

Elucidation of the epigenetic code of the chromosomal region 13q14.3 in Chronic Lymphocytic Leukemia (CLL)

vom Fachbereich Biologie der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades
eines

Doctor rerum naturalium (Dr. rer. nat.)
genehmigte Dissertation

von
Dipl. Biol. Melanie Ruppel
aus Vreden

Referenten: Prof. Dr. H. Ulrich Göringer
Prof. Dr. Peter Lichter
Prof. Dr. Markus Löbrich

Tag der Einreichung: 06.01.2009
Tag der mündlichen Prüfung: 19.02.2009

Darmstadt, Januar 2009

to Stephan

Science never solves a problem without creating ten more.
George Bernard Shaw

The most exciting phrase to hear in science, the one that heralds new discoveries,
is not 'Eureka!' but 'That's funny...'
Isaac Asimov

Contents

Abbreviations.....	iii
1 Introduction.....	1
1.1 Cancer	1
1.1.1 Molecular biology and genetics of cancer cells	3
1.1.2 Chromatin remodeling in normal and malignant development	6
1.1.3 Epigenetic events in cancer	13
1.2 Leukemias and lymphomas	14
1.2.1 Hematopoiesis	16
1.2.2 B-cell development.....	17
1.2.3 Aberrant B-cell development	19
1.3 CLL	20
1.3.1 Genomic aberrations in CLL and prognostic relevance	22
1.3.2 Chromosomal band 13q14.3.....	23
1.3.3 Mechanisms of regulating gene expression	27
1.3.4 Epigenetic aberrations in CLL	30
1.4 Aims of this work	31
2 Material & Methods	33
2.1 Biological material	33
2.2 Chemicals and biochemicals.....	34
2.3 Enzymes.....	36
2.4 Antibodies.....	36
2.5 Primer sequences	37
2.6 Cell culture material.....	37
2.7 Kits	37
2.8 Other materials	38
2.9 Standard solutions.....	38
2.10 Instruments	42
2.11 Software.....	43
2.12 Preparation of proteins.....	44
2.13 Immunoprecipitation of proteins	45
2.14 Chromatin immunoprecipitation (ChIP) assay	45
2.15 Protein analysis	47
2.16 DNA analysis	48
2.17 BioCOBRA	54
2.18 Fluorescence in situ hybridization (FISH)	54
2.19 Microarray based methods.....	56
2.20 Statistical analysis.....	60
2.21 Experimental set up.....	63
3 Results	65
3.1 Establishment and validation of chromatin immunoprecipitation.....	65
3.1.1 Preparation of chromatin – nChIP versus xChIP	66
3.1.2 To crosslink or not to crosslink?.....	68
3.1.3 Chromatin fragmentation – Native nucleosomes versus sonicated chromatin.	69
3.1.4 Optimization of the precipitation procedure	70
3.1.5 Performance of antibodies – Which antibodies precipitate chromatin?.....	71
3.1.6 Optimization of ChIP – How to analyze precipitated DNA?.....	73
3.1.7 Validation of chromatin precipitation in H3K4me2- and macroH2A-ChIPs.....	74
3.2 Chromatin modifications at 13q14.3 in non-malignant cells	78
3.2.1 Are there differences in enrichment of H3K4me2 between genes and promoters of the 13q14.3 region?.....	78
3.2.2 Promoter-restricted H3K4me2-pattern correlates with monoallelic expression.	79
3.2.3 MacroH2A-enrichment at 13q14.3 in non-malignant cells.....	81

3.2.4	Comparison of histone code and DNA-methylation of chromosomal region 13q14.3 .	82
3.2.5	Allelic modifications of CpG-islands D and E – Is there a regulatory element in CGI D or CGI E?	85
3.2.6	Histone modifications in different age groups	86
3.2.7	Differences in H3K4me2-enrichment between B- and T-cells	88
3.3	Aberrant chromatin of chromosomal region 13q14.3 in CLL.	89
3.3.1	H3K4me2-enrichment at 13q14.3 in CLL	89
3.3.2	Promoter-restricted H3K4me2-pattern correlates with monoallelic expression and is not changed in CLL	90
3.3.3	MacroH2A-enrichment at 13q14.3 in CLL.....	90
3.3.4	Histone code and DNA-methylation of chromosomal region 13q14.3 in CLL.....	91
3.3.5	Aberrant chromatin at 13q14.3 in CLL?	94
3.3.6	Correlation of macroH2A- and H3K4me2-enrichment at 13q14.3 with different karyotypes in CLL.....	95
3.3.7	Identification of a locus control element in CGI D or E.....	99
4	Discussion	103
4.1	Epigenetic regulation of 13q14.3?	104
4.2	Analysis of epigenetic modifications at 13q14.3	104
4.2.1	Establishment of the ChIP protocol	105
4.2.2	Why of all histone modifications quantify H3K4me2 and macroH2A at 13q14.3?	109
4.3	Epigenetic modifications at 13q14.3 in non-malignant cells	111
4.3.1	The epigenetic code of 13q14.3 in non-malignant hematopoietic cells.....	111
4.3.2	Histone code of monoallelic expression at 13q14.3.....	115
4.3.3	Are epigenetic marks at 13q14.3 changed during aging?	116
4.3.4	Tissue-specificity of epigenetic marks at 13q14.3. Are there differences between B- and T-cells?	117
4.3.5	Differential epigenetic marks at the two 13q14 copies? Model of the epigenetically regulated tumor suppressor mechanism in the critical region in 13q14.3.	118
4.3.6	Consequences of the epigenetic code at 13q14.3 for the tumor suppressor mechanism in the critical region 13q14.3.....	120
4.4	The aberrant epigenetic code of 13q14.3 in CLL	123
4.4.1	Aberrant chromatin modifications at 13q14.3 in CLL.....	124
4.4.2	Epimutations or aging-related changes at 13q14.3?	128
4.4.3	Correlation of epigenetic aberrations at 13q14.3 and genomic aberrations in five CLL patients.....	129
4.4.4	Identification of a potential epimutation in the postulated LCR of the critical region in 13q14.3	131
4.4.5	Model of epigenetic deregulation of the tumor suppressor mechanism in 13q14.3 in CLL.....	132
	Summary	135
	Zusammenfassung.....	137
	Bibliography	139
	Appendix A – Additional tables.....	157
	Appendix B – Primer sequences	162
	Publications.....	169
	Curriculum vitae	171
	Acknowledgments	173
	Eidesstattliche Erklärung	175

Abbreviations

A	Adenine
AML	Acute Myeloid Leukemia
APS	Ammonium persulfate
C	Cytosin
BCA	Bicinchonin acid
<i>BCMS</i>	Deleted in lymphocytic leukemia 1
<i>BCMSUN</i>	BCMS upstream neighbor
BCR	B-cell receptor
bp	Base pairs
β-ME	Beta-mercapto ethanol
BSA	Bovine serum albumin
C	Cytosine
<i>C13ORF1</i>	Chromosome 13 open reading frame 1
CD	Cluster of differentiation
cDNA	Complementary DNA
CGI	CpG-island
ChIP	Chromatin immunoprecipitation
CLL	B-cell Chronic Lymphocytic Leukemia
DAPI	4',6-Diamidino-2'-phenylindol-dihydrochloride
ddNTP	Didesoxy-nucleotide triphosphate mix
<i>DLeu2</i>	Deleted in Lymphocytic Leukemia 2
DMR	Differentially methylated region
DNA	Desoxy-ribonucleic acid
DNMT	DNA methyl-transferase
dNTP	Desoxy-nucleotide triphosphate mix
<i>E. coli</i>	<i>Escherichia coli</i>
e.g.	For example (from Latin: exempli gratia)
ELISA	Enzyme-linked immunosorbent assay
FISH	Fluorescence <i>in situ</i> hybridization
G	Guanine
g	Unit of acceleration
Granta-519	Mantle cell lymphoma cell line
h	hour(s)
HAT	Histone acetyl-transferase
HDAC	Histone deacetylase
HL60	Acute myeloid leukemia cell line
HMT	Histone methyl-transferase
HRP	Horseradish peroxidase
ICE	Imprinting control element
i.e.	That is (from Latin: id est)
IF	Immunofluorescence
Ig	Immunoglobulin
IP	Immunoprecipitation
Jurkat	T-cell leukemia cell line
kb	Kilo base pairs
KD	Knowledge Base
kDa	Kilo Dalton
<i>KPNA3</i>	Karyopherin alpha 3
LCR	Locus control region

Abbreviations

mA	Milli Ampere
MCL	Mantle Cell Lymphoma
min	Minute(s)
MM	Multiple Myeloma
MNase	<i>MICROCOCCUS</i> nuclease
mRNA	Messenger RNA
miRNA	Micro RNA
NAMALWA	Burkitt's lymphoma cell line
NCBI	National Center for Biotechnology Information
ncRNA	Noncoding RNA
OD	Optical density
ORF	Open reading frame
PB	Peripheral blood
PBL	Peripheral blood lymphocytes
PCR	Polymerase chain reaction
PVDF	Polyvinylidene fluoride
qPCR	Quantitative polymerase chain reaction
<i>RFP2</i>	Ret finger protein 2
<i>RFP2OS</i>	Ret finger protein 2 opposite strand
RNA	Ribonucleic acid
RT	Room temperature
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
s	Second(s)
<i>SETDB2</i>	SET domain, bifurcated 2
SNP	Single nucleotide polymorphism
T	Thymine
TBE	Tris-Borate-Buffer
TSG	Tumor suppressor gene
U	Uracil
U	Units
UTR	Untranslated region
UV light	Ultraviolet light
V	Variable
V	Volt
V(H); V _H	Variable heavy chain
v/v	Volume per volume
WB	Western blot analysis
w/v	Weight per volume

1 Introduction

1.1 Cancer

Cancer is the second most common cause of death in Germany and the Western world after cardio-vascular diseases. According to estimations by the World Health Organization (WHO), cancer causes around 25 % of all deaths worldwide. In Germany in 2006, 22 % and 28 % of all deaths in women and men, respectively, were cancer-related (Figure 1-1). More recent health statistics show an increase in cancer-related deaths in Western countries. Jemal et al. (2008) estimated that cancer will be the most frequent cause of death in the next years.

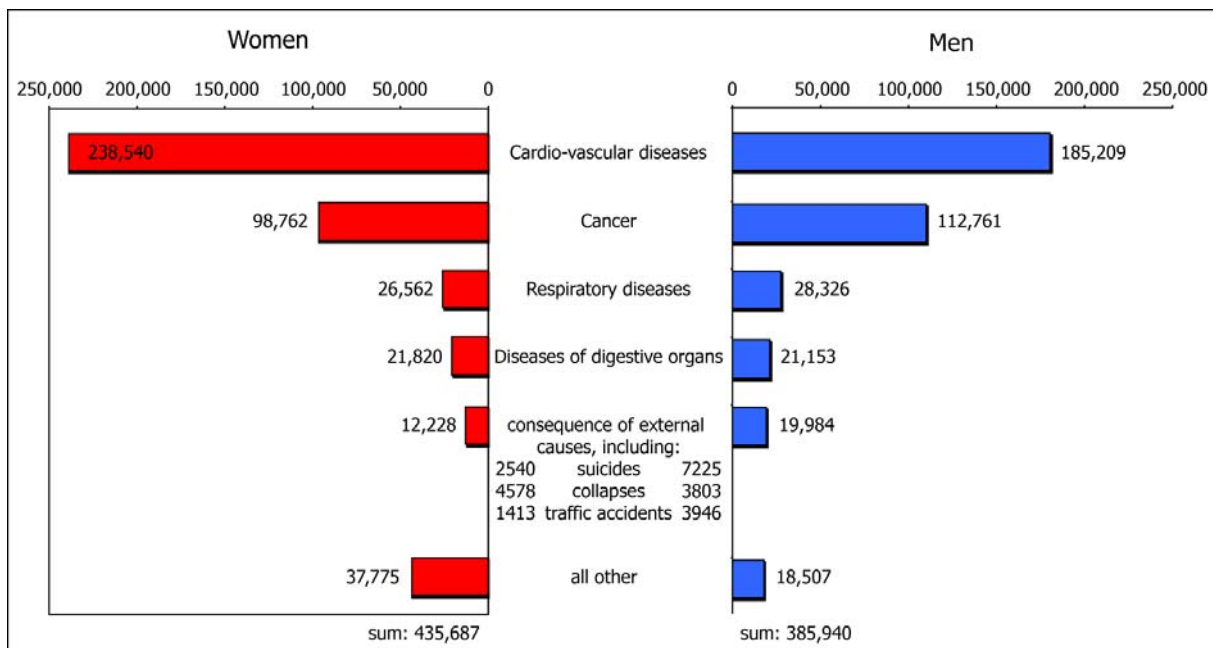


Figure 1-1: Most frequent causes of death among women and men in Germany in 2006 (modified from Statistisches Bundesamt Deutschland (2008); <http://www.destatis.de>)

The development of cancer is the result of microevolution of an initial cell clone towards malignancy. Such aberrant development is suppressed in healthy organisms by multiple defense systems (Klein et al., 2007). Clearly, all tumors arise from normal tissue, which is due to the great tissue plasticity created by autonomous and versatile cells. Every cell of an organism carries in its genome all the information required for the function of that organism, but the shape and function of a single differentiating cell is tightly regulated by expression of a distinct set of genes required for the function of the cell in its tissue environment. Non-malignant cells are programmed to construct and maintain the function of the tissues they are part of. However, malignant cells ignore this cellular program and acquire uncontrolled growth, prevention of apoptosis and/or tissue-invasion. These phenotypic abilities are acquired by alterations of the genomic sequence or the epigenome or by loss of the regulated expression in malignant cells. This can lead to a malignant phenotype and tumorigenic deregulation of tissue growth. Thus, versatility and autonomy of individual cells bear the danger of loss of tissue integrity, especially when individual cells gain access to information in their genomes which is normally denied to them (Weinberg, 2007).

Alterations of the genomic information in malignant cells can be point mutations, deletions, gains of genomic material or translocations as well as epimutations. It has been discussed that tumorigenesis involves loss other than gain of functions, which affects one or several of the numerous cellular surveillance mechanism that are normally involved in cancer prevention (Klein et al., 2007). Cellular pathways that are affected in cancer by (epi)mutations are e.g. loss of function of DNA repair and cell cycle check-point controls, apoptosis, chromatin structure integrity and imprinting stringency as well as tumor microenvironment (Klein et al., 2007).

To date, the process of cancer development can solely be explained from an evolutionary point of view. The evolution of cells led to development of very effective defense mechanisms that prevent transformation of malignant cells and therefore cancer. Cancer cells must go through a process very similar to natural selection. Within tumors, only those cells with mutations that best promote growth and enhance survival are selected to multiply, whereas malignant cells with less efficient mutations are outcompeted within the tumor; this process is called clonal evolution (Nowell, 1976; Gao et al., 2007).

Tumorigenesis of cells from certain tissues is more frequent than of others, which is e.g. depending on exposure of this tissue to carcinogens. The genetic abnormalities in malignant cells can be caused by a variety of carcinogens, e.g. chemicals, radiation, infectious agents and tobacco smoke. Those carcinogens mostly cause errors in the DNA repair machinery (Spadari et al., 1987; Maher and McCormick, 1986; Ahmed, 1980). Additionally, mutations and abnormalities of the genetic material can be randomly acquired through errors in DNA replication. The consequence of these inherited mutations is their presence in every cell of an organism. Thus, every cell with such a mutation is prone to malignant transformation (Weinberg, 2007).

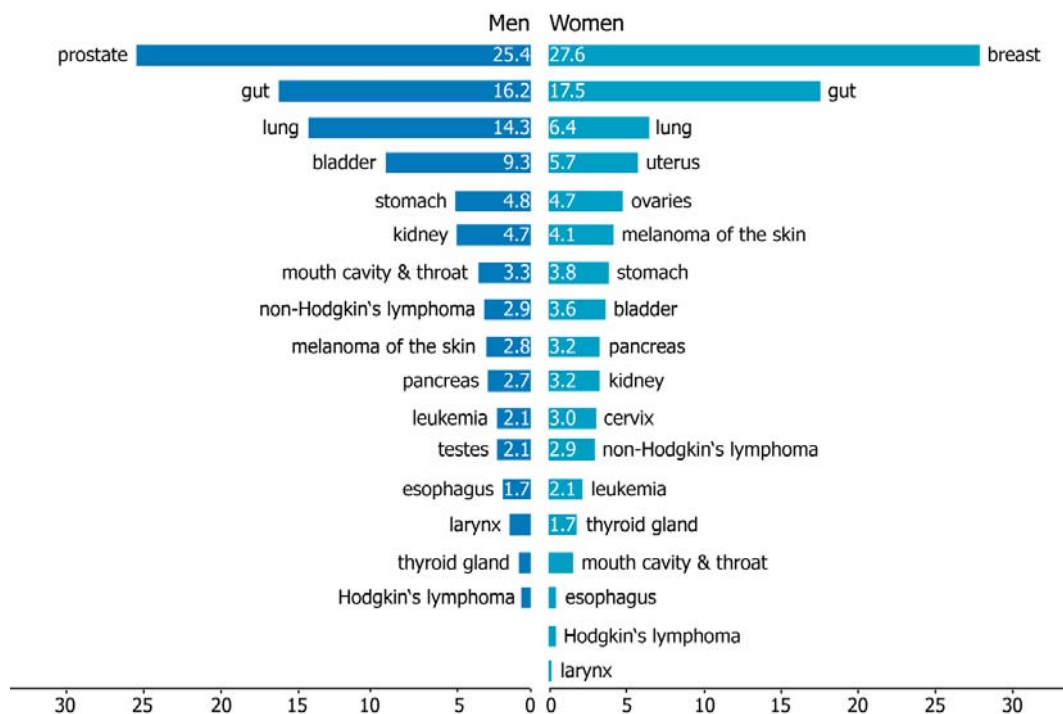


Figure 1-2: Frequency of of new tumor cases according to tumor localization. Shown are percentages of selected tumor loci of all new diagnosed cancer cases, in Germany in 2004 by the *Dachdokumentation Krebs* within the Robert Koch-Institut (modified from Robert-Koch-Institut, 2008).

Nonetheless, almost every tissue of an organism can be affected by cancer. The most frequently affected tissues are given in Figure 1-2. The frequency of these tumor entities is different in women and men. However, frequently affected tissues in both genders are gut, lung, stomach and bladder. In women, the most commonly affected tissue is breast tissue, whereas in men prostate tissue is most commonly affected by tumorigenic transformation (Robert-Koch-Institut, 2008). Often tumors continue the evolutionary process that initially led to its formation. The tumor cells undergo natural selection in response to chemotherapy whereby they can acquire resistance to particular therapeutic agents (O'Dwyer et al., 2002). Finally, cancer is a disease that is common in many multicellular organisms, however, the risk of tumor development increases with age (see chapter 1.1.2 for more details).

1.1.1 Molecular biology and genetics of cancer cells

In the year 2000, Douglas Hanahan and Robert A. Weinberg summarized six universal "hallmarks of cancer", which describe six crucial steps of the transformation process of a normal to a malignant cell. These steps are functions that are gained during the process of malignant transformation of a cell (Hanahan and Weinberg, 2000). The mechanistic pathways vary from cell to cell and not all six hallmarks have to be fulfilled in each malignant cellular transformation. Nonetheless, most tumor cells have acquired this set of tumorigenic functional capabilities (Table 1-1).

Table 1-1: The six capabilities or hallmarks of cancer (Hanahan and Weinberg, 2000)

No.	Hallmark of cancer	Cellular effects
1	Evasion of apoptosis	Allows cell growth despite of genetic mutations and internal or external anti-growth signal(s)
2	Self-sufficiency in growth signals	Unchecked growth
3	Insensitivity to anti-growth signals	Unchecked growth
4	Sustained angiogenesis	Allows tumors to grow beyond limitations of passive nutrition
5	Limitless replicative potential	Immortality of cancer cells
6	Tissue invasion	Metastasis

These hallmarks have been described in carcinomas, but neither in carcinomas nor in other tumor entities all hallmarks need to be achieved. For instance, properties of normal non-malignant blood cells are tissue invasion and displacement to distant tissue sites, so there is no need of acquiring these functions in the development of leukemia. The capabilities described in malignant cells are gained during the transformation process by genomic and/or epigenetic mutations. This does not imply that every hallmark needs to be caused by an individual mutation. So-called mutator changes in genes with diverse control functions can cause gain of several functions at once. These mutator changes can be point mutations, microsatellite instability or loss of heterozygosity, and they affect genes that control DNA replication, check-points of the cell cycle and the efficacy of DNA repair. Reducing the efficacy of DNA repair, leads to an increased mutation rate and genomic instability and thereby accelerates the rate of genomic change in malignant cells. A prominent example is the *TP53* gene, which is lost in a large number of cancers. Loss of P53 protein leads to genomic instability, evasion of apoptosis and increased angiogenesis as it is involved in control of all of these processes (Klein et al., 2007).

The most essential components of tumorigenic transformation are resistance to apoptosis and cell cycle dysregulation (Klein et al., 2007). The prevention of apoptosis in cancer cells is achieved on several levels. The first level can frequently be overexpression of apoptosis inhibitors, like for instance BCL2. The second can be inactivation of executive caspases by mutations. The third level of apoptosis prevention can be the downregulation and mutation of proapoptotic genes, for example of Bax, Apaf1 or the death receptor CD95. The fourth and last level is the alteration of PI3K/AKT signaling pathways (Klein et al., 2007). Tumor cells are more resistant to apoptosis than non-malignant cells. However, a cell that is completely apoptosis-resistant is not known. As many tumors can be treated by X-ray irradiation or by cytotoxic drugs, a residual activity of apoptotic pathways can still be triggered in cancer cells (Klein et al., 2007).

In general, cancer cells acquire their functional capabilities by mutations that typically affect two classes of genes: oncogenes and tumor suppressor genes (TSG). Cancer-promoting oncogenes are activated through gain-of-function mutations and are mostly involved in pathways that confer a variety of properties to the transformed cells including proliferation, apoptosis prevention, loss of tissue boundaries and invasion. Tumor suppressor genes are inactivated in tumors by loss-of-function mutations. They are normally involved in regulation of DNA replication, DNA repair, cell cycle control, cellular orientation within a tissue and adhesion within tissues or promotion of apoptosis.

The first oncogene was discovered in 1970 by G. Steve Martin from an avian Rous sarcoma virus and was named *src* for sarcoma (Martin, 1970). Six years later J. Michael Bishop and Harold E. Varmus could demonstrate that oncogenes are defective proto-oncogenes (Stehelin et al., 1976b), which can originate from retroviruses (Stehelin et al., 1976a). Proto-oncogenes give rise to a number of protein classes including transcription factors, proteins involved in signal transduction or cell cycle regulators. Thus, in normal cells proto-oncogenes have diverse functions and are critical in cell growth, control of differentiation, and in different signal transduction pathways. Proto-oncogenes are quiescent counterparts of oncogenes. They are critical for growth, repair and homeostasis of an organism, whereas oncogenes promote uncontrolled cell growth. Activation of proto-oncogenes leads to conversion into an oncogene with tumor-promoting activity, e.g. due to imbalance of cell cycle regulation. This activation is often based on point mutations, so called "gain-of-function" mutations of proto-oncogenes that directly lead to increased expression of the oncogenic product. Other mechanisms of oncogene-activation are loss of regulation either by alteration of gene expression, activity of the protein product or by chromosomal translocations, which may lead to overexpression of oncogenes in the wrong cell type or expression at the wrong time. A prominent example for a proto-oncogene is *Ras*, which functions as molecular switch between apoptosis and survival of normal cells. *Ras* is a proapoptotic oncogene, like *MYC*, *E2F1* or *E1A*, which all normally upregulate components of the apoptosome that are associated with the mitochondrial pathway. *Ras* is frequently activated through point mutations in a diverse range of tumors. It promotes survival of these tumor cells because oncogenic *Ras* is constitutively active and therefore insensitive to cellular negative control mechanisms (Klein et al., 2007).

