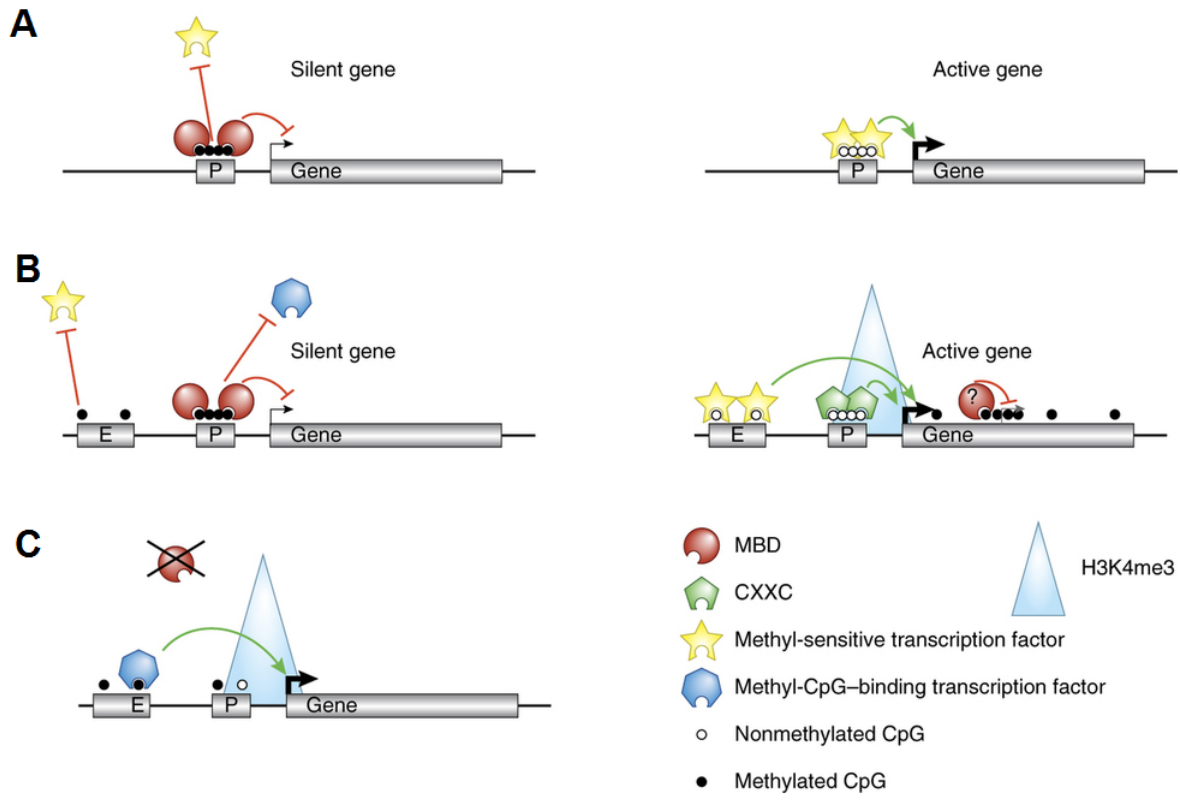
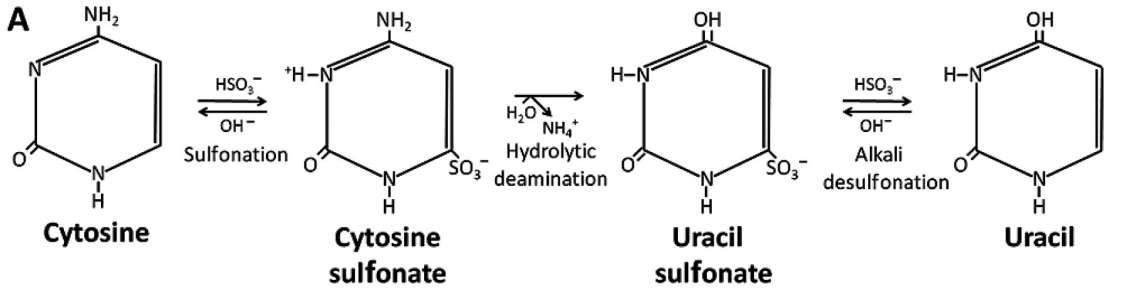


Figure 8. Possible methylation-dependent mechanisms regulating transcription. **(A)** MBPs bind methylated CpGs in promoter regions resulting in silencing of downstream genes (left), while non-methylated CpGs are associated with TFs allowing gene transcription (right). **(B)** Association of methylated CpGs with MBPs in promoters and methylation of enhancer elements block transcription factor binding and gene expression (left), while association of TFs with enhancers and proteins containing a CXXC domain with promoters permit access of activating histone modifications and allow gene expression. **(C)** Low-CpG-dense but methylated promoters are observed not to be bound by MBPs resulting in transcription initiation (129).

1.1.5 Methods to detect DNA methylation

Notable breakthrough in DNA methylation analysis was achieved by applying sodium bisulfite (SB) on DNA for further analytical procedures (Fig. 9). SB treatment converts cytosines into uraciles but leaves methylated cytosines unaffected. Upon PCR uraciles are further turned into thymines while the originally methylated cytosines remain cytosines (130-132).



B

	Original sequence	Sequence after bisulfite treatment
Unmethylated DNA	A-T-**C**-**G**-G-T-C-A-T-**C**-**G**-C-A-T	A-T-**U**-**G**-G-T-U-A-T-**U**-**G**-U-A-T
Methylated DNA	A-T-**C**-**G**-G-T-C-A-T-**C**-**G**-C-A-T	A-T-**C**-**G**-G-T-U-A-T-**C**-**G**-U-A-T

Figure 9. Sodium bisulfite treatment of DNA. **(A)** The chemical reaction starts with sulfonation of C5 followed by hydrolytic deamination to generate uracil sulfonate, which is desulfonated to uracile. **(B)** SB converts cytosines into uraciles if orginially unmethylated but does not affect methylated cytosines (133).

1.1.5.1 Gene-specific approaches

1.1.5.1.1 Methylation-specific Polymerase chain reation (MSP)

MSP depends on the conversion of DNA with SB and the design of two primer pairs that discriminate between methylated and unmethylated DNA. To analyse methylation of a specific genomic region, the SB-treated DNA is applied for two independent PCR reactions using either the primer pair complementing methylated DNA or the primer pair complementing unmethylated DNA. If the sequence was originally methylated, amplification is only observed with methylation-specific primers but not with primers specific for unmethylated DNA

and vice versa. Thus, MSP allows to claim whether a sequence is methylated or not but forbids quantitative methylation statements (134).

1.1.5.1.2 Combined Bisulfite Reaction Analysis (COBRA)

COBRA is a quantitative assay to measure DNA methylation of specific gene loci by applying genomic DNA to SB treatment and PCR amplification with sequence-specific but methylation insensitive primers. Based on the original methylation status of the sequence, the resulting amplicons might have lost or retained CpG-containing restriction sites (CGCG) upon PCR. The fragments are then digested with a restriction enzyme (e.g. BstUI) sensitive for CGCG and separated by gel electrophoresis. The band pattern and their intensities reveal information about the degree of methylation by comparing the intensities of the undigested sequence with that of digested sequences (135).

1.1.5.1.3 Methylation-sensitive high resolution melt analysis (MS-HRM)

MS-HRM is a sensitive, fast and quantitative method to determine DNA methylation levels of a specific gene locus. It is based on SB treatment of genomic DNA and a PCR reaction using methylation insensitive primers. During PCR, the fluorescent dye EvaGreen intercalates into double-stranded PCR amplicons exclusively. After PCR completion, the temperature is raised which leads to denaturation of the amplicons. Due to decreasing fluorescence levels, a melting curve is obtained. The melting of each amplicon depends on the initial methylation state that is reflected by different AT and GC compositions after SB treatment. While AT base pairs are bound by just 2, GC dinucleotides are connected with 3 H-bonds revealing a higher melting temperature for sequences with elevated GC content. Thus, melting profiles accurately differentiate between strong and low methylated sequences (Fig.10). To calculate methylation levels, MS-HRM has to be simultaneously performed with DNA standards with known methylation degree (136, 137).

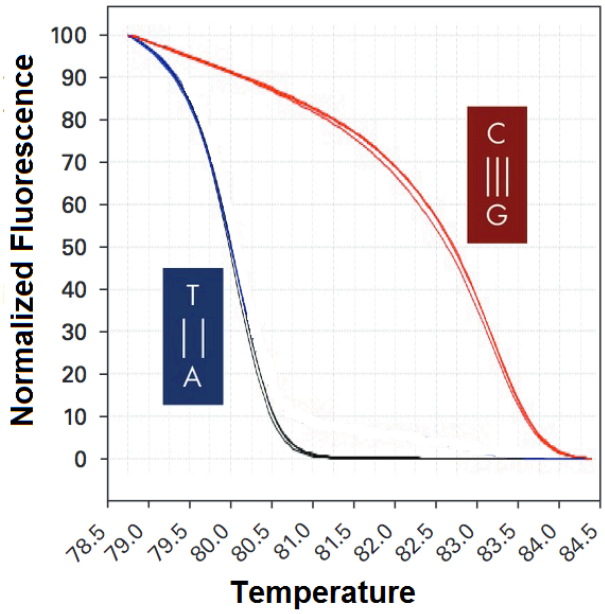


Figure 10. Melting curves of methylated and unmethylated samples. Depending on the initial methylation level, SB treatment leaves different amounts of GC and AT base pairs behind. Since GC and AT bases are bound by 3 and 2 H-bonds respectively, melting of PCR amplicons with elevated GC content requires higher melting temperatures than melting of amplicons with elevated AT content. This results in clearly distinguishable melting curves. Adapted from (137).

1.1.5.2 Genome-wide approaches

1.1.5.2.1 Methylated DNA immunoprecipitation (MeDIP) in combination with array- or sequencing based approaches (MeDIP-chip/MeDIP-seq)

Before precipitation of methylated DNA, genomic DNA has to be purified and sonicated to obtain fragments with a length between 200-1000bp. These fragments are denatured, incubated with an anti-5-methylcytidine antibody and subsequently precipitated with magnetic beads or salmon sperm DNA blocked agarose beads. After removal of the unbound DNA, the antibodies are digested and the DNA is amplified in a Whole Genome Amplification (WGA) reaction. For analysis with an array-based technology (MeDIP-chip), the input DNA obtained after sonication is labelled with cyanine-5 (red) while the precipitated DNA is labelled with cyanine-3 (green). Both fractions are then co-hybridized onto a high-density microarray platform to identify the presence of specific probes and to obtain quantitative data (138). On the other hand, the precipitated DNA can be sequenced using high throughput sequencing approaches (MeDIP-seq). Sequencing of methylated DNA generates short fragments that can be aligned to a

reference genome and enable estimations of methylation levels within specific genomic regions (98).

1.1.5.2.2 Illumina 450K Bead Chip

The Infinium HumanMethylation450 Bead Chip is frequently used to profile genome-wide DNA methylation changes and covers 485.000 CpG sites. Two different chemistry technologies can be applied (Infinium I and II) which cover 99% of RefSeq genes and 96% of CGIs if they are combined. At first, genomic DNA is BS converted, amplified in a WGA reaction, digested enzymatically and purified. By using Infinium I technology, the DNA fragments are hybridized onto the chip where they anneal to CpG locus-specific oligomers that are linked to 2 different bead types (M-bead, U-bead). The M-bead is associated with the methylated query probe with a 3' terminus matching the originally methylated and not SB-converted cytosine while the U-bead is associated with the unmethylated query probe with a 3' terminus matching a thymine that represents an originally unmethylated cytosine. After hybridization, single base pair extension is performed using Biotin-labelled ddNTPs. Single base pair extension is inhibited if there is a mismatch of the query site. The array is stained in order to measure the fluorescent intensities of both bead types and to calculate the methylation ratio.

By using Infinium II technology, the DNA fragments anneal to locus-specific probes associated with only one bead type. Each allele-specific probe is designed to anneal one base upstream from the query site. Depending on the original methylation state of the query site, single base pair extension results in addition of a biotin-labelled G-base on a methylated query site or of a labelled A-base on a unmethylated query site. In contrast to Infinium I assay which uses a single color channel for detection of the fluorescent signals, Infinium II assay uses a red channel for methylated signals and a green channel for unmethylated signals (139-141).

1.1.5.2.3 Reduced Representation Bisulfite Sequencing (RRBS)

RRBS specifically enriches for CpG dense sequences, whose methylation state can be analysed with single base pair resolution. It is based on restriction digestion with MspI. According to its recognition site CCGG, the methylation-

insensitive enzyme generates short fragments ending with a CpG dinucleotide. Thus, the large amount of fragments is mainly produced by digestion of CGIs and CGI shores. These fragments are further modified during library preparation for the Illumina sequencing platform. Beginning with end-repair, the sequences are tailed with adenosines and bilaterally ligated to methylated adapters specific for Illumina sequencing. Subsequently, the CpG-rich sequences are selected by size to reveal fragments ranging from approximately 40 to 220 bp. Before sequencing, the fragments are converted with SB and enriched in a PCR reaction with primers complementary to the adapters (Fig.11). Alignment and methylation analyses require specific software that enables the characterization of methylomes within cancer tissues or between different developmental stages and allows for their comparison (142).

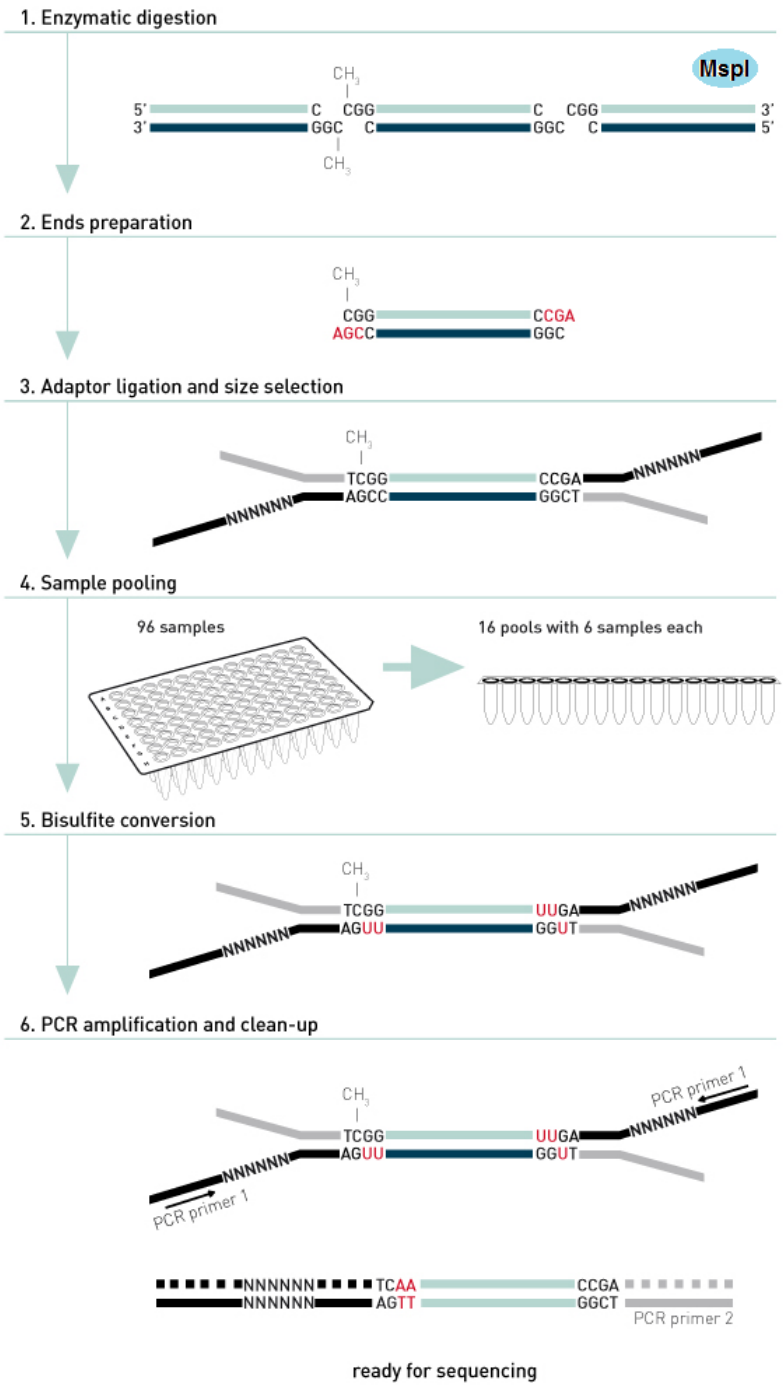


Figure 11. Reduced Representation Bisulfite Sequencing (RRBS) workflow. DNA is enriched after MspI digestion, end-repair, A-tailing and size-selection and is finally SB converted before sequencing. Adapted from (143).

1.2 Chronic myeloid leukemia

Chronic myeloid leukemia (CML) is a hematologic malignancy with elevated numbers of malfunctioning granulocytes and its precursors. The majority of CML patients (95%) share a chromosomal abnormality described as the Philadelphia chromosome which is suggested as the causing event of disease by leading to the expression of a hybrid oncoprotein, a kinase termed BCR/ABL1. The worldwide incidence rate is considered to be 1-2 cases in 100.000 people per year (144). While CML is a rare disease in children, it counts for 10-15% leukemia cases in adults. Men are more frequently diagnosed than women with an average age of 64 (145).

CML is a bi- or triphasic disease starting with a chronic phase (CP), followed by a potential intermediate accelerated phase (AP) and ending in a terminal phase called blast crisis (BC). The initial CP can last for several years showing no or only weak symptoms and is characterized by a blast count less than 10%. In this phase, the majority of blasts still differentiate resulting in a large amount of mature but malfunctioning granulocytes. CP may progress into AP with increased differentiation arrest and genomic instability indicated by 10-20% blasts in blood or bone marrow. About 20% of CML patients skip AP and directly progress into BC with elevated levels (>20%) of poorly differentiated myeloid blasts (Fig.12) (146). BC-CML resembles the morphological characteristics of an acute leukemia and drastically increases mortality due to rising therapy-resistance (147, 148).