Tumor suppressor genes (TSGs) normally protect cells from cancer; only if TSG functions are lost, cells can undergo malignant transformation. Normal functions of TSGs are regulation of cell cycle, induction of apoptosis or cellular adhesion. TSGs repress other genes that are responsible for progression in cell cycle and they link DNA damage to cell cycle and to

apoptosis. Some tumor suppressors are involved in cell adhesion and prevent normal cells from dispersion, block loss of contact inhibition and thus inhibit metastases.

The first TSG, *RB1*, was discovered by its protein product, the tumor suppressor retinoblastoma (pRb) in human retinoblastoma (Friend et al., 1986). The function of pRb is lost in retinoblastomas through germline and somatic mutations of the *RB1* TSG in retinal cells, leading to formation of retinoblastomas. The most famous example of a TSG is *TP53*, which is affected in nearly all cancers. Homozygous loss of *TP53* is found, for example, in 70 % of colon cancers, in 30-50 % of breast cancers and in 50 % of lung cancers. Additionally, mutated *TP53* is involved in the pathophysiology of leukemias, lymphomas, sarcomas and neuronal tumors (Weinberg, 2007). Normal *TP53* has many cellular functions. It acts as a transcription factor, is involved in the cell cycle control upon DNA damage, and contributes to genomic integrity, DNA repair and recombination (Vogelstein et al., 2000; Goi et al., 1997; Guillouf et al., 1995). A second important example for a TSG is *P TEN* (phosphatase and tensin homolog deleted on chromosome 10), a tumor suppressor that is part of the PI3K (phosphatidylinositol 3-kinase)/AKT signaling pathway that regulates signaling of cellular processes such as apoptosis, cell proliferation and cell growth. *P TEN* encodes a dual protein/lipid phosphatase, whose main substrate is phosphatidylinositol-3,4,5-triphosphate (PIP3) which is a product of the PI3K. An increase in PIP3 within the cytoplasm recruits AKT to the cellular membrane where it is further activated by other kinases that are similarly dependent on PIP3. Loss of activity of the *P TEN* TSG by either mutations, deletions or silencing by promoter hypermethylation is frequently observed in many human cancers (Blanco-Aparicio et al., 2007) and leads to deregulated AKT-signaling in the malignant cells. The most common mechanisms of inactivation of TSGs are loss-of-function mutations and inactivation of signaling components that are involved in TSG-activation pathways. Besides this, recessive germline mutations can occur and are inherited by the offspring, which predisposes to cancer without producing a phenotype.

The two-hit-hypothesis by Alfred Knudson states that an inherited germline mutation in a TSG only causes cancer if a second mutation occurs in the second allele of the same TSG (Knudson, 1971). The first hit does not necessarily lead to cancer as the function of TSGs can counteract on the mutated oncogene. Only a second hit, which inactivates the second copy of the TSG, leads to malignant transformation. Alfred Knudson performed statistical analyses on inherited and sporadic retinoblastomas to explain age differences between the two types of retinoblastomas. Inherited retinoblastoma develops early in children, as the first mutation of the *RB1* gene is inherited and therefore a second mutation rapidly leads to cancer. In contrast, sporadic retinoblastomas occur later in life because two mutations have to take place before development of the cancer (Knudson, 1971). The two hit hypothesis by Knudson can be extended to more recent findings of epigenetic aberrations in cancer, as the second hit may be the epigenetic silencing of a TSG or the epigenetic activation of an oncogene.

Nonetheless, the two-hit-hypothesis had also been challenged by the finding that a mutation in one allele of a TSG is sufficient to cause cancer, even if it is a recessive mutation. This phenomenon occurs in haploinsufficient TSGs and tumors caused by haploinsufficiency are characterized by an earlier age of onset (Cabelof et al., 2006). Haploinsufficiency is described as a dominant loss of function mutation of a haploinsufficient gene of which one copy does not produce sufficient protein to generate the wild type condition. Thus, haploinsufficiency leads to an abnormal disease state since a haploinsufficient gene needs expression from both alleles to establish the wild type state. A wide range of cancers are caused by haploinsufficiency, including for example lymphomas, leukemias, prostate, breast,

colorectal, and skin cancers (Santarosa and Ashworth, 2004; Le Toriellec et al., 2008). After discussion of the chromatin remodeling in normal and malignant development, a short overview of described epigenetic events in cancer will be given at the end of chapter 1.1.2.

1.1.2 Chromatin remodeling in normal and malignant development

The cellular plasticity in non-malignant cells and disease is generally achieved by a complex regulatory network that involves modulation of chromatin structure and thereby genome function. The molecular basis of a part of this network is chromatin remodelling, which describes all chemical and biological modifications to DNA and histone proteins. The mechanisms by which epigenetic changes determine how the genetic information is used across the highly variable background of developmental stages, tissue types and disease states is not yet fully understood (Bernstein et al., 2007).

A simplified view of decisions of cellular fate that are made in the nucleus is presented in Figure 1-3. External signals are translocated by signal transduction cascades into the nucleus and lead to changes of genome organization. In turn, gene expression patterns are altered and decisions are provoked that lead to differentiation, proliferation, senescence, apoptosis or specific metabolic states of the cell (Rippe, 2007). Thus, the tight regulation of epigenetic states and chromatin remodeling potential is essentially indispensable for the regulation of any cell.

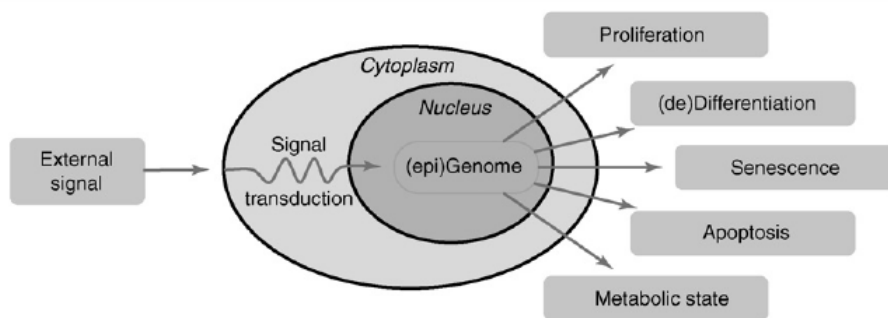


Figure 1-3: Schematic view of cell fate decisions in the nucleus (Rippe, 2007)

Many proteins are involved in the different chromatin remodeling processes, for example DNA-methyltransferases (DNMTs) and DNA-demethylases, histone-methyltransferases (HMTs), histone-acteyltransferases (HATs), histone-deacetylases (HDACs) and members of the chromatin remodeling protein-complexes. Here, a short introduction to general aspects of epigenetics will discuss DNA-methylation, histone modifications and the histone code as well as epigenetic changes during aging. The function of DNA-methylation in non-malignant cells and its role in monoallelic expression will be described below (chapter 1.3.3).

DNA-methylation in vertebrates occurs almost exclusively at cytosines within CpG-dinucleotides, whereby most CpGs in the genome are methylated (Bird, 2002; Goll and Bestor, 2005).

The four core histones, H2A, H2B, H3 and H4, build the nucleosomes and are subject to over 100 posttranslational modifications of residues in their histone tails. These posttranslational modifications include acetylation, methylation, phosphorylation and ubiquitination and they regulate chromatin accessibility and transcriptional activity (Kouzarides, 2007). Distinct histone modifications influence each other and may interact with DNA-methylation (Figure 1-4). The epigenomic state of a particular cell is the result of genetic determinants, lineage

specific cues and the environment (Figure 1-4). Epigenetic modifications are read by protein complexes that can bind to modified histones or to methylated cytosines (Li et al., 2007). The term epigenome sums up all epigenetic modifications of a given cell, but it has been shown that although DNA-methylation patterns are epigenetically inheritable, only a subset of histone modifications are (Bernstein et al., 2007).

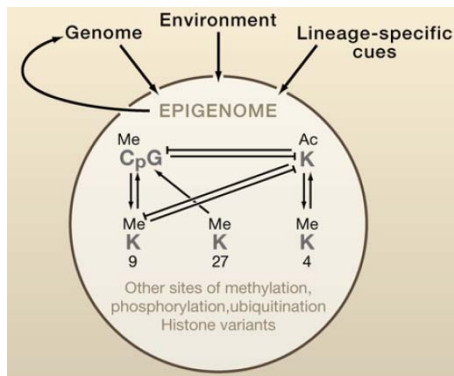


Figure 1-4: Schematic representation of the epigenetic network that modulates chromatin structure and genome function (modified from Bernstein et al., 2007).

DNA-methylation is the covalent modification of cytosine bases by methyl-groups that are transferred from the donor S-adenosylmethionine to the C5-position of the cytosine by DNA-methyltransferase enzymes (DNMTs). In mammals, three different DNMT enzymes with different specificities and functions have been described. DNMT1 has the function of maintaining DNA-methylation and during replication it copies the DNA-methylation patterns from the template strand to the daughter strands. DNMT1 was found to be critical for proper embryonic development, for genomic imprinting and also for inactivation of the female X chromosome (Cervoni et al., 1999; Bhattacharya et al., 1999; Wolffe et al., 1999). The two other mammalian DNMTs, DNMT3a and DNMT3b, are *de novo* methyltransferases that set up the DNA-methylation pattern in early development. Both DNA-methylation and DNA demethylation are essential for the regulation of gene expression. Several enzymatic activities have been described that demethylate DNA *in vivo* and *in vitro* (Cervoni et al., 1999; Bhattacharya et al., 1999; Szyf and Bhattacharya, 2002; Hattori et al., 2001). However, it is still highly debated whether these functions are present *in vivo* or if DNA-demethylation occurs during replication as a passive process after loss of DNMT1 activity (Suzuki and Bird, 2008). The regulatory role of DNA-methylation is complex, includes a variety of cellular functions and its deregulation results in a number of pathologies. It regulates tissue-specific gene expression, differentiation, genomic imprinting, regulation of chromatin structure, aging and carcinogenesis (Bird, 2002). DNA-methylation is crucial for normal development and for the survival of differentiated cells (Li et al., 1992; Okano et al., 1999; Jackson-Grusby et al., 2001). CpG-dinucleotides of the mammalian genome are globally methylated, while specialized structures, called CpG-islands, are unmethylated. Especially repetitive sequences and transposable elements, but also gene bodies and intergenic regions are highly methylated in order to maintain chromosomal stability (Kouzarides, 2007). DNA-methylation patterns are highly dynamic throughout development and disease. It has been shown that in normal early development, paternal and maternal genomes are demethylated. While the paternal genome is actively demethylated, the maternal genome is demethylated later passively (Reik et al., 2001; Santos et al., 2002). In the blastocyst state, genomewide levels of DNA-methylation increase rapidly to produce the

majority of the methylation patterns that are later found in adult tissues. In addition to these coordinated changes during normal development, the totality of DNA-methylation (the so-called DNA methylome (Esteller, 2007)) undergoes characteristic changes in malignant transformation (see chapter 1.1.3). DNA-methylation in normal and diseased cells affects recruitment of regulatory proteins. Two mechanisms have been described so far. The first one is the exclusion of DNA-binding proteins from their target sites by methylated DNA. One well investigated example is the imprinted H19 locus. A differentially methylated region (DMR) at the locus leads to inhibition of CTCF-binding and activates gene expression at this allele (Hark et al., 2000). In contrast, the methylated CpG-dinucleotides of the inactive allele recruit HDACs to methylated genomic loci that further deacetylate the histone tails leading to transcriptional repression (Bird, 2002). This leads to further silencing of inactive genomic loci by DNA-methylation.

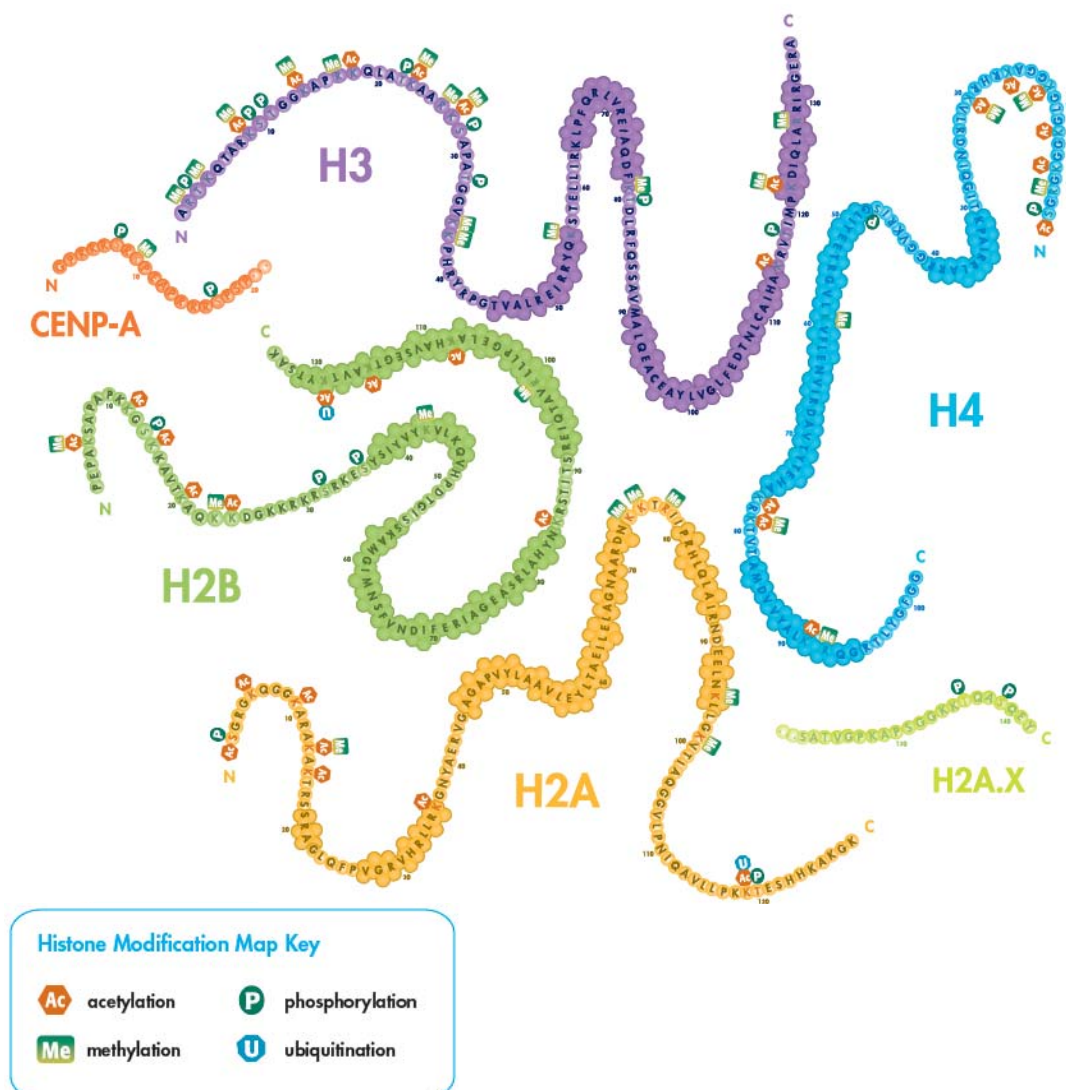


Figure 1-5: Schematic representation of the sequence of the four human core histone tails and of the variants CENP-A and H2A.X and all of their published post-translational modifications. (<http://www.millipore.com/techpublications/tech1/pb1014en00-74xp9g>).

Modifications of histone tails are involved in regulation of the chromatin structure and the regulation of accessibility and expression of genomic loci. The DNA-helix is wrapped around

nucleosomes that are formed by histone proteins and regulated by modifications of the DNA and the histone tails. The nucleosomes of the chromatin fiber are functional and structural subunits and form a nucleosomal chain. Each nucleosome consists of four core histones, H2A, H2B, H3 and H4, that form an octamer, the nucleosomal core particle, which is generated by one H3-H4-tetramer and two H2A-H2B-dimers. 146 base pairs of the DNA-helix are wrapped slightly less than two turns around this core particle. The nucleosomes are connected by a linker DNA sequence that is variable in length (15 to 55 base pairs), and the linker histone H1 that is bound to the DNA-helix can attach one nucleosome to the other (Lodish, 2000).

The histones themselves are small basic proteins each consisting of a globular domain, which is part of the nucleosomal core particle and of a more flexible and charged N-terminus of 20 to 140 amino acid residues, the histone tail. The globular domains of the four core histones constitute the nucleosomal core whereas the histone tails protrude from the nucleosomes and are subject to over 100 known chemical posttranslational modifications including acetylation, methylation, phosphorylation and ubiquitination (see Figure 1-5). Those histone tail modifications have regulatory function in modulating chromatin accessibility and transcriptional activity (Kouzarides, 2007). It was already recognized in the beginning of epigenetic studies that particularly the histone acetylation and methylation has direct effects on a great variety of nuclear processes including gene transcription, DNA repair, DNA replication and the organization of chromosomes (Esteller, 2008; Lodish, 2000).

The combinations of several histone modifications are recognized by distinct proteins and thus, the modifications can influence gene expression in normal and malignant cells. At first, histone modifications were only recognized to have an effect on the structure and condensation of chromatin. The methylation of lysine residues in histone tails was considered to only result in compaction of chromatin, while acetylation leads to opening of chromatin. This view was supported by findings that the heterochromatin protein 1 (HP1) selectively binds to methylated lysine 9 residues in the tail of histone H3, which promotes heterochromatin formation. Euchromatinization, on the other hand, could be linked to the acetylation of histones (Lodish, 2000). Since then, many studies have shown that this is only a simplified view of how chromatin structure regulates cellular functions.

Apart from the canonical histones (H1, H2A, H2B, H3 and H4), a variety of histone variants has been described; a short selection is listed in Table 1-2. The most extensively studied histone variant is histone H3.3, which is a variant of the canonical histone H3 (H3.1). The exchange of H3.1 with the variant H3.3 was the clue to the puzzle of how histone tails are demethylated and deacetylated outside of replication. All modifications on the histone tail of H3 are erased at once by replacing H3.1 with its variant H3.3 thereby making the chromatin at this site more or less accessible for transcription depending on the preceding modifications of the canonical H3 (Ahmad and Henikoff, 2002; Wirbelauer et al., 2005).

Another interesting histone variant is macroH2A, which contains an unusual additional globular domain in place of the histone tail (Pehrson and Fried, 1992). If macroH2A instead of H2A is incorporated into the nucleosomal core the chromatin is transcriptionally inactivated. MacroH2A was shown to have repressive effects on gene expression and has been shown to be enriched at the human female inactivated X chromosome (Changolkar and Pehrson, 2006; Costanzi and Pehrson, 1998). Additionally, it was shown that macroH2A is allele-specifically enriched at imprinting control regions within imprinted loci (Choo et al., 2006). This monoallelic enrichment of macroH2A was suggested to be present at all monoallelically enriched loci, which is why the enrichment of macroH2A at the monoallelically expressed locus within 13q14 was analyzed in this work.

1 Introduction

Table 1-2: Description of selected nucleosomal histones, their variants and potential unique functions of the histone variants (modified from Culhane and Cole, 2007)

Histone protein/ variant	Property/ Proposed function (prop. fct.)	% of identity to major isotype
H2A, H2B, H3 and H4	Canonical core histones, encoded by replication-coupled genes	-
MacroH2A	Vertebrate-specific H2A variant with a C-terminal globular domain; enriched on mammalian inactive X-chromosome; proposed function: X-chromosome inactivation; gene expression	64 (in histone region)
H2A.X	H2A variant with a four amino acid SQ[E or D]-C-terminal motif that becomes serine-phosphorylated at sites of double-strand breaks	-
H3.3 (H3.2 in plants)	H3 variant that replaces H3 and differs at position 31 and at a few residues on helix 2 that enable deposition outside of replication	-
CENP-A	Essential for viability in mouse; proposed function: kinetochore assembly	62

Brian D. Strahl and C. David Allis proposed the histone code hypothesis in the year 2000, which was later refined by Allis together with Thomas Jenuwein (Strahl and Allis, 2000; Jenuwein and Allis, 2001). According to this hypothesis, the histone code consists of distinct histone modifications on one or several histone tails that together form a code, which is established and read by specialized proteins.

Table 1-3: Differential distribution of histone modification marks in interphase nuclei of *Drosophila*, mammals and *Arabidopsis* (modified from Henikoff et al., 2004).

Position	Modification	<i>Drosophila</i>	Mammals	<i>Arabidopsis</i>
H3K4		EU (IB)	EU	EU
		EU (B)	EU	EU
H3K9		HET, [EU (B)]	EU	HET
		HET, [EU (B)]	EU, fac. HET	HET
		HET, [EU (B)]	HET	EU
H3S10		EU (IB)	EU	HET
H3K14		EU (B)	EU	EU
		HET, EU (B)	HET, EU	HET
H3K27		HET, EU (B)	EU	HET
		HET, EU (B)	EU, fac. HET	EU
H3K36		EU (IB)	unknown	EU
		HET, EU (B)	EU, fac. HET	HET
H4K20		HET, EU (B)	EU	EU
		HET, EU (B)	HET	EU

EU – euchromatin, HET – heterochromatin, B – bands, IB – interbands

The read-out of the histone code by specialized proteins at specific chromatin sites leads to distinct downstream events. Jenuwein and Allis (2001) postulated three features of the histone code: Firstly, distinct histone tail modifications induce the interaction with chromatin associated proteins; secondly, modifications on the same or on different histone tails can be interdependent and thus generate various combinations or codes on the nucleosomes, and

thirdly the concentration and combination of differently modified nucleosomes is responsible for distinct qualities of higher order chromatin, which simply means that the combinations of histone modifications generate euchromatic and heterochromatic sites (Jenuwein and Allis, 2001). In addition, single histone marks can be linked to heterochromatic and euchromatic sites, like for example mono-, di- or tri-methylation of lysine 4 of histone H3 (H3K4me1-3), which is a mark of euchromatin regardless of methylation states in all species that are depicted in Table 1-3. Other single histone marks for euchromatin in mammals include H3K9 acetylation, H3S10 phosphorylation and mono- and dimethylation of H4K20. Mammalian marks of heterochromatin are deacetylated histones, and methylation of H3K9 and H3K27. Table 1-3 shows that the histone marks themselves seem to be highly conserved among species (in plants, insects and mammals), while the suggested functions, i.e. the coding, can vary from species to species (Henikoff et al., 2004).

Not only histone modifications are conserved among species, but also the protein domains that establish and read those marks. The protein domains that interact with histone modifications are mainly bromo-, chromo- and SET-domains (Jenuwein and Allis, 2001). These are found in two classes of antagonizing chromatin regulators, the Polycomb (PcG) and trithorax (trxG) groups. Generally, PcG and trxG genes are important for the maintenance of expression boundaries of homeotic selector genes (Hox genes) and several other key developmental regulator genes. The regulation by PcG and trxG mostly occurs through the modulation of chromatin structure of their target loci (Orlando et al., 1998; van Lohuizen, 1999). The mammalian trithorax proteins contain bromodomains that bind for example to acetylated histones (Strahl and Allis, 2000). In contrast to that, the chromodomains of Polycomb and HP1 proteins were found to preferentially bind to methylated histone lysine residues. Finally, the SET-domains are shared within both classes of chromatin regulators, PcG and trxG, and these domains have a methylating activity (Jenuwein et al., 1998; Xiao et al., 2003). No methylating activity can be assigned to the bromo- and chromodomains, which are solely reading and translating the histone signals through their binding to specific modified epitopes. Those protein domains are therefore the basis for establishing and reading the histone code (Strahl and Allis, 2000). Another group of enzymes that regulate the site-specific methylation are homologs of the *Drosophila* Su(var)3-9. The Su(var)3-9 protein and its homologs, which are Clr4 in yeast and SUV39H1 in human, were shown to encode a histone methyltransferase (HMT) with specificity for the methylation of histone H3 at the lysine 9 residue (H3K9) (Rea et al., 2000). The methylation activity of these enzymes comes from the highly conserved SET-domains of these HMTs. The methylation of the H3K9 epitope induces further binding of HP1 by its chromodomains, and thus a heterochromatic mark is stably established (Bannister et al., 2001; Lachner et al., 2001). The sum of these findings generates a strong link between Su(var)-function, gene-silencing and the assembly of heterochromatin.

The first protein module for which direct and selective interaction with acetylated lysines of the histone tails could be shown was the bromodomain. Sequentially, a class of proteins was found to recognize and bind to these bromodomain-containing proteins, which fulfills the second prediction of the histone code hypothesis of a combinatorial readout (Strahl and Allis, 2000).

The histone code hypothesis has been challenged since it was published in 2000. However, it has been confirmed in a number of studies that distinct combinations of several histone modifications mark chromatin for transcriptional activity or transcriptional silence. In addition, histone modifications distinctively mark functional and structural chromatin domains

like transcriptional start sites or enhancer regions (Barski et al., 2007). Moreover, features like the binding of chromatin remodeler proteins and DNA-methylation contribute as epigenetic marks to chromatin. It has been shown that the histone code is not comparable to the genetic code in its universality and stringency. Thus, the original model of a histone code has been displaced by the view of a more general epigenetic code that includes all chromatin modifications that lead to distinct functions and transcriptional regulation of the genome. The term "epigenome" has been used to describe all epigenetic modifications and regulation patterns within a given cell, and this term is used analog to the term genome (Fraga and Esteller, 2007).

The epigenome is described to include all inheritable changes and marks of distinct gene functions that occur without changes to the DNA nucleotide sequence itself. The epigenetic modifications that constitute the epigenome were found to be variable, both intraindividually and interindividually, and tissue-specific (Fraga and Esteller, 2007). It was shown in monozygotic twin studies (Fraga et al., 2005a) that epigenetic variations accumulate during aging independently of changes in the genetic sequence. This so-called epigenetic drift between twins during aging is associated with phenotypic discordance, probably because of a different environmental input. According to this study, it is strongly suggested that epigenetic modulation is the response to environmental factors. However, there are also stochastic events of epigenetic changes occurring during ontogenic development that cannot be related to environmental effects alone. These stochastic epigenetic recombination events are the basis for the concept of epigenotypes, which explain the differences of epigenetic patterns in young individuals (Fraga et al., 2007).

Aberrant epigenetic silencing has been established to be a hallmark in tumorigenesis (see chapter 1.1.3). In addition, epigenetic silencing is also important in cellular aging and senescence (Fraga et al., 2005a; Issa, 2003). Altogether, four epigenetic events could be associated with the aging process. The first event includes environmental and stochastic changes of the epigenetic landscape during the lifetime of an organism and these changes have been described for example in the twin study by Fraga et al. (2005a). Secondly, the aging process itself leaves a mark in the cellular DNA-methylation pattern. The third epigenetic event associated with aging includes genetic and epigenetic disturbances of progeroid genes, which converge in aged and cancer cells. Lastly, the histone marks H4K16ac and H4K20me3 are associated with the aging process (Fraga and Esteller, 2007). Manel Esteller postulated in 2007 that in the future a DNA methylome and histone map of aging will be available, analog to that established for cancer cells (Esteller, 2007), and this map will make it possible to define a young versus an old (aged) cell on the basis of their distinct epigenetic pattern (Fraga and Esteller, 2007).

The DNA-methylation pattern is propagated after each cell division by DNMTs, such as DNMT1 in human cells. However, aging cells undergo an epigenetic drift, which includes for example, a global decrease of DNA-methylation in many tissue types. This global DNA-hypomethylation is probably the result of passive demethylation, which in turn is the effect of progressive loss of DNMT1-efficacy and of false targeting of DNMT1 by other cofactors. Another mechanism that may lead to aberrant DNA-methylation patterns in aged cells is the overexpression of the DNMTs with a *de novo* DNA methylating activity, i.e. DNMT3b, as a response to the DNA-hypomethylation in repetitive sequences (Casillas et al., 2003). The resulting outcome of DNMT3b overexpression is aberrant hypermethylation of promoter CpG-islands of tumor suppressor genes (Issa, 2003).

Interestingly, epigenetic events associated with the aging process are reminiscent of the epigenetic events that lead to tumorigenesis. The links between aging and cancer are global

DNA hypomethylation, aberrant promoter hypermethylation and modest DNMT overexpression (Esteller, 2003). This observed accumulation of epigenetic aberrations during aging might contribute to tumorigenic transformation. Many examples of genes are known that show increased promoter methylation during aging in non-tumorigenic tissues (Issa, 2003; Oakes et al., 2003). Thus there is a potential association of the accumulation of promoter hypermethylation of TSGs during aging with the predisposition to develop cancer. However, a direct relationship between the methylation status of these genes and the aging process itself has not been proven, except for the notion that most cancer types only arise late in adult life.

1.1.3 Epigenetic events in cancer

Epigenetics is defined as regulation of gene expression by modifications of DNA and chromatin structure that is not reflected by differences in the DNA nucleotide sequence itself. Both DNA hypo- and hypermethylation has been described to be part of malignant transformation of cancer cells (reviewed in Esteller, 2006). DNA hypomethylation frequently occurs as global loss of DNA-methylation in tumors and is most often found at repetitive DNA sequences. Aberrant DNA-hypomethylation could result in chromosomal instability, translocations and gene disruptions caused by reactivation of transposable elements (Esteller, 2005; Feinberg and Tycko, 2004). A second major epigenetic event in cancer is hypermethylation of CpG-islands of tumor suppressor gene promoters that leads to gene silencing (Esteller, 2006; Jones and Baylin, 2002). The first described TSG found to be inactivated by hypermethylation of its promoter-CpG-island was *RB1* (Greger et al., 1989; Sakai et al., 1991). Since then, many other TSGs have been reported to be inactivated by promoter-CpG-island-hypermethylation, including *VHL* (Herman et al., 1994), p16^{INK4a} (Merlo et al., 1995; Herman et al., 1995; Gonzalez-Zulueta et al., 1995), *hMLH1* and *BRCA1* (Herman and Baylin, 2003). More recently, the role of histone modifications in epigenetic control of gene expression and chromatin structure has been recognized. Histone tail modifications have been found to be closely involved in the epigenetic regulation, and at some loci have been shown to recruit the DNA-methylation machinery (Fraga and Esteller, 2005). Only little is known about the role of histone modifications in cancer cells. However, it has been shown that histone modifications together with CpG-island-hypermethylation can silence TSGs (Ballestar et al., 2003; Fahrner et al., 2002). Conversely, it has also been described that in some cases, for example in the TSG p21^{WAF1}, histone modifications can promote silencing independently of CpG-island-hypermethylation (Fahrner et al., 2002). Frequently, the hypermethylation of promoter-CpG-islands of TSGs is linked to a distinct combination of histone marks. Several histone modifications have been defined as markers for cancer. These marks are global deacetylation of histones H3 and H4, loss of trimethylation at lysine 4 of histone H3 (H3K4me3) and gain of methylation and trimethylation at lysine 9 (H3K9) and lysine 27 (H3K27) of histone H3, respectively (Esteller, 2007). Fraga et al. (2005b) showed that many human tumors are subject to overall loss of monoacetylation of lysine 16 and of trimethylation of lysine 20 in the tail of histone H4. These two distinct histone marks are considered as universal epigenetic markers for malignant transformation (Fraga et al., 2005b). Finally, loss of H3K4 dimethylation (H3K4me2) and of H3K18 acetylation (H3K18ac) has been proposed as a marker of high risk recurrence in prostate cancer (Seligson et al., 2005). Thus the global pattern of distinct histone modifications may even have a prognostic value.

In summary, epigenetic events that have been related to malignant transformation in cancer are a global DNA-hypomethylation, CpG-island-hypermethylation and overall loss of H4K16ac

and of H4K20me3. Finally, two recent examples of genetic lesions that have also an impact on the epigenetic landscape of cancer cells should be mentioned. The first is the expression pattern of histone and DNA modifying enzymes, such as DNA-methyltransferases (DNMTs), histone-methyltransferases (HMTs) and histone-acetyltransferases (HATs), which can vary between cancer and normal tissues and also between different tumor types (Ozdogan et al., 2006). In leukemia and sarcoma entities characteristic chromosomal translocations that create aberrant fusion proteins frequently involve the genes encoding histone modifying enzymes like HATs and HMTs (Esteller, 2006). The second example for genetic lesions with an epigenetic impact is the deregulation of microRNAs (miRNA) that has been linked to cancer development. Downregulation of miRNAs that degrade or inhibit translation of mRNAs is frequently found in cancer and implies a tumor suppressive function for these miRNAs (Chen, 2005; Calin and Croce, 2006). Furthermore, hypermethylation of the miRNA 5'-regulatory regions can account for the downregulation of miRNAs just as for coding genes (Saito et al., 2006; Lujambio et al., 2007). Two examples of miRNAs that have been found to be downregulated in distinct tumors are the let-7 miRNA, targeting the *RAS* oncogene in lung tumors, and miR15/miR16, targeting *BCL2* in leukemia (Johnson et al., 2005; Cimmino et al., 2005).

1.2 Leukemias and lymphomas

Generally, tumors are classified according to the tissue they originate from. The majority of human cancers arise from epithelial tissues; these carcinomas are responsible for more than 80 % of cancer-related deaths in the Western world. Three groups of tumors with nonepithelial origin can be classified in addition to carcinomas. The first group consists of tumors that are derived from various connective tissues like bones or muscles; these entities are called sarcomas. Sarcomas only account for 1 % of all human tumors. The second group comprises all tumors that arise from various cell types of the central and peripheral nervous system. Tumors of this group are classified as neuroectodermal tumors according to their origin from the embryonic ectodermal layer, including gliomas, glioblastomas, medulloblastomas or neuroblastomas. Interestingly, only 1.3 % of all diagnosed cancers are neuroectodermal tumors, but they are responsible for 2.5 % of cancer-related deaths (Weinberg, 2007). The third group comprises tumors that can arise from every hematopoietic tissue including cells of the immune system. Leukemias and lymphomas are the two main subtypes of hematological neoplasias. Literally, leukemia means "white blood", which refers to nonpigmented malignant progeny of several hematopoietic cell lineages. Leukemias form dispersed, single-cell populated tumor cells in the blood where the tumor populations can move freely. Lymphomas are tumors arising in the lymphoid organs, like lymph nodes, spleen and thymus. This category includes all tumors originating from the lymphoid cell lineage, for example from B- and T-lymphocytes. Aberrant lymphoid cells aggregate to form solid tumor masses, most frequently found in lymph nodes (Weinberg, 2007). The most common hematopoietic malignancies are Acute Lymphoblastic Leukemias (ALL), Acute Myelogenous Leukemias (AML), Chronic Myelogenous Leukemias (CML), Chronic Lymphocytic Leukemias (CLL), Multiple Myeloma (MM), Non-Hodgkin's Lymphoma (NHL) and Hodgkin's Disease (Weinberg, 2007).

Leukemias can be roughly subdivided into acute and chronic leukemias according to the course of the disease. The characteristic of acute leukemias is a rapid increase of immature blood cells that crowd the bone marrow, which is in turn unable to produce healthy blood

cells. Acute leukemias occur mainly in children and young adults and are the most common cause of cancer-related deaths in children in the U.S. (Jemal et al., 2005). Acute forms of leukemia usually need immediate treatment as they are rapidly progressing with malignant cells accumulating in the blood and spreading to other organs. In comparison, chronic leukemias are progressing very slowly and abnormal cells of chronic leukemias can mature regardless of their lack of functionality. Moreover, the leukemic cells are usually long-lived and essentially outlive other healthy blood cells. In chronic leukemias, aberrant cells are excessively accumulating, but it can take months or even years for the progression of the disease. The malignant cells are produced at a much higher rate than non-malignant cells resulting in a slow but steady accumulation of aberrant leukemic cells in the blood and sometimes secondary lymphoid organs. Chronic leukemias occur mostly but not exclusively in older people and are often only diagnosed in a routine blood test, as no symptoms manifest until the disease reaches an advanced stage. In contrast to the requirement of immediate treatment of acute leukemias, chronic lymphocytic leukemia is monitored before treatment, which maximizes overall survival.

Leukemias can be subdivided further into lymphoblastic (or lymphocytic) and myeloid (or myelogenous) leukemias depending on the type of the affected blood cells. In lymphocytic leukemias, lymphocytes or lymphocytic progenitor cells are transformed. In myelogenous leukemia, a progenitor cell that would normally differentiate into red or white blood cells or into platelets has transformed to malignancy.

Two types of acute leukemias have been classified according to the affected blood cell type: AML and ALL. Acute Myelogenous Leukemia (AML) involves myeloid lineage cells including granulocytes, monocytes and platelets, while in Acute Lymphoblastic Leukemia (ALL) cells of the lymphoid lineage are affected, including B- and T-lymphocytes (Catovsky, 1982). Analog to acute leukemias, the two categories of chronic leukemias are Chronic Lymphocytic Leukemias (CLL) and Chronic Myelogenous Leukemia (CML). ALL is the most common leukemia in children, but it affects also adults at an age of 65 years and older. The survival rates of ALL patients vary between groups of different age, but approximately 85 % of children and 50 % of adults survive the disease (Zippin et al., 1971). CLL is the most common leukemia in adults, and patients have a median age at onset of 72 (Seiler et al., 2006). Men are slightly more frequently affected with an incidence ratio of 2:1, in men:women. The five-year survival rate is 75 %, depending on the genomic alterations that are present (see chapter 1.3). CLL patients are frequently without symptoms for a prolonged period of time. In addition, the majority of CLL patients respond well to treatment, with current first-line regimens resulting in overall response rates of 95 %. However, all patients finally relapse and the disease remains incurable outside of transplantation procedures (Boyiadzis et al., 2007). The myeloid leukemias AML and CML occur also predominantly in adults. While AML patients need treatment with several chemotherapeutica (Burnett, 2005), first-line treatment of CML is imatinib (Gleevec) (Avery et al., 2004; Hirose et al., 2002). In AML, the five-year survival rate is only 40 %, while it is 90 % in CML. Hairy cell leukemia (HCL) is not categorized to one of these groups. HCL is sometimes considered a subset of CLL, but shows somewhat different features (Gine et al., 2002). About 80 % of affected patients are adult men; the disease has not been reported in young children. HCL is incurable, but easily treatable with a ten-year-survival rate of 96-100 % (Jansen et al., 1984).

1.2.1 Hematopoiesis

Hematopoiesis is the process by which all blood lineages arise from pluripotent hematopoietic stem cells (HSC; Figure 1-6). Blood lineages include T- and B-cells that constitute the lymphoid lineage, and neutrophils, eosinophils, basophils, monocytes, macrophages, megakaryocytes, platelets and erythrocytes that are part of the myeloid lineage (Zhu and Emerson, 2002). All these components of the blood are derived from HSCs through asymmetric division that produces a HSC and a hematopoietic progenitor cell (HPC). HPCs are multipotent cells that partially retain the self-renewal capacity of HSCs and give rise to lineage committed progenitor cells that can differentiate into all cells of the hematopoietic system (in Figure 1-6). First, HPCs generate either lymphoid or myeloid progenitor cells (CMP or CLP, respectively) that further differentiate into cell types of the hematopoietic system (depicted in Figure 1-6). In the myeloid lineage, CMPs differentiate into a large variety of human blood cells like erythrocytes, platelets, neutrophils, eosinophils, basophils and macrophages (Figure 1-6). In contrast, lymphoid progenitor cells differentiate into B- and T-lymphocytes or natural killer (NK) cells (Figure 1-6). A more detailed view of the development of B-cells will be given in the following chapter.

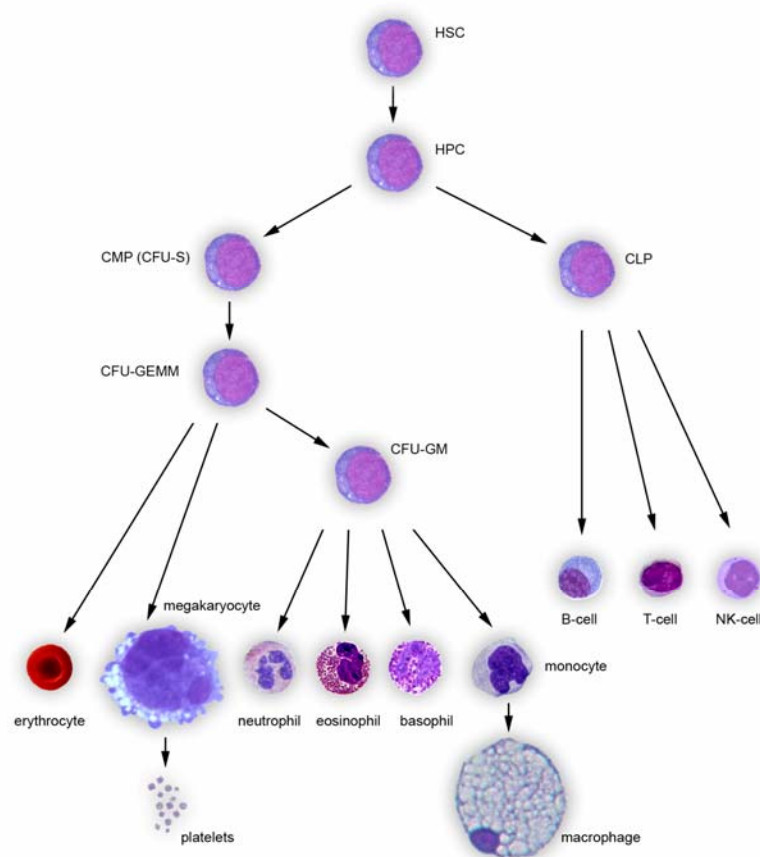


Figure 1-6: Ontogeny of human blood cells.
(<http://daley.med.harvard.edu/assets/Willy/haematopoiesis.jpg>)

1.2.2 B-cell development

B-cell development starts in the bone marrow, where immature B-cells are produced, and proceeds through different stages of rearrangement of the genomic immunoglobulin loci (Edry and Melamed, 2004). Immunoglobulins (Ig) are composed of light (L) and heavy (H) chains. These genes are found in distinct genomic clusters. Both, H and L chains consist of a variable (V) and a constant (C) region. In a process called V(D)J recombination the variable regions of H and L chains are assembled to produce a unique complete V-domain in the immunoglobulin of a B-cell. The genetic locus of the variable region is organized into several gene segments, two segments in the V_L -region (V and J) and three gene segments (V, D and J) in the V_H -region (Matsuda and Honjo, 1996). Recombinations between the gene segments of the V_L and V_H locus take place in a stepwise fashion during the development of B-lymphocytes. After V(D)J-recombination, precursor B-cells have functional immunoglobulin genes, which constitute the B-cell receptor (Bassing et al., 2002). Only those B-cells with a functional BCR on their surface survive and can develop into mature naïve B-cells, leave the bone marrow and participate in the immune response.

In the germinal center of the lymph node, B-cells mature and are activated by antigen presenting cells. Mature naïve B-cells are activated by binding of an antigen to the BCR, which upon binding triggers specific cytoplasmic signaling cascades depending on the maturation stage of the B-cell (Gauld et al., 2002). The activation of B-cells often depends on antigen-presenting T-cells and takes place in the germinal centers of lymph nodes (Figure 1-7). A T-cell-independent maturation of B-cells is also possible, and this takes place in marginal zones around lymphoid follicles (depicted in Figure 1-9). T-cell-activated B-lymphocytes undergo clonal expansion and genetic changes (see below) as part of their affinity maturation process within the germinal centers (Küppers, 2005).

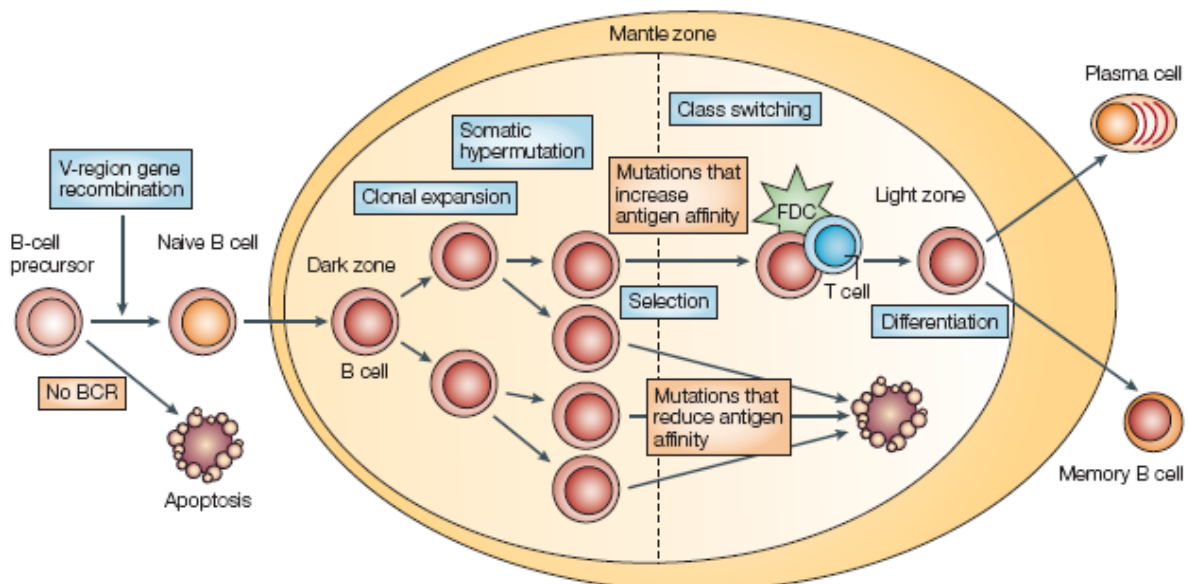


Figure 1-7: B-cell differentiation in the germinal-centre reaction (Küppers, 2005).

In the dark zone of the germinal centers, B-cells undergo clonal expansion, and in a process called somatic hypermutation the immunoglobulin genes of B-cell clones are mutated in order to increase antigen affinity. However, only if affinity to the antigen is increased, mature B-cells carrying these optimized BCRs are positively selected within the germinal centers and transit from the dark to the light zone of this tissue. Mature B-cells with a

reduced antigen affinity of the BCR after somatic hypermutation do not receive sufficient survival signals and go into apoptosis (Figure 1-7).

Somatic hypermutations in the variable regions of immunoglobulin genes mark the first step of the affinity maturation process of B-cells in the germinal centers of the lymph nodes. This process introduces point mutations into the variable regions of recombined V_H and V_L genes. The mechanism of somatic hypermutation is not fully understood (Janeway, 2002), but some aspects have been elucidated. Mutations in the variable regions are induced by double-strand breaks of the DNA, which are repaired inaccurately and thus produce sequence variations (Neuberger et al., 1998). Expression of hypermutated IgV_H genes stochastically alters the binding specificity and affinity of resulting antibodies, i.e. of the BCR, produced by the B-cell clones (Janeway, 2002).