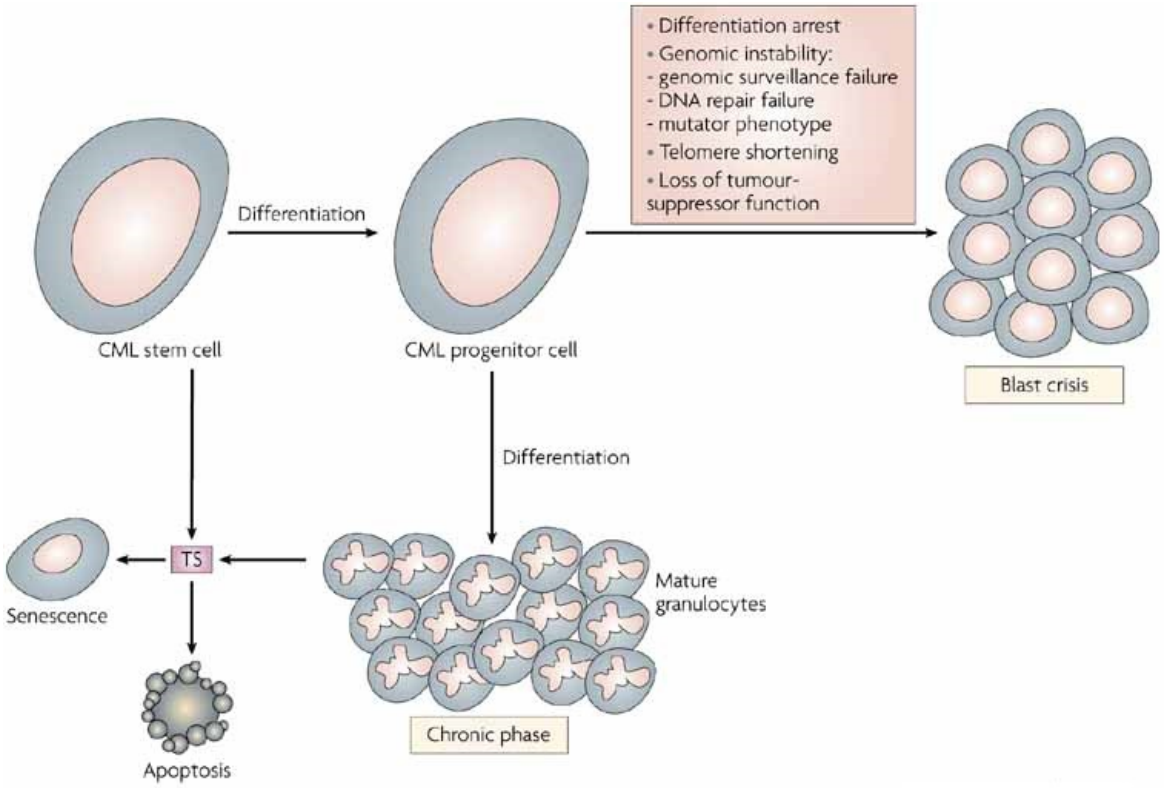


Figure 12. CML evolution from chronic phase to blast crisis. Cell differentiation gets lost during transition from CML-CP to advanced phases, which is thought to happen at the stage of CML progenitor cells. This leads to elevated levels of myeloid blasts in blood and bone marrow. In addition, genetic and molecular alterations in the leukemic clone are believed to drive disease evolution resulting in loss of tumor suppressors, telomere shortening and increased genomic instability. Adapted from (149).

1.2.1 Molecular pathogenesis of CML

1.2.1.1 Philadelphia chromosome formation and the leukemic stem cell model

In 1960 Peter C. Nowell and David Hungerford detected a unique small chromosome in CML leukocytes which they designated as Philadelphia chromosome (Ph1). This discovery constituted the first link between a chromosomal abnormality and a specific malignancy and revolutionized the understanding for CML development (150). In 1973, Ph1 was found to arise through a reciprocal translocation event between chromosome 9 and chromosome 22 shown in Fig.13 (151). This chromosomal rearrangement (t(9;22)(q34;q11)) juxtaposes *c-ABL* (Abelson murine leukemia viral oncogene homolog 1) and *BCR* (Breakpoint Cluster Region gene) on the truncated chromosome 22 resulting in

expression of a BCR/ABL1 fusion protein (152). The *BCR/ABL1* fusion gene encodes a constitutively active kinase which is involved in several oncogenic pathways and promotes the leukemic phenotype. Depending on the breakpoint within the *BCR* gene, its size can range from 190 to 230kDa. However, the most common one in CML is a burst in the major breakpoint cluster region (*M-bcr*) between *BCR*-exons 12 and 16 (153). Upon fusion to *c-ABL*, this results in expression of a 210kDa protein ($p210^{BCR/ABL1}$) (154).

Interestingly, 5-10% of patients displaying clinical features of CML are Ph1-deficient but 30-50% of these cases show a *BCR* rearrangement on chromosome 22. Thus, these patients appear to have a similar course of disease as Ph1-positive CML patients, while those lacking the mutation show different clinical and genomic characteristics (155).

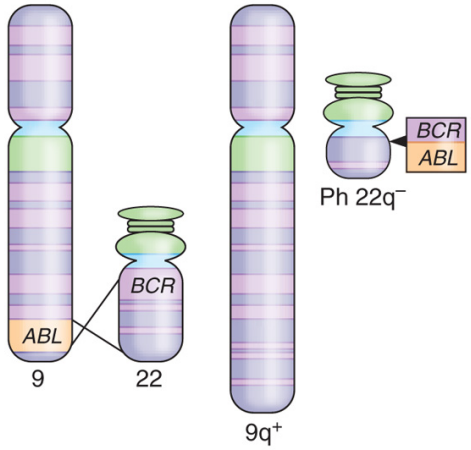


Figure 13. Formation of Ph1**.** Reciprocal translocation of the long arms of chromosome 9 and chromosome 22 result in the formation of Ph1 carrying the fusion gene of *c-ABL* and *BCR* (156).

In 1990 BCR/ABL1 was shown to be sufficient of inducing a CML-like disease in a murine model system by Daley et al. (157). Thus, the occurrence of Ph1 is believed to be a primary and causative event in CML genesis. However, since there are few CML cases without Ph1 or related translocations, there might be additional events relevant for disease initiation.

Ph1 formation was determined to take place at a very early stage within the hematopoietic cell lineage (158). The leukemic stem cell model postulates that gain of the BCR/ABL1 mutation in hematopoietic stem cells (HSCs) results in aberrant differentiation and production of so-called leukemic stem cells (LSCs).

Latter share features with normal HSCs and produce a blast population which is highly proliferative and shows inhibition of differentiation and apoptosis (159, 160). Since different studies demonstrated the presence of BCR/ABL1 in all blood lineages of patients in CP-CML, the LSCs are thought to derive from either HSC- or multipotent progenitor cell-origin (158, 161). Additional studies suggest that LSCs from BC-CML patients derive from more differentiated lineage-restricted progenitors which gain a self-renewal capacity through the course of disease (162).

1.2.1.2 CML disease progression

The causes for progression of CML from CP to AP/BC are mainly unknown. Although BCR/ABL1 triggers CML development and drives cell proliferation and survival by deregulating certain signal transduction pathways (e.g. RAF/MAPK, PI3K/AKT, JNK/STAT) and by preventing DNA damage repair, secondary lesions independent of BCR/ABL1 activity were connected to disease progression (163-169). These lesions include different karyotypic aberrations which can be detected in 50-80% of BC-CML patients. The most common karyotypic aberrations are trisomy 8, double Ph1, isochromosome i(17q) and trisomy 19 (170-172). Moreover, several mutations have been described that inactivate certain tumor suppressor genes (e.g. *p53, p16/ARF, Rb*) or lead to increased expression of oncogenes or proto-oncogenes (e.g. *c-MYC, RAS, AML-EVI-1, SET, MDM2*). Similarly to the chromosomal aberrations, these mutations are considered to be additional hits responsible for disease progression (173-176).

In addition, CML progression was linked to a decreased average length of telomeres in AP-/BC-CML patients compared to CP-CML patients. The shortened telomere length significantly correlated with a reduced duration for developing AP-CML (177-179).

Epigenetic changes during CML progression have not been extensively studied so far. However, it has been suggested that they contribute to disease progression. An increased methylation frequency in AP/BC-CML compared to CP-CML was reported for some genes (e.g. *TFAP2A, p15^{INK4A}, EBF2*) (180-183) (see chapter 3.3).

1.3 DNA methylation in CML

Cancer cells have been associated with an aberrant epigenetic landscape comprised of global loss of DNA methylation but increased CGI promoter methylation and deregulated histone modifications (184, 185). While genome-wide hypomethylation results in genomic instability by activating transposable elements, satellite repeats and oncogenes, CGI methylation contributes to the transcriptional inactivation of tumor suppressor genes, cell cycle regulators and genes involved in cell adhesion and DNA repair (186, 187).

In CML aberrant DNA methylation (referred to as methylation) has been documented for a small number of genes which are summarized in Table 1. Allele-specific methylation of the translocated *c-ABL* promoter is an early event during hematopoietic precursor and mononuclear bone marrow transformation and can be observed in a majority of CML patients. Although *c-ABL* promoter methylation does not favour transcription of the *BCR/ABL1* fusion gene, it might prevent a competition for enhancers and transcription factors with the native *BCR* promoter which enables efficient expression of the *BCR/ABL1* transcript (188, 189).

Genome-wide methylation studies in CML are very rare to date. In a recent study, BCR/ABL1 associated DNA methylation changes were determined using a murine model system with conditional BCR/ABL1 expression. Induced BCR/ABL1 expression lead to genome-wide CGI methylation changes which were reversible upon BCR/ABL1 repression (190). Another study compared DNA methylation profiles of different forms of leukemia and found similar methylation patterns between CML patients and Ph1-negative APL (acute promyelocytic leukemia) patients but distinct patterns between CML patients and Ph1-positive ALL (acute lymphoid leukemia) patients (191).

Although the cause for DNA methylation changes in tumorigenesis is not known to date, deregulation of methylation mediating or demethylating proteins might be involved. DNMT expression levels were shown to be elevated in advanced CML compared to normal hematopoiesis (192). Moreover, Mancini et al. reported that BCR/ABL1 might lead to defects in DNA demethylation by TET2 inactivation (193).

Table 1. Methylated genes in CML patients and their molecular function

Gene	Function	Frequency of methylation (%)
ABL1	Cell adhesion, cell division, differentiation, stress response	20-80
p15^{INK4A}	Cell cycle regulation	~18
p16^{INK4A}	Cell cycle regulation	~8
TFAP2A	Transcriptional regulation	63
EBF2	Cell differentiation	57
DAPK1	Programmed cell death	23
ATG16L2	Autophagy	69
OSCP1	Molecule transporter	~30
PDLIM4	Bone development	~20
NPM2	Chromatin reprogramming	~70
TRPC4	Cell proliferation, neurotransmitter release, vasodilation	>50
JunB	Cell growth	N/A
HOXA	Embryonic development	26-79
CDH13	Cell adhesion	79

(181-183, 188, 189, 194-197)

2 Aims

Due to limited data on epigenetic abnormalities in CML, our main goal was to elucidate genome-wide DNA methylation changes involved in CML pathogenesis. This thesis served as a preliminary study to test RRBS and RNA-seq for leukemia cell lines before the study of patient samples. We therefore aimed to:

- analyse the methylome and transcriptome of a BCR/ABL1-positive cell line and controls
- compare DNA methylation patterns with gene expression patterns
- identify genes transcriptionally regulated by methylation
- investigate identified genes according to enrichment in terms of molecular function, biological processes, disease phenotypes, signalling pathways and other GO terms
- develop gene-specific MS-HRM assays for selected genes to confirm RRBS-results and to analyse patient samples

We hypothesized to find methylation and gene expression differences between leukemia cell lines and control samples and between a BCR/ABL1-positive and a BCR/ABL1-deficient leukemia cell line. Moreover, we expected to correlate methylation with gene silencing of specific genes. By analysing these genes according to their function and involvement in biological processes, we hoped to identify genes which are critical for CML pathogenesis.

3 Material & Methods

3.1 Origin of tumor cell lines

The human megakaryoblastic cell line MO7e and its BCR/ABL1 transfected subline MO7e/p210 were kindly provided by Ao. Univ.-Prof. Dr. Peter Valent (Medical University of Vienna). All cell lines were stored in liquid nitrogen until use. Mononuclear cells (MNCs) of normal bone marrow (nBM) obtained from 2 lymphoma patients without BM involvement and peripheral blood (nPB) from 4 healthy donors served as controls and were isolated using Ficoll (Sigma). For gene-specific analyses, MNCs of BM from 6 CML patients (2 CP-CML, 3 AP-CML, 1 BC-CML) were used. All donors gave written informed consent. This study was approved by the ethics committee of the Medical University of Vienna, Austria (#1594/2015).

3.2 Cell culture

Material:

Growth medium RPMI 1640+Glutamax	(Gibco; #145238)
Fetal Bovine Serum - FCS	(Gibco; #179822)
Penicillin/Streptomycin (50mg/mL)	(Gibco; #10378016)
GM-CSF Human 20UE	(Peprotech; 300-03)
DMSO	(Sigma; D2650)
Cryotubes	(Nunc)

All materials were prewarmed in a 37°C water bath before complementing RPMI1640 with 10% FCS and 1% Pen/Strep (antibiotic). For MO7e cells, the medium was supplemented with 10ng/mL recombinant human GM-CSF.

The cell lines were stored in cryotubes in the liquid nitrogen tank and were defrosted in a 37°C warm water bath. After thawing, cells were transferred into 50mL BD falcons containing 5mL prepared growth medium and centrifuged for 5

minutes and 1000 rounds per minute (rpm). Before repeating centrifugation, the supernatant was discarded and the pellet was resolved in additional 5mL growth medium. For culturing in a 37°C warm incubator with conditions of 96% humidity and 5% CO_2, the cell pellet was dissolved in 15mL growth medium and transferred to a T75 cell culture flask (BD falcon). Cells were passaged twice a week by addition of prewarmed fresh medium to maintain a confluence of $2x10^5$cells/mL. The surplus of cells was freezed in cryotubes in a 1:10 dilution of FCS and DMSO, which were stored in a cell freezing box containing methanol. The box was put on -20°C for 30min, before incubation in -80°C over night and final storage in liquid nitrogen.

"Bürker-Türk" cell counting chamber was used to obtain the number of cells in 10µL cell suspension. Cells were counted in the 4 large squares (each containing 16 smaller squares) at the corners of the chamber. The mean value of counted cells was inserted into following formula:

$$Cells\ per\ mL = \frac{Number\ of\ cells}{counted\ area * chamber\ depth * dilution} * 10^3$$

Counted area = $1mm^2$

Chamber depth = 0,1mm

3.3 Reduced Representation Bisulfite Sequencing (RRBS)

Material:

1xDPBS	(Lonza; #BE17-512F)
Wizard SV Genomic DNA Purification kit	(Promega; #A2360)
Proteinase K (20ng/µL)	(New England Biolabs; #nPB1025)
Elution buffer (EB)	(Qiagen; #19086)
Qubit High Sensitivity Assay	(Life Technologies; #Q32854)
MspI restriction enzyme (20 000units/mL)	(NEB; #R0106L)
10x NEB Buffer	(NEB; #2)
Klenow Polymerase	(NEB; #M0210L)

dNTPs	(GE Healthcare; #28-4065-51)
Epitect control DNA (unmethylated)	(Qiagen; #59665)
Epitect control DNA (methylated)	(Qiagen; #59655)
Quick Ligation Kit	(New England Biolabs; #E0542L)
Illumina TruSeq DNA LT kit (Set A)	(Illumina; #FC-121-2001)
Ampure XP SPRI beads	(Beckman Coulter; #A63880)
100% EtOH	(Merck; #204315)
2x Gotaq qPCR Mastermix	(Promega, #A6002)
EZ DNA Methylation Direct kit	(Zymo Research; #D5020)
PfuTurbo Cx Hotstart DNA Polymerase	(Agilent; #600410)
Experion DNA 1K Analysis Kit for 10 Chips	(Biorad; #700-7107)

3.3.1 Isolation of genomic DNA from cultured cells

Cell Lysis

2-5x10^5 cells were pelleted by centrifugation for 5 minutes and 200 relative centrifugal force (rcf). After removal of the supernatant, each pellet was resuspended in 200µL 1xDPBS before repeating the centrifugation step. For cell lysis, the supernatant was discarded and each pellet was resuspendend in 150µL of Lysis Buffer from the Wizard SV Genomic DNA Purification kit.

Digestion of RNA and proteins

Cell lysates were mixed with RNase A (4µg/µL, Wizard SV Genomic DNA Purification kit) to obtain a final concentration of 50ng/µL. After incubation for 10 minutes at room temperature, 1µg/µL proteinase K was added and heated to 55°C at the thermoblock for 90 minutes.

DNA clean-up and quantification

Genomic DNA was purified according to protocol instructions of Wizard SV Genomic DNA Purification kit with elution in 50µL EB buffer.

DNA was quantified using a Qubit 3.0 Fluorometer according to the manufacturer's instructions.

3.3.2 RRBS library preparation

Restriction digestion

MspI digestion was set up over night at 37°C on a thermocycler according to Tab. 2.

Table 2. MspI digestion.

Component	Amount	Final Amount
10x NEB buffer	2-3µL	1x
Sample	100ng	100ng
MspI	1µL	20U
Nuclease-free H_2O	up to 30µL	final volume 30µL

End-repair and A-tailing

MspI digested samples were end-repaired and A-tailed without any previous clean-up or inactivation of the enzyme according to Tab. 3 and 4.

Table 3. End-repair/A-tailing

Component	Amount	Final Amount
Klenow Polymerase	1µL	5 U
dNTPs (10mM dATP, 1mM dCTP, 1mM dGTP)	1µL	300uM, 30uM
Unmethylated control	1µL	1/1000 of library
Methylated control	1µL	1/1000 of library

Table 4. Cycling conditions during End-repair/A-tailing

	Temperature	Time
End repair	30°C	20 min
A-Tailing	37°C	20 min
Heat inactivation of enzymes	75°C	20 min
Final	4°C	∞

Adapter ligation

For adapter ligation, samples were mixed with components shown in Tab. 5, incubated at 25°C for 20 minutes and enzymes were finally heat inactivated at 65°C for 10 minutes. Six adapters from the Illumina TruSeq DNA LT kit (Set A: 2,4,5,6,7,12) were diluted 1:25 each in EB buffer before combining them for the reaction set up.

Table 5. Adapter Ligation

Component	**Amount**	**Final Amount**
2x Quick Ligase Buffer	32µL	1x
Quick Ligase	1µL	N/A
Adapter Index #2 (1:25 diluted)	1.25µL	N/A

Size Selection

Samples were filled up to 110µL with nuclease-free H_2O and incubated with 82.6µL of 1:5 diluted AMPure XP beads for 15 minutes at room temperature. After the DNA had bound to the beads, the tubes were put on a magnet and the supernatant was discarded. Each pellet was washed twice with freshly prepared 80% EtOH for 5 seconds, air-dried and eluted in 21µL EB buffer.

Quantification and Pooling

Mastermix for qPCR was prepared for samples and controls according to Tab 6. Primer cocktail was used from the Illumina TruSeq kit.

Table 6. qPCR Mastermix

Samples	**4**	**8**	**12**	**16**
Gotaq qPCR Master Mix	70µL	110µL	150µL	190µL
PCR Primer Cocktail	14µL	22µL	30µL	38µL
Nuclease-free H_2O	49µL	77µL	105µL	133µL

For each library 19µL Mastermix and 1µL sample were added to one well of a 96-well plate. All reactions were performed in duplicates. An old library served as positive control, a duplicate without sample as negative control. The plate was

sealed with an optical seal and incubated in a Step One Plus Real-Time PCR Cycler (Applied Biosystems) using following program:

Temperature	Time	
98°C	3 min	
95°C	15 sec	25 cycles
65°C	30 sec	
72°C	30 sec	
72°C	5 min	
4°C	∞	

For library pooling, 17µL of the sample with the maximum Ct value (Ct_{max}) were used and combined with 5 other samples, whose volumes were calculated using the formula 17µL x 2^{-dCT} (dCt = Ct_{max} – $Ct_{sample)}$. The samples were pooled in combinations of six according to the calculated volumes. Then 2.5 volumes of 1:5 diluted AMPure XP beads were added, each reaction was mixed and incubated at room temperature for 15 minutes. After putting it on a magnet for 5 minutes, the supernatant was carefully removed and each reaction was washed twice with 100µL 80% EtOH. The beads were air-dried for 10 minutes and resuspended in 34µL EB buffer.

Bisulfite Conversion

EZ DNA Methylation Direct kit was used for bisulfite conversion with deviations from the manufacturer's protocol regarding bisulfite concentration, incubation temperature, denaturation steps and desulphonation time (Tab. 7).

Table 7. Bisulfite conversion.

Component	Amount	Final Amount
Pool of 6 RRBS libraries	33µL	< 600ng
CT conversion reagent	117µL	0.9x

Reactions were set up and incubated in a thermocycler with following program:

Temperature	Time	
95°C	1 min	20 cycles
60°C	10 min	
4°C	∞	

Column clean-up was performed according to the manufacturer's protocol with elution in 20µL M-elution buffer.