In the light zone of germinal centers, B-cells undergo the next maturation step, clonal selection, which is taken out by follicular dendritic cells (FDC) and T-cells in the light zone of germinal centers. These cells present antigens to the B-cells after somatic hypermutation and only those B-cells with the highest affinities for the antigens are selected and in turn survive and can differentiate further (Küppers, 2005). The following last step of affinity maturation is class switch recombination, through which the constant regions of immunoglobulin heavy chains can be switched to other classes of constant immunoglobulin C-regions. The C-regions determine the Ig classes of the antibody and thus the different effector functions of the B-cells. In humans, five immunoglobulin classes exist: IgM, IgD, IgG, IgA, and IgE. After class switch recombination, the resulting B-cells have different cellular functions and can develop into mature B-cells, into memory B-cells or plasma cells (Figure 1-7).

Allelic exclusion in the immune system

Discussed above are several mechanisms that exist to generate the required repertoire of diverse and adaptable B-lymphocytes that are part of the immune system. Another interesting mechanism essential in B-cell development is called allelic exclusion. This is a form of monoallelic expression of immunoglobulin genes in lymphocytes (Goldmit and Bergman, 2004). The diversity of the antigen-receptor repertoire is created from a relatively small number of V, D and J segments by their shuffling in DNA rearrangement in somatic cells (=V(D)J recombination, see above). Allelic exclusion is required to ensure exclusive production of a single antigen receptor per B- or T-cell. If both copies of the immunoglobulin loci would be expressed, BCR (and TCR) rearrangement would lead to two different BCR (or TCR) molecules with different affinities on the same lymphocyte and thereby make positive and negative selection of the cell impossible.

Allelic exclusion is achieved by differential chromatin marks on the two alleles. These marks are established in the early embryo and lead to asynchronous replication of the maternal and the paternal allele of immune receptor loci (Figure 1-8). The choice of alleles for early or late replication is a stochastic event. However, the resulting replication timing is an instructive mark for the two alleles and directs further allele-specific epigenetic events in the immune system. Both alleles are packaged into a closed chromatin structure and are DNA methylated until the pre-B-cell stage. At this stage, early replicating alleles undergo programmed histone modification, specific nuclear localization and DNA demethylation (Cedar and Bergman, 2008), which make these alleles accessible to rearrangement (Figure 1-8). The result of this epigenetically regulated selective monoallelic access to rearrangement is that only one allele of the antigen receptor is expressed, whereas the second one is silenced. The first step of allelic exclusion is the rearrangement of the IgV_H genes on one allele, which results in shut-

down of the rearrangement process on the second allele. If no successful rearrangement occurs on the first allele, it is silenced and rearrangement takes place on the second chromosome copy. If this rearrangement is also not successful the cell will go into apoptosis. This process of allelic exclusion thus results in Ig antibodies on the surface of individual B-lymphocytes that are all the same (Goldmit and Bergman, 2004). Recent studies show that both the initial selection of one allele for rearrangement and the maintenance of silencing of the other allele are controlled by epigenetic mechanisms, but these are not yet fully understood (Bergman et al., 2003).

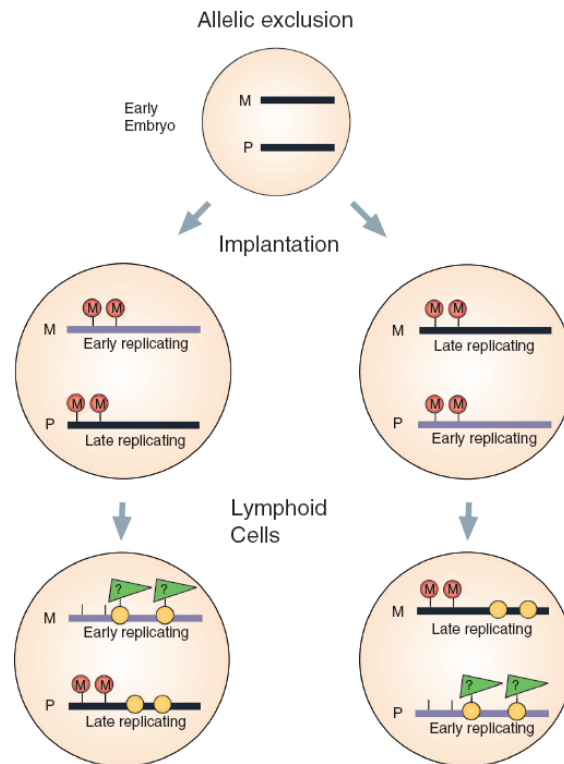


Figure 1-8: Allelic exclusion in the immune system (Goldmit and Bergman, 2004).

1.2.3 Aberrant B-cell development

As detailed above, distinct stages of B-cell development are characterized by distinct rearrangement of Ig-genes and by the structures of antigen receptors on the surface of B-cells. These features can be used to determine the cellular origin of several human B-cell lymphomas (Küppers et al., 1999). As depicted in Figure 1-9, B-cell lymphomas can arise from B-cells at virtually every step of the differentiation process.

The most immature cells, the naïve B-cells, can give rise to Splenic Marginal Zone Lymphomas, while antigen-experienced but not somatically hypermutated B-cells could be the origin of an aggressive form of B-cell Chronic Lymphocytic Leukemia (CLL). Mature and somatically hypermutated B-cells (plasma cells), can give rise to Multiple Myeloma, Diffuse Large B-cell Lymphoma (DLBCL), Hairy-cell Leukemia, Prolymphocytic Leukemia and a more indolent form of CLL (memory B-cells; see Figure 1-9). All B-cell lymphomas can thus be correlated according to their phenotype to a distinct differentiation state of normal hematopoiesis (Küppers et al., 1999). Several tumors are thought to originate from different B-cell origins, among them DLBCL and CLL (Figure 1-9). B-cell origins determine the

subtypes of DLBCL and CLL diseases that have different outcome and prognosis depending on the maturity of those cells the tumor originates from. Often tumors that are similar to more immature developmental stages are more aggressive. In CLL, for example, patients with CLL cells with mutated V_H genes have a better prognosis than those with unmutated V_H genes (Hamblin et al., 1999).

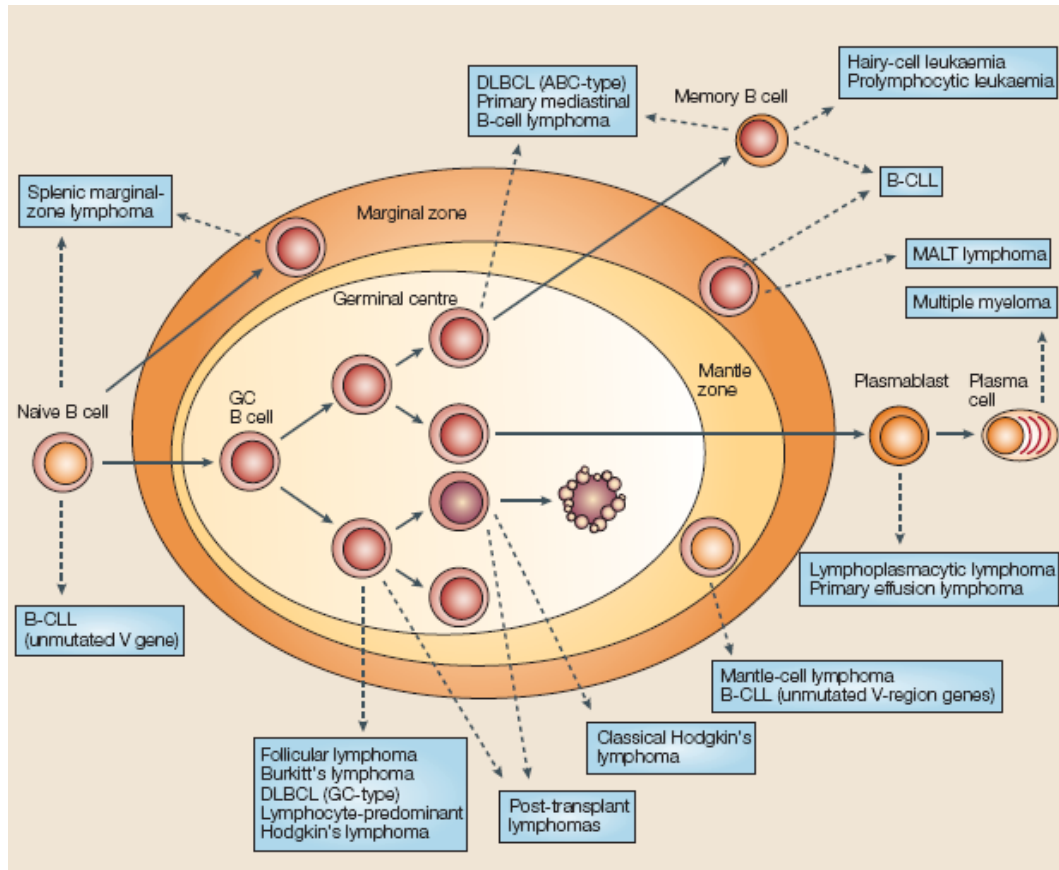


Figure 1-9: Human B-cell lymphomas are assigned to their proposed normal B-cell counterpart (Küppers et al., 1999).

1.3 CLL

B-cell Chronic Lymphocytic Leukemia (CLL) is the most common leukemia among adults in the Western world. It affects elderly people with a median age of 72 years at diagnosis. However, a third of all patients is less than 60 years old at diagnosis (Seiler et al., 2006). Frequently, the diagnosis of CLL is based on the initial diagnosis of a blood lymphocytosis. As a diagnostic cut-off, patients with a B-cell count of greater than 5000 cells/ μ L of blood are diagnosed with CLL while elevated lymphocyte counts that are below this threshold are called Monoclonal B-cell Lymphocytosis (MBL). The diagnosis further requires a specific makeup of CLL/MBL specific surface antigens. CLL cells present a specific combination of surface markers (CD5, CD19 and CD23), which are usually suppressed during B-cell differentiation. Furthermore, CLL cells are clonal and originate from one single cell, which is detectable because the normally diverse population of antibody L chains, kappa or lambda, are homogeneous in the clonal CLL cell population. CLL is a prognostically very

heterogeneous disease and the clinical course of the disease can be highly divergent (see chapter 1.3.1).

CLL is currently not treated in patients with an early stage of the disease and markers that correlate with a good prognosis (Rai 0-II; Binet A and B). In low-risk CLL, no treatment is given before progression ("watch-and-wait") because no survival benefit has been associated with early treatment of asymptomatic CLL (Zenz et al., 2008). However, if the disease becomes symptomatic, CLL has been treated with a variety of chemotherapeutica like chlorambucil, the purine analogs fludarabine, pentostatin and cladribine, and the monoclonal antibodies rituximab and alemtuzumab (Tam and Keating, 2007; Begleiter et al., 1996; Kay et al., 2006; Byrd et al., 2006; Cheson, 2006). The current recommended first-line treatment is a combination of fludarabine, cyclophosphamid and rituximab ("FCR"). In younger patients whose disease shows markers associated with a bad prognosis, bone marrow, i.e. stem cell transplantation, is the therapeutic strategy (Boyiadzis et al., 2007; Gine et al., 2007). Unfortunately, all patients outside of transplantation relapse and CLL remains an incurable disease (Boyiadzis et al., 2007) with a yet unknown pathomechanism.

The origin of CLL cells is a highly debated question. Evidence suggests that CLL is derived from an antigen-experienced B-cell (Ghia and Caligaris-Cappio, 2006). IgV_H gene mutations can be detected in a portion of CLL, which indicates their encounter with an antigen (Ghia and Caligaris-Cappio, 2006; Kröber et al., 2002; Damle et al., 1999). Furthermore, it was shown that both, IgV_H unmutated and mutated CLL cells show a common surface phenotype, which is similar to that of antigen-experienced B-cells. In addition, the BCR in CLL cells resembles the BCR of normal B-cells upon antigen interaction. As shown above CLL cells share similarities with mutated memory B-cells as well as with naïve unmutated B-cells (Figure 1-7, Küppers et al., 1999). The rearrangements of V_HDJ_H gene segments are strikingly similar in unmutated and in mutated forms of CLL and this similarity suggests an antigenic selection in both disease subsets (Ghia and Caligaris-Cappio, 2006).

Clonal evolution is a common process in cancer cells. As described in chapter 1.1.1, the malignant cells that form a tumor are clonal, as they are derived from a single tumor initiating cell. Clonal evolution in CLL is described as the acquisition of additional chromosomal aberrations. Stilgenbauer et al. (2007) showed that 17 % of analyzed patients showed additional aberrations within 42 months of observation. Interestingly, deletion of the first copy of 13q14 is never observed as secondary aberration in clonal evolution. The most frequent acquired secondary aberration in CLL is deletion of 17p13 and the second most frequent one is biallelic loss of 13q14. Furthermore, Stilgenbauer et al. (2007) found that clonal evolution correlates with a poor outcome of CLL, which can be explained on the one hand by the most frequent loss of 17p13, which itself is correlated with a bad prognosis (Döhner et al., 2000). On the other hand, these observations point to a selective advantage of those CLL cells that have lost both copies of 13q14. The estimated survival times of 21.7 months was much shorter than before the occurrence of clonal evolution in CLL cells. Interestingly, clonal evolution only occurred in patients with unmutated V_H genes, which suggests that their karyotypic chromosomal instability is a potential molecular pathomechanism.

CLL patients can be roughly subdivided into two subgroups of different prognosis according to the mutational status of immunoglobulin variable heavy chain genes (IgV_H) of the CLL cells. Both CLL cell types, i.e. with mutated and unmutated IgV_H genes, are mature and most likely not antigen experienced. However, only mutated CLL cells may have undergone somatic hypermutation in germinal centers (Figure 1-9). Patients with mutated IgV_H gene

CLL cells have a far better prognosis (median survival 10-20 years) than those patients with unmutated IgV_H genes (median survival only 5-10 years) (Küppers et al., 2001; Kröber et al., 2002). A surrogate marker for the V_H mutational status is the differential expression of a BCR associated gene, the zeta-associated protein 70 (ZAP-70; (Kienle et al., 2005)). ZAP-70 is a tyrosine kinase, which normally functions in T-cell signaling (Isakov et al., 1995; Madrenas et al., 1995). However, the correlation between V_H mutational status and ZAP-70 expression does not hold true in every patient (Del Principe et al., 2006; Kröber et al., 2006; Rassenti et al., 2004).

1.3.1 Genomic aberrations in CLL and prognostic relevance

In CLL, prognostic subgroups have been identified defined by genomic aberrations (Stilgenbauer et al., 2002; Kröber et al., 2002; Döhner et al., 2000; Cotter and Auer, 2007; Hamblin et al., 2002). Interestingly, translocations are rare in CLL, and the most common genetic aberrations in CLL are deletions of chromosomal material and trisomy 12 (Döhner et al., 1997). Chromosomal aberrations are detectable in over 80 % of CLL cases. Table 1-4 depicts the most common aberrations and their frequency in CLL patients.

Table 1-4: Incidence of chromosomal abnormalities in 325 patients with chronic lymphocytic leukemia. *175 patients had one aberration, 67 had two aberrations, and 26 had more than two aberrations (modified from Döhner et al., 2000).

Aberration	No. of Patients (%)*
13q deletion	178 (55)
11q deletion	58 (18)
12q trisomy	53 (16)
17p deletion	23 (7)
6q deletion	21 (6)
8q trisomy	16 (5)
t(14q32)	12 (4)
3q trisomy	9 (3)
Clonal abnormalities	268 (82)
Normal karyotype	57 (18)

The most frequent aberration is deletion of 13q14, which occurs in 55 % of CLL patients. Additional frequent aberrations are deletion of 11q22-23 (in 18 % of patients), trisomy 12 (in 16 %), deletion of 17p13 (in 7 %) and deletion of 6q21 in 6 % of patients (Döhner et al., 2000). These aberrations are prognostically linked to survival of CLL patients. Patients with a deletion of 13q as single aberration have the best prognosis with a median survival time of 133 months (Figure 1-10). The worst prognosis in CLL is associated with loss of the *TP53* gene in 17p-deleted patients, where the median survival time is only 32 months. The deletion of 11q leads to the loss of function of the tumor suppressor gene *ATM* in a subset of patients, and correlates with a median survival of only 79 months (Figure 1-10). Median survival times of patients with a normal karyotype and trisomy 12 are 111 and 114 months, respectively (Döhner et al., 2000). Despite the good characterization of CLL karyotypes the pathomechanism of CLL remains still unclear. While distinct genotypes have been linked to distinct prognosis of CLL subtypes they have not identified one tumor suppressor gene that is responsible for the development of CLL. Loss of function of *ATM* on 11q and *TP53* on 17p can serve only as an explanation for the worse outcome of these karyotypic subgroups, but are most likely not causative events responsible for the original development of the malignant CLL cell clone. In contrast, the chromosomal band 13q14 and the genes localized in this region are strong candidates for playing a role in the development of CLL cells for

several reasons. The critical region in 13q14.3 is lost in more than 54 % of CLL cases, making it the most common aberration in CLL. In addition, while deletions of 17p and 11q can be detected as secondary effects in clonal evolution of the disease, loss of the first copy of 13q14 is always the primary event (Stilgenbauer et al., 2007).

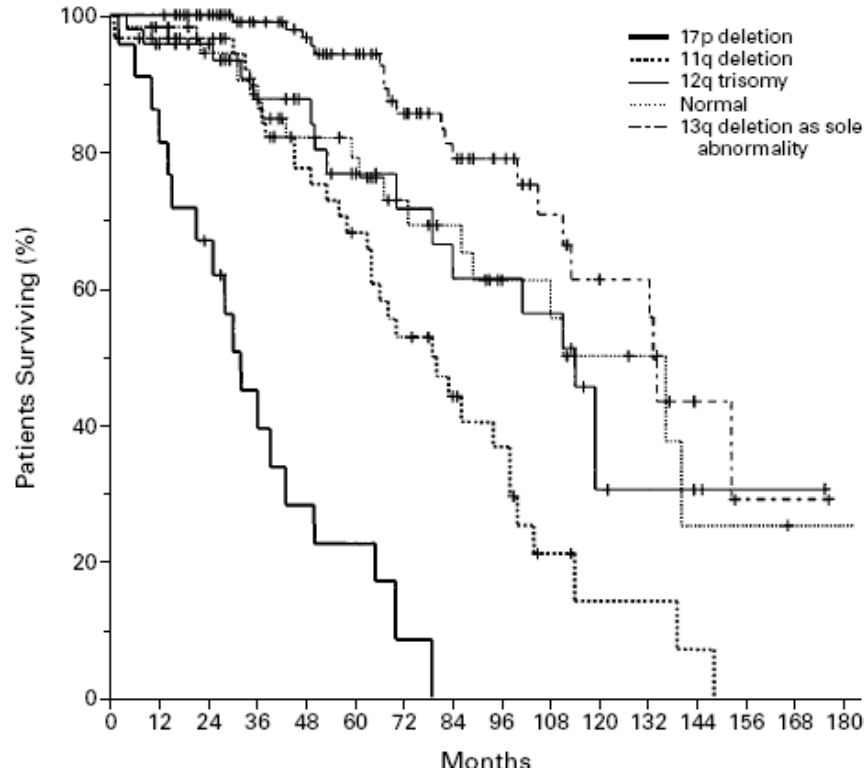


Figure 1-10: Kaplan-Meier-Plots showing survival of CLL patients from the date of diagnosis among patients in the five genetic categories (Döhner et al., 2000).

1.3.2 Chromosomal band 13q14.3

Deletion of chromosomal material from 13q14 is the most common aberration in CLL and correlates with a good prognosis (Döhner et al., 2000). As it is the most common genomic aberration in CLL, it is possible that despite the good prognosis of del(13q) patients, 13q14 might harbor the tumor suppressor mechanism responsible for development of CLL. However, no single tumor suppressor gene could be identified so far despite extensive studies (Chena et al., 2008; Mertens et al., 2002; Stilgenbauer et al., 1998; Hammarlund et al., 2004; Mabuchi et al., 2001). Figure 1-11 represents a schematic overview of the most commonly deleted region in 13q14.3 telomeric to the *RB1* locus that is located about one mega base pairs centromeric to the gene *KPNA3* of 13q14.3. The region contains several protein-coding genes, *KPNA3*, *C13ORF1*, *KCNRG* and *RFP2*, as well as two large noncoding RNA genes *BCMS* (*DLeu1*) and *BCMSUN* (*DLeu2*, *RFP2OS*), which will be discussed in more detail below. Two microRNAs miR15a and miR16-1 are located within an intron of *BCMSUN/DLeu2* (Lagos-Quintana et al., 2001) and have been postulated to play an important role for the pathogenesis of CLL (Calin et al., 2007). The majority of these genes (but not the microRNA genes) have been assigned to promoters each consisting of a CpG-

island that most likely regulate expression of these 13q14 genes (CpG-islands A, B, C, D and E; see Figure 1-11).

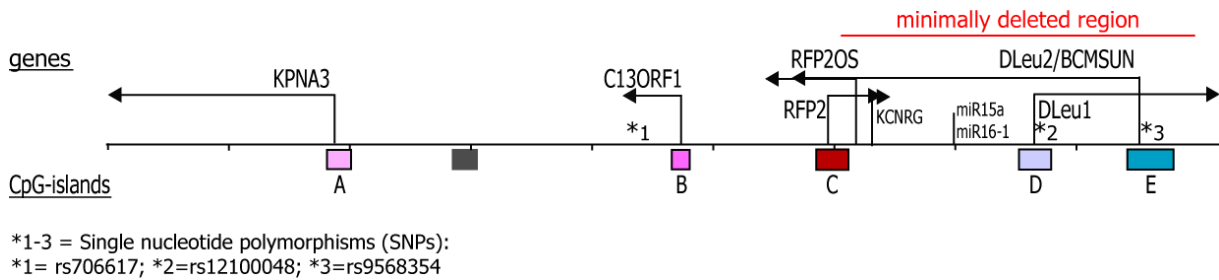


Figure 1-11: Schematic view of chromosomal region 13q14.3. The depicted minimally deleted region is shown as defined in (Stilgenbauer et al., 1998).

In close vicinity to the locus in 13q14 represented in Figure 1-11, the *RB1* gene is located. The function of *RB1* led to the Knudson two-hit-model of tumor suppressor genes (Knudson, 1971, see chapter 1.1.1). The proximity to the minimally deleted locus in CLL made it reasonable that *RB1* might be involved in the tumor suppressor mechanism of 13q14. However, several analyses could show that at least one intact copy of *RB1* and along with it the tumor suppressive function of *RB1* can be found in a subset of CLL patients with loss of 13q14 (Stilgenbauer et al., 1993; Liu et al., 1993) formally showing that a del(13q14) affects an additional tumor suppressor mechanism distinct from *RB1*. Consecutively, it was shown that a region telomeric to the *RB1* gene was more frequently deleted than the TSG itself (Liu et al., 1993). Much effort was invested to determine the minimally deleted region of 13q14.3 in CLL patients. Several minimally deleted regions have been proposed (see Wolf et al., 2001), but a segment flanked by chromosomal markers D13S273 and D13S25 was defined to be the critical region of loss of 13q14.3 (Wolf et al., 2001; Stilgenbauer et al., 1998). This critical minimally deleted region (Figure 1-11) comprises the noncoding RNA genes *BCMS*, *RFP2OS* and *BCMSUN/DLeu2*, the two microRNA genes miR15a and miR16-1 and the protein-coding genes *RFP2* and *KCNRG*.

The candidate genes of 13q14

The protein coding genes of 13q14 have been characterized extensively in mutational and functional analyses as they were the most likely candidates for the tumor suppressor in the region (Stilgenbauer et al., 1998; Wolf et al., 2001; Mertens et al., 2002; Hammarsund et al., 2004; Tschuch, 2006; Serra-Barrionuevo, 2008). These protein coding genes are highly conserved across species and downregulated in CLL patients (Kapanadze et al., 2000; Mertens et al., 2002). However, no point mutation in the coding sequence of any of these candidate genes could be determined so far that would lead to an obvious loss of function. Thus, other possible mechanisms have been considered relevant for the development of CLL due to 13q14 loss. The candidate genes of 13q14.3 are downregulated in CLL patients (Mertens et al., 2002), this suggests that the pathomechanism of CLL involves a deregulation of gene expression of 13q14-genes. Downregulation of the two microRNAs, miR15a and miR16-1, probably leads to downstream deregulation of the apoptosis inhibitor *BCL2* (Calin et al., 2008; Cimmino et al., 2005). *BCL2*-deregulation by the 13q14-microRNAs thus links 13q14-deletions to *BCL2*-overexpression, which is frequently observed in CLL-patients and may be one cause for the inhibition of apoptosis in CLL cells (Kitada et al., 1998).

Aside from the microRNAs, it was also proposed that the large noncoding RNAs might play a role in the pathomechanism of the critical region (Kapanadze et al., 1998; Wolf et al., 2001).

Recently, diverse epigenetic regulatory mechanisms of 13q14 have been postulated to play a role for CLL pathogenesis (Yu, 2006; Mertens et al., 2006; Calin et al., 2005). In 2006 it was shown by our group that the critical region 13q14 is monoallelically expressed due to epigenetic regulation of the candidate genes (Mertens et al., 2006). An epigenetic regulation of the region could explain i) the lack of genetic aberrations in gene copies and ii) the high level of downregulation detected for candidate genes upon loss of only one gene copy (Mertens et al., 2002). As one allele is epigenetically silenced in normal cells, loss of the single active copy in CLL cells would lead to complete inactivation of the tumor suppressor locus. Although many possible pathomechanisms for CLL have been postulated and are discussed, the mechanism has not yet been elucidated. This is why in the present study the epigenetic modifications of the tumor suppressor locus in 13q14 and their possible regulation were analyzed.

A short introduction into proposed molecular functions of 13q14 candidate genes

The candidate tumor suppressor gene *RFP2* (LEU5; TRIM13) is located in the minimally deleted region of 13q14.3, and expression of *RFP2* is downregulated in both patients with monoallelic deletion of 13q14 and retention of both copies of 13q14 (Mertens et al., 2002). It is the only protein-coding gene within the minimally deleted region and therefore the most likely candidate for the TSG of 13q14. The organization of the *RFP2* gene locus is conserved in human and mouse (Baranova et al., 2003). The human *RFP2* promoter region was described to have a unique structure and unusual strength (Skoblov et al., 2006). In human and murine cells, several mRNA isoforms of *RFP2* are expressed, and the *RFP2OS* antisense RNA-transcript is unique in humans (Baranova et al., 2003). Furthermore, the *RFP2* locus contains a second open reading frame (ORF) encoding KCNKG that functions homologous to potassium channel proteins. KCNKG probably has regulatory function for other potassium channels by binding to them (Ivanov et al., 2003). The *RFP2*-promoter region contains a CpG-island (CGI C; Figure 1-11), is TATA-less and contains GGGGA-repeats. Additionally, the promoter was found to have low potential in formation of nucleosomes (Skoblov et al., 2006). Sequence analyses are negative for mutations in the coding sequence of the *RFP2*-locus that was previously regarded as the candidate tumor suppressor gene in the critical region in 13q14.3 (Migliazza et al., 2001). *RFP2* encodes a protein of 510 amino acids, which by homology belongs to a family of TRIM proteins ("TRIPartite Motif"; (Reymond et al., 2001; van Everdink et al., 2003)). TRIM-proteins contain three domains: a RING finger domain, a B-box domain and a coiled-coil domain. Recently, *RFP2* was described to be a transmembrane E3 ubiquitin ligase involved in the endoplasmic reticulum associated degradation, ERAD (Lerner et al., 2007). It was shown that the transmembrane domain of *RFP2* is localized within the ER membrane. Biochemical characterization of *RFP2* is difficult because it is a very unstable protein due to a self-degradation property (Lerner et al., 2007). Previous studies in our group indicate involvement of *RFP2* in apoptosis (Ruppel, 2003; Tschuch, 2006) and in NFκB signalling (Tschuch, 2006).

In addition to the molecular function of *RFP2*, it is essential to determine the epigenetic regulation of the locus in normal and malignant B-cells in order to understand how 13q14 is involved in the pathomechanism of CLL. To this end, the epigenetic status of the critical region in 13q14 was characterized.

Two protein-coding genes of 13q14.3 are localized just proximal to the minimally deleted region, *KPNA3* and *C13ORF1*. Their relevance for the pathomechanism of CLL is not known as their molecular function is unclear.

Karyopherin α -3 (KPNA3) consists of 521 amino acids, a basic N-terminus and a strongly acidic C-terminus (Takeda et al., 1997). The functional domains of the protein are an importin-beta binding domain and several armadillo repeats that promote protein-protein interactions. KPNA3 belongs to a family of importin proteins that mediate transport from the cytoplasm into the nucleus. KPNA3 was shown to be specifically involved in the nuclear transport of RCC1, which is a guanine exchange factor for Ran (Talcott and Moore, 2000). It is further proposed that KPNA3 is involved in the nuclear import of NF κ B, however, direct interaction with NF κ B was only shown for KPNA4 (Fagerlund et al., 2005). Involvement of KPNA3 in import of either NF κ B or RCC1 would generate an interesting link between the 13q14 locus and cell cycle control. Therefore, 13q14 was epigenetically analyzed in this study including the potentially relevant KPNA3 locus and its possible impact on CLL pathogenesis.

C13ORF1 (*CLLD6*) encodes a protein with 196 amino acids. The gene shows high conservation between human and mouse with a nucleotide sequence similarity of 97 % (Baranova et al., 2004) that might highlight the functional relevance of *C13ORF1*. The single homology domain of *C13ORF1*, the SPRY domain, is only poorly characterized. In proteins that contain a SPRY-domain it has been shown that the domain mediates protein-protein interaction and homo-oligomerization (Blayney et al., 2004). A common feature of SPRY-domains is to bind partner proteins in target specific manner. Interestingly, the RFP protein, a repressive transcription factor and homolog of RFP2, also contains a SPRY domain in addition to the TRIM domain. Considering the co-evolved genomic localization of *RFP2* and *C13ORF1* next to each other in 13q14, it is possible that these two proteins interact with each other and together might have a similar function as RFP.

The two microRNA genes localized in 13q14.3, miR15a and miR16-1, are encoded within an intron of *BCMSUN/DLeu2*, and this localization is conserved in mouse (Cimmino et al., 2005). These two microRNAs possibly target the anti-apoptotic gene *BCL2*. Calin et al. proposed in 2005 a direct correlation between the frequent loss of miR15/16 by 13q14-deletion and the previously shown upregulation of *BCL2* in CLL patients.

Not much is known about the two large noncoding RNA (ncRNA) genes, *BCMS* (*DLEU1*) and *BCMSUN* (*DLEU2*) that are both localized in the critical region in 13q14.3. However, their importance in the regulatory mechanism of the region is stressed by the high degree of conservation of the ncRNA *DLeu2/BCMSUN* itself, of their nucleotide sequence and of the topological organization in human and mouse (Corcoran et al., 2004). The *BCMS* RNA was found to be heavily spliced, but no molecular function could be assigned to the ncRNA transcript (Wolf et al., 2001). Recently, Stelzl et al. (2005) performed a large yeast two-hybrid screen with consecutive co-immunoprecipitation assays. In this large screen, *DLeu1* was expressed as a fusion protein, and intriguingly the product of this short predicted ORF interacted with other proteins. The second large noncoding RNA of 13q14, *BCMSUN/DLeu2*, spans a large part of the minimally deleted region. The orientation of the *BCMSUN* gene is opposite to that of the genes *BCMS* and *RFP2* (Corcoran et al., 2004) and consequently, the two large noncoding RNA loci are organized in opposite orientation to each other and actually overlap without sharing coding sequence (Figure 1-11). The organization and conservation of the two large noncoding RNA genes in 13q14 suggests that they are part of a complex regulatory mechanism. Other human ncRNA genes, for example *Air* or *Xist*, have been described to regulate the expression of neighboring genes in imprinted and monoallelically expressed loci (Rougeulle and Heard, 2002).

As *BCMS* and *BCMSUN* are likewise long ncRNA genes, it is possible that they have similarly a role in the regulation of expression of the genes localized in *cis* in 13q14 and that they take

part in the transcriptional control of the whole region. In this model, the aberrant expression of *BCMSUN* and *BCMS* would lead to the deregulated expression of the 13q14 protein coding genes that has been described (Calin et al., 2002; Mertens et al., 2002).

A subset of CpG-dinucleotides of the human genome are clustered in distinct regions termed CpG-islands that cover about 1 % of the human genome (Bird, 2002; Lander et al., 2001). These CpG-islands are found to be associated with 60 % of human gene promoters. Although most of the cytosines within CpG dinucleotides are unmethylated, the subset contained in promoter-CpG-islands has been shown to be differentially methylated during development in a tissue-specific manner (Bird, 2002). The CpG-island promoters are not exclusively subjected to regulation by differential DNA-methylation. However, some studies could show that promoters that lack CpG-islands according to the original strict definition, also have tissue-specific methylation patterns that correlate with transcriptional activity (examples in Blleloch et al., 2006; Hattori et al., 2004). Therefore, recently it has been suggested to amend the definition of a CpG-island (Illingworth et al., 2008). Five CpG-islands in the 13q14-region meet the stringent criteria (see Figure 1-11). These CpG-islands (CGIs A to E) have been analyzed in detail in this work to characterize their epigenetic status. In prior studies, only the *RFP2* promoter region was analyzed in great detail (Skoblov et al., 2006), and in our group it was shown that CpG-island E is DNA methylated in cultured cells (Mertens et al., 2006). Here, a detailed analysis of five promoter-CpG-islands with respect to chromatin modifications (i.e. histone modifications and DNA-methylation) was performed in healthy and in CLL cells in order to gain insight in the epigenetic regulatory pattern of candidate genes in the critical region 13q14.

To date, several effects resulting in deregulated expression of 13q14 candidate genes can be related – at least in theory – to epigenetic regulation of the locus: Downregulation of several genes in the critical region, namely of *RFP2*, *BCMSUN*, *CLLD7*, *C13ORF1* and also of *RB1* in 13q14-deleted CLL patients, could be due to epigenetic silencing. The only gene found to be significantly down-regulated also in patients with retention of both copies of the critical region was *RFP2*, which was found along with significant upregulation of *BCMS* (Mertens et al., 2002). The fact that no point mutations have been identified in the protein-coding and microRNA genes in CLL patients also suggests epigenetic regulation being responsible for deregulated expression of the genes. Recently, asynchronous replication of the two 13q14 copies and monoallelic expression of most genes localized in the critical region was shown by our group, and a possibly epigenetic pathomechanism of CLL localized in 13q14 was postulated (Mertens et al., 2006). According to Knudsons “two-hit hypothesis”, epigenetic silencing of one allele in normal cells would require a “second hit” for complete loss of function. This second hit might be the deletion of 13q14 in CLL patients.

1.3.3 Mechanisms of regulating gene expression

Gene expression can be regulated by distinct epigenetic modifications. The epigenetic code determines the function of distinct genomic sites in different tissues by regulating the extent of transcriptional activity (see chapter 1.1.2). Monoallelic expression has recently been shown to be more common in the human genome than expected (Verlaan et al., 2008; Gimelbrant et al., 2007). Different mechanisms have been found to contribute to several forms of monoallelic expression of genes. The mechanisms of autosomal monoallelic expression and imprinting and the associated epigenetic marks are described in this chapter.

One of the best described examples for monoallelic expression is genomic imprinting. Imprinted genes are expressed in a parent-of-origin specific manner, where always either the maternal or the paternal allele is monoallelically expressed (Goldmit and Bergman, 2004).

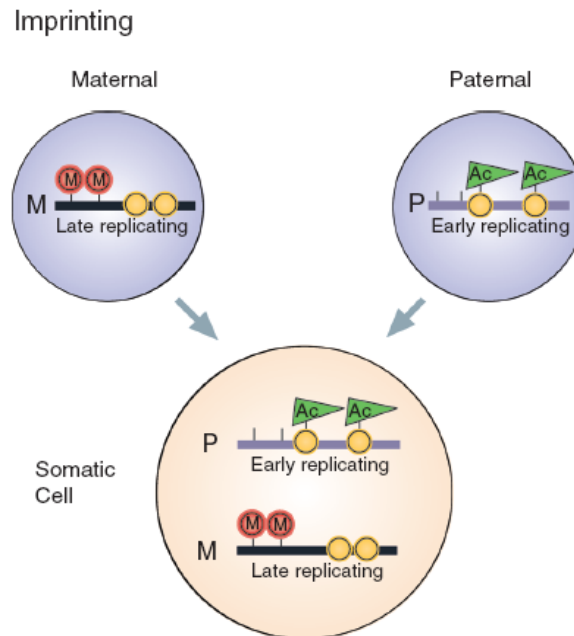


Figure 1-12: Schematic representation of the establishment of epigenetic imprinting marks during development. DNA-methylation = red balloons, replication timing: purple ribbon = early replicating, black ribbon = late replicating, chromatin structure: yellow circles = nucleosomes, green flags = acetylated (Ac) histones (Goldmit and Bergman, 2004).

The process of establishment and maintenance of genomic imprinting is well characterized and implies several steps of epigenetic control. The first epigenetic marks on imprinted genes are established very late during gametogenesis to identify each allele according to its origin as paternal (P) or maternal (M; Figure 1-12). The epigenetic marks include DNA-methylation, replication timing (Figure 1-12; see also allelic exclusion, chapter 1.2.2) and chromatin structure. These epigenetic marks are stably maintained through somatic cell divisions after fertilization. Maintenance of imprints is achieved by differentially methylated regions (DMR) or imprinting control elements (ICE) that are located within or near the imprinted loci. The DMRs are differentially methylated at their two parental copies; however, the parental copy carrying the methylated DMR is not necessarily silenced, as the DMRs are only one part in the regulatory system of imprinted loci. Often DMRs are involved in the regulation of expression of noncoding RNAs that are frequently localized at imprinted genomic loci. One example for a large noncoding RNA that is regulated by a DMR and in turn regulates expression and silencing of the imprinted locus is *Air* at the *Igf2*-locus (Rougeulle and Heard, 2002). Physiological functions of imprinted genes are frequently the control of embryonic growth and development, development of the placenta and also post-natal development (Tycko and Morison, 2002). According to the parental-conflict hypothesis, paternally expressed (maternally imprinted) genes are promoting growth of the embryo/fetus, whereas maternally expressed (paternally imprinted) genes are growth limiting (Moore and Haig, 1991).

Loss of imprinting (LOI) is associated with several human syndromes like Beckwith-Wiedemann syndrome, which is caused by loss of imprinting on chromosome 11 (Grundy et al., 1998). The two different diseases Angelman's and Prader-Willi syndrome are caused by deletion of a region of chromosome 15q. Which of the two syndromes develops depends on whether maternal or paternal imprinting is lost, i.e. whether the mutation is inherited from the father or the mother (Zeschnigk et al., 1997). Finally, it was shown that global LOI can lead to widespread tumorigenesis in adult mice, which proves a direct predisposition of cells to tumorigenesis caused by LOI (Holm et al., 2005). Thus, imprinting can also play a role in the development of cancer by its homeostatic function that provides tumor suppressor activity in adult tissues (Klein et al., 2007). Most of the genes in the human genome of diploid cells are expressed from both maternally and the paternally derived copies of each gene. However, recently it has been shown that more genes than previously known are expressed from only one of the two gene copies (Gimelbrant et al., 2007; Verlaan et al., 2008). Monoallelically expressed genes are grouped into three categories. The first category comprises imprinted genes that were described above. The second category of monoallelically expressed genes consists of X-inactivated genes that are mostly randomly inactivated early in development. The third category comprises autosomal genes that are subject to random monoallelic expression. The third category was considered to consist only of a very small number of genes, however, Gimelbrant et al. (2007) assessed in a genome-wide screen allele-specific transcription of about 4000 human genes and could show that more than 5 % of these are subject to random monoallelic expression. Thus it becomes clear that monoallelic expression is a quite frequent phenomenon independent of gene function or tissue environment. Autosomal monoallelically expressed genes that were known before this study, include odorant receptor genes, immunoglobulin genes (chapter 1.2.2), T-cell receptor genes, natural killer cell receptor genes and interleukins. The genes that were found in the study by Gimelbrant et al. to be monoallelically expressed encode proteins with various functions and tissue-specificities, except for a large group of genes that encode cell surface proteins (Gimelbrant et al., 2007). The most interesting genes identified in the study are those that have been previously assigned roles in human leukemias, like for example *ABR* (active BCR related gene), *EBF* (early B-cell factor), *TCL1A* (T-cell leukemia/lymphoma 1A) and *ZNFV1A1* (Zinc finger protein, subfamily 1A, Ikaros) (Gimelbrant et al., 2007). These examples clearly show the other side of monoallelic expression that can rapidly lead to cancer.

An additional epigenetic mark of monoallelically expressed genes was determined by Rougeulle et al. (2003). In their study it was shown that all monoallelically expressed genes have a distinct histone modification pattern, consisting of promoter-restricted dimethylation of histone H3 at lysine 4 (H3K4me2). Promoter-restricted H3K4me2 was proposed as a distinct epigenetic mark for monoallelic expression. Since the genes of 13q14 are monoallelically expressed (chapter 1.3.2), this epigenetic mark was analyzed at the genomic region 13q14 in this work.

Monoallelic expression is often regulated by locus control regions (LCR), which are cis-regulatory elements that have the ability to enhance gene expression of linked genes to physiological levels. They influence gene expression depending on the tissue and may vary in their composition and location relative to regulated genes (Li et al., 2002). The first LCR was identified about 20 years ago in the mouse β -globin locus (Jimenez et al., 1992; Moon and Ley, 1990). Since then, many LCRs have been described in mammalian gene systems and they are now considered to be common functional elements in mammalian developmental

and cell-lineage specific regulation of gene expression (Li et al., 2002). This regulation not only depends on promoters, enhancers and silencers but also on interactions of diverse cis-regulatory elements, including LCRs, and on dynamic chromatin interactions (Li et al., 2002). The functionally and structurally conserved composition of the region 13q14 points to an epigenetically regulated pathomechanism that may involve a control element comparable to the LCRs explained above.

1.3.4 Epigenetic aberrations in CLL

The epigenetic alterations described in CLL include the deregulation of miR15/16 (Calin et al., 2002), global DNA-hypomethylation (Wahlfors et al., 1992) and DNA-hypermethylation of for instance the death-associated protein kinase 1 (DAPK1) gene (Raval et al., 2007). Also genomewide promoter hypermethylation has been described in CLL (Raval et al., 2006) and DNA-methylation profiles specific for CLL have been established by several groups (Rush et al., 2004; Plass et al., 2007). The two miRNA genes of 13q14, miR15a and miR16-1, are frequently down-regulated in CLL resulting in removal of their tumor suppressive influence on their target *BCL2* (Cimmino et al., 2005). However, to date no mechanisms for this observed deregulation has been described, i.e. it is unclear how these candidate tumor suppressor genes are downregulated in CLL. Further epigenetic changes described in CLL concerning miRNAs is description of a miRNA signature that distinguishes CLL cells from normal B-cells (Calin et al., 2004) and is supposed to be associated with prognosis and disease progression. Several findings suggest an epigenetic pathomechanism that is localized in the critical region 13q14 in the development of CLL. The absence of pathogenic point mutations that would result in a functional loss within the candidate genes of the critical region suggests that the downregulation of these genes in monoallelically deleted patients is either due to haploinsufficiency or epigenetic suppression (Cotter and Auer, 2007). Epigenetic suppression of the single retained 13q14 copy in CLL patients is suggested by data from several studies (Kienle et al., 2005; Mertens et al., 2002; Dickinson et al., 2006; Haslinger et al., 2004; Mertens et al., 2006). Asynchronous replication timing of the 13q14 critical region was demonstrated by Mertens et al. (2006), which points to differential packaging of chromatin of the two alleles. Also, monoallelic expression of almost all candidate genes in 13q14 was detected in B- and T-cells of healthy donors, which could be the result of the differential allelic chromatin packaging. Both findings explain the nearly complete downregulation of candidate 13q14 genes (Mertens et al., 2002) and microRNAs (Calin et al., 2002) frequently observed in CLL patients with a monoallelic deletion of 13q14. Thus, it becomes clear that epigenetic modifications seem to play a major role for the pathogenesis of CLL.

In the present study, a detailed analysis of several epigenetic modifications of the critical region 13q14 was performed in order to elucidate the epigenetic regulatory pathomechanism of 13q14 that we propose to be involved in the development of CLL.

1.4 Aims of this work

Deletion of genetic material from a critical region in 13q14.3 is the most frequent chromosomal aberration in CLL. Several candidate genes from the critical region including protein-coding and noncoding RNA genes have been proposed to participate in the pathomechanism of CLL. Since neither pathogenic point mutations nor a single tumor suppressor gene were identified that would explain the downregulation of these candidate genes in CLL, in addition to the finding that most of them are monoallelically expressed, it is proposed that the regulatory mechanism of 13q14.3 is epigenetically controlled.

In the present work, the genes and CpG-islands of the critical region in 13q14 were epigenetically characterized. The aim was to identify an epigenetic regulatory mechanism that controls gene expression in the region 13q14.3 and to describe how its function is lost in CLL. Thus, the wild type epigenetic code of 13q14.3 was determined and subsequently, CLL cells were screened for aberrant epigenetic modifications.

Chromatin immunoprecipitation (ChIP) was used to study epigenetic modifications at the critical region in different tissues. ChIP methodology was established and optimized to work with hematopoietic tissue. ChIP was used to quantify two epigenetic modifications, H3K4me2- and macroH2A-enrichment at 13q14.3 in non-malignant cells and CLL cells.

The critical region in 13q14.3 shares similarities with imprinted regions including the regulatory large noncoding RNA genes, monoallelic expression of the genes and regulation of gene expression by distinct epigenetic modifications. The transcription of imprinted ncRNA genes is often dependent on differentially methylated imprinting control elements (ICE) located in the vicinity of the genes. Thus, the existence of a locus control region (LCR) in the critical region in 13q14.3 is postulated to control the transcription of the ncRNA genes in the critical region analogous to ICEs in imprinted regions. The aim of this work was to identify such an epigenetically regulated region control element in chromosomal region 13q14.3, which might be functionally inactivated in CLL cells either together with the minimally deleted region or by epigenetic silencing (epimutation). The characterization of the epimutation in 13q14.3 in this work is an essential step towards elucidation of the underlying molecular cause for the development of CLL.