Enrichment PCR

To determine the optimal cycle number (typically Ct-1), qPCR was repeated as described before.

Enrichment PCR Mastermix was prepared according to Tab. 8.

Table 8. Mastermix for Enrichment PCR

Sample Pools	**1**
10 x PfuTurbo C_x reaction buffer	30µL
PCR Primer Cocktail	30µL
dNTPs (25mM each)	3µL
PfuTurbo C_x Hotstart DNA Polymerase (2.5U/L)	6µL
H_2O	203µL

To enrich the RRBS libraries, 181µL mastermix and 19µL bisulfite-converted sample pool were mixed in a 96-well plate. Each reaction was split in 8 aliquots of 25µL per well and incubated on a thermocycler:

Temperature	Time	
95°C	2 min	
95°C	30 sec	Cycles (optimal number)
65°C	30 sec	
72°C	45 sec	
72°C	7 min	
4°C	∞	

Clean-up and Quantification

For the final clean-up, the 8 aliquots of each reaction were combined and mixed with 360µL AMPure XP beads. After 15 minutes at room temperature, tubes were put on a magnet for 5 minutes and washed twice with 500µL 80% EtOH for 5 seconds after removal of the supernatant. The beads were air-dried for 10 minutes and DNA was eluted with 15µL EB buffer with an incubation time of 5 minutes at room temperature. Thereafter, the tubes were put on a magnet and the DNA solution was transferred into new 1.5mL Eppendorf tubes.

DNA was quantified using a Qubit 3.0 Fluorometer according to the manufacturer's instructions. For sequencing, 30ng of each DNA library were mixed with 7.2µL Experion Loading Buffer and quality was checked on an Experion DNA 1K Chip for gel electrophoresis.

3.4 RNA Sequencing (RNA-Seq)

Material:

RNeasy Mini Kit	(Qiagen; #74104)
TruSeq RNA Sample Preparation Kit	(Illumina; #RS-122-2001)

Total RNA from MO7e and MO7e/p210 cell lines as well as MNCs from nBM and nPB was isolated and prepared for sequencing according to the manufacturer's protocols of RNeasy Mini Kit and TruSeq RNA Sample Preparation Kit.

3.5 Next-Generation Sequencing (NGS)

Sequencing of RRBS library pools as well as of adaptor-ligated cDNA was performed on an Illumina HiSeq 2000 machine by the group of Christoph Bock in the Biomedical Sequencing Facility CeMM.

3.6 Methylation-sensitive high resolution melt analysis (MS-HRM)

3.6.1 DNA isolation for MS-HRM

Material:

1M Tris (pH 8,0)	(Sigma; #T3038-1L)
0.5M EDTA (pH 8,0)	(Sigma; #03690)
SDS 20%	(Biorad; #161-0418)
Proteinase K (10mg/mL)	(Ambi; #2546)
Phenol	(Merck; #100206)
Chloroform:Isoamylalcohol (24:1)	(Sigma; #C0549)
100% EtOH	(Merck; #100983)
5M NaCl	(Sigma; #SIG71386)

DNA from MO7e, MO7e/p210, patient samples and controls was extracted using phenol-chloroform. Each pellet was resuspended in 800µL Proteinase K (PK) buffer (Tab. 9) and kept on 50°C for at least 2 hours. Next, every sample was split into two Eppendorf tubes and mixed with 400µL phenol. After centrifugation at 14.000 rpm for 10 minutes, the upper phase containing genomic DNA was transferred into a new tube. Then 480µL chloroform and 20µL isoamylalcohol were added and tubes were centrifuged at same conditions. After transferring the upper phase to a new tube, 1mL of 100% EtOH and 20µL of 5M NaCl were added and samples were put on -80°C over night. Defrosted samples were centrifuged at 14.000 rpm for 5 minutes, supernatant was discarded and 500µL 70% EtOH were added before the tubes were centrifuged a second time for 10 minutes and 14.000 rpm. Afterwards, supernatant was removed and residual ethanol was evaporated by keeping the tubes open for several minutes. The pellets were resuspended in approximately 100µL nuclease-free H_2O, put on 37°C and 700 rpm for 30 minutes and stored at 4°C.

Table 9. Proteinase K buffer.

Component	10mL PK buffer
1M Tris (pH 8,0)	0,5mL
0.5M EDTA (pH 8,0)	20µL
SDS 20%	50µL
Proteinase K 10mg/mL	200µL
ddH_2O	9,23mL

3.6.2 Bisulfite conversion of genomic DNA

Material:

EpiTect Bisulfite Kit (Qiagen; #591049)

To analyse DNA methylation within specific genes of leukemia cell lines and patient samples, genomic DNA was sodium bisulfite treated prior to MS-HRM analysis. Bisulfite conversion was performed on 1µg DNA using the EpiTect Bisulfite Kit according to the manufacturer's instructions.

3.6.3 MS-HRM

Material:

EpiTect HRM PCR Kit (100) (Qiagen; #59445)

EpiTect control DNA (unmethylated) (Qiagen; #59665)

EpiTect control DNA (methylated) (Qiagen; #59655)

Primers were designed with a length between 90-180nt using the sequences from ENSEMBL database (GRCh37 release 76) and MethylPrimer Express® software (Tab. 10). Methylation insensitive primers were generated using following settings:

- Primer length: 18-27bp
- Tm reaction: 56-64°C
- Tm forward-Tm reverse: <7°C

Methylation independence of primers was achieved by setting their location into regions without CpG dinucleotides or by randomly incorporating Cs or Ts (presented as Y in the primer sequences) and Gs or As (presented as R in the primer sequences) into the forward and reverse primers, respectively. Primers were diluted in ddH_2O to obtain a concentration of 100pmol/µL and mixed 1:10. Each reaction was performed by mixing 12.5µL EpiTect HRM Master Mix, 1.9µL primer mix, 9.6µL nuclease-free H_2O and 1µL genomic DNA. DNA standards were maintained by diluting 0% and 100% EpiTect control DNAs to obtain 0%, 10%, 25%, 50%, 75% and 100% methylated DNA controls. Water blanks served as negative controls. Real-Time PCR and melt analysis were performed on a RotorGene®Q cycler (Qiagen) with following cycling conditions:

Stage	**Temperature**	**Time**
Hold	95°C	5 min
40 cycles	95°C	10 sec
	55°C	30 sec
	72°C	10 sec
Melting	65°C - 99°C	Hold for 1 sec in 1st step, hold for 2 sec in 2nd step

Results were analysed using the provided RotorGene®Q software by normalising and differencing all melting curves against the 0% methylated DNA standard. Using the fluorescence maximum of all DNA control standards, a regression line was generated enabling methylation calculation (in %) of all tested samples by use of the linear equation formula.

Table 10. MS-HRM primer sequences.

Gene	**Forward primer (5' → 3')**	**Reverse Primer (5' → 3')**
TGFBI	TGTTGTAGTATAGTAGGTTTYGGG	CCAATTCAATACCCCCAAC
SGK1	GAAGAGTAGTTAGGTAGTTGGGG	AACATATACATCACCRCTACAA
ST3GAL6	GGTGATTTTTYGGTTTYG	AACCAAACRAAAACACAACC

3.7 Statistical Analyses

3.7.1 Analyses of RRBS data

Quality control of raw RRBS data was performed using TrimGalore software (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).
Reads were aligned to a virtual, bisulfite treated genome using Bismark software. The following commands were performed on a UNIX machine:

1. perl trim_galore –RRBS *.fastq
2. perl bismark /~/bismark/HG19/ *_trimmed.fastq
3. perl bismark_methylation_extractor –s *.sam

For analysis and visualization of data the following softwares were used: Bismark MethylKit, Seqmonk (http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/), Clustvis, Ontologizer, The Panther Classification System, Great Annotation and Genemania (198-203).

3.7.1.1 Analysis of RRBS data using Methylkit software

First, .sam files generated by Bismark were sorted according to chromosomal regions using the following command:

sort -k 3,3 -k 4,4n hits.sam > hits.sam.sorted

The sorted Bismark output-files were imported into MethylKit software v0.9.2 (198, 199). The following R commands were used to analyse differential methylation (≥25%) and correlation coefficients between samples, the distribution of differentially methylated CpGs along chromosomes and their location within genomic regions:

```
> # read methylation call files
> filelist <- list(system.file("extdata", "BMA4_CpGout.txt",
package = "methylKit"), system.file("extdata",
"Mo7p210_CpGout.txt", package = "methylKit"))
```

```
> # read the files to a methylRAwList object:myobj
> myobj = read( filelist, sample.id = list("BMA",
"Mo7e_p210"), assembly = "hg19", treatment = c(0, 1), context
= "CpG")

> # calculate correlation between samples
> getCorrelation(meth,plot=T)

> # find differentially methylated bases
> myDiff=calculateDiffMeth(meth)

> # get hypermethylated bases with q-value<0.01 and %
methylation difference>25%
> myDiff25pHyper=get.methylDiff(myDiff,difference=25,
qvalue=0.01,type="hyper")

> # get hypomethylated bases with q-value<0.01 and %
methylation difference>25%
> myDiff25pHypo=get.methylDiff(myDiff,difference=25,
qvalue=0.01,type="hypo")

> # annotate differentially methylated
(hypermethylated/hypomethylated) Cs with promoter/exon/intron
> gene.obj=read.transcript.features(system.file("extdata",
"refseq.hg19.txt", package = "methylKit"))
> diffAnn = annotate.WithGenicParts(myDiff25pHyper,gene.obj)

> plotTargetAnnotation(diffAnn, precedence = TRUE, main =
"BMA vs MO7e_p210")
> diffAnn = annotate.WithGenicParts(myDiff25pHypo,gene.obj)
> plotTargetAnnotation(diffAnn, precedence = TRUE, main =
"BMA vs MO7e_p210")
> # annotate differentially methylated
(hypermethylated/hypomethylated) Cs with CGIs/shores/other
```

```
> cpg.obj <- read.feature.flank(system.file("extdata",
"cpgi.hg19.txt", package="methylKit"), feature.flank.name =
c("CpGi", "shores"))
> diffCpGann <annotate.WithFeature.Flank(myDiff25pHyper,
cpg.obj$CpGi, cpg.obj$shores, feature.name = "CpGi",
flank.name = "shores")
> plotTargetAnnotation(diffCpGann, col = c("green", "gray",
"white"), main = "BMA vs MO7e_p210")
> diffCpGann <annotate.WithFeature.Flank(myDiff25pHypo,
cpg.obj$CpGi, cpg.obj$shores, feature.name = "CpGi",
flank.name = "shores")
> plotTargetAnnotation(diffCpGann, col = c("green", "gray",
"white"), main = "BMA vs MO7e_p210")
```

3.7.1.2 Analysis of differential methylation using Seqmonk software

Output files of Bismark Methylation extractor were imported into SeqMonk v0.27.0 software using the generic text import with following settings:

- chromosome to Col: 3
- Start/End to Col: 4
- Strand to Col: 2

Data of BMA and BMB were grouped editing the replicate set to generate one nBM group.

To quantitate each cytosine separately, the “Define Probes” pipeline was applied using the “Read Position Probe Generator” with a minimum read count of 4. Methylation of single CpGs within all matched reads or of CGIs was quantified using the “Bisulphite methylation over features” pipeline.

To quantitate each cytosine, following settings were used:

- features to quantitate: existing probes
- minimum count to include position: 4
- minimum observations to include feature: 1
- value for features with no data: -1
- combined value to report: Mean

The filter on “values for individual probes” was used to get the valid observations in all samples (value>0 for all data stores). Probes were filtered by a statistical test using the Windowed Replicate filter to compare methylation values for each individual cytosine between samples:

- p-value cutoff: 0.01
- probes overlapping features
- select feature type: CGIs
- multiple testing correction

An absolute difference threshold was applied using a windowed value difference filter with at least 25% difference over correction of a 1000bp window. To obtain the location of all differentially methylated CpGs, the “Annotated Probe Report” option was used. In order to find their association with genes, the “Feature Report” option was used. Genes were analysed according to GO terms using Ontologizer software (Parent-Child Intersection, Benjamin Hochberg Correction).

To quantitate CGIs, following settings in the “Bisulphite methylation over features” pipeline were used:

- features to quantitate: CGI
- minimum count to include position: 4
- minimum observations to include feature: 3
- value for features with no data: -1
- combined value to report: Mean

To remove CGIs with insufficient data, the filter on values for individual probes was applied with a value greater 0 for all data stores. To find CGIs with a certain level of difference in methylation between samples, the filter on value differences was used to analyse changes of at least 25%. To find differentially methylated CGIs located 2000bp upstream, downstream or overlapping a TSS, the features filter was applied to select probes which are close to a TSS. For annotation of the reported probes, the “Annotated Probe Report Option” was used with annotation closest to CpG islands.

3.7.1.3 Visualization of methylation data

Methylation data were visualized in Seqmonk using DataStore Tree, Box Whisker plot and Scatter plot options. Distances of hyper- and hypomethylated CpGs/CGIs to TSS was illustrated using Great Annotation software by importing chromosomal positions of each CpG/CGI. In addition, enriched GO/MSigDB output generated using Great Annotation was illustrated in Graphpad Prism 6 using the –log10 values of Hyper FDR q-values of each term (202). Heatmaps of genes associated with differentially methylated CpGs or CGIs were visualized using Clustvis software by importing methylation values of all samples (200). Clustering options were set as following:

- clustering distance for rows: Manhattan
- clustering method for rows: McQuitty
- tree ordering for rows: tightest cluster first
- clustering distance for columns: no clustering

3.7.2 Analyses of RNA-Seq data

Raw RNA-seq reads were aligned to HG19 using TopHat2 algorithm and Galaxy server (204-207). TopHat2 output was imported into SeqMonk using the option “split spliced reads” to produce multiple reads for each mapped read to cover the different spliced portions. Data sets were quantified using the “RNA-Seq quantitation” pipeline with following settings:

- transcript features: mRNA
- library type: non-strand specific
- merge transcript isoforms: yes
- log transform: yes

Following this, data were normalized using the Percentile Normalisation Quantitation (Percentile: 75) and the Match Distribution Quantitation options. To filter expression differences between samples the Intensity Difference Filter was applied using a p-value<0.01. Results were annotated using the Annotate Probe

Report with an annotation distance cutoff of 2000bp closest to gene and illustrated using Clustvis software with same settings for clustering as described in 3.7.1.3 (200). Moreover, enrichment for GO terms was investigated using Gene Ontologizer calculating raw p-values by applying Parent-Child-Intersection and adjusting same according to Benjamin-Hochberg correction.

3.7.3 Comparison of RRBS and RNA-Seq

To find correlations between methylation and transcription data using SeqMonk, all quantified and normalised RNA-Seq probes were compared with data sets with a Benjamin-Hochberg adjusted p-value<0.01 and differential methylation above 25% in MO7e/p210 versus nBM using the Filter by Probe names option. All matched probes were filtered using the Filter on Value Differences to obtain probes with a fold-change of at least 2. Genes were annotated using the Feature Report option and investigated according to GO term enrichment using Gene Ontologizer (Parent-Child-Intersection, Benjamin-Hochberg adjusted p-values).

3.7.3.1 Visualization of merged data

Seqmonk software was used to generate Box plots and Scatter plots of merged data sets. Expression of methylated genes was illustrated in heat maps using Clustvis software with same settings for clustering as described in 3.7.1.3 (200). Panther Classification System was applied to visualize “Biological Process” and “Molecular Function” of annotated genes (201).

4 Results

4.1 Genome-wide methylation analyses

This master thesis was aimed to establish Next-Generation Sequencing (NGS) approaches for methylome and transcriptome analyses and served as preliminary study for subsequent NGS analyses of CML patients. Therefore, we initially investigated the methylomes of the leukemia cell lines MO7e, MO7e/p210 and of two BM control samples (BMA, BMB) using RRBS. A mean of $4{,}5x10^6$ cytosines located in a CpG context were analysed per sample. On average 52,4% of cytosines located in a CpG context were methylated per sample. By comparison of the content of methylated CpG sites in MO7e, MO7e/p210 cells with that in control samples, we observed a higher frequency of methylated CpG sites in MO7e cells versus controls (62,1% vs. 42,2%) and in MO7e/p210 cells versus controls (63,2% vs. 42,2%). Moreover, a slightly higher frequency of methylated CpG sites was also seen in MO7e/p210 cells compared to MO7e cells (63,2% vs. 62,1%).

Hierarchichal clustering based on CpG site methylation clearly separated MO7e/p210 and MO7e from nBM (Fig. 14A). By comparison of their overall CpG site methylation levels, we found a much higher median CpG site methylation level in cell lines compared to nBM (Fig. 14B). This was further confirmed by calculation of pairwise Pearson's correlation coefficients: While the correlation coefficient between nBM samples (BMA, BMB) was 0.94, we observed clearly lower values between cell lines versus nBM (R=0.64) and, moreover, even between MO7e/p210 and MO7e cells (R=0.88) (Fig. 14C).

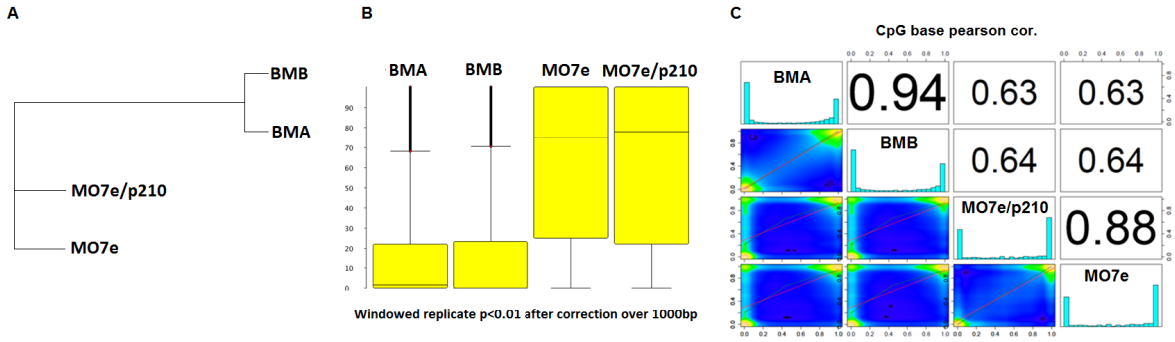


Figure 14. (A) DataStore Tree clustering based on methylation of all samples generated using Seqmonk software. **(B)** Boxplot of median expression in nBM and cell lines using Seqmonk software is shown. **(C)** Scatter plots of percentage of CpG methylation for sample pairs generated using MethylKit software are shown (lower left corner). Numbers in the upper right corner are pairwise Pearson´s correlation scores. Diagonal histograms demonstrate percentage of methylation per CpG site per sample.

4.1.1 Identification of differentially methylated CpG sites

Overall, 903.150 CpG sites were commonly detected in all samples analysed. 230.636 of these CpG sites were found to be differentially methylated (adjusted p<0.01, above 25% methylation difference) between MO7e/p210 and nBM (increased methylation: 219.552, decreased methylation: 11.084) and 231.960 of these CpG sites were found to be differentially methylated between MO7e and nBM (increased methylation: 221.450, decreased methylation: 10.510). In addition, we found 2.149 differentially methylated CpG sites in MO7e/p210 cells compared to MO7e (increased methylation: 494, decreased methylation: 1.655). The chromosomal distributions of these differentially methylated CpG sites are shown in Figure 15.