2 Material & Methods

Material

2.1 Biological material

2.1.1 Cultured cell lines

Table 2-1: Cultured cell lines

Cell line (DSMZ no.)	Origin/tumor tissue	Culture conditions characteristics	Reference
GRANTA-519 (ACC 342)	Human B-cell lymphoma; high-grade B-NHL; established from peripheral blood of 58-year-old female patient	90 % Dulbecco's MEM (4.5 g/L glucose) + 10 % FBS + 2 mM L- glutamine; spheroid cells in suspension, single or in clumps	(Jadayel et al., 1997)
HL60 (ACC 3)	Human acute myeloid leukemia; (AML FAB M2); established from peripheral blood of 35-year-old female patient	90 % RPMI 1640 + 10 % FBS; round, single cells in suspension	(Collins et al., 1977)
JURKAT (ACC 282)	Human T-cell leukemia; acute lymphoblastic leukemia (ALL); established from peripheral blood of 14-year-old male patient	90 % RPMI 1640 + 10 % FBS; round cells in suspension; single or in clumps	(Schneider et al., 1977)
NAMALWA (ACC 24)	Human Burkitt lymphoma; established from tumor mass of an African child in 1967	90 % RPMI 1640 + 10 % FBS; lymphoblastoid, single cells or small clusters in suspension	(Nadkarni et al., 1969)

2.1.2 Primary cell samples

Table 2-2: Tumor samples from the peripheral blood of B-CLL patients (n.k. = not known).

Sample No.	Tumor No.	sample age at income [years]	Binet stage	Karyotype	Tumor cell content [%]	V _H status	Gender
CLL1	02PB2757	51	A	17p deletion, +12q13	74 (17p), 14 (12q)	Unmutated	Male
CLL2	05PB3784	76	n.k.	13q deletion	85	Mutated	Female
CLL3	04PB3493	54	B	Normal	n.k.	Unmutated	Male
CLL4	05PB3867	81	C	13q single deletion	93.5	Unmutated	Male
CLL5	05PB3937	63	B	13q single deletion	83	Unmutated	Male

Coding of tumors in anonymous and assigned by the Universitätsklinik Ulm.

2 Material & Methods

Table 2-3: Primary wild type samples from the peripheral blood of healthy probands. (PBL = peripheral blood lymphocytes)

Sample No.	Tissue type	Sample age at income [years]	Gender	Sample type
P1	Wild type PBL	33	Male	Young subgroup
P2	Wild type PBL	23	Female	Young subgroup
P3	Wild type PBL	37	Male	Young subgroup
P4	Wild type PBL	25	Female	Young subgroup
P5	Wild type PBL	43	Male	Young subgroup
PB5	Wild type PBL	52	Male	Old subgroup
PB6	Wild type PBL	48	Male	Old subgroup
PB7	Wild type PBL	48	Male	Old subgroup
PB17	Wild type PBL	54	Male	Old subgroup
PB18	Wild type PBL	49	Male	Old subgroup
PB19	Wild type PBL	50	Male	Old subgroup
PB9	CD19+ B-cells	70	Male	Wild type B-cells
PB9	CD19- T-cells	70	Male	Wild type T-cells
PB12	CD19+ B-cells	47	Male	Wild type B-cells
PB12	CD19- T-cells	47	Male	Wild type T-cells
PL1	Wild type PBL	27	Female	E6 (rs9568354) heterozygote
PU1	Wild type PBL	35	Female	E6 (rs9568354) heterozygote
PU2	Wild type PBL	40	Female	E6 (rs9568354) heterozygote
PU3	Wild type PBL	39	Female	E6 (rs9568354) heterozygote
PU4	Wild type PBL	33	Female	E6 (rs9568354) heterozygote
PU5	Wild type PBL	24	Female	D6 (rs12100048) heterozygote
PU6	Wild type PBL	42	Female	D6 (rs12100048) heterozygote
PU7	Wild type PBL	37	Male	D6 (rs12100048) heterozygote
PU8	Wild type PBL	38	Male	D6 (rs12100048) heterozygote
PU9	Wild type PBL	25	Male	D6 (rs12100048) heterozygote
PU10	Wild type PBL	29	Female	D6 (rs12100048) heterozygote
PU11	Wild type PBL	30	Male	D6 (rs12100048) heterozygote
PU12	Wild type PBL	26	Female	D6 (rs12100048) heterozygote

2.2 Chemicals and biochemicals

Table 2-4: Chemicals and biochemicals

Chemical	Supplier
2-mercaptoethanol (thioethylene glycol)	Sigma-Aldrich, Munich
2-Propanol (isopropanol)	Merck, Darmstadt
Acetic acid	Roth, Karlsruhe
Acrylamide/Bis-Acrylamide (30 % w/v)	Bio-Rad Laboratories, Munich
Agarose	Sigma-Aldrich, Munich
Ampicillin	Roche Diagnostics, Mannheim
Aprotinin	Sigma-Aldrich, Munich
Benzamidin	Sigma-Aldrich, Munich
Betaine	Sigma-Aldrich, Munich
Bicinchoninic acid (BCA)	Sigma-Aldrich, Munich
Bovine serum albumine (BSA)	New England Biolabs, Frankfurt am Main
Chloroform	Merck, Darmstadt
Complete EDTA-free protease inhibitor cocktail tablets	Roche Diagnostics, Mannheim
Copper-(III)-sulfate	Sigma-Aldrich, Munich
CotI human DNA	Roche Diagnostics, Mannheim
Cyanin3-dUTP	Amersham Bioscience, Freiburg

Chemical	Supplier
Cyanin5-dUTP	Amersham Bioscience, Freiburg
Deoxynucleotide (dNTP) set	Amersham Bioscience, Freiburg
DNA 10 bp ladder	Invitrogen Karlsruhe
Dithiothreitol (DTT)	Sigma-Aldrich, Munich
Ethylenediaminetetraacetic acid (EDTA)	Merck, Darmstadt
Ethanol	Merck, Darmstadt
Ethanolamine	Sigma-Aldrich, Munich
Ethidium bromide	Sigma-Aldrich, Munich
Ficoll 400	Biochrom AG, Berlin
Fluorescein Avidin D	Vector Laboratories, Burlingame, USA
Formamide	Merck, Darmstadt
Glycerine	Roth, Karlsruhe
Glycine	Roth, Karlsruhe
Karyomax Colcemid	Invitrogen, Karlsruhe
Hepes	Sigma-Aldrich, Munich
Hydrochloride acid (HCl)	Merck, Darmstadt
Isoamylalcohol	Merck, Darmstadt
Leupeptin	Sigma-Aldrich, Munich
Lithium chloride (LiCl)	Merck, Darmstadt
Magnesium chloride	Merck, Darmstadt
Methanol	Merck, Darmstadt
Milk powder	Sigma-Aldrich, Munich
Nonidet P-40	Amersham Bioscience, Freiburg
Paraformaldehyde	Sigma-Aldrich, Munich
Pepsin	Sigma-Aldrich, Munich
Phenol-chloroform-isoamylalcohol	Carl Roth, Karlsruhe
Phenol	Roth, Karlsruhe
Phenylmethylsulphonyl fluoride (PMSF)	Fluka/Sigma-Aldrich, Munich
Potassium chloride (KCl)	Sigma-Aldrich, Munich
Potassium dihydrogen phosphate (KH ₂ PO ₄)	Carl Roth, Karlsruhe
Protein A sepharose	Amersham Bioscience, Freiburg
Protein G sepharose	Amersham Bioscience, Freiburg
Rainbow molecular weight marker	Amersham Bioscience, Freiburg
Random primer (random hexamers)	Roche Diagnostics, Mannheim
Salmon sperm DNA	Invitrogen, Karlsruhe
Salmon sperm DNA/protein A agarose	Millipore, Schwalbach
SeaBlue Plus2 pre-stained standard	Invitrogen, Karlsruhe
Sodium acetate (NaAc)	Merck, Darmstadt
Sodium chloride (NaCl)	Merck, Darmstadt
Sodium citrate	Merck, Darmstadt
Sodium dodecyl sulfate (SDS)	Carl Roth, Karlsruhe
Sodium hydrogen phosphate (Na ₂ HPO ₄)	Sigma-Aldrich, Munich
Sodium hydroxide (NaOH)	Merck, Darmstadt
Spermidine	Sigma-Aldrich, Munich
Spermine	Sigma-Aldrich, Munich
T4 gene 32 protein	USB, Cleveland, USA
TEMED	Bio-Rad Laboratories, Munich
Tricine	Carl Roth, Karlsruhe
Tris-Base	Sigma-Aldrich, Munich
Triton X-100	Sigma-Aldrich, Munich
Tween-20	Merck, Darmstadt
ULTRAhyb hybridization buffer	Ambion, Austin, USA
VECTASHIELD with DAPI	Linaris, Wertheim-Bettingen
Yeast tRNA	Invitrogen, Karlsruhe

2.3 Enzymes

Table 2-5: Enzymes

Enzyme	Supplier
Advantage cDNA polymerase mix	BD Bioscience Clontech, Heidelberg
DNase I (10 U/ μ L)	Roche Diagnostics, Mannheim
ExoKlenow (1 U/ μ L)	Invitrogen, Karlsruhe
Proteinase K (20 mg/mL)	Quiagen, Hilden
Restriction enzymes	New England Biolabs, Beverly, USA
BstUI	
HpaII	
MseI	
McrBcr	
SuperScript II reverse transcriptase (200 U/ μ L)	Invitrogen, Karlsruhe
T4 DNA ligase (1 U/ μ L)	Roche, Grenzach-Whylen

2.4 Antibodies

Table 2-6: Commercial antibodies. WB = Western Blot; "-" = not used in ChIP.

Antibody	Supplier	Dilution (WB)	Amount per ChIP
Anti-CTCF	Upstate/Millipore, Schwalbach	1:1000	5 μ g
Anti-Digoxigenin-Fluorescein Fab Fragment	Roche Diagnostics, Mannheim	-	-
Anti-E2F4	Santa Cruz, Heidelberg	-	7 μ g
Anti-histone H3 dimethylated Lys 4 (anti-H3K4me2)	Upstate/Millipore, Schwalbach	1:2000	4 μ g
Anti-histone H3 trimethylated Lys 4 (anti-H3K4me3)	Abcam, Cambridge, UK	1:1000	5 μ g
Anti-histone H3 trimethylated Lys 9 (anti-H3K9me3)	Abcam, Cambridge, UK	1:500	5 μ g
Anti-histone H3 unmodified	Upstate/Millipore, Schwalbach	1:7500	4 μ g
Anti-histone H3 trimethylated Lys 27 (anti-H3K27me3)	Upstate/Millipore, Schwalbach	1:500	5 μ g
Anti-histone H4 unmodified	Upstate/Millipore, Schwalbach	1:2000	4 μ g
Anti-HP1 γ	Upstate/Millipore, Schwalbach	1:1000	6 μ g
Anti-Ikaros	Santa Cruz, Heidelberg	1:1000	5 μ g
Anti-macroH2A1.2	Upstate/Millipore, Schwalbach	1:1000	
HRP-conjugated anti-goat IgG	Dianova, Germany	1:5000	-
HRP-conjugated anti-mouse IgG	Amersham Bioscience, Freiburg	1:5000	-
HRP-conjugated anti-rabbit IgG	Abcam, Cambridge, UK	1:5000	-

2.5 Primer sequences

All primer sequences designed during this work are listed in the appendix (Appendix B).

Table 2-7: Additional primers used in aCGH and aPRIMES. Primer sequences designed by S. Pfister; B. Radlwimmer.

Primer name	Sequence (5' → 3')
pGEM-forward	GGCCGCGGGATATCACTA-
pGEM-reverse	CTCAAGCTATG CATCCAACG
ddMse11	TAACTGACAG
Lib1	AGTGGGATTCTGCTGTCACT

2.6 Cell culture material

Table 2-8: Cell culture material

Material	Supplier
Dimethyl sulfoxid (DMSO)	Sigma-Aldrich, Munich
DMEM (Dulbecco's Modified Eagle Medium)	Gibco BRL/Invitrogen, Karlsruhe
RPMI	Gibco BRL/Invitrogen, Karlsruhe
FCS (fetal cow serum)	Biochrom AG, Berlin
Penicillin (10,000 U/mL)	Gibco BRL/Invitrogen, Karlsruhe
Streptomycin (100 µg/mL)	Gibco BRL/Invitrogen, Karlsruhe
L-Glutamine (200 mM)	Gibco BRL/Invitrogen, Karlsruhe
Streptomycin (10,000 U/mL)	Gibco BRL/Invitrogen, Karlsruhe

2.7 Kits

Table 2-9: Kits

Kit	Supplier
Absolute™ QPCR SYBR Green ROX dUTP mix	ABgene, Epsome, UK
BCA Protein Assay	Pierce/Thermo Scientific, Rockford, USA
Concert Rapid Gel Extraction System	Marligen Biosciences, Ijamsville, USA
ECL Western Blot detection Kit	Amersham Bioscience, Freiburg
ECL ⁺ Western Blot detection kit	Amersham Bioscience, Freiburg
Plasmid Maxi Kit	Quiagen, Hilden
PCR rapid purification kit	Marligen Biosciences, Ijamsville, USA
PRISM Big DyeDeoxy Terminator Cycle	PE Applied Biosystems, Darmstadt

2.8 Other materials

Table 2-10: Other materials

Material	Supplier
3MM paper	Whatman, Maidstone, UK
Cover slips	Sigma-Aldrich, Munich
Falcon™ 50 mL conical tubes	BD Biosciences, San Jose, USA
Falcon™ 15 mL conical tubes	BD Biosciences, San Jose, USA
Immobilon™-P-PVDF membrane	Millipore, Schwalbach
Hyperfilm ECL	Amersham Bioscience, Freiburg
MicroAmp plates, optical 384-well reaction plate	PE Applied Biosystems, Weiterstadt
MicroAmp plates, optical 96-well reaction plate	PE Applied Biosystems, Weiterstadt
MicroAmp, optical caps	PE Applied Biosystems, Weiterstadt
Microcentrifuge tubes, PCR clean, 0.5 mL	Eppendorf, Hamburg
Microcentrifuge tubes, PCR clean, 1.5 mL	Eppendorf, Hamburg
Microcentrifuge tubes, PCR clean, 2 mL	Eppendorf, Hamburg
Poly-Prep L-lysine coated slides	Sigma-Aldrich, Munich
TOPO TA vector	Invitrogen, Karlsruhe

2.9 Standard solutions and buffers

2.9.1 Nuclei extraction (ChIP/IP)

<u>Isotonic buffer</u>	150 mM	NaCl
	1.5 mM	MgCl ₂
	10 mM	Tris-HCl, pH 7.5
	0.5 %	Nonidet P-40
		ddH ₂ O

<u>Lysis buffer</u>	250 mM	NaCl
	20 mM	Na-phosphate
	30 mM	Na-pyrophosphate
	0.5 mM	EDTA
	10 mM	NaF
	0.1 % (v/v)	Nonidet P-40
	10 % (v/v)	Glycerol
	ddH ₂ O	

Protease inhibitors, 0.1 M PMSF and 0.1 M DTT were added freshly prior to use of buffers.

2.9.2 ChIP-buffers – xChIP

<u>SDS-Lysis buffer</u>	1 % (w/v)	SDS
	50 mM	Tris, pH 8.1
	10 mM	EDTA
		ddH ₂ O

<u>Low salt buffer</u> (ChIP wash buffer 1)	0.1 % (w/v)	SDS
	1 %	Triton X-100
	2 mM	EDTA
	20 mM	Tris-HCl, pH 8.0
	150 mM	NaCl
		ddH ₂ O
<u>High salt buffer</u> (ChIP wash buffer 2)	0.1 % (w/v)	SDS
	1 %	Triton X-100
	2 mM	EDTA
	20 mM	Tris-HCl, pH 8.0
	500 mM	NaCl
		ddH ₂ O
<u>LiCl buffer</u> (ChIP wash buffer 3)	0.25 M	LiCl
	1 % (v/v)	Nonidet P-40
	1 % (w/v)	Na-deoxycholate
	1 mM	EDTA
	10 mM	Tris-HCl, pH 8.0
		ddH ₂ O

2.9.3 ChIP-buffers – nChIP

<u>CIB buffer</u>	5 mM	Tris, pH 7.4
	0.5 mM	EDTA
	20 mM	KCl
		ddH ₂ O
freshly added:	0.5 mM	DTT
	0.2 mM	PMSF
	0.05 mM	Spermine
	0.125 mM	Spermidine
<u>WASH buffer</u>	20 mM	Hepes, pH 8.0
	20 mM	KCl
	0.5 mM	EDTA
		ddH ₂ O
freshly added:	0.5 mM	DTT
	0.1 mM	PMSF
<u>WASH 0.3 M buffer</u>	20 mM	Hepes, pH 8.0
	20 mM	KCl
	0.5 mM	EDTA
	0.3 M	NaCl
		ddH ₂ O
freshly added:	0.5 mM	DTT
	0.05 mM	PMSF
<u>NET buffer</u>	150 mM	NaCl
	50 mM	Tris, pH 7.5
	1 mM	EDTA
	0.1 % (v/v)	Nonidet P-40
	0.25 % (w/v)	Gelatine
		ddH ₂ O

2.9.4 Acid extraction of histones

<u>Lysis buffer</u>	10 mM	Hepes, pH 7.9
	1.5 mM	MgCl ₂
	10 mM	KCl
		ddH ₂ O
freshly added:	0.5 mM	DTT
	1.5 mM	PMSF

2.9.5 Western Blot

<u>Blocking solution</u>	10 g	Milk powder
	3 g	BSA
	1 % (v/v)	Tween-20
		ddH ₂ O
<u>2x Laemmli loading buffer</u>	50 mM	Tris-HCl, pH 8.0
	2.4 % (w/v)	SDS
	8 % (v/v)	Glycerin
	0.2 % (w/v)	Bromophenolblue
		ddH ₂ O
<u>5x SDS-PAGE running buffer</u>	25 mM	Tris-Base
	200 mM	Glycine
	10 % (w/v)	SDS
		ddH ₂ O
<u>1x TricinePAGE running buffer</u>	17.9 g	Tricine
	1 g	SDS
	12.11 g	Tris-Base
	1 L	ddH ₂ O
<u>10x Blotting buffer</u>	25 mM	Tris-Base
	192 mM	Glycine
		ddH ₂ O
Added at preparation of 1x buffer:	20 % (v/v)	Methanol
<u>TBS-Tween</u>	20 mM	Tris-Base
	137 mM	NaCl
	3.8 mM	HCl
	0.1 % (v/v)	Tween-20
		ddH ₂ O

2.9.6 CHIP-on-chip (mCGH)

<u>Blocking solution</u>	50 mM	Ethanolamine
	0.1 % (v/v)	SDS
	0.1 M	Tris, pH 9.0
		ddH ₂ O

<u>Wash A</u>	50 % (v/v) 0.1 % (v/v)	Formamide Tween 20 2x SSC pH 7.0
<u>2x SSC</u>	0.05 % (v/v)	Tween 20 2x SSC pH 7.0
<u>PBS-T</u>	0.05 % (v/v)	Tween 20 1x PBS pH 7.0

2.9.7 FISH

<u>Wash A</u>	50 % (v/v) 30 mL 120 mL	Formamide pH 7.0 20x SSC ddH ₂ O
<u>Wash B</u>		0.5x SSC
<u>Wash C</u>	0.1 % (v/v)	Tween-20 4x SSC
<u>Blocking</u>	3 % (w/v)	BSA 4x SSC
<u>Detection</u>	1% (w/v)	BSA 4x SSC
<u>DAPI</u>	60 mL 45 µL	2x SSC, pH 7.0 DAPI
<u>DAPI-Wash</u>	0.05% (v/v)	Tween-20 2xSSC
Column buffer (Sephadex columns)	10 mM 1 mM 0.1 % (w/v)	Tris-HCl pH 8.0 EDTA pH 8.0 SDS ddH ₂ O

2.9.8 Standard solutions

<u>5x Loading buffer</u>	100 mM 30 % (v/v) 0.25 % (w/v)	EDTA Glycerine Bromophenolblue ddH ₂ O
--------------------------	--------------------------------------	--

2 Material & Methods

<u>5x TBE</u>	455 mM	Tris-Borate
	10 mM	EDTA
		ddH ₂ O
<u>TE</u>	20 mM	Tris-HCl, pH 8.0
	1 mM	EDTA, pH 8.0
		ddH ₂ O
<u>20x SSC</u>	0.3 M	NaCl
	0.03 M	Sodium citrate
		ddH ₂ O
<u>Hepes</u>	50 mM	Hepes
	150 mM	NaCl
	10 % (v/v)	Glycerine
	0.2 % (v/v)	Nonidet P-40
	20 mM	NaF
	1.5 mM	MgCl ₂
	0.2 mM	EDTA
		ddH ₂ O
<u>10x PBS</u>	137 mM	NaCl
	27 mM	KCl
	100 mM	NaH ₂ PO ₄
	17 mM	KH ₂ PO ₄
		ddH ₂ O
<u>RIPA</u>	150 mM	NaCl
	50 mM	Tris-HCl, pH 8.0
	1 mM	EGTA
	1 mM	EDTA
	0.25 % (w/v)	Na-deoxycholate
	1 % (v/v)	Nonidet P-40
	0.1 % (w/v)	SDS
		ddH ₂ O

2.10 Instruments

Table 2-11: Instruments

Instrument	Supplier
ABI PRISM 3100 Genetic Analyzer, 16 Capillary DNA Sequencer	PE Applied Biosystems, Langen
ABI PRISM 7900 Sequence Detection System	Applied Biosystems, Forster City, USA
Agilent DNA microarray scanner	Agilent Technologies, Palo Alto, USA
Automatic developing machine	Amersham, Freiburg
Axioplan microscope	Carl Zeiss, Jena
Biofuge Fresco refrigerated tabletop centrifuge	Heraeus/Kendro, Hanau
Centrifuge 5810 R	Eppendorf, Hamburg
Cytospin 3	Shandon Life Sciences International, England
Digital camera Hamamatsu ORCA-ER	Hamamatsu Photonics, Herrsching am Ammersee
EAS Gel Documentation System	Herolab, Wiesloch

Instrument	Supplier
EL800 Universal Microplate Reader	BIO-TEK, Neufahrn
Gel electrophoresis power supply	E-C Apparatus Corporation, USA
GeneAmp PCR-System 2400, 9600 and 9700	PE Applied Biosystems, Weiterstadt
Gene Pulser Xcell	Bio-Rad Laboratories, Munich
Gene Tac hybridization station	Genomic Solutions (GeneMachines), Ann Arbor, USA
GS 6, GS 6KR, Centrifuge	Beckmann, Wiesloch
GSA- and SS34-Rotor	DuPont, Boston, USA
Heating block QBT2	Grant Instruments/CLF, Emersacker
HMT 702 C Microwave oven	Robert Bosch GmbH, Stuttgart
Mastercycler PCR-Maschine	Eppendorf, Köln
Micro-centrifuge	NeoLab Laborbedarf, Heidelberg
Mini-Protean 3 gel and electrophoresis system	Bio-Rad Laboratories, Munich
Mini Trak CSS	Packard
Multifuge 3 SR	Heraeus/Kendro, Hanau
NanoDrop ND-1000 Spectrophotometer	NanoDrop Technologies, San Diego
Sigma 4K10-Coldcentrifuge	Sigma-Aldrich, Heidelberg
Sonopuls HD2070	Bandelin Electronic GmbH, Berlin
Thermomixer compact	Eppendorf, Köln
Ultrospec 2000 Photometer	Amersham, Freiburg
Unimax 1010 Shaker	Heidolph Instruments, Schwabach
Varifuge 3.0/3.0R Laborcentrifuge	Heraeus/Kendro, Hanau
Waterbad SW22	Julabo Labortechnik, Seelbach
Waterfiltrationsmaschine Millipore-Q UF Plus	Millipore, Eschborn
Weighing machine BL 610 und BL150S	Sartorius AG, Göttingen

2.11 Software

Table 2-12: Software

Software	Supplier
Adobe Illustrator CS	Adobe, USA
AnalysisIS	Olympus, Germany
Bioconductor 1.7	Open source (http://www.bioconductor.org/)
Endnote X	Thomson Reuters, New York, USA
Excel 2003	Microsoft, USA
Gene Pix Pro 6.0	Axon Instruments, Burlingame, USA
Ingenuity Pathway Analysis	Ingenuity Systems, San Diego, USA
Photoshop 7.0	Adobe, USA
Primer3	Open source (http://frodo.wi.mit.edu/)
R 2.2.1	Open source (http://www.r-project.org/)
SDS 2.2.2	Applied Biosystems, Forster City, USA
Sigma Plot	Systat System Inc., USA
Vector NTI	Invitrogen, USA
Word 2003	Microsoft, USA

Methods

2.12 Preparation of proteins

2.12.1 Isolation of peripheral blood lymphocytes from whole blood

Human lymphocytes can be isolated from peripheral blood by density centrifugation over a step gradient consisting of a mixture of the carbohydrate polymer Ficoll™ and the density iodine-containing compound metrizamide. This yields a population of mononuclear cells, lymphocytes, at the interface that has been depleted of erythrocytes and most polymorphonuclear leukocytes and monocytes. The lymphocytes were isolated from whole blood according to the manufacturer's instructions (Biochrom AG, Berlin).

2.12.2 Cell lysis for whole protein extraction

Cell lysates were prepared from cultured cells using RIPA-buffer, which was freshly supplemented with the Complete EDTA-free protease inhibitor cocktail-tablets (Roche Diagnostics, Mannheim) prior to lysis. Buffers were used at 4°C and 150 µL buffer per $5 \cdot 10^5$ - $1 \cdot 10^6$ cells were used for the cell lysis. Cells were lysed in 1x RIPA buffer for 30 min on ice and lysates were used for Western blot detection of proteins. After lysis, all cell extracts were centrifuged at 13.000 rpm for 10 min at 4°C. The supernatants were then mixed with 3x Laemmli loading buffer for acrylamid gel analysis, or stored at -80°C.

2.12.3 Isolation of nuclei from whole cell extracts

Nuclei extraction was used for two purposes, the first was to generate input-fractions for ChIP analyses and the second was to gain proper controls for ChIP-fractions in Western blot analyses. Nuclei extracts were prepared from cultured cells or frozen cell material using isotonic and lysis buffer (chapter 2.9.1). 1 mL isotonic buffer was added to $1 \cdot 10^8$ cells and incubated at 4°C for 5 min. Then these samples were centrifuged for 5 min at 13.000 rpm and 4°C. Afterwards the lysis buffer was added (1 mL/ $1 \cdot 10^8$ cells) following incubation at 4°C for 5 min. Subsequently, lysed nuclei extracts were centrifuged at 13.000 rpm for 5 min at 4°C again, and thereafter supernatants included nuclei for further use.

2.12.4 Acid extraction of histones

Another control tissue for use in Western blot analyses to control ChIP-fractions including precipitated modified histones, were histone proteins. Histone extracts were prepared from $1 \cdot 10^8$ cells, which were pelleted by centrifugation at 200 g for 10 min. The cell pellet was resuspended with 10-15 volumes of 1x PBS and then again centrifuged for 10 min at 200 g. After the PBS-washing step the cell-pellet was resuspended in 5-10 volumes of lysis buffer (chapter 2.9.4) and hydrochloric acid was added to a final concentration of 0.2 M. Following incubation for 30 min at 4°C, the cells were centrifuged for 10 min at 11.000 g and 4°C. The supernatant thus contained the acid soluble proteins, for example histones, whereas the acid-insoluble pellet was discarded. The histone-containing supernatant was then dialyzed twice against 200 mL of 0.1 M acetic acid for 1-2 h each. Afterwards the supernatant was dialyzed against 200 mL 1x PBS for 1 h, 3 h and over night, respectively. The yielded histone extracts could be quantified and lyophilized or stored at -70°C.

2.13 Immunoprecipitation of proteins

Cultivated cells were harvested and nuclei were extracted according to the protocol described in chapter 2.12.2. These nucleic extracts were then incubated with the precipitating antibody overnight at 4°C on a rotator. Afterwards, 10 % protein A sepharose (Amersham Bioscience, Freiburg) was added and precipitation of antibody-protein complexes was allowed for 2-4 h at 4°C while rotating. Precipitated sepharose-antibody-protein complexes were washed with 1x PBS and further analyzed by Western blot.

2.14 Chromatin immunoprecipitation (ChIP) assay

Two distinct standard approaches for the ChIP assay were tested prior to establishment of the adjusted protocol. A scheme of the general procedures during ChIP assays is shown in Figure 2-1:

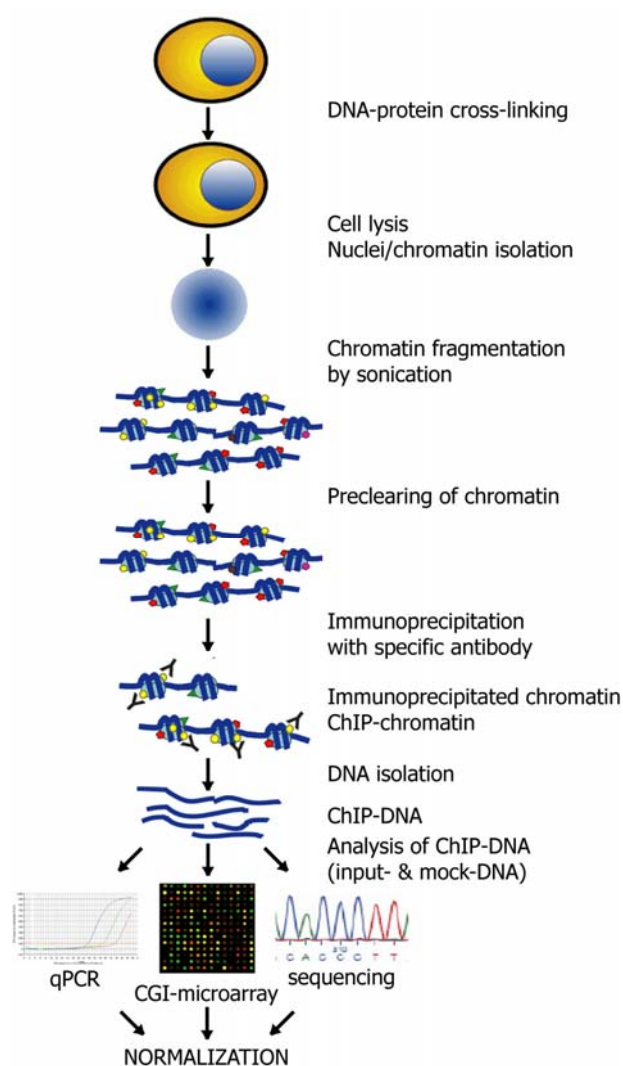


Figure 2-1: Schematic overview of the chromatin immunoprecipitation procedure; adapted from (Collas and Dahl, 2008).

A brief introduction to the characteristic features of native ChIP (nChIP) and of cross-linking ChIP (xChIP) will be given before description of the protocol established during the work of this thesis.

2.14.1 nChIP assay

The native ChIP or nChIP protocol includes preparation of native chromatin fragments by Micrococcal nuclease enzymatic digestion after cell lysis (O'Neill and Turner, 2003). Digestion with nuclease enzymes lead to a mixture of nucleosomal chains with different lengths. Micrococcal nuclease cleaves chromatin at linker DNA sequences between the nucleosomes and thus generates a mixture of mono- to heptanucleosomes. Digestion was checked by agarose gel electrophoresis, by which the nucleosomal chain mixture is visible as nucleosomal ladder according to separation by size of nucleosomal DNA in multiples of 146 bp on agarose gels.

The native ChIP protocol was applied according to the protocol established in the lab of Kinya Yoda and described in (Ando et al., 2002) with instructions by Alicia Alonso.

2.14.2 xChIP assay

The cross-linking ChIP protocol was originally established to efficiently precipitate proteins that are bound directly to the DNA by specific antibodies. Here, the instructions and buffers described in (Orlando, 2000 and; Spencer et al., 2003) were followed for xChIP-protocol and the stringent washing procedure including immune complex wash buffers, to ensure removal of unspecifically bound chromatin, was adapted in the self-developed ChIP protocol.

2.14.3 ChIP protocol established during this thesis

At first, 1×10^8 cells, which were either freshly prepared from cell cultures, from lymphocyte extractions from peripheral blood (PBL) or from cell pellets stored at -70°C , were resuspended in 1 mL 1x PBS. Reversible cross-linking of proteins and DNA was taken out by incubation with 2 % (v/v) formaldehyde solution at 4°C for 20 min. The cross-linking process was stopped by addition of 0.125 M glycine prior to centrifugation of cells for 5 min at 5.000 rpm and 4°C . Isotonic and lysis buffers used for the following steps were supplemented with diverse proteinase inhibitors (aprotinin, benzamidin, leupeptin, spermidine and spermine) and with DTT and PMSF, to avoid degradation of proteins and to stabilize DNA-protein bounds. The cross-linked cell pellets were resuspended in isotonic buffer and incubated for 5 min at 4°C . The buffer was discarded after centrifugation for 5 min at 5.000 rpm and 4°C and nuclei containing pellets were resuspended in 1 mL lysis buffer. Nuclei were ultra-sonicated in 5 pulses for 40 s at 70 % output each, with cooling of samples at 4°C between pulses. After ultra-sonication the fragmented chromatin was soluble in the lysis buffer. Therefore remaining nuclei and cell debris was pelleted and removed by two steps of centrifugation at 13.000 rpm and 4°C for 15 min each. The supernatant contained prepared chromatin-fragments and an aliquot of this "input"-fraction was set aside. The input chromatin was subjected to three steps of preclearing with 30 μL of protein A agarose/salmon sperm DNA beads at 4°C for 15 min each on a rotator to ensure equal mixing of the solutions and the beads. In between the preclearing steps the solutions were centrifuged at 5.000 rpm for 3 min at 4°C , pelleted preclearing beads were set aside and another volume of fresh beads was applied for the subsequent preclearing step. After this procedure, an aliquot of "cleared input" chromatin was set aside, and the prepared cleared input chromatin was used for precipitation. Chromatin prepared for each ChIP reaction was divided into two halves; one half served as "mock precipitate"-fraction to detect unspecific binding and the other half was combined with around 5 μg of the precipitating antibody. Both fractions were incubated overnight at 4°C on a rotator to allow binding of the antibody to target chromatin fragments. Afterwards, the antibody-chromatin complexes were precipitated by addition of 80 μL protein A agarose/salmon sperm DNA mix beads for 4 h at 4°C on a rotator. The same amount of beads was added to the mock precipitate fraction. After the incubation for 4 h, fractions were centrifuged for 5 min at 5.000 rpm and 4°C . The

supernatant of pelleted beads was saved as “unbound”-fractions for both, the antibody and the mock precipitate fractions, to check for unbound target chromatin. The pelleted bead fractions were washed stringently three times by addition of 100 μ L 1x High salt immune complex wash buffer, 100 μ L of 1x Low salt immune complex wash buffer and 100 μ L of 1x LiCl immune complex wash buffer, respectively. Centrifugation to pellet beads and remove wash buffer at 5000 rpm and 4°C for 3 min was performed after each addition of the respective buffer. The pelleted beads of precipitate and mock precipitate fractions were then resuspended in 150 μ L TE buffer. All aliquots taken during the ChIP procedure (fractions: input, cleared input, unbound, precipitate, mock unbound and mock precipitate) were split for separate analysis of protein and DNA results. The protein parts of all fractions were quantified in BCA-assays and the precipitation of target protein was specifically analyzed by Western Blot analysis using the same antibody as for precipitation.

The DNA-parts of the six ChIP-fractions were first incubated for 4-6 h at 65°C with 20 μ L 4 M NaCl to reverse cross-links, subsequently, remaining protein debris, protein A and antibody was destroyed by incubation with 3 μ L of proteinase K (20 mg/mL) and 3 μ L of 20 % SDS overnight at 56°C. DNA was extracted by phenol-chloroform extraction protocol described in chapter 2.16.1 and precipitated by addition of 2.5 volumes of absolute ethanol, 1/10 volume of 3 M NaAc (pH 5.2) and 1 μ g glycogen as carrier and incubation at -80°C for 30 min. DNA was then centrifuged for 20 min at 13.000 rpm and 4°C, washed with 70 % ethanol as described below and after drying, DNA was resuspended in 100 μ L of TE buffer. The DNA amounts in all fractions were quantified by spectrometric measurements (NanoDrop; see below) and diluted 1:1000 or 1:500 (for input and unbound fractions) or 1:3 (precipitate fractions) prior to analysis by qPCR.

2.15 Protein analysis

2.15.1 Protein quantification

The BCA-Assay is a Biuret reaction where peptide bonds react with copper ions in an alkaline solution. The reacted ions in turn react with bicinchonic acid (BCA) to give a violet colour which varies in intensity depending on the amount of protein present. Because peptide bonds occur with approximately the same frequency per gram of material in every protein, it can be used for estimation of protein concentration (Rehm, 2002). Standard curves were made by measuring BSA with standard amounts ranging from 2000 μ g/mL to 25 μ g/mL. The Biuret reaction was performed with copper-(II)-sulfate and bicinchonic acid (1:50 copper-(II)-sulfate:bicinchonic acid). 10 μ L of protein solution was added to 200 μ L of reaction mixture, incubated at 37°C for 30 min and then read out in an ELISA reader at 550 nm wavelengths.

2.15.2 Tris-glycine SDS-PAGE: Tricine-PAGE

SDS polyacrylamid gel electrophoresis (SDS-PAGE) is a method to separate proteins according to their size (Schagger and von Jagow, 1987). Proteins are coated with the detergent sodium-dodecyl-sulfate (SDS) that confers a linear shape and a negative charge to the proteins. Because of similar shape of the denatured proteins and the fact that the net charge depends only on the size of the proteins, the polypeptides can be separated according to their molecular weight. In PAGE, the matrix used to separate the proteins is a polyacrylamide gel, a polymer that gives a highly reproducible pore size upon polymerisation. In subsequent electrophoresis, the charged molecules migrate in an electric field. A discontinuous system was used, with a non-restrictive large pore gel stacked on top of the resolving gel. Samples were prepared by mixing with 3x Laemmli solution, boiling at 95°C for 5-10 min and cooling them before loading together with SeaBlue Plus2 pre-stained protein

size standard (Invitrogen, Karlsruhe). The gels for SDS-PAGE were prepared either with Tris-glycine or with tricine and run in SDS-PAGE running buffer or in tricine running buffer, respectively, at 100 mA for 1-2 h depending on the size of the proteins and the percentage of polyacrylamide in the gel (histones 1.5 h in 15 %-gels; macroH2A 2-3 h in 15 %-gels). For casting and running SDS-PAGE gels the Mini-Protean 3 gel system (BIORAD, Munich) was used.

2.15.3 Protein staining in SDS-PAGE gels

After separation with SDS-PAGE, the gels were either further processed for Western blot or stained with Coomassie brilliant blue R 250 (Merck, Darmstadt). During establishment of optimal Western blotting conditions to analyze ChIP-fractions, gels were analyzed after blotting by Coomassie staining. Gels were soaked in staining solution (40 % (v/v) Methanol + 10 % (v/v) glacial acetic acid + 50 % (v/v) H₂O + 0.05 % (w/v) Coomassie brilliant blue R 250). Unbound Coomassie was removed by immersion in 250 mL ddH₂O/gel and boiling in a microwave oven until staining of the gel was optimal for imaging.

2.15.4 Western blotting

Western blotting is a method to detect target proteins in a sample with specific antibodies, whereby proteins have been previously separated by SDS-PAGE. In Western blotting (Burnette, 1981; Renart et al., 1979) the separated proteins are transferred to a membrane where the proteins are better accessible for antibodies and can be further analyzed. After separating the samples by SDS-PAGE, the Western blot was performed wet; the assembled blot was embedded in the blotting buffer and run at 250 mA for 2-4 h on ice. To verify transfer of proteins, membranes were stained in Ponceau S (Sigma-Aldrich, Munich) for 5 min and de-stained by washing 5 min in TBS/Tween washing buffer. Membranes were blocked in blocking solution for 1 h at room temperature (RT) or overnight at 4°C. After blocking the primary antibody diluted in 1 % (w/v) BSA solution was applied (at concentrations depicted in Table 2-6) and incubated 1 h or overnight at 4°C. After incubation of the primary antibody, membranes were washed twice for 5 min and once for 10 min at RT in 1x TBS-Tween. The secondary antibody diluted in blocking solution was applied and incubated 1 h at 4°C. Membranes were washed as above and the blot was developed with either ECL or ECL⁺ Western blot detection kits (Amersham Bioscience, Freiburg) according to the instructions by the manufacturer.

2.16 DNA analysis

2.16.1 DNA isolation

DNA was isolated according to the standard phenol-chloroform extraction protocol, which is based on the immiscibility of phases produced by phenol, chloroform and isoamylalcohol. Homogenization of these liquids is only transient, and after centrifugation, three phases are produced: the upper aqueous phase contains all water-soluble molecules including nucleic acids, the lower fatty phase contains fatty-soluble proteins and the interphase captures proteins and molecules that are neither fatty nor water soluble. Thus a separation of nucleic acids from the cellular proteins is achieved and DNA can be isolated.

After the cellular digestion by proteinase K, cell pellets were combined with an equal volume of 25:24:1 phenol-chloroform-isoamylalcohol mixture (around 100 µL) and vortexed until homogenized. After a following centrifugation step at 13.000 rpm for 10 min the upper aqueous phase was taken and combined with an equal volume 24:1 chloroform-isoamylalcohol mixture (around 200 µL) for higher purity and further depletion of phenol.

After homogenization and centrifugation for 10 min at 13.000 rpm the aqueous upper phase was taken and combined with 2.5 volumes of ethanol, 1/10 volume of 3 M NaAc (pH 5.2) and 1 µg glycogen as a carrier. This was homogenized and DNA was allowed to precipitate at -80°C for 30 min. The precipitated DNA was pelleted by centrifugation for 20 min at 13.000 rpm and 4°C, and then the pellet was depleted of salts by washing with 200 µL 70 % ethanol. After air-drying of the DNA-pellet, it was resuspended in 30 µL ddH₂O or TE-buffer.

2.16.2 Genomic DNA preparation

Isolation of genomic DNA from cultivated cell lines was prepared according to the phenol-chloroform extraction protocol described above. Alternatively, genomic DNA was prepared according to the manufacturer's protocol for genomic DNA isolation with columns (Qiaagen genomic tip 500, Qiaagen, Hilden).

2.16.3 DNA quantification

The concentration of DNA solutions was determined by spectrophotometry. This procedure is based on the absorption maximum of nucleic acids at 260 nm, which is caused by the aromatic ring structures of the nucleotide bases. Standardized, the absorption is usually measured in quartz cuvettes with a gage of 1 cm, which facilitates estimation of concentrations via the law of Lambert-Beer:

$$A = \epsilon * c * d$$

Following calibration with the appropriate solvent, sample readings should range between 0.05 and 1 OD (optical density), which is the linear range of common photometric devices. The concentration was then determined using the following formula:

$$[\text{dsDNA}] = \text{OD}_{260} * 50 \mu\text{g/mL} * \text{dilution factor}$$

The absorption maximum of proteins is at 280 nm due to the absorbance of aromatic amino acids. The ratio $\text{OD}_{260}/\text{OD}_{280}$ can be used as an indication for the purity of nucleic acid solutions. Pure DNA solutions have a ratio of about 1.8, whereas contaminations by phenol or proteins cause a decrease of this value. DNA solutions with a ratio smaller than 1.5 were considered inappropriate for further analysis.

The OD of DNA solutions was measured in an automated spectrometric reader (NanoDrop).

2.16.4 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA-fragments in an electric field according to their size. The used concentration of agarose was dependent on the size of DNA fragments. Usually, 1 % (w/v) agarose gels were used for separating DNA-fragments of 0.5-4 kb in size. Agarose was melted by boiling in 1x TBE buffer. The solution was filled into a gel sled and was allowed to polymerize. For electrophoresis, the agarose-gel was placed into a gel chamber and covered with 1x TBE buffer. Probes were mixed with loading buffer and loaded into the slits of the gel. Electrophoresis was run for 0.5-1 h at 120-150 V and 500 mA. Afterwards, the separated DNA was stained by soaking the gel for 10 min in 1 µg/mL ethidium bromide at RT. Images were taken with a digital camera device that is part of the Gel Documentation System.

2.16.5 PCR purification and gel extraction purification

PCR products and plasmids were purified to remove enzymes, salts and other contaminants. Nucleic acids were purified with a PCR purification kit following the manufacturer's protocol (PCR rapid purification kit, Marligen Biosciences, Ijamsville, USA) and usually eluted in 30 μ L TE-buffer or ddH₂O. In order to isolate nucleic acids of distinct sizes, the bands were cut out from agarose gels after electrophoresis. To avoid DNA damages of the DNA fragments by gel extraction purification, nucleic acids were detected by a low wavelength UV-light. Gel slices were processed with the Concert Gel Extraction Kit following the manufacturer's instructions (Marligen Biosciences, Ijamsville, USA).

2.16.6 Primer design

For the design of oligonucleotide primers, the conformation of primer dimers was generally avoided by considering that the 3' end of the forward primer and the 3' end of the reverse primer are not able to hybridize by homologous base pairing. The five first and five last residues of each primer should contain 2-3 G- or C-residues; if possible the last residue at the 3' end should also be a G or a C to ensure a stable starting position for the enzyme. To avoid wrong primer annealing, stretches of more than four residues of the same type were prevented. Primers were designed on the basis of genomic sequences of the genes and CpG-island stretches of the chromosomal region 13q14.3, which were available on the NCBI homepage (<http://www.ncbi.nlm.nih.gov/>). The open source software Primer3 was used to design primers to be used in quantitative PCR analyses, as these PCR-products should not exceed a size of maximal 200-250 bp.

2.16.7 PCR

Chromatin immunoprecipitation analyses were performed to enrich DNA-sequences that are specific for a distinct chromatin mark. For example, in ChIP-assays using the antibody anti-H3K4me2 only DNA sequences bound only to these modified histones H3 were enriched in precipitate fractions. Characterization of DNA sequences could be carried out by microarray- or PCR-analyses. Analyses of ChIP-DNA with microarrays ("ChIP on chip") generated a genome wide profile of H3K4me2-enrichment. Another possibility to analyze ChIP-DNA was to amplify specific sequences of interest by PCR to identify those that show the modification analyzed. The latter analysis was only possible because a known chromosomal region (13q14.3) should be characterized for epigenetic modifications. In PCR analyses, tested ChIP-fractions could be divided into controls that should contain the positive control sequence (input-, cleared input- and unbound-fractions) and those controls that should not contain any positive control sequence (preclearing beads; mock precipitate).

The polymerase chain reaction (PCR) is a technique to amplify distinct DNA fragments *in vitro* (Mullis et al., 1986). It is based on the ability of DNA-polymerases to generate complementary strands of DNA out of a single-stranded DNA template. The enzymes used can elongate existing short oligonucleotide primers of 15-25 bp in length. The complementarity of the primers to a specific DNA sequence determines the fragment of DNA that will be amplified. A standard PCR reaction contains:

100 ng	template (DNA/cDNA)
10 pmol	forward primer
10 pmol	reverse primer
4 μ L	dNTP-mix (10 mM of each dNTP)
5 μ L	10x polymerase buffer
1 μ L	DNA polymerase (1 U/ μ L)
ad 50 μ L	ddH ₂ O

Standard cycling conditions used for amplification were 96°C for 5 min; 30 cycles of 96°C for 20 s, 60°C for 20 s and 72°C for 1 min; 72°C for 5 min, hold at 4°C. These standard conditions were modified if DNA fragments were only weakly amplified or showed no product at all.