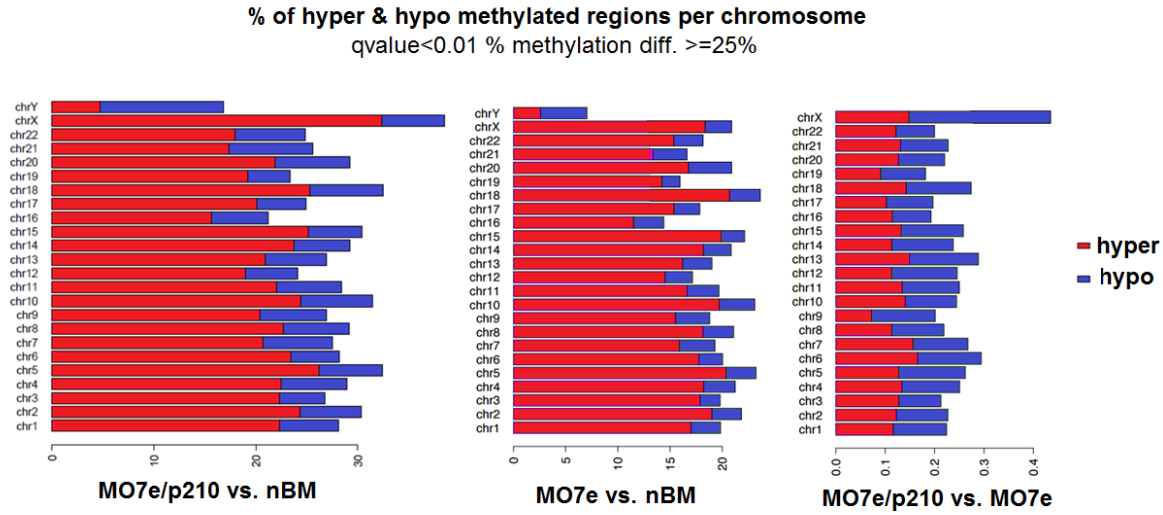


Figure 15. Percentage of hyper- and hypomethylated regions per chromosome in MO7e/p210 versus nBM, MO7e versus nBM and MO7e/p210 versus MO7e is shown. nBM, normal bone marrow.

4.1.1.1 Genomic annotation of differentially methylated CpG sites in leukemia cells compared to controls

We determined the distance of differentially methylated CpG sites to TSS of genes in MO7e/p210 versus nBM. While CpG sites with increased methylation were mainly located in close proximity to TSS, CpG sites with decreased methylation were predominanty located far away (>5000 bp) from the TSS (Fig. 16A). In addition, differentially methylated CpG sites between MO7e/p210 versus nBM were analysed according to their distribution within exons, introns, promoters and intergenic regions as well as according to their location within CGIs and CGI shores. CpG sites with increased methylation in MO7e/p210 compared to nBM were found to be located predominantly in promoter regions and CGIs, while CpG sites with decreased methylation were mainly located in introns and intergenic regions rather than in promoters (Fig. 16B, C). Furthermore, we identified the association of CpG sites with increased methylation with 6398 genes, while CpG sites with decreased methylation were associated with 854 genes. Among genes with increased methylation, we found regulators important for timing cell differentiation (*ASCL1, DLL1, FGF9, HES3*, *NODAL, NR2E1, PAX6, RBPJ*), positive regulators for cell adhesion (e.g. *CDH13, COL16A1, FN1, KIF26B*) and for apoptosis (e.g. *TP53INP1, TP73, CASP8*).

Finally, we analysed CpG sites with at least 25% increase in methylation in MO7e/p210 compared to MO7e to detect possible BCR/ABL1 regulative effects. In total, 494 CpG sites were found to be associated with 79 genes (Fig. 16D). Interestingly, we found positive regulators of immune system processes (*IGF2, KCNN4, PAK1, ZP3*), of cell adhesion (*IGF2, NRG1, PAK1, ZP3*) and of cell communication (*ARHGAP6, IGF2, KCNN4, NEK6, NRG1, PAK1, PARP1, SASH1, ZP3*) and moreover regulators for accurate and complete genome replication and chromosome segregation (*CEP126, CEP76, CHMP6, FMN2, IGF2, NEK6, TUBG1*).

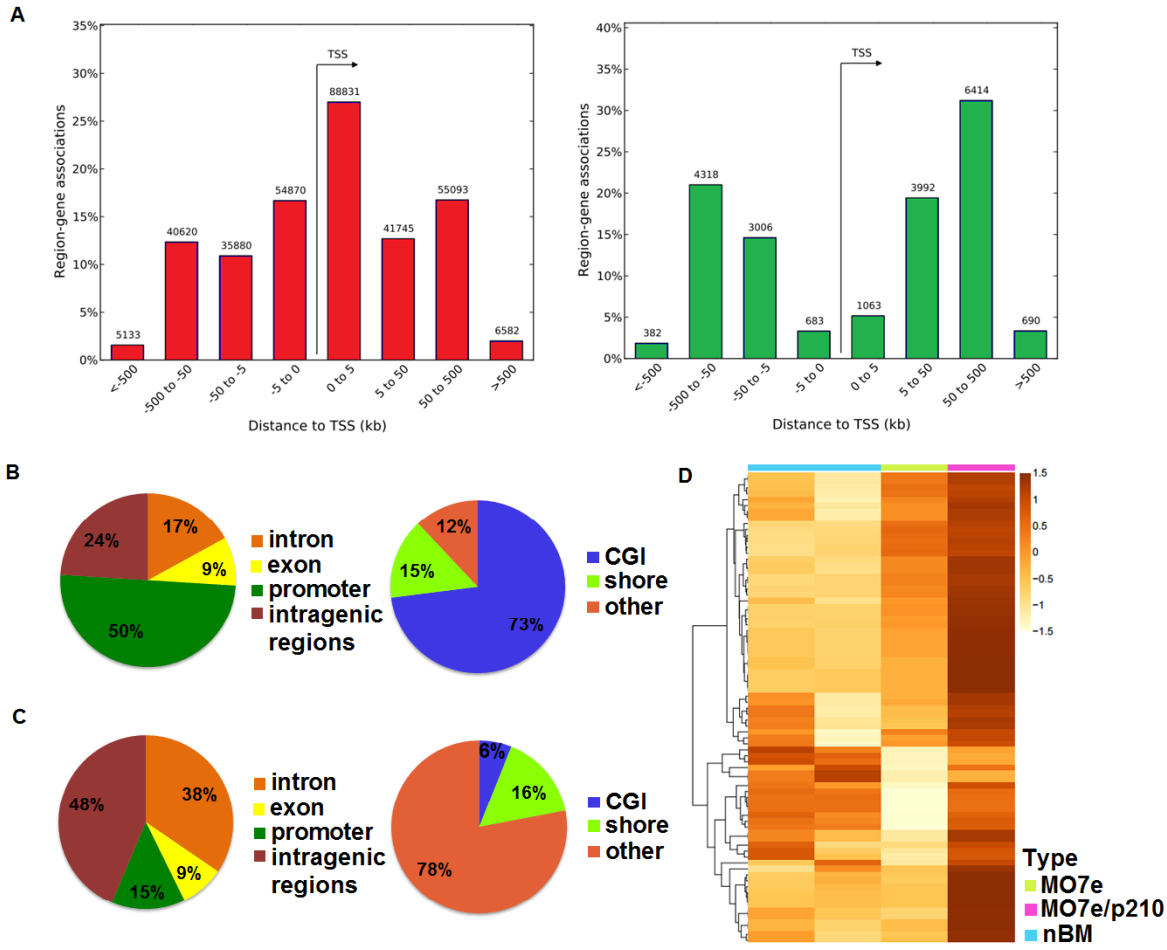


Figure 16. **(A)** Location of CpG sites with increased methylation (red) or decreased methylation (green) in MO7e/p210 compared to nBM samples relative to TSS of genes is shown. Error bars indicate the standard errors of the mean. **(B/C)** Location within introns, exons, promoters and intergenic regions (left) and distribution within CGIs, CGI shores and others are shown (right) for CpG sites with increased methylation **(B)** and decreased methylation **(C)** in MO7e/p210 (methylation difference compared to nBM ≥25%). **(D)** Heatmap demonstrates 79 genes associated with CpG sites with increased methylation in MO7e/p210 compared to MO7e. Colors ranging from light (low methylation) to dark brown (strong methylation). nBM, normal bone marrow, CGI, CpG island.

4.1.2 Identification of differentially methylated CGIs

We further analysed differential methylation of CGIs using Seqmonk and ClustVis software. Overall, 19.167 CGIs were commonly detected in all samples analysed. We found highly comparable CGI methylation between the two nBM samples and between the two leukemia cell lines (Fig. 17A, upper panel). When we compared CGI methylation between nBM versus MO7e/p210 cells and nBM versus MO7e cells, we found dramatically increased CGI methylation in the leukemia cell lines (Fig. 17A, lower panel). In total, 6.240 CGIs were differentially methylated between MO7e/p210 cells versus nBM (increased methylation: 6088, decreased methylation: 152) and 6.454 CGIs were differentially methylated between MO7e cells versus nBM (increased methylation: 6112, decreased methylation: 342). CGIs with increased methylation are associated with 5275 and 5304 genes in MO7e/p210 and MO7e versus nBM respectively. CGIs with highest differential methylation values (≥80%) are illustrated in Figure 17B (left/middle panel).

In addition, we found 815 differentially methylated CGIs between MO7e/p210 and MO7e (increased methylation: 512, decreased methylation: 303) (Fig. 17B, right panel). CGIs with increased methylation are associated with 409 genes, while CGIs with decreased methylation are associated with 250 genes. Among genes associated with CGIs with increased methylation, we identified negative regulators of cell migration (*ABR1, NOTCH1*), positive regulators of apoptosis (e.g. *CTSH, DNM2, TPD52L1, TRAF7*) and of cell adhesion (*AP3D1, DMD*) and genes involved in DNA damage repair (*ATRX, CSNK1D, FAN1, MUM1, PARP4, XAB2*).

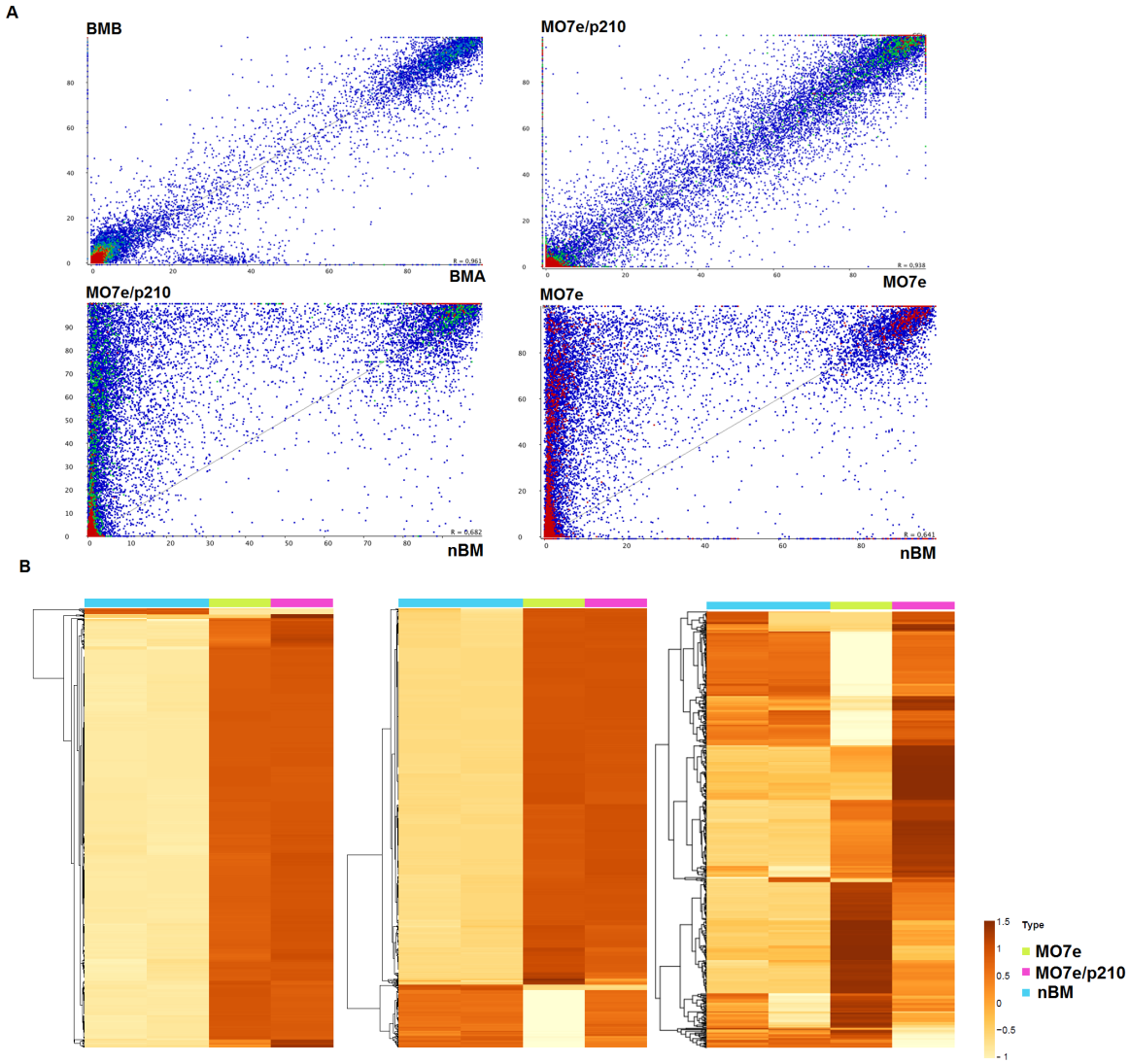


Figure 17. **(A)** Scatter plots summarizing differentially methylated CGIs between BMA versus BMB, MO7e/p210 versus MO7e, MO7e/p210 versus nBM and MO7e versus nBM. **(B)** Heatmaps demonstrate top hits with ≥80% differential CGI methylation between MO7e/p210 versus nBM (left), MO7e versus nBM (centre) and ≥25% differential CGI methylation between MO7e/p210 versus MO7e (right). Colors ranging from light (low methylation) to dark brown (strong methylation). nBM, normal bone marrow.

4.1.2.1 Genomic annotation of differentially methylated CGIs in leukemia cells compared to controls

In addition, we determined the location of the differentially methylated CGIs in MO7e/p210 cells compared to nBM relative to the TSS of genes using GREAT software. These analyses revealed that CGIs with increased methylation are located close to TSS. By contrast, CGIs with decreased methylation were not

found to be in close vicinity to TSS (Fig. 18). This observation was confirmed by analyses using Seqmonk software. While 4224 CGIs with increased methylation were found to be located 2000bp upstream, downstream or overlapping a TSS, only 29 CGIs with decreased methylation were located 2000bp upstream, downstream or overlapping a TSS.

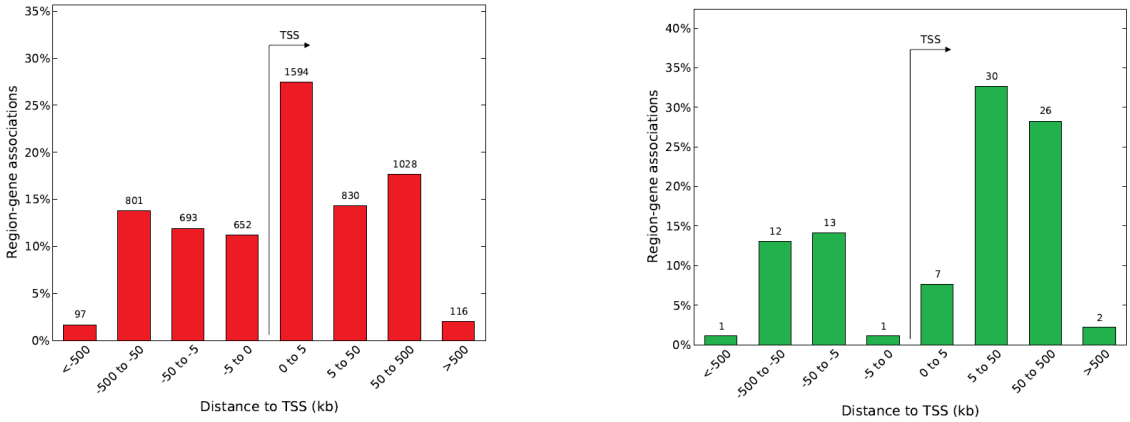


Figure 18. Location of CGIs with increased methylation (red) or decreased methylation (green) in MO7e/p210 compared to nBM samples relative to TSS of genes is shown. Error bars indicate the standard errors of the mean.

Overall, these results demonstrate that CpG site methylation and CGI methylation are extensively increased in leukemia cell lines compared to controls and that a decrease in CpG site/CGI methylation is an infrequent event in leukemia cells. Moreover, our findings demonstrate that hypermethylated regions are mainly located within promoters close to TSS, while hypomethylated regions are distantly situated.

4.1.3 In silico functional analyses of differentially methylated CpG sites

To obtain a better insight into cellular and molecular functions of differentially methylated CpG sites/CGIs, we used GREAT Annotation software, which investigates genomic regions according to published gene-based analyses and enrichments from DAVID database. Our analyses covered ≥25% differentially methylated CpG sites between MO7e/p210 versus nBM resulting in highly

enriched GO/MSigDB terms predominantly associated with developmental processes, cell differentiation, different abnormal mouse and human phenotypes and H3K27 trimethylation for CpG sites with increased methylation in MO7e/p210 cells (Fig.19, Tab. 11). In contrast, MO7e/p210 CpG sites with decreased methylation correlated with cell organization, membrane channel activity and, moreover, showed highly significant association with angiogenesis (Fig.19, Tab. 12). Analyses of MO7e/p210 CGIs with increased methylation revealed enrichment of similar functions/processes as described above.

Overall, these data indicate strong disturbances of many cellular and biological processes by hyper- and hypomethylation events in leukemia cell lines.

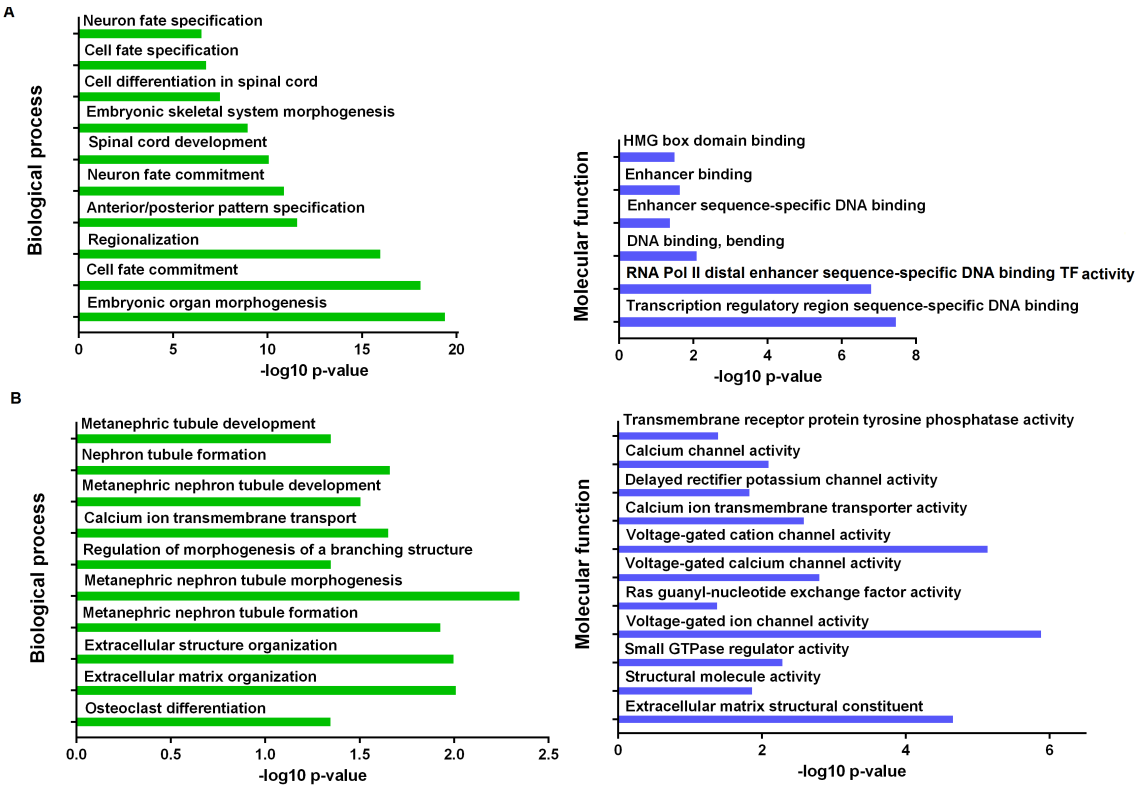


Figure 19. GO enrichment analyses of genes associated with CpG sites with increased methylation **(A)** and decreased methylation **(B)** in MO7e/p210 using GREAT Annotation software. Maximal 10 most significant (hyper fdr q-values, shown as -log10 p-values) categories of "Biological process" and of "Molecular function" are listed.

Table 11. Over-represented GO and MSigDB categories for MO7e/p210 CpGs with increased methylation analysed using GREAT.

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
Mouse Phenotype	abnormal tectum morphology	1	0	2,45	210	1,53E-07	1,98
	abnormal corpora quadrigemina morphology	1	0	2,58	225	2,87E-07	2,06
	abnormal vertebral body morphology	1	0	2,28	227	3,72E-07	1,85
	abnormal metacarpal bone morphology	1	0	2,23	257	2,25E-06	1,89
	abnormal spinal cord interneuron morphology	1	0	2,17	294	7,05E-06	1,87
	abnormal neuromere morphology	1	0	4,25	328	1,49E-05	2,03
	abnormal rhombomere morphology	1	0	4,41	353	2,83E-05	2,02
	abnormal pancreas development	1	0	2,06	364	3,43E-05	1,81
	abnormal inferior colliculus morphology	1	0	2,45	367	3,61E-05	2,09
	abnormal metanephros morphology	1	0	3,12	459	2,15E-04	1,89
Human Phenotype	Abnormality of the incisor	19	1,49E-319	2,40	137	0,02	1,80
	Ectopia lentis	27	9,36E-297	2,52	118	0,02	1,73
	Synostosis of carpals/tarsals	41	1,04E-257	2,38	157	0,03	1,58
	Abnormality of the carpal bones	44	1,94E-251	2,07	129	0,02	1,52
	Transposition of the great arteries	53	1,81E-234	2,89	114	0,02	1,87
	Carpal synostosis	65	3,53E-214	2,39	185	0,05	1,65
	Small nail	68	2,50E-213	2,05	179	0,04	1,52
	Synostosis involving bones of the hand	75	2,29E-207	2,20	106	0,02	1,70
	Lumbar hyperlordosis	172	3,07E-144	2,26	134	0,02	1,63

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
MSigDB Pathway	Genes involved in Class B/2 (Secretin family receptors)	1	0	2,08	8	1,55E-04	1,57
	Maturity onset diabetes of the young	1	0	2,63	28	2,10E-02	1,76
	Wnt signaling network	3	1,03E-293	2,48	4	2,90E-06	2,12
MSigDB Perturbation	Genes with high-CpG-density promoters (HCP) bearing histone H3 trimethylation mark at K27 (H3K27me3) in MEF cells (embryonic fibroblast),	1	0	2,05	6	2,61E-104	1,95
	Genes with high-CpG-density promoters (HCP) bearing the tri-methylation mark at H3K27 (H3K27me3) in MCV6 cells (embryonic fibroblasts trapped in a differentiated state),	1	0	2,46	7	2,06E-79	1,97
	Genes with high-CpG-density promoters (HCP) bearing histone H3 trimethylation mark at K27 (H3K27me3) in neural progenitor cells (NPC),	1	0	2,30	8	3,33E-73	2,03
	Genes with high-CpG-density promoters (HCP) bearing histone H3 dimethylation mark at K4 (H3K4me2) and trimethylation mark at K27 (H3K27me3) in neural precursor cells (NPC),	1	0	2,30	9	4,67E-65	1,97

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
MSigDB Perturbation	Genes with high-CpG-density promoters (HCP) bearing the H3K27 tri-methylation (H3K27me3) mark in brain,	1	0	4,72	10	5,68E-61	2,06
	Genes up-regulated in NB4 cells (acute promyelocytic leukemia, APL) in response to tretinoin [PubChem=444795]; based on Chip-seq data,	1	0	2,41	16	1,20E-24	1,41
	Genes with high-CpG-density promoters (HCP) bearing the H3K27 tri-methylation (H3K27me3) mark in neural precursor cells (NPC),	1	0	2,89	19	1,93E-18	2,09
	Genes with high-CpG-density promoters (HCP) bearing the tri-methylation mark at H3K27 (H3K27me3) in MCV8,1 (induced pluripotent cells, iPS),	1	0	2,05	21	7,26E-16	1,90
	Genes bearing H3K27me3 mark or whose promoters are bound by the polycomb proteins SUZ12 or EED [GeneID=23512;8726]; their DNA is methylated de novo in cancer,	1	0	2,08	22	2,63E-14	1,92
	Genes within amplicon 17q21-q25 identified in a copy number alterations study of 191 breast tumor samples,	1	0	2,73	110	4,29E-04	1,26

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
InterPro	Homeobox, conserved site	1	0	3,85	1	2,88E-30	1,93
	Homeobox domain	1	0	3,03	2	9,41E-23	1,73
	Homeodomain, metazoa	1	0	6,47	3	4,76E-22	2,10
	Homeodomain-like	1	0	2,57	4	3,60E-18	1,57
	Transcription factor, fork head, conserved site	1	0	2,09	27	6,38E-04	1,74
	Transcription factor, fork head	1	0	2,07	29	8,57E-04	1,72
	Homeobox protein, antennapedia type, conserved site	1	0	33,71	47	8,44E-03	1,98
	OAR domain	22	1,18E-281	2,86	35	1,96E-03	2,20
	Wnt	41	9,26E-216	2,84	54	1,46E-02	1,97
	Wnt protein, conserved site	41	9,26E-216	2,84	54	1,46E-02	1,97
	Frizzled domain	96	1,79E-116	2,01	48	8,97E-03	1,91
HGNC Gene Families	NKL	1	0	5,89	1	9,65E-13	2,15
	HOXL	1	0	10,61	2	1,41E-11	2,07
	FOX	1	0	2,27	4	3,00E-05	1,84
	LIM	1	0	3,62	9	4,12E-03	2,20

Table 12. Over-represented GO and MSigDB categories for MO7e/p210 CpGs with decreased methylation analysed using GREAT.

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
GO Cellular Component	extracellular matrix part	16	1,82E-66	2,18	17	2,63E-04	2,03
	basement membrane	23	2,36E-55	2,61	26	1,66E-02	2,17
	voltage-gated calcium channel complex	104	2,01E-16	2,38	31	4,66E-02	2,80
Mouse Phenotype	impaired osteoblast differentiation	6	3,36E-146	37,95	228	4,50E-02	9,17
	abnormal choroid plexus morphology	18	2,29E-115	4,57	208	3,96E-02	2,47
	Leydig cell hypoplasia	24	3,52E-104	5,94	255	4,88E-02	2,86
	decreased chemically-elicited antinociception	43	1,53E-77	5,33	85	4,07E-03	4,17
	abnormal chemically-elicited antinociception	57	8,74E-65	4,37	122	1,25E-02	3,36
	abnormal afterhyperpolarization	58	1,72E-64	5,06	77	3,33E-03	5,24
	abnormal neurotransmitter secretion	69	6,93E-59	2,76	160	2,00E-02	2,29
	abnormal neurotransmitter level	88	1,08E-50	2,51	170	0,02	2,16
	persistence of hyaloid vascular system	100	1,79E-46	3,85	161	0,02	3,59
	decreased dopaminergic neuron number	112	3,74E-44	3,83	212	0,41	3,49
PANTHER Pathway	Angiogenesis	4	4,56E-46	2,12	1	0,03	1,94

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
MSigDB Perturbation	Genes up-regulated in NB4 cells (acute promyelocytic leukemia, APL) in response to tretinoin [PubChem=444795]; based on Chip-seq data.	1	0	4,59	1	0,00	2,31
	Genes with high-CpG-density promoters (HCP) bearing the tri-methylation mark at H3K27 (H3K27me3) in MCV6 cells (embryonic fibroblasts trapped in a differentiated state).	1	0	2,92	2	0,00	2,54
	Genes within amplicon 7p22 identified in a copy number alterations study of 191 breast tumor samples.	1	0	23,11	16	0,00	5,12
	Genes within amplicon 16q24 identified in a copy number alterations study of 191 breast tumor samples.	1	0	26,70	22	0,00	3,60
	Genes within amplicon 16p13 identified in a study of 191 breast tumor samples.	5	2,99E-320	28,72	34	0,00	2,25
	Genes with high-CpG-density promoters (HCP) bearing histone H3 trimethylation mark at K27 (H3K27me3) in neural progenitor cells (NPC).	6	7,65E-244	2,76	7	1,24E-14	2,50

Ontology	Term Name	Binom Rank	Binom FDR Q-Val	Binom Fold Enrichment	Hyper Rank	Hyper FDR Q-Val	Hyper Fold Enrichment
MSigDB Perturbation	Genes with high-CpG-density promoters (HCP) bearing histone H3 trimethylation mark at K27 (H3K27me3) in MEF cells (embryonic fibroblast).	8	1,04E-221	2,20	4	5,70E-17	2,21
	Genes with high-CpG-density promoters (HCP) bearing histone H3 dimethylation mark at K4 (H3K4me2) and trimethylation mark at K27 (H3K27me3) in neural precursor cells (NPC).	9	3,91E-215	2,58	3	1,29E-17	2,67
	Genes within amplicon 5p15 identified in a copy number alterations study of 191 breast tumor samples.	10	7,15E-205	17,62	18	3,28E-07	5,73
	Genes within amplicon 20q12-q13 identified in a copy number alterations study of 191 breast tumor samples.	11	9,07E-192	6,43	11	9,59E-10	3,03
InterPro	Laminin G domain	64	8,57E-47	2,33	1	1,49E-02	3,16
MSigDB Oncogenic Signatures	Genes down-regulated in HUVEC cells (endothelium) by treatment with PIGF [Gene ID=5281].	1	3,58E-86	2,64	6	5,00E-02	1,68

4.2 Transcriptome analyses

We investigated the transcriptome of the same samples which were also used for methylome analyses. In addition, because RNA from only one nBM control was available, we included three additional peripheral blood (nPB) samples from healthy donors. Aligned .bam files were imported into Seqmonk software for further processing and data analyses. To be able to compare RNA-seq data between samples, data were normalised using the "Percentile Normalisation Quantitation" pipeline and the "Match Distribution" option. Data before and after normalisation are shown in Figure 20. Overall, 30.951 transcripts were analysed by RNA-Seq.

Afterwards, expression differences (intensity difference with a non-adjusted p-value<0.01) between nBM and MO7e/p210 were calculated. These analyses lead to the identification of 442 differentially expressed genes. By filtering for clusters with a minimum number of 10 probes per cluster, we were able to identify 4 big clusters within the 442 gene cohort (Fig. 21A). Overall, 162 of all differentially methylated genes were found to be downregulated in leukemia cell lines but not in controls (nBM, nPB) (cluster 1), while only 64 genes turned out to be upregulated in leukemia cell lines but not or less in controls (cluster 4). Moreover, 194 genes were found to be upregulated in nBM only (cluster 2), whereas only 19 genes turned out to be downregulated in nBM in comparison to all other samples (cluster 3).

Correction for multiple testing (adjusted p-value<0.1) of differential expression between MO7e/p210 and nBM reduced the number of genes to 59. Interestingly, 54 of these genes turned out to be at least 4-fold downregulated in leukemia cell lines compared to nBM. Some of these genes are assumed a function as potential tumor suppressors (e.g. *SPARC, XIST, TIMP2, TSC22*), regulators of apoptosis (e.g. *DUSP-1, MAGED1, CD36*), cell growth (e.g. *SPARC*), cell cycle (e.g. *CDK6*) or cell adhesion (e.g. *CDH11, COLL11A1, COL3A1, IBSP, FN1*).

The comparative gene expression analysis (intensity difference with a non-adjusted p-value<0.01) between MO7e/p210 and nPB identified 507 differentially

expressed genes which are clustered in 2 clusters (Fig. 21B). We identified 148 genes with a lower expression (cluster 2) and 361 genes with a higher expression in nPB compared to leukemia cell lines (cluster 1). Correction for multiple testing (adjusted p-value<0.1) reduced the number of genes to 96. Among them 13 genes were expressed in leukemia cell lines but downregulated in nPB (e.g. *BCAT1, KIT, FASN*). Among the downregulated genes in leukemia cell lines, we identified genes important for positively regulating cell adhesion (e.g. *CCL5, CD36, KLRK1, TRAC, HLA-DPA1*) and apoptosis (e.g. *DUSP1, ANXA1, CASP1, MNDA, IGF2R*) and negatively regulating cell proliferation (e.g. *BTG1, ETS1, IRF1*).

In order to identify genes which might be transcriptionally regulated by BCR/ABL1, we additionally analysed expression differences between MO7e/p210 and MO7e (intensity difference with a non-adjusted p-value<0.01). We found 491 genes with at least 2-fold expression difference. Expression of 279 of them was upregulated, while expression of 212 genes was downregulated in MO7e/p210 compared to MO7e (Fig. 21C). Among the 212 downregulated genes, we identified several tumor suppressors (e.g. *BRCA2, E2F2, KLF10, KLF11, KLF7, PARP1, PCDH9*). After correction for multiple testing (adjusted p-value<0.1) 83 differentially expressed genes remained. 44 of them were downregulated and included positive regulators of cell death (e.g. *BMF, GRIN2A, OSGIN1*) or of cell adhesion (e.g. *COL6A3, LAMA2, PVRL4*).

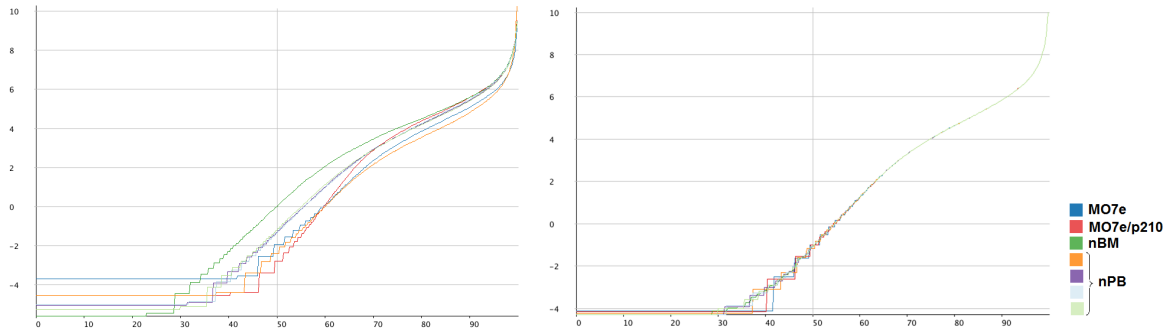


Figure 20. RNA-Seq data before (left) and after (right) normalization using Seqmonk software.

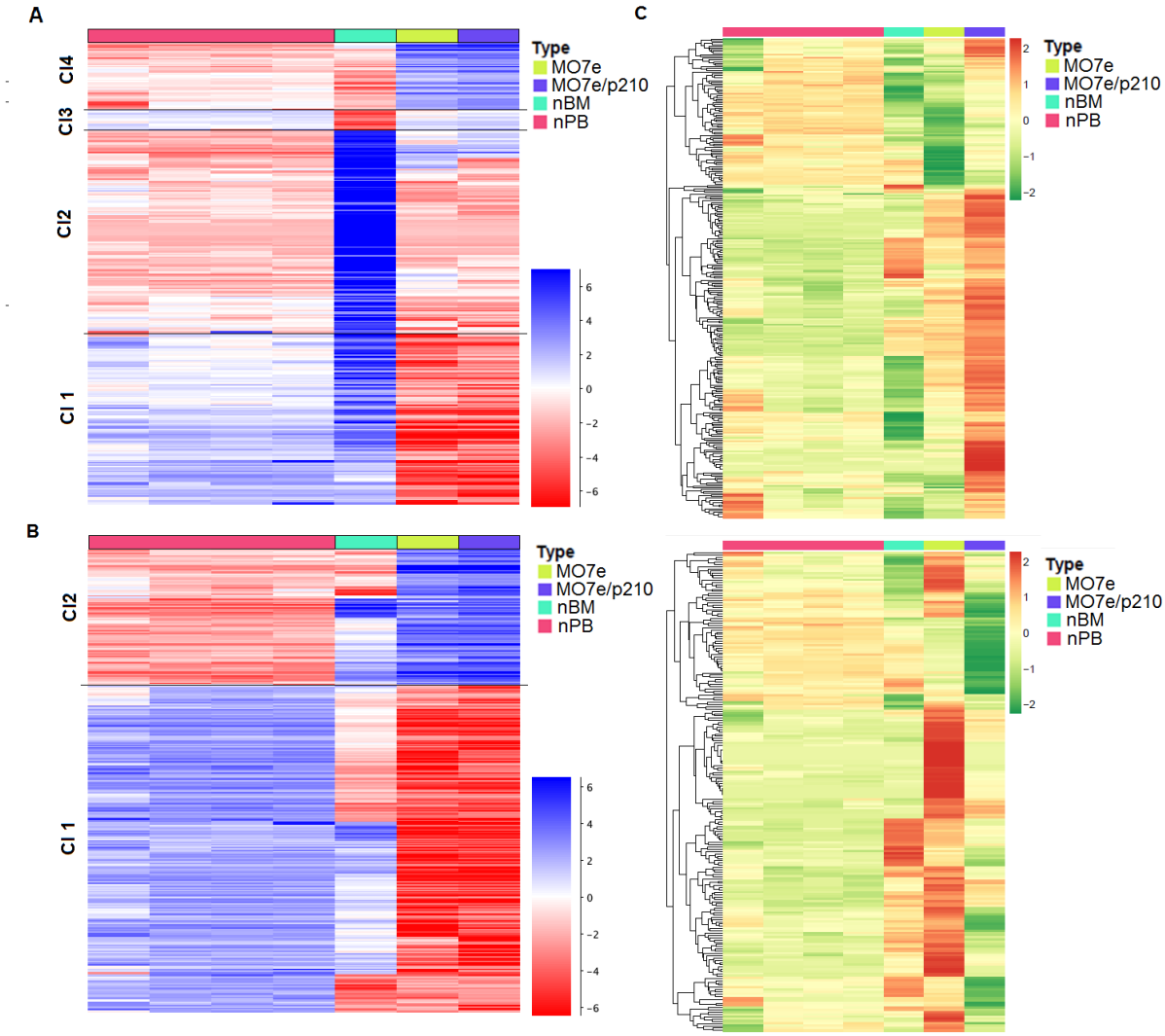


Figure 21. Heatmaps show expression values of genes with at least 2-fold expression difference between MO7e/p210 and nBM **(A)** and MO7e/p210 and nPB **(B)** and expression values of genes with at least 2-fold upregulation (upper panel) and 2-fold downregulation (lower panel) in MO7e/p210 compared to MO7e and nBM **(C)**. nBM, normal bone marrow, nPB, peripheral blood monocytes, Cl, Cluster.

4.3 Correlation between DNA methylation and gene expression

To identify genes which might be transcriptionally regulated by methylation in leukemia cell lines, we determined expression levels of the genes found to be methylated in MO7e/p210 cells by RRBS. We investigated the expression levels of genes associated with CpGs with increased methylation (≥25%) in MO7e/p210

compared to nBM. The analysis covered 6359 methylated genes. By comparison of their overall expression levels, a lower median expression was observed in leukemia cell lines compared to nBM (Fig. 22). Moreover, we identified 2403 methylated genes with at least 2-fold downregulation and 1598 methylated genes with at least 2-fold upregulation in MO7e/p210 compared to nBM. Figure 23A shows top hits with at least 4-fold up- and downregulation.

In order to identify genes that might be regulated by BCR/ABL1, we additionally searched for genes with at least 2-fold downregulation in MO7e/p210 compared to MO7e within the 2403 gene cohort. We detected 667 methylated genes with at least 2-fold downregulation in MO7e/p210 compared to nBM and MO7e (Fig.23B). Table 13 summarizes enriched GO terms for these genes. Panther Classification System was used to analyse molecular function and biological processes of 611 of the 667 genes (Fig. 23C).

Overall, these results indicate that methylation of many genes leads to their transcriptional silencing.

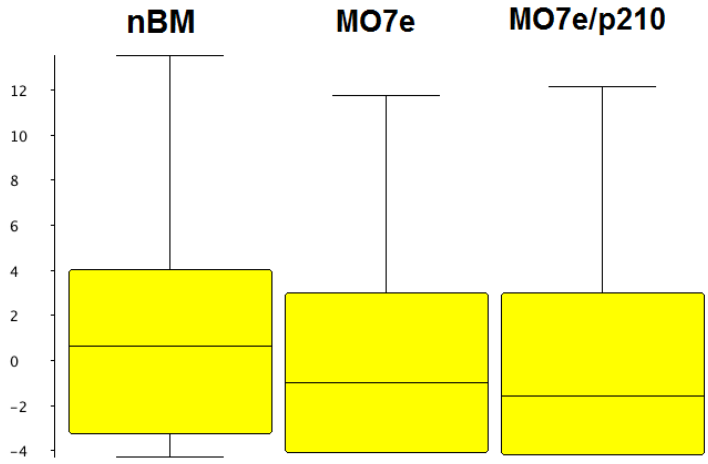


Figure 22. Median expression levels of 6359 genes with increased methylation in MO7e/p210 compared to nBM. nBM, normal bone marrow.

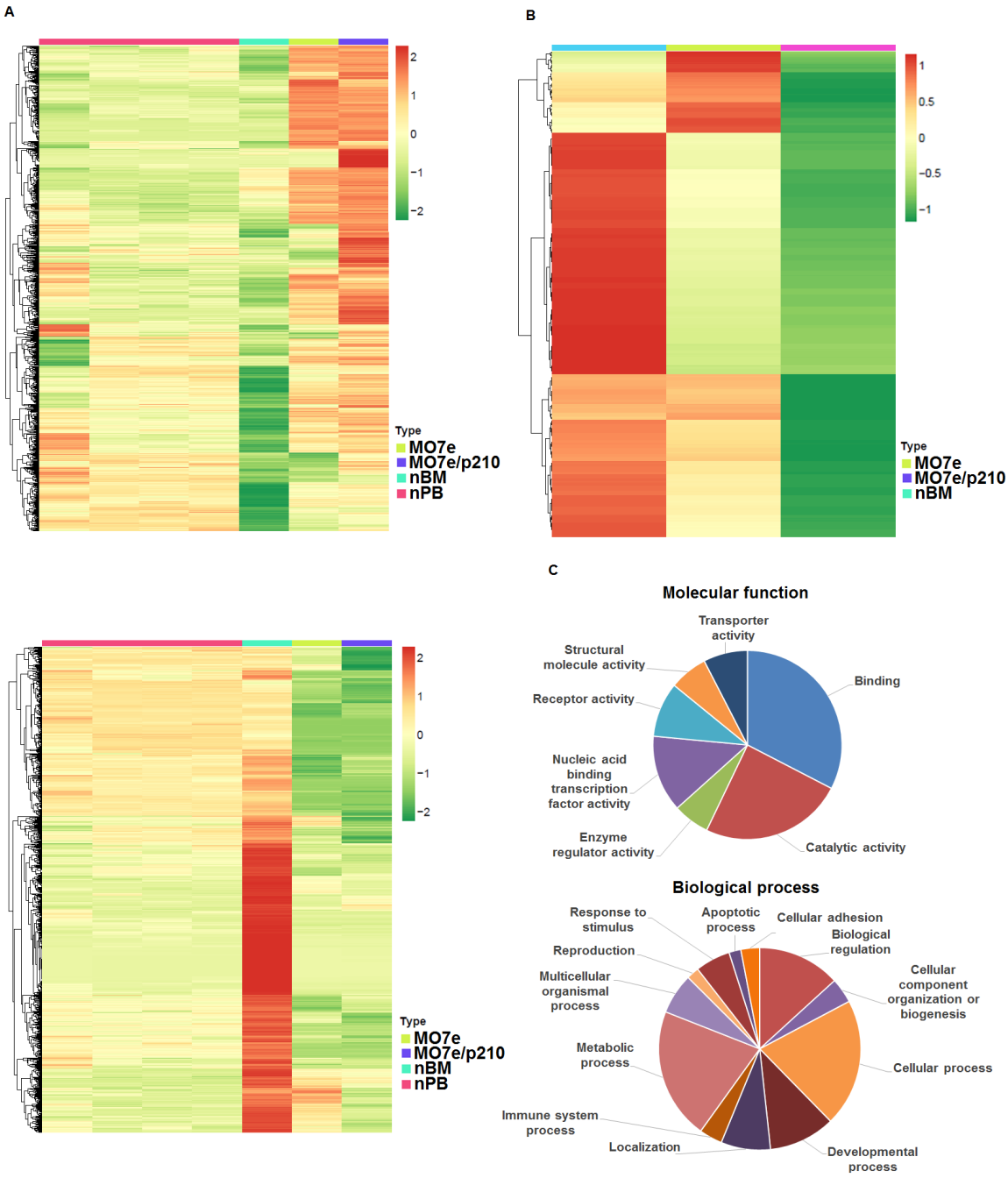


Figure 23. (A) Heatmaps illustrate expression values of methylated genes with high expression (upper panel) and low expression (lower panel) in MO7e/p210 compared to nBM. **(B)** Heatmap shows expression values of methylated genes with at least 2-fold downregulation in MO7e/p210 compared to MO7e and nBM. **(C)** Panther classification system matched 611 of 667 MO7e/p210 methylated and 2-fold downregulated genes according to molecular function and biological process. nBM, normal bone marrow, nPB, peripheral blood.

Table 13. GO enrichment analyses of methylated and downregulated genes in MO7e/p210 compared to MO7e and nBM.

GO ID	Name	Genes
GO:2000146	negative regulation of cell motility	*ARAP3, CYP1B1, HOXA7, MAGI2, PLCB1, PODN, PTPRG, TBX5, THBS1, THY1, TP53INP1*
GO:0008285	negative regulation of cell proliferation	*ALDH1A2, ATP8A2, BCL11B, BCL6, BMP4, BMPR1B, CASK, CAV1, CYP1B1, ERBB2, ERBB4, EREG, GLI3, GPC3, INHBA, KLF11, MAGI2, MSX1, MYO16, NFIB, NKX3-1, NOX4, PODN, PPARG, PTH1R, ROBO1, SCIN, SOX9, TBX5, TGFB2, TGIF1, THBS1, TIMP2, TOB1, TP53INP1, TP73, WNT5A, ZNF503*
GO:0040013	negative regulation of locomotion	*ARAP3, CYP1B1, HOXA7, MAGI2, PLCB1, PODN, PTPRG, ROBO1, SEMA5A, TBX5, THBS1, THY1, TP53INP1, WNT5A*
GO:0002684	positive regulation of immune system process	*BCL6, BLK, C3, CACNB4, CADM1, CAV1, CD14, CD74, COLEC12, EREG, FOS, FOXC1, FOXF1, GLI3, HCK, HLX, IGF2, IL1RL2, INHBA, IRAK2, IRF4, IRS2, LEP, MYO10, PDE4B, PDE4D, PELI2, PRKCH, SCIN, SPON2, TGFB2, THBS1, THY1, WNT5A*
GO:0043068	positive regulation of programmed cell death	*ADRB1, AIFM2, ALDH1A2, ATF3, BCL2L14, BCL6, BMP4, BMPR1B, CAV1, CD14, CYP1B1, EGLN3, F3, FOXO1, GRIK2, GRIN2A, HMGA2, IL20RA, INHBA, INHBB, KLF11, MMP9, MSX1, NGFR, NKX3-1, NOX4, NR4A3, PPARG, ROBO1, SCIN, SQSTM1, TGFB2, THBS1, TP53INP1, TP73, WNT5A*
GO:0022409	positive regulation of cell adhesion	*ANK3, CAV1, CD74, GLI3, HLX, IGF2, IL1RL2, TBX18, THY1, WNT5A*

4.4 Gene-specific methylation analyses

Based on our NGS data, we selected 3 genes (*TGFBI, SGK1, ST3GAL6*) for further methylation analyses in MO7e/p210, MO7e, nBM, nPB and CML patient samples (CP, n=2; AP, n=3; BC, n=1) using MS-HRM. We were able to confirm methylation observed by RRBS of all 3 genes in MO7e/p210 and MO7e cells by MS-HRM. Moreover, methylation of these genes was increased in BC compared to CP or AP samples (Fig. 24).

To obtain more information about the potential function of the genes, Genemania software was used (Fig. 25). TGFBI was found to interact with several proteins involved in cell adhesion (FN1, POSTN, ITGA3), while SGK1 interacts with the tumor suppressor FOXO3. Moreover, we found that SGK1 co-localizes with the phosphatase DUSP4 which is a negative regulator of proliferation-controlling kinases.

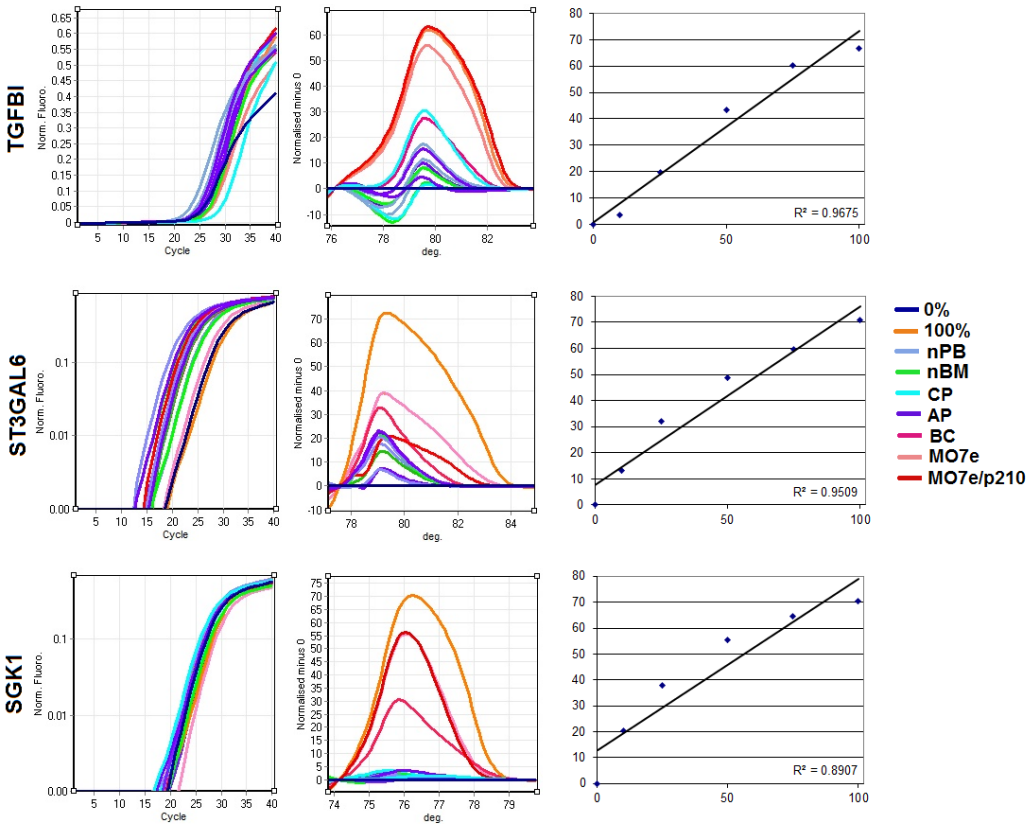


Figure 24. MS-HRM analyses of *TGFBI, ST3GAL6* and *SGK1*. Shown are amplification plots (left panel), differential graphs with maximum peak values (middle panel) of samples and DNA methylation standards and regression lines of standards (right panel). nPB, peripheral blood, nBM, normal bone marrow, CP, chronic phase, AP, accelerated phase, BC, blast crisis.

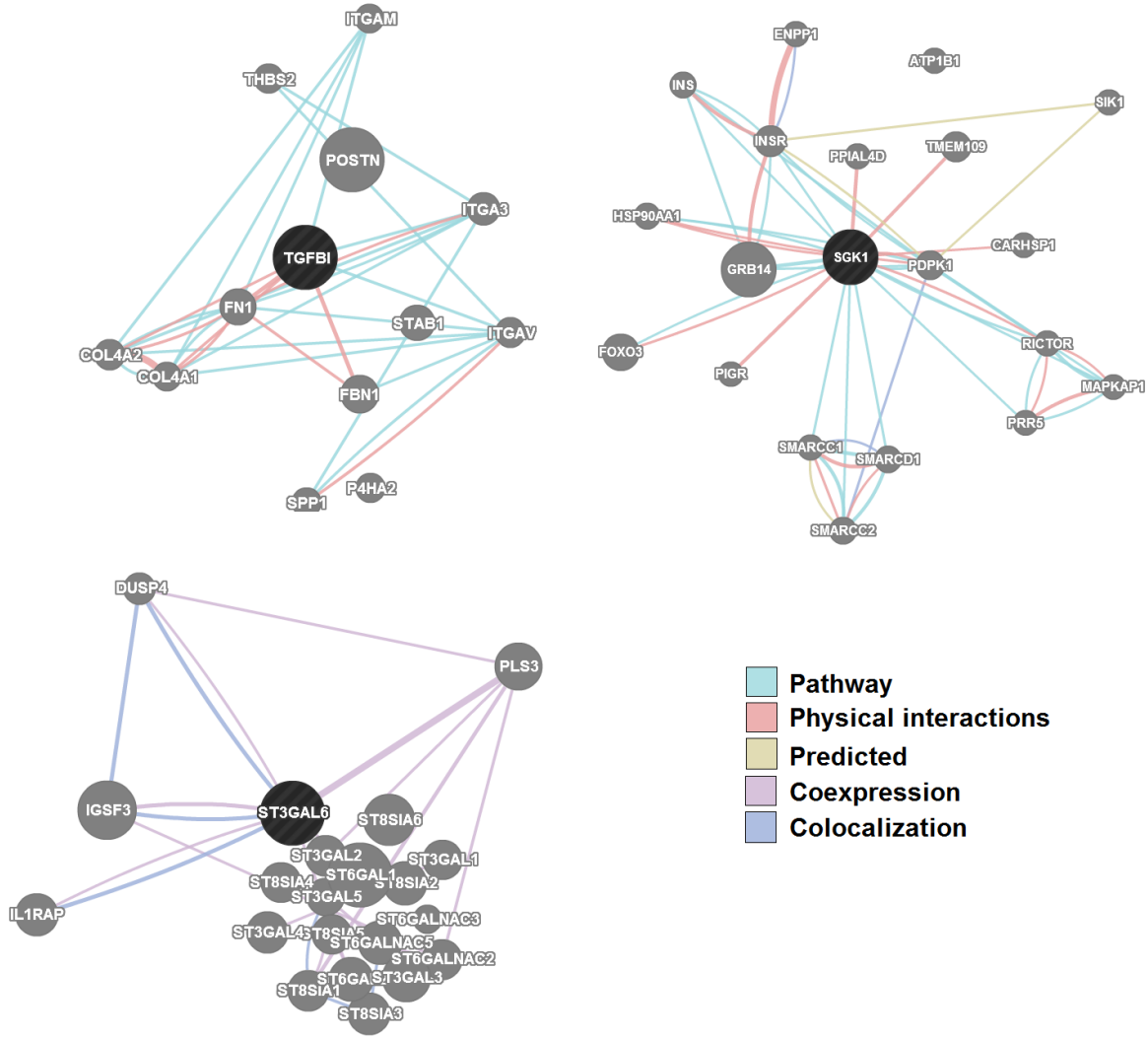


Figure 25. Genemania network analyses of selected genes (*TGFBI, SGK1, ST3GAL6)*. Shown are associations with pathways, physical interactions, predicted interactions, and proteins with coexpression and –localization.

5 Discussion

DNA methylation changes are known to be involved in the development of various malignant diseases. However, only little is known about the impact of methylation in CML pathogenesis. So far, studies were mainly restricted to analyses of single genes, while whole-genome studies investigating DNA methylation patterns in CML are rare. To elucidate the role of DNA methylation in CML, we established a method for genome-wide methylation analyses in our lab. This technique is called RRBS and was chosen because of its high sensitivity and enrichment for CpG-rich regions, which makes it cost-effective in comparison to Whole Genome Bisulfite Sequencing (WGBS) (142). We tested RRBS using leukemia cell lines and nBM controls.

By comparing CpG site and CGI methylation between leukemia cell lines and nBM, we observed heavily increased methylation levels in leukemia cell lines. Overall, the frequency of CpG sites with increased methylation was much higher than the frequency of CpG sites with decreased methylation indicating that hypomethylation is an infrequent event in these cells. The majority of methylated CpG sites in MO7e/p210 was found to be located in promoter regions within CGIs and close to TSS. Genes associated with CpG sites with increased methylation in MO7e/p210 were found to be involved in the regulation of developmental processes, cell differentiation and the occurrence of different disease phenotypes while genes associated with CpG sites with decreased methylation were significantly enriched for angiogenesis. In comparison to methylation differences between leukemia cell lines and nBM, we observed methylation differences between the BCR/ABL1 expressing cell line (MO7e/p210) and its untransformed counterpart (MO7e) in a lower degree. These differences might be caused by BCR/ABL1. However, additional analyses are necessary to prove this hypothesis. In particular, genome-wide analyses of human BCR/ABL1 inducible cells should be performed. Genes associated with methylated CpG sites in MO7e/p210 compared to MO7e were found to be involved in cell adhesion, cell division, cell communication and immune defense.

Moreover, we analysed mRNA expression by RNA-Seq and identified a large number of genes where expression was downregulated in MO7e/p210 compared to nBM. These genes included tumor suppressors, regulators of

apoptosis, cell growth, cell cycle and cell adhesion. In addition, several tumor suppressors were found to be downregulated in MO7e/p210 compared to MO7e. These data suggest that gene expression is strongly deregulated in MO7e/p210 compared to controls.

To analyse whether DNA methylation is associated with transcriptional gene silencing, we compared RRBS and RNA-Seq data of leukemia cell lines. We found that the majority of methylated genes was also downregulated in MO7e/p210 indicating that these genes are transcriptionally regulated by DNA methylation. In comparison to MO7e and nBM, we identified 667 methylated and downregulated genes in MO7e/p210. Many of them were found to play roles in apoptosis, immune defense and cell adhesion.

We selected 3 genes (*TGFBI, SGK1, ST3GAL6*) for methylation analyses using a gene-specific approach. Increased promoter methylation of these genes was observed in MO7e/p210 and MO7e cell lines and in a BC-CML patient sample. While promoter methylation of TGFBI and ST3GAL6 was also observed in AP- and CP-CML patient samples, promoter of SGK1 was not found to be methylated in these samples. To get information about the involvement in molecular networks, we performed network analyses using Genemania software. We found physical interactions of TGFBI with proteins involved in cell adhesion such as ITGA3, COL4A2, COL4A3 and FN1. SGK1 was found to interact with the tumor suppressor FOXO3A, while ST3GAL6 was found to co-localize with the potential tumor suppressor DUSP4. So far, it remains unclear whether these genes are involved in the pathogenesis of CML and further experiments are necessary to elucidate the function of these genes in CML cells. However, based on the available literature, there is evidence that these genes might be involved in the carcinogenesis of certain malignancies. TGFBI (Transforming Growth Factor, Beta-Induced) is suggested to participate in cell-collagen interactions and has been shown to be methylated in ovarian cancer, multiple myeloma, neuroblastoma, prostate cancer, lung cancer and MLL. (208-213). SGK1 (serum- and glucocorticoid-regulated kinase 1) is known to be involved in many cellular processes such as stress response by regulating ion channels and cell volume but was not frequently linked to cancer development (214). ST3GAL6 (ST3 Beta-Galactoside Alpha-2,3-Sialyltransferase 6) is important for adhesion-mediating

selectin ligand synthesis and was demonstrated to be methylated in colorectal and gastrointestinal cancer (215-217).

Overall, our findings demonstrate a high frequency of CpG site and CGI methylation in leukemia cell lines that likely regulates transcription of a large number of genes. These genes frequently participate in cellular processes essential for maintaining a healthy cell phenotype. Thus, promoter methylation of the identified genes is suggested to result in strong deregulation of signalling pathways and cellular networks and might be involved in the pathogenesis of CML.

6 References

1. Waddington CH. Introduction to Modern Genetics. New York: The Macmillan Company; 1939.

2. Waddington CH. The Strategy of the Genes. London: Allen&Unwin; 1957.

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

4. Jones PA, Taylor SM. Cellular differentiation, cytidine analogs and DNA methylation. Cell. 1980;20(1):85-93.

5. Wallis JW, Rykowski M, Grunstein M. Yeast histone H2B containing large amino terminus deletions can function in vivo. Cell. 1983;35(3 Pt 2):711-9.

6. Peterson CL, Herskowitz I. Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell. 1992;68(3):573-83.

7. Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science. 2008;322(5908):1717-20.

8. Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B. Xist RNA and the mechanism of X chromosome inactivation. Annual review of genetics. 2002;36:233-78.

9. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007;129(7):1311-23.

10. Sanchez-Elsner T, Gou D, Kremmer E, Sauer F. Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. Science. 2006;311(5764):1118-23.

11. Wang P, Yin S, Zhang Z, Xin D, Hu L, Kong X, et al. Evidence for common short natural trans sense-antisense pairing between transcripts from protein coding genes. Genome biology. 2008;9(12):R169.

12. Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, et al. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature. 2008;454(7200):126-30.

13. Lanz RB, McKenna NJ, Onate SA, Albrecht U, Wong J, Tsai SY, et al. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell. 1999;97(1):17-27.

14. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes & development. 2009;23(7):781-3.

15. Holde KEV. Chromatin. New York: Springer-Verlag; 1988.

16. Thomas JO, Kornberg RD. An octamer of histones in chromatin and free in solution. Proceedings of the National Academy of Sciences of the United States of America. 1975;72(7):2626-30.

17. Prunell A, Kornberg RD. Variable center to center distance of nucleosomes in chromatin. Journal of molecular biology. 1982;154(3):515-23.

18. Finch JT, Klug A. Solenoidal model for superstructure in chromatin. Proceedings of the National Academy of Sciences of the United States of America. 1976;73(6):1897-901.

19. Paulson JR, Laemmli UK. The structure of histone-depleted metaphase chromosomes. Cell. 1977;12(3):817-28.

20. Heitz E. Das Heterochromatin der Moose1928. 762-818 p.

21. Brown SW. Heterochromatin. Science. 1966;151(3709):417-25.

22. Brockdorff N. X-chromosome inactivation: closing in on proteins that bind Xist RNA. Trends in genetics : TIG. 2002;18(7):352-8.

23. Orlando V. Polycomb, epigenomes, and control of cell identity. Cell. 2003;112(5):599-606.

24. Oberdoerffer P, Sinclair DA. The role of nuclear architecture in genomic instability and ageing. Nature reviews Molecular cell biology. 2007;8(9):692-702.

25. Jenuwein T, Allis CD. Translating the histone code. Science. 2001;293(5532):1074-80.

26. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975;187(4173):226-32.

27. Ohno S, Kaplan WD, Kinosita R. Formation of the sex chromatin by a single X-chromosome in liver cells of Rattus norvegicus. Experimental cell research. 1959;18:415-8.

28. Ahmad K, Henikoff S. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Molecular cell. 2002;9(6):1191-200.

29. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Molecular cell. 2011;44(4):667-78.

30. Dulac C. Brain function and chromatin plasticity. Nature. 2010;465(7299):728-35.

31. Jones PA. Altering gene expression with 5-azacytidine. Cell. 1985;40(3):485-6.

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

33. Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature. 2000;406(6796):593-9.

34. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proceedings of the National Academy of Sciences of the United States of America. 1964;51:786-94.

35. Nowak SJ, Corces VG. Phosphorylation of histone H3 correlates with transcriptionally active loci. Genes & development. 2000;14(23):3003-13.

36. Grunstein M. Histone acetylation in chromatin structure and transcription. Nature. 1997;389(6649):349-52.

37. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, et al. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell. 1996;84(6):843-51.

38. Cheung P, Allis CD, Sassone-Corsi P. Signaling to chromatin through histone modifications. Cell. 2000;103(2):263-71.

39. Kim JK, Samaranayake M, Pradhan S. Epigenetic mechanisms in mammals. Cellular and molecular life sciences : CMLS. 2009;66(4):596-612.

40. Sarma K, Reinberg D. Histone variants meet their match. Nature reviews Molecular cell biology. 2005;6(2):139-49.

41. Hassler MR, Egger G. Epigenomics of cancer - emerging new concepts. Biochimie. 2012;94(11):2219-30.

42. Saha A, Wittmeyer J, Cairns BR. Chromatin remodelling: the industrial revolution of DNA around histones. Nature reviews Molecular cell biology. 2006;7(6):437-47.

43. Saxena A, Carninci P. Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. BioEssays : news and reviews in molecular, cellular and developmental biology. 2011;33(11):830-9.

44. Fuks F, Burgers WA, Brehm A, Hughes-Davies L, Kouzarides T. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nature genetics. 2000;24(1):88-91.

45. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic acids research. 2003;31(9):2305-12.

46. Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, et al. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. The Journal of biological chemistry. 2010;285(34):26114-20.

47. Cervoni N, Szyf M. Demethylase activity is directed by histone acetylation. The Journal of biological chemistry. 2001;276(44):40778-87.

48. Robertson KD, Wolffe AP. DNA methylation in health and disease. Nature reviews Genetics. 2000;1(1):11-9.

49. Zangi R, Arrieta A, Cossio FP. Mechanism of DNA methylation: the double role of DNA as a substrate and as a cofactor. Journal of molecular biology. 2010;400(3):632-44.

50. Klimasauskas S, Kumar S, Roberts RJ, Cheng X. HhaI methyltransferase flips its target base out of the DNA helix. Cell. 1994;76(2):357-69.

51. Bestor TH. The DNA methyltransferases of mammals. Human molecular genetics. 2000;9(16):2395-402.

52. Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. The Journal of biological chemistry. 1999;274(46):33002-10.

53. Bestor T, Laudano A, Mattaliano R, Ingram V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. Journal of molecular biology. 1988;203(4):971-83.

54. Robertson KD, Keyomarsi K, Gonzales FA, Velicescu M, Jones PA. Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. Nucleic acids research. 2000;28(10):2108-13.

55. Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science. 1997;277(5334):1996-2000.

56. Leonhardt H, Page AW, Weier HU, Bestor TH. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell. 1992;71(5):865-73.

57. Ha K, Lee GE, Palii SS, Brown KD, Takeda Y, Liu K, et al. Rapid and transient recruitment of DNMT1 to DNA double-strand breaks is mediated by its interaction with multiple components of the DNA damage response machinery. Human molecular genetics. 2011;20(1):126-40.

58. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nature genetics. 1998;19(3):219-20.

59. Chen T, Ueda Y, Xie S, Li E. A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. The Journal of biological chemistry. 2002;277(41):38746-54.

60. Jones PA, Liang G. Rethinking how DNA methylation patterns are maintained. Nature reviews Genetics. 2009;10(11):805-11.

61. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 2013;38(1):23-38.

62. Beard C, Li E, Jaenisch R. Loss of methylation activates Xist in somatic but not in embryonic cells. Genes & development. 1995;9(19):2325-34.

63. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature. 1993;366(6453):362-5.

64. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992;69(6):915-26.

65. Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R, et al. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes & development. 2008;22(12):1607-16.

66. Sado T, Fenner MH, Tan SS, Tam P, Shioda T, Li E. X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. Developmental biology. 2000;225(2):294-303.

67. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.

68. Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, et al. Epigenetic reprogramming in mouse primordial germ cells. Mechanisms of development. 2002;117(1-2):15-23.

69. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. Nature. 2000;403(6769):501-2.

70. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature. 2012;484(7394):339-44.

71. Ji H, Ehrlich LI, Seita J, Murakami P, Doi A, Lindau P, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature. 2010;467(7313):338-42.

72. Szyf M, Rouleau J, Theberge J, Bozovic V. Induction of myogenic differentiation by an expression vector encoding the DNA methyltransferase cDNA sequence in the antisense orientation. The Journal of biological chemistry. 1992;267(18):12831-6.

73. Guo JU, Ma DK, Mo H, Ball MP, Jang MH, Bonaguidi MA, et al. Neuronal activity modifies the DNA methylation landscape in the adult brain. Nature neuroscience. 2011;14(10):1345-51.

74. Waterland RA, Kellermayer R, Rached MT, Tatevian N, Gomes MV, Zhang J, et al. Epigenomic profiling indicates a role for DNA methylation in early postnatal liver development. Human molecular genetics. 2009;18(16):3026-38.

75. Swami M. Epigenetics: Demethylation links cell fate and cancer. Nature reviews Cancer. 2010;10(11):740.

76. Rai K, Sarkar S, Broadbent TJ, Voas M, Grossmann KF, Nadauld LD, et al. DNA demethylase activity maintains intestinal cells in an undifferentiated state following loss of APC. Cell. 2010;142(6):930-42.

77. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. Leukemia. 2014;28(3):485-96.

78. Muto H, Sakata-Yanagimoto M, Chiba S. TET2 as a gatekeeper for hematologic malignancies. [Rinsho ketsueki] The Japanese journal of clinical hematology. 2015;56(6):651-6.

79. Patra SK, Patra A, Rizzi F, Ghosh TC, Bettuzzi S. Demethylation of (Cytosine-5-C-methyl) DNA and regulation of transcription in the epigenetic pathways of cancer development. Cancer metastasis reviews. 2008;27(2):315-34.

80. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324(5929):930-5.

81. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature. 2011;477(7366):606-10.

82. Iqbal K, Jin SG, Pfeifer GP, Szabo PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(9):3642-7.

83. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nature communications. 2011;2:241.

84. Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood. 2011;118(17):4509-18.

85. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer cell. 2011;20(1):11-24.

86. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science. 2011;333(6047):1303-7.

87. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. 2011;333(6047):1300-3.

88. Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer research. 2007;67(3):946-50.

89. Inoue A, Shen L, Dai Q, He C, Zhang Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell research. 2011;21(12):1670-6.

90. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. The Journal of biological chemistry. 2011;286(41):35334-8.

91. Nabel CS, Jia H, Ye Y, Shen L, Goldschmidt HL, Stivers JT, et al. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nature chemical biology. 2012;8(9):751-8.

92. Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, et al. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic acids research. 1982;10(8):2709-21.

93. Hellman A, Chess A. Gene body-specific methylation on the active X chromosome. Science. 2007;315(5815):1141-3.

94. Wolf SF, Jolly DJ, Lunnen KD, Friedmann T, Migeon BR. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. Proceedings of the National Academy of Sciences of the United States of America. 1984;81(9):2806-10.

95. Aran D, Toperoff G, Rosenberg M, Hellman A. Replication timing-related and gene body-specific methylation of active human genes. Human molecular genetics. 2011;20(4):670-80.

96. Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nature biotechnology. 2009;27(4):361-8.

97. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. Trends in genetics : TIG. 1997;13(8):335-40.

98. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature. 2010;466(7303):253-7.

99. Portela A, Esteller M. Epigenetic modifications and human disease. Nature biotechnology. 2010;28(10):1057-68.

100. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. Journal of molecular biology. 1987;196(2):261-82.

101. Bird AP. DNA methylation and the frequency of CpG in animal DNA. Nucleic acids research. 1980;8(7):1499-504.

102. Bird AP, Taggart MH. Variable patterns of total DNA and rDNA methylation in animals. Nucleic acids research. 1980;8(7):1485-97.

103. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(6):3740-5.

104. Edwards CA, Ferguson-Smith AC. Mechanisms regulating imprinted genes in clusters. Current opinion in cell biology. 2007;19(3):281-9.

105. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature. 2007;447(7143):425-32.

106. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nature genetics. 2009;41(2):178-86.

107. Tazi J, Bird A. Alternative chromatin structure at CpG islands. Cell. 1990;60(6):909-20.

108. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. Nature reviews Molecular cell biology. 2010;11(9):607-20.

109. Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nature genetics. 2006;38(6):626-35.

110. Cuadrado M, Sacristan M, Antequera F. Species-specific organization of CpG island promoters at mammalian homologous genes. EMBO reports. 2001;2(7):586-92.

111. Illingworth RS, Bird AP. CpG islands--'a rough guide'. FEBS letters. 2009;583(11):1713-20.

112. Ehrlich M, Ehrlich KC. Effect of DNA methylation on the binding of vertebrate and plant proteins to DNA. Exs. 1993;64:145-68.

113. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature. 1998;393(6683):386-9.

114. Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, et al. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nature genetics. 1999;23(1):58-61.

115. Wade PA, Gegonne A, Jones PL, Ballestar E, Aubry F, Wolffe AP. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nature genetics. 1999;23(1):62-6.

116. Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, et al. CDD: specific functional annotation with the Conserved Domain Database. Nucleic acids research. 2009;37(Database issue):D205-10.

117. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic acids research. 2011;39(Database issue):D225-9.

118. Zhang Y, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes & development. 1999;13(15):1924-35.

119. Nakao M, Matsui S, Yamamoto S, Okumura K, Shirakawa M, Fujita N. Regulation of transcription and chromatin by methyl-CpG binding protein MBD1. Brain & development. 2001;23 Suppl 1:S174-6.

120. Clouaire T, Stancheva I. Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cellular and molecular life sciences : CMLS. 2008;65(10):1509-22.

121. Hendrich B, Guy J, Ramsahoye B, Wilson VA, Bird A. Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes & development. 2001;15(6):710-23.

122. Lin RK, Hsu HS, Chang JW, Chen CY, Chen JT, Wang YC. Alteration of DNA methyltransferases contributes to 5'CpG methylation and poor prognosis in lung cancer. Lung cancer. 2007;55(2):205-13.

123. Meklat F, Li Z, Wang Z, Zhang Y, Zhang J, Jewell A, et al. Cancer-testis antigens in haematological malignancies. British journal of haematology. 2007;136(6):769-76.

124. Pancione M, Sabatino L, Fucci A, Carafa V, Nebbioso A, Forte N, et al. Epigenetic silencing of peroxisome proliferator-activated receptor gamma is a biomarker for colorectal cancer progression and adverse patients' outcome. PloS one. 2010;5(12):e14229.

125. Park HY, Jeon YK, Shin HJ, Kim IJ, Kang HC, Jeong SJ, et al. Differential promoter methylation may be a key molecular mechanism in regulating BubR1 expression in cancer cells. Experimental & molecular medicine. 2007;39(2):195-204.

126. Zhuravel E, Shestakova T, Glushko N, Soldatkina M, Pogrebnoy P. Expression patterns of murine beta-defensin-2 mRNA in Lewis lung carcinoma cells in vitro and in vivo. Experimental oncology. 2008;30(3):206-11.

127. McGough JM, Yang D, Huang S, Georgi D, Hewitt SM, Rocken C, et al. DNA methylation represses IFN-gamma-induced and signal transducer and activator of transcription 1-mediated IFN regulatory factor 8 activation in colon carcinoma cells. Molecular cancer research : MCR. 2008;6(12):1841-51.

128. Villa R, Morey L, Raker VA, Buschbeck M, Gutierrez A, De Santis F, et al. The methyl-CpG binding protein MBD1 is required for PML-RARalpha function. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(5):1400-5.

129. Spruijt CG, Vermeulen M. DNA methylation: old dog, new tricks? Nature structural & molecular biology. 2014;21(11):949-54.

130. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(5):1827-31.

131. Hayatsu H, Wataya Y, Kai K, Iida S. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. Biochemistry. 1970;9(14):2858-65.

132. Wang RY, Gehrke CW, Ehrlich M. Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. Nucleic acids research. 1980;8(20):4777-90.

133. Kristensen LS, Hansen LL. PCR-based methods for detecting single-locus DNA methylation biomarkers in cancer diagnostics, prognostics, and response to treatment. Clinical chemistry. 2009;55(8):1471-83.

134. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(18):9821-6.

135. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic acids research. 1997;25(12):2532-4.

136. Wojdacz TK, Dobrovic A. Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. Nucleic acids research. 2007;35(6):e41.

137. Wojdacz TK, Dobrovic A, Hansen LL. Methylation-sensitive high-resolution melting. Nature protocols. 2008;3(12):1903-8.

138. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nature genetics. 2005;37(8):853-62.

139. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011;98(4):288-95.

140. Dedeurwaerder S, Defrance M, Bizet M, Calonne E, Bontempi G, Fuks F. A comprehensive overview of Infinium HumanMethylation450 data processing. Briefings in bioinformatics. 2014;15(6):929-41.

141. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011;3(6):771-84.

142. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic acids research. 2005;33(18):5868-77.

143. Diagenode. Premium RRBS Kit. 2015.

144. Hehlmann R, Hochhaus A, Baccarani M, European L. Chronic myeloid leukaemia. Lancet. 2007;370(9584):342-50.

145. AmericanCancerSociety. Cancer Facts&Figures 20142014. Available from: http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf.

146. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes2009. 937-51 p.

147. Schiffer CA, Hehlmann R, Larson R. Perspectives on the treatment of chronic phase and advanced phase CML and Philadelphia chromosome positive ALL(1). Leukemia. 2003;17(4):691-9.

148. Apperley JF. Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia. The Lancet Oncology. 2007;8(11):1018-29.

149. Melo JV, Barnes DJ. Chronic myeloid leukaemia as a model of disease evolution in human cancer. Nature reviews Cancer. 2007;7(6):441-53.

150. Nowell PC, Hungerford DA. Chromosome studies on normal and leukemic human leukocytes. Journal of the National Cancer Institute. 1960;25:85-109.

151. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature. 1973;243(5405):290-3.

152. de Klein A, van Kessel AG, Grosveld G, Bartram CR, Hagemeijer A, Bootsma D, et al. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature. 1982;300(5894):765-7.

153. Groffen J, Stephenson JR, Heisterkamp N, de Klein A, Bartram CR, Grosveld G. Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell. 1984;36(1):93-9.

154. Konopka JB, Watanabe SM, Witte ON. An alteration of the human c-abl protein in K562 leukemia cells unmasks associated tyrosine kinase activity. Cell. 1984;37(3):1035-42.

155. Morris CM, Reeve AE, Fitzgerald PH, Hollings PE, Beard ME, Heaton DC. Genomic diversity correlates with clinical variation in Ph'-negative chronic myeloid leukaemia. Nature. 1986;320(6059):281-3.

156. Lydon N. Attacking cancer at its foundation. Nature medicine. 2009;15(10):1153-7.

157. Daley GQ, Van Etten RA, Baltimore D. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. Science. 1990;247(4944):824-30.

158. Fialkow PJ, Jacobson RJ, Papayannopoulou T. Chronic myelocytic leukemia: clonal origin in a stem cell common to the granulocyte, erythrocyte, platelet and monocyte/macrophage. The American journal of medicine. 1977;63(1):125-30.

159. Dameshek W. Some speculations on the myeloproliferative syndromes. Blood. 1951;6(4):372-5.

160. Goto T, Nishikori M, Arlin Z, Gee T, Kempin S, Burchenal J, et al. Growth characteristics of leukemic and normal hematopoietic cells in Ph' + chronic myelogenous leukemia and effects of intensive treatment. Blood. 1982;59(4):793-808.

161. Jonas D, Lubbert M, Kawasaki ES, Henke M, Bross KJ, Mertelsmann R, et al. Clonal analysis of bcr-abl rearrangement in T lymphocytes from patients with chronic myelogenous leukemia. Blood. 1992;79(4):1017-23.

162. Jamieson CH, Ailles LE, Dylla SJ, Muijtjens M, Jones C, Zehnder JL, et al. Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. The New England journal of medicine. 2004;351(7):657-67.

163. Coluccia AM, Vacca A, Dunach M, Mologni L, Redaelli S, Bustos VH, et al. Bcr-Abl stabilizes beta-catenin in chronic myeloid leukemia through its tyrosine phosphorylation. The EMBO journal. 2007;26(5):1456-66.

164. Xie S, Lin H, Sun T, Arlinghaus RB. Jak2 is involved in c-Myc induction by Bcr-Abl. Oncogene. 2002;21(47):7137-46.

165. Kirchner D, Duyster J, Ottmann O, Schmid RM, Bergmann L, Munzert G. Mechanisms of Bcr-Abl-mediated NF-kappaB/Rel activation. Experimental hematology. 2003;31(6):504-11.

166. Cortez D, Reuther G, Pendergast AM. The Bcr-Abl tyrosine kinase activates mitogenic signaling pathways and stimulates G1-to-S phase transition in hematopoietic cells. Oncogene. 1997;15(19):2333-42.

167. Goga A, McLaughlin J, Afar DE, Saffran DC, Witte ON. Alternative signals to RAS for hematopoietic transformation by the BCR-ABL oncogene. Cell. 1995;82(6):981-8.

168. Nieborowska-Skorska M, Wasik MA, Slupianek A, Salomoni P, Kitamura T, Calabretta B, et al. Signal transducer and activator of transcription (STAT)5 activation by BCR/ABL is dependent on intact Src homology (SH)3 and SH2 domains of BCR/ABL and is required for leukemogenesis. The Journal of experimental medicine. 1999;189(8):1229-42.

169. Sattler M, Mohi MG, Pride YB, Quinnan LR, Malouf NA, Podar K, et al. Critical role for Gab2 in transformation by BCR/ABL. Cancer cell. 2002;1(5):479-92.

170. Alimena G, De Cuia MR, Diverio D, Gastaldi R, Nanni M. The karyotype of blastic crisis. Cancer genetics and cytogenetics. 1987;26(1):39-50.

171. Bernstein R. Cytogenetics of chronic myelogenous leukemia. Seminars in hematology. 1988;25(1):20-34.

172. Cervantes F, Ballesta F, Mila M, Rozman C. Cytogenetic studies in blast crisis of Ph-positive chronic granulocytic leukemia: results and prognostic evaluation in 52 patients. Cancer genetics and cytogenetics. 1986;21(3):239-46.

173. Ahuja H, Bar-Eli M, Advani SH, Benchimol S, Cline MJ. Alterations in the p53 gene and the clonal evolution of the blast crisis of chronic myelocytic leukemia. Proceedings of the National Academy of Sciences of the United States of America. 1989;86(17):6783-7.

174. Calabretta B, Perrotti D. The biology of CML blast crisis. Blood. 2004;103(11):4010-22.

175. Feinstein E, Cimino G, Gale RP, Alimena G, Berthier R, Kishi K, et al. p53 in chronic myelogenous leukemia in acute phase. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(14):6293-7.

176. Sill H, Goldman JM, Cross NC. Homozygous deletions of the p16 tumor-suppressor gene are associated with lymphoid transformation of chronic myeloid leukemia. Blood. 1995;85(8):2013-6.

177. Boultwood J, Peniket A, Watkins F, Shepherd P, McGale P, Richards S, et al. Telomere length shortening in chronic myelogenous leukemia is associated with reduced time to accelerated phase. Blood. 2000;96(1):358-61.

178. Brummendorf TH, Holyoake TL, Rufer N, Barnett MJ, Schulzer M, Eaves CJ, et al. Prognostic implications of differences in telomere length between normal and malignant cells from patients with chronic myeloid leukemia measured by flow cytometry. Blood. 2000;95(6):1883-90.

179. Drummond M, Lennard A, Brummendorf T, Holyoake T. Telomere shortening correlates with prognostic score at diagnosis and proceeds rapidly during progression of chronic myeloid leukemia. Leukemia & lymphoma. 2004;45(9):1775-81.

180. Bodoor K, Haddad Y, Alkhateeb A, Al-Abbadi A, Dowairi M, Magableh A, et al. DNA hypermethylation of cell cycle (p15 and p16) and apoptotic (p14, p53, DAPK and TMS1) genes in peripheral blood of leukemia patients. Asian Pacific journal of cancer prevention : APJCP. 2014;15(1):75-84.

181. Dunwell T, Hesson L, Rauch TA, Wang L, Clark RE, Dallol A, et al. A genome-wide screen identifies frequently methylated genes in haematological and epithelial cancers. Molecular cancer. 2010;9:44.

182. Nguyen TT, Mohrbacher AF, Tsai YC, Groffen J, Heisterkamp N, Nichols PW, et al. Quantitative measure of c-abl and p15 methylation in chronic myelogenous leukemia: biological implications. Blood. 2000;95(9):2990-2.

183. Uehara E, Takeuchi S, Yang Y, Fukumoto T, Matsuhashi Y, Tamura T, et al. Aberrant methylation in promoter-associated CpG islands of multiple genes in chronic myelogenous leukemia blast crisis. Oncology letters. 2012;3(1):190-2.

184. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature. 1983;301(5895):89-92.

185. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nature reviews Cancer. 2011;11(10):726-34.

186. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nature reviews Genetics. 2007;8(4):286-98.

187. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis. 2010;31(1):27-36.

188. Zion M, Ben-Yehuda D, Avraham A, Cohen O, Wetzler M, Melloul D, et al. Progressive de novo DNA methylation at the bcr-abl locus in the course of chronic myelogenous leukemia. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(22):10722-6.

189. Asimakopoulos FA, Shteper PJ, Krichevsky S, Fibach E, Polliack A, Rachmilewitz E, et al. ABL1 methylation is a distinct molecular event associated with clonal evolution of chronic myeloid leukemia. Blood. 1999;94(7):2452-60.

190. Amabile G, Di Ruscio A, Muller F, Welner RS, Yang H, Ebralidze AK, et al. Dissecting the role of aberrant DNA methylation in human leukaemia. Nature communications. 2015;6:7091.

191. Byun H, Eshaghian, S., Douer, D., Trent, J., Garcia-Manero, G., Bhatia, R., Siegmund, K., Yang, A.S. DNA methylation profiling of chronic myelogenous leukemia in relationship to genomic translocation. Epigenetics in Cancer. 2013;1(1):1-9.

192. Mizuno S, Chijiwa T, Okamura T, Akashi K, Fukumaki Y, Niho Y, et al. Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. Blood. 2001;97(5):1172-9.

193. Mancini M, Veljkovic N, Leo E, Aluigi M, Borsi E, Galloni C, et al. Cytoplasmatic compartmentalization by Bcr-Abl promotes TET2 loss-of-function in chronic myeloid leukemia. Journal of cellular biochemistry. 2012;113(8):2765-74.

194. Jelinek J, Gharibyan V, Estecio MR, Kondo K, He R, Chung W, et al. Aberrant DNA methylation is associated with disease progression, resistance to imatinib and shortened survival in chronic myelogenous leukemia. PloS one. 2011;6(7):e22110.

195. Qian J, Wang YL, Lin J, Yao DM, Xu WR, Wu CY. Aberrant methylation of the death-associated protein kinase 1 (DAPK1) CpG island in chronic myeloid leukemia. European journal of haematology. 2009;82(2):119-23.

196. Yang MY, Liu TC, Chang JG, Lin PM, Lin SF. JunB gene expression is inactivated by methylation in chronic myeloid leukemia. Blood. 2003;101(8):3205-11.

197. Strathdee G, Holyoake TL, Sim A, Parker A, Oscier DG, Melo JV, et al. Inactivation of HOXA genes by hypermethylation in myeloid and lymphoid malignancy is frequent and associated with poor prognosis. Clinical cancer research : an official journal of the American Association for Cancer Research. 2007;13(17):5048-55.

198. Akalin A, Garrett-Bakelman FE, Kormaksson M, Busuttil J, Zhang L, Khrebtukova I, et al. Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. PLoS genetics. 2012;8(6):e1002781.

199. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome biology. 2012;13(10):R87.

200. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. Nucleic acids research. 2015;43(W1):W566-70.

201. Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic acids research. 2013;41(Database issue):D377-86.

202. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology. 2010;28(5):495-501.

203. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic acids research. 2010;38(Web Server issue):W214-20.

204. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, et al. Galaxy: a web-based genome analysis tool for experimentalists. Current protocols in molecular biology / edited by Frederick M Ausubel [et al]. 2010;Chapter 19:Unit 19 0 1-21.

205. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: a platform for interactive large-scale genome analysis. Genome research. 2005;15(10):1451-5.

206. Goecks J, Nekrutenko A, Taylor J, Galaxy T. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome biology. 2010;11(8):R86.

207. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology. 2013;14(4):R36.

208. Kaiser MF, Johnson DC, Wu P, Walker BA, Brioli A, Mirabella F, et al. Global methylation analysis identifies prognostically important epigenetically inactivated tumor suppressor genes in multiple myeloma. Blood. 2013;122(2):219-26.

209. Kang S, Dong SM, Park NH. Frequent promoter hypermethylation of TGFBI in epithelial ovarian cancer. Gynecologic oncology. 2010;118(1):58-63.

210. Li Z, Luo RT, Mi S, Sun M, Chen P, Bao J, et al. Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias. Cancer research. 2009;69(3):1109-16.

211. Shah JN, Shao G, Hei TK, Zhao Y. Methylation screening of the TGFBI promoter in human lung and prostate cancer by methylation-specific PCR. BMC cancer. 2008;8:284.

212. Caren H, Djos A, Nethander M, Sjoberg RM, Kogner P, Enstrom C, et al. Identification of epigenetically regulated genes that predict patient outcome in neuroblastoma. BMC cancer. 2011;11:66.

213. Kim JE, Kim SJ, Lee BH, Park RW, Kim KS, Kim IS. Identification of motifs for cell adhesion within the repeated domains of transforming growth factor-beta-induced gene, betaig-h3. The Journal of biological chemistry. 2000;275(40):30907-15.

214. Lang F, Shumilina E. Regulation of ion channels by the serum- and glucocorticoid-inducible kinase SGK1. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2013;27(1):3-12.

215. Chen W, Xiang J, Chen DF, Ni BB, Chen H, Fan XJ, et al. Screening for differentially methylated genes among human colorectal cancer tissues and normal mucosa by microarray chip. Molecular biology reports. 2013;40(5):3457-64.

216. Kawamura YI, Toyota M, Kawashima R, Hagiwara T, Suzuki H, Imai K, et al. DNA hypermethylation contributes to incomplete synthesis of carbohydrate determinants in gastrointestinal cancer. Gastroenterology. 2008;135(1):142-51 e3.

217. Yang WH, Nussbaum C, Grewal PK, Marth JD, Sperandio M. Coordinated roles of ST3Gal-VI and ST3Gal-IV sialyltransferases in the synthesis of selectin ligands. Blood. 2012;120(5):1015-26.

7 Abstract

Aberrant DNA methylation is a frequent event in the pathogenesis of various cancer types and is associated with reduced gene expression. Due to the limited knowledge about DNA methylation landscapes in chronic myeloid leukemia (CML), we conducted a genome-wide methylation study in the leukemia cell lines MO7e, MO7e/p210 and in normal bone marrow controls (nBM) by reduced-representation bisulfite sequencing (RRBS). In addition, we investigated gene expression using RNA-Sequencing (RNA-Seq) and compared RRBS data and RNA-Seq data. Overall, we observed higher CpG site methylation levels in cell lines compared to nBM and identified ~230.000 methylated CpG sites between MO7e/p210 and NBM. Moreover, we found ~2.000 differentially methylated CpG sites between MO7e/p210 and MO7e. While CpG sites with increased methylation were mainly located in CpG islands (CGIs) and CGI shores associated with promoter regions, CpG sites with decreased methylation were predominantly located within introns and intragenic regions in large distance to transcription start sites (TSS). By searching for enriched GO/MSigDB terms, we observed that genes with increased methylation in MO7e/p210 were mainly associated with developmental processes, cell differentiation and abnormal disease phenotypes. By contrast, genes with decreased methylation were mainly associated with angiogenesis. RNA-Seq analyses revealed that 442 genes are differentially expressed between MO7e/p210 and nBM. The majority of these genes was found to be downregulated in MO7e/p210 and included known/putative tumor suppressors and regulators of apoptosis, cell growth and cell adhesion. By comparison of DNA methylation and gene expression data, we identified genes with increased methylation, of which the majority was downregulated in MO7e/p210 compared to nBM. Within this gene cohort, we identified ~660 genes that were strongly downregulated in MO7e/p210 compared to MO7e and nBM. Many of these genes were found to be involved in apoptosis, cell adhesion and immune defense. Gene-specific methylation analyses of *TGFBI*, *SGK1* and *ST3GAL6* revealed elevated promoter methylation of these genes in CML patient samples.

Overall, our findings demonstrate that RRBS is an efficient approach to investigate methylation genome-wide. In addition, we identified many genes which

are methylated in CML cell lines and which are transcriptionally downregulated in these cells.

8 Zusammenfassung

Die aberrante DNA Methylierung ist ein häufig auftretendes Ereignis in der Pathenogenese verschiedener Krebsarten und ist mit einer reduzierten Genexpression verbunden. Aufgrund des bisher limitierten Wissens über DNA Methylierungsmuster in chronischer myeloischer Leukämie (CML), führten wir mittels der Reduced Representation Bisulfite Sequencing (RRBS) Methode eine genomweite Methylierungsstudie in den Leukämiezelllinien MO7e und MO7e/p210 sowie normalem Knochenmark (nKM) als Kontrolle durch. Zusätzlich untersuchten wir unter Verwendung von RNA-Sequencing (RNA-Seq) die genomweite Genexpression und verglichen RRBS Daten mit RNA-Seq Daten. Insgesamt konnten wir eine deutlich höhere CpG Methylierung in den Zelllinien im Vergleich zu nKM feststellen und ~ 230.000 unterschiedlich methylierte CpGs zwischen MO7e/p210 und nKM identifizieren. Weiters identifizierten wir ~ 2.000 unterschiedlich methylierte CpGs zwischen MO7e/p210 und MO7e. Methylierte CpGs waren vorwiegend in CpG islands (CGIs) und CGI shores in der Nähe von Genpromotorregionen zu finden, während unmethylierte CpGs vermehrt mit Introns und intragenischen Abschnitten abseits von Transkriptionsstarts (TSS) assoziiert werden konnten. Gene mit erhöhter Methylierung in MO7e/p210 konnten mit Entwicklungsprozessen, Zelldifferenzierung und abnormen Krankheitsphänotypen in Verbindung gebracht werden. Dagegen wurden hypomethylierte Gene mit Angiogenese assoziiert. Mit der RNA-Seq Datenanalyse wurden 442 Gene mit unterschiedlicher Expression zwischen MO7e/p210 und nKM identifiziert. Der Mehrheit dieser Gene zeigte eine verminderte mRNA Expression und einige der Gene wurden als bekannte/potentielle Tumorsuppressoren und Regulatoren von Apoptose, Zellwachstum und Zelladhäsion beschrieben. In einem umfassenden Vergleich von Methylierungs- und Genexpressionsdaten konnten Gene mit erhöhtem Methylierungsgrad, dessen Mehrheit mit vermindeter mRNA Expression korrelierte, in MO7e/p210 gefunden werden. Die mRNA Expression von 660 dieser Gene waren im Vergleich zu MO7e und nKM vermindert. Einige dieser Gene konnten mit Apoptose, Zelladhäsion und Immunabwehr assoziiert werden. Gen-spezifische Methylierungsanalysen von *TGFBI, SGK1* und *ST3GAL6* ergaben, dass diese Gene in der Mehrheit der Proben von CML-Patienten methyliert waren.

Zusammenfassend zeigen unsere Ergebnisse, dass RRBS eine effiziente Methode zur Untersuchung genomweiter DNA Methylierung darstellt. Wir konnten Gene identifizieren, die in den Leukämiezelllinien durch Methylierung transkriptionell reguliert werden.

9 Curriculum vitae

Personal data

Name	Thais Topakian
Year of birth	1990
Place of residence	Vienna
E-mail	thais.topakian@gmx.at
Nationality	Austrian

Education

2012-2015	Master Studies Genetics and Developmental Biology
09/2012	Bachelor Degree Biology
2008-2012	Bachelor Studies Biology
2000-2008	High school (AHS Waidhofen/Thaya, Austria) graduated summa cum laude

Work experience

10/2013-12/2014	Master Thesis “Methylome and Transcriptome Analyses in MO7e and MO7e^{p210} cells”; supervised by Univ. Prof. Dr. Zöchbauer-Müller, Division of Oncology, Department of Medicine I, Medical University of Vienna
03-04/2012	Internship “Effects of UV irradiation on RNA Pol II and splicing patterns in Physcomitrella patens”; supervised by Ao. Univ.-Prof. Mag. Dr. rer. nat. Andrea Barta, Max F. Perutz Laboratories, Medical university of Vienna
07-08/2012	Volunteering in an archaeological excavation for copper mining, Tyrol, Austria
07-08/2011	Volunteering in an archaeological excavation for copper mining, Tyrol, Austria

Lab skills

Epigenetic research techniques: sodium-bisulfite treatment of genomic DNA, methylation-sensitive high-resolution melt analysis (MS-HRM), genomic bisulfite sequencing (BGS), and preparation of genomic DNA for Reduced Representation Bisulfite Sequencing

Others: isolation of nucleic acids and proteins, PCR, RT-PCR, Western blot, cell culture, cloning, colony formation assay, transformation, immunostaining, WGA, Cell Titer Blue Assay, Tunnel Assay, FACS

Computer skills

Microsoft Office, IBM SPSS, Graphpad Prism, Methyl Primer Express Software, Seqmonk Software, Methylkit Software, Clustvis Software, Ontologizer, RotorGene Q Software, Image Lab

Other skills

- Languages: German (mother tongue), English (fluently), Latin, Russian (basics), Spanish (basics)
- Good presentation skills
- Driving License B