2.16.8 Nested PCR, gradient PCR, colony PCR

If no DNA product is amplified after the first PCR reaction, a second primer set lying within the first amplicon, was used to nest the first PCR. Usually after optimization of the hybridization temperature, a specific target product was visible on the agarose gel.

For the optimization of hybridization temperatures a gradient PCR reaction was used. A central temperature together with a temperature gradient was chosen to perform the single reactions. In case the temperature was chosen too high, there was no PCR product because the primers are not able to anneal to the target sequence. In case the temperature was too low, a DNA smear or several bands were visible, due to unspecific binding of the used primer pair.

Colony PCR reactions were used to quickly screen for plasmid inserts of *E. coli* colonies. This technique can be used to determine insert size and/or orientation in the vector depending on the choice of primers. PCR reaction mixtures consisted of 5 µL 10x polymerase-buffer, 2 µL dNTP-mix (2.5 mM of each dNTP), 0.1 µL forward and reverse primer (100 µM each), 0.5 µL of Taq-polymerase (2 U/µL) and 42.3 µL of sterile distilled water in a total volume of 50 µL. To each PCR tube containing the PCR reaction a small amount of colony was added by a fine yellow pipette tip. The amount of cells should be small, as just a touch will do it. Thermal cycling conditions were 96°C for 4 min; 30 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 1 min; 72°C for 5 min, hold at 4°C.

2.16.9 Reverse transcription

Reverse transcription is the transcription of RNA into single-stranded complementary DNA (cDNA). cDNA is much more stable than RNA and can be used for a number of approaches like PCR, real-time quantitative PCR or sequencing. For a final volume of 20 µL cDNA solution, 1 µg of total RNA (8 µL) was used together with 2 µL of 5x 1st strand buffer and 1 µL DNase I (10 U/µL) and incubated for 20 min at RT to allow digestion of DNA. 3 µL of the mastermix I, which contains in equal amounts EDTA (25 mM), dNTP mix (10 mM each) and 300 ng/µL dT primer or random hexamers, were added to the mix and incubated in a thermo-cycler for 10 min at 65°C and for 10 min at 25°C to get rid of secondary structures and allow the annealing of the primer to the RNA, respectively. After adding 5 µL of the second mastermix containing 5x 1st strand buffer, 0.1 M DTT and nuclease-free water at a ratio of 2:2:1, the solution is shifted to 42°C for 2 min before adding 1 µL of reverse transcriptase (200 U/µL) and 0.2 µL of T4 gene 32 protein. After the reverse transcription reaction for 50 min at 42°C, the enzyme was inactivated for 10 min at 95°C.

2.16.10 Real-time PCR analysis, quantitative PCR

Real-time PCR analysis was used in this study to quantify DNA of the fractions generated during ChIP-experiments. It was used to quantify the precipitated target sequences in different samples and to quantify the amount of precipitated sequences at different targets.

The method of real-time PCR analysis is based on the detection of bound fluorescent dye SYBR-Green I to the minor groove of double stranded DNA. In the unbound state, SYBR-Green I does not emit any fluorescence, but during the elongation step of the PCR, it can bind to the double stranded PCR products, whereby it is activated and emits fluorescence. At the end of every elongation step, the emitted fluorescence was measured, which allows the detection and quantification of PCR products in real time. To get a conclusion about the

specificity and sensitivity of the primer used in this reaction, the analysis of dissociation curves is indispensable. In these dissociation curves side products and primer dimers would be visible as peaks with a low melting temperature.

PCR reaction mixtures consisted of 6 μL SYBR-Green I mixture, 100 nM forward and 100 nM reverse primer and 2 μL template in a reaction volume of 12 μL . Thermal cycling conditions comprised an initial enzyme activating step of 15 min at 95°C, 40 cycles of 15 s at 95°C and 1 min variable annealing/elongation temperature (usually 60°C), depending on the respective set of primers. Double-stranded-DNA-specific fluorescence was measured at the end of each extension phase. For creation of the dissociation curve data, the mix was hold for 15 s at 95°C for denaturation, 15 s at 60°C and was then slowly (within 20 min) heated to 95°C, hold for 15 s and cooled down to 4°C. Product-specific amplification was confirmed by analysis of dissociation curves. A standard curve consisting of a dilution series of genomic DNA was used to determine the efficiency of the PCR reaction. The amount of PCR product was plotted against the cycle number in a logarithmic scale. This plot shows at which cycle number in the exponential phase a defined threshold is reached by the amplification curves of the samples. The intersection of the amplification curve with the defined threshold gives a defined C_T value for each sample.

2.16.11 Quantification and normalization of ChIP-DNA by qPCR

To determine the amount of precipitated DNA at a specific target sequence by qPCR, the measured C_T -values of different ChIP-fractions were taken to calculate fraction of input values. Fraction of input values were calculated from the C_T -values determined for input, precipitate and mock-precipitate fractions by this equation:

$$\text{Fraction of input} = (C_T(\text{precipitate}) - C_T(\text{mock})) / C_T(\text{input})$$

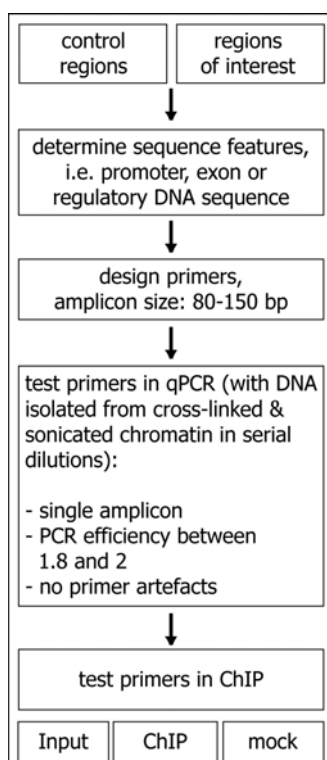


Figure 2-2: Outline for setting up qPCR to analyze ChIP DNA.

To validate H3K4me2-precipitation of ChIP by qPCR, the amount of DNA was measured at the promoter of the housekeeping gene *GAPDH* and at the promoter of the inactive gene *MYH1* (i.e. *MYH1* is inactive in hematopoietic cells).

QPCR for *GAPDH*-promoter served as positive control, because it is an H3K4me2-enriched-sequence, which should be frequently contained in precipitate-fractions. QPCR for *MYH1*-promoter served as negative control as it shows no H3K4me2-enrichment. The same principle was used to validate macroH2A-precipitation, using the promoter of the gene *IL8* as positive control and again the *MYH1*-promoter as negative control.

Further normalization of fraction of input values was carried out by dividing the fraction of input value of each analyzed 13q14-sequence with the fraction of input value of a second positive control sequence, both measured in parallel for every analyzed sample. The second positive control sequence was localized in the *CDH12* promoter for H3K4me2-ChIPs and in the *HK2*-promoter for macroH2A-ChIPs. Figure 2-2 schematically depicts the general outline for the set up of qPCR analysis for ChIP assays.

The analysis of 39 sequences of 13q14 in each of 27 samples was partially automated. Three template plates in 96-well format containing the ChIP-fractions of 27 samples were produced in order to automatically pipet the templates along with the PCR-reaction reagents (pre-loaded in 96-well plates) into a 384-well plate with a pipetting roboter.

2.16.12 DNA-sequencing

Sequencing of DNA was done according to the Dideoxy-Sequencing method (Sanger et al., 1977). This method is a direct, statistically distributed termination of the complementary strand synthesis during the cyclic sequencing reaction, in which the amplification takes part in a linear matter with only one primer. The reaction mix contains fluorescently labelled dideoxy-derivates for all of the four nucleotides (Rosenblum et al., 1997). The detection of fragments occurs via laser excitation. Due to initial denaturation steps during the reaction, sequencing of dsDNA is possible. The sequencing reaction was done according to the manufacturer's protocol using the PRISM Big DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Darmstadt). The reaction contains an initial denaturation step for 2 min at 96°C and 25 cycles of 5 s at 96°C, 10 s at 55°C and 4 min at 60°C. After this PCR-reaction, samples were precipitated with 2.5 volumes of ethanol and 1/10 volume of 3 M NaAc, washed with 70 % ethanol, dried and resuspended in 10 µL formamide. Subsequently, this DNA was used for sequencing in an automatic capillary sequencing machine.

2.16.13 Detection of monoallelic expression by SNP-analysis

The SNP-analysis was performed in collaboration with Angela Philippen. First, SNPs within the open reading frame (ORF) of genes or within CpG-islands were identified using the corresponding web-pages of the National Center of Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/>). To test monoallelic expression of genes or monoallelic histone modifications at CpG-islands, healthy probands that are heterozygous for one or more of these SNPs were identified: In a PCR reaction, the sequence containing the SNP was amplified using genomic DNA of healthy probands as template (see 2.16.7 and 2.16.8.). Subsequently, the DNA-amplicon was sequenced (see 2.16.12), whereby heterozygosity of probands could be detected by double peaks at the SNP location. For analysis of monoallelic expression of *C13ORF1*, B- and T-cells were isolated from probands showing these double-peaks in their SNP-sequences. These B- and T-cells were used for reverse transcription of RNA into cDNA (see 2.16.9), amplification of the SNPs by PCR and subsequent sequence analysis.

2.16.14 Cloning of PCR-fragments into TOPO TA vectors

PCR amplification of CpG-islands is limited and amplification efficiency varies in each single reaction. Therefore, the 13q14-CpG-islands were divided into 1 kb parts, and primers were designed to amplify these parts in PCR reactions. These 1 kb-PCR-fragments of the CpG-islands C, D and E subsequently were cloned into TOPO TA-vectors for stable reamplification. The cloned PCR-fragments of the three CpG-islands were also included in a CpG-island array produced by B. Radelwimmer (DKFZ, Division Molecular Genetics).

TOPO TA cloning is a patented technology from Invitrogen (Invitrogen, Karlsruhe) for inserting PCR products directly into a TOPO vector. These vectors have topoisomerases covalently linked at their ends that twist the insert-DNA into the vector within minutes. TOPO TA cloning was used to directly clone the gene-specific PCR amplified products into vectors for sequencing and for re-amplification. Typically, 6 μL of cloning reaction consisted of 2 μL of purified PCR product, 1 μL of 1.2 M NaCl; 0.06 M MgCl_2 , 0.5 μL of TOPO TA-vector and 2.5 μL H_2O . Cloning reactions were incubated at RT for 20 min. Afterwards, 3 μL of the completed cloning reactions were used to transform one vial of TOP10 chemically competent cells (Invitrogen).

2.17 BioCOBRA

DNA-methylation at 13q14.3 was determined by BioCOBRA, a modified protocol for Combined Bisulfite Restriction Analysis (COBRA) that incorporates an electrophoresis step in microfluidics chips. All BioCOBRA experiments discussed in this thesis were carried out by Angela Philippen (University Hospital Ulm, AG "Mechanism of Leukemogenesis") and data presented here is courtesy of A. Philippen.

In order to support or exclude the DNA-methylation data obtained by aPRIMES (see below) data generated by BioCOBRA was included in the results and discussion chapters. The BioCOBRA experiments were performed as described in (Brena et al., 2006) using the REPLI-g Kit, the QIAamp DNA Mini Kit (Qiagen, Hilden) and SSI methylation.

2.18 Fluorescence in situ hybridization (FISH)

Fluorescent *in situ* hybridization (FISH) is a cytogenetic technique that is used to detect and localize the presence or absence of specific DNA sequences on chromosomes. The fluorescent probes bind to those parts of the chromosome to which they show a high degree of sequence similarity. Afterwards, fluorescence microscopy is used for detection of probe hybridization to chromosomes. The probes used in FISH are mostly derived from DNA-fragments that are cloned as part of bacterial artificial chromosomes (BACs). Short fragments from these BACs are amplified by PCR and further translated into FISH-probes. Construction of probes needs to fulfil several requirements; which include that the probe has to be long enough to hybridize specifically to its target and not to any similar sequence; while it should not be too large to impede the hybridization process, furthermore it has to be tagged directly with fluorophores that can be targeted by antibodies or biotin (Lichter and Ried, 1994).

For preparation of probes, nick translation and PCR with labelled nucleotides was used. Then, an interphase or metaphase chromosome preparation is produced. The chromosomes are firmly attached to glass slides and repetitive DNA sequences are blocked by addition of short DNA-sequences (CotI DNA) to the sample. The probe is applied to the chromosomal DNA on the slides and incubated for ~ 12 hours while hybridizing. Several wash steps remove all unhybridized and partially hybridized probes. The results are then visualized and quantified using a microscope that is capable of exciting the dye and recording images.

2.18.1 Preparation of FISH probes

The conventional method for generation of labelled DNA probes is the so-called nick-translation. DNA is incubated with a limited amount of DNase I, which introduces on the one hand single strand breaks (i.e. "nicks") into the DNA. On the other hand, digestion with DNase I leads to delimited double strand breaks, which in turn generates short DNA fragments. The enzyme DNase I generates free 3'- and 5'-ends of DNA, which are necessary for DNA polymerase action. The latter enzyme is also included in a nick translation reaction, because it mediates the incorporation of labelled nucleotides into the DNA sequence. In nick translations, template DNA sequences are used to generate short labelled DNA probes for FISH, carried out by the combined action of DNase I and DNA polymerase.

DNase I, polymerase digestion and labelling of template DNA was allowed for at least 1 h at 15°C in a 100 µL reaction volume including 10 µL NT-buffer, 10 µL of the fluorophores, 10 µL of 0.1x Mercaptoethanol, 4 µL of DNA-polymerase (1 U/µL) and 4 µL of 1:4000 diluted DNase I (10 U/µL). To control digestion efficiency, the DNA was denatured for 2 min at 100°C and analyzed by agarose gel electrophoresis. The DNA-fragments can be further used as probes if they are digested to a size of around 300 to 600 bp. DNase I was inactivated by addition of 2 µL 0.5 M EDTA and 1 µL 10 % SDS and incubation at 68°C for 10 min. Nonetheless, the remaining enzymes and buffers had to be removed and DNA probes were isolated by custom prepared sephadex-columns. These columns were filled with glass wool soaked with sephadex buffer, which is prepared by 30 g of sephadex in 500 mL 1x column buffer. The probes were applied to the prepared sephadex columns and centrifuged. Only DNA-fragments can move through the packed sephadex glass wool within the columns and thus DNA was separated from the enzymes. The FISH-probes were transferred to fresh tubes and stored at -20°C until further use. For FISH experiments in this work a centromeric X-chromosome probe was prepared and labelled with the fluorophore Rhodamine and a centromeric Y-chromosome probe was labelled with FITC.

2.18.2 Preparation of interphase chromosomes

Around 5 to 10 million cells were extracted from cell cultures and remaining cell culture medium was removed by washing in 1x PBS. After centrifugation at 1200 rpm at RT for 10 min, cell pellets were resuspended in 1 mL 1x PBS. 300 µL of these suspensions were applied to poly-L-lysine coated glass slide by a Cytospin3 centrifuge (Shandon). Cells were subjected by 5 min Cytospin3-centrifugation at 500 rpm at RT and pre-fixed by incubation in dehydrating ethanol dilution series. The interphase chromosomes were incubated in 70 %, 90 % and absolute ethanol for 5 min each and stored at -70°C.

2.18.3 Preparation of metaphase chromosomes

In order to prepare metaphase chromosomes, cultured cell lines are arrested in cell cycle by addition of 0.6 % (v/v) colcemid and incubation for 90 min at 37°C and 5 % CO₂. Colcemid was removed by gentle centrifugation at 1010 rpm and discarding of the supernatant. The cell pellets were carefully resuspended in 12 mL pre-warmed 1x hypotonic buffer and incubated at 37°C for 30 min. The hypotonic buffer was removed by gentle centrifugation at 860 rpm at RT for 10 min, the supernatant was discarded and pellets were carefully resuspended in 2 mL of fixative, which was freshly prepared by combination of methanol and acetic acid in a 3:1 ratio. These prepared metaphase chromosomes were incubated for at least 30 min to a maximum of 4 h at -20°C followed by gentle centrifugation at 860 rpm and RT for 5 min. Pelleted metaphase chromosomes were washed with fixative and centrifuged as described above for three times before they were finally resuspended in around 2 mL of the fixative and brought onto glass slides in single drops coming from a disposable pipette. The glass slides were prepared prior to metaphase dropping by incubation in absolute

ethanol and subsequent washing in ddH₂O to prepare the surface in such that metaphase chromosomes are enabled to spread on the surface of the glass slides.

2.18.4 FISH protocol

The pre-fixed metaphase or interphase chromosome slides were dehydrated by incubation in ethanol dilutions, starting with 70 % ethanol, to 90 % ethanol to absolute ethanol for 5 min each at RT. Pepsin digestion was taken out by incubation of slides in 0.01 N HCl including 1 mg/mL pepsin at 37°C for 3 min. Residual pepsin or HCl was washed away by incubation in 1x PBS for 10 min. Subsequently, slides were post-fixed in 1x PBS containing 10 % (v/v) PFA for 5 min at 4°C and washed again in 1x PBS for 10 min. Another step of dehydration in three ethanol dilutions at RT as described above was followed by air-drying of slides for 5-10 min and final incubation at 63°C for 20 min. In order to denature the DNA fixed on the slides, they were incubated in 49 mL formamide, 7 mL 20x SSC and 14 mL ddH₂O at pH 7.0 for 2 min and 10 s at 65°C. Prior to application of the respective probes, a third dehydrating dilution series with 70 %, 90 % and absolute ethanol for 5 min in each dilution at 4°C was carried out. Probes were applied onto the slides and both were incubated overnight at 37°C in a humid chamber to allow hybridization.

Centromeric probes for X and Y chromosomes were precipitated together with 3 µL human CotI DNA and 7 µL salmon sperm DNA in 1/20 volume 3 M NaAc and 2.5 volumes absolute ethanol for 30 min at -80°C. The probe was precipitated by centrifugation at 13.000 rpm and 4°C for 20 min, washed with 70 % ethanol and finally dried for 10-15 min at 37°C. In order to denature the probe-DNA, de-ionized formamide was added as well as HybMix buffer and denaturing was enabled by incubation at 75°C for 6 min. After short incubation at 4°C the denatured single-stranded probe DNA was applied to the denatured single-stranded DNA on the slide and hybridization was allowed at 37°C overnight.

Hybridized slides were incubated in Wash A buffer at 42°C for 10 min and three times for 5 min, whereby after each incubation step the buffer was exchanged. Afterwards, the hybridized slides were incubated three times for 5 min at 42°C in pre-warmed Wash B buffer. Prior to probe-detection the single stranded DNA on the slides was blocked in blocking buffer for 30 min at 37°C in a humid chamber. Subsequently, probe hybridization was detected by addition of detection buffer, of the dyes FITC coupled to avidin and antiDig-Rhodamin and by incubation for 30 min 37°C in the humid chamber. Unbound dyes and detection buffers were removed with Wash C buffer added three times for 5 min at 42°C. The nuclei were dyed by DAPI, which was contained in the used mounting medium (Vectashield). The results were then visualized and quantified using a fluorescence detecting microscope that was capable of exciting the dyes, and images were recorded.

2.19 Microarray based methods

2.19.1 The CpG-island microarray includes 13q14-CGI clone set

The CpG-island microarray was designed and constructed by Bernhard Radlwimmer and colleagues and included in addition to the described clone library, clones from CpG-islands of chromosomal arms 13q and 17p. For the construction of the CpG-island-(CGI)-microarray a human library consisting of 10,560 bacterial clones harboring DNA sequences enriched for CGIs was purchased from the UK HGMP (<http://www.hgmp.mrc.ac.uk/>). This clone set is part of the original CGI-library, which was generated by Cross et al. (1994). To amplify the CGI sequences, library aliquots were grown in 1 mL LB-media including ampicillin (50 mg/mL) overnight at 37°C. Subsequently, PCRs were performed in a 96-well format (MWG, Ebersberg, Germany) using 2 mL of bacterial culture supplemented with 10 % DMSO, 180 mM dNTPs, 150 nM of each primer, 1.8 mM MgCl₂ and 2 U Eurotaq Polymerase (Biotac,

Heidelberg, Germany). Thirty-five amplification cycles were performed including denaturing at 94°C for 30 s, annealing at 58°C for 30 s and elongation at 72°C for 60 s. The primers used for amplification were pGEM-forward and pGEM-rev. All CpG-island-sequences were reamplified by using 5 mL of product of the first PCR as a template for a second PCR under identical conditions, because the second PCR resulted in more homogenous product quantities for microarray printing (Pfister et al., 2007).

The CpG-island-clones of 13q14.3 were generated by PCR-amplification specific for 1 kb sized parts of the three CpG-islands C, D and E in the region. The CGIs were of variable sizes, which resulted in different numbers of fragments: six parts for CGI C; eight parts for CGI D and 12 parts for CGI E; all parts were of a size of about 1 kb (range: 900 bp to 1.3 kb). These PCR-generated 1 kb-CGI-parts of 13q14 had to be cloned into TOPO TA vectors to be able to re-amplify them by PCR, which was difficult when using genomic DNA. The CGI-parts were amplified under the same PCR conditions described above to generate the clones that were finally printed onto the microarray. These CGI-microarrays that included the clones for the 13q14-CGIs C, D and E were used together with hybridization conditions of the matrixCGH protocol (see below) for all ChIP-on-chip analysis.

The CGI-microarray production itself was carried out by B. Radlwimmer and colleagues. Therefore, only a brief description of the printing procedure should be included here: For microarray printing, 30 mL of PCR products was dried in a vacuum centrifuge and resuspended in 12 mL spotting buffer (3x SSC, 1.5 M betaine). The PCR products were printed onto amino-silane-coated Corning™ Gaps II slides (Corning, Acton, USA) in triplicates using a printing robot (OmniGrid, GeneMachines, San Carlos, USA) and 48 (4x12 configuration) Telechem SMP3 pins (Telechem International, Sunnyvale, USA) at 20°C and 40 % humidity. After printing, slides were UV cross-linked and baked for 2 h at 80°C (Pfister et al., 2007).

2.19.2 ChIP on chip - (input vs precipitate)

ChIP-on-chip differs from ChIP and qPCR only in the method of analysis of ChIP-DNA (see also Figure 2-3). ChIP-DNA is eluted after reversal of cross-links and the 5'- and 3'-ends are repaired with DNA polymerase to generate blunt ends. A linker is applied to each DNA fragment to enable PCR amplification. Then, a fluorescent label (Cy3) is incorporated during PCR amplification of ChIP precipitate-fractions. Similarly, an aliquot of input DNA is labelled with a second fluorophore, Cy5. The two samples are mixed and hybridized onto the CGI-microarray. In this dual-color approach, enrichment of a specific histone modification is established when intensity of precipitate-DNA significantly exceeds that of the input-DNA on the array. Statistical analysis software and personal evaluation determine the significance of enrichment of the precipitated modification to a specific CGI-regulated gene/motif in the region examined.

The first step in ChIP-on-chip analysis was the labelling of the ChIP-DNA according to the matrixCGH-protocol. For the analysis on the CGI-microarray the ChIP-fractions input and precipitate of one ChIP-experiment were hybridized onto the CGI-microarray. The labelling of genomic DNA-fragments was carried out in a three step protocol. At first, a reaction volume of 75 µL per fraction/sample including 500 ng ChIP-DNA, 30 µL 2.5x random primer and 22 µL 5 M betaine (ad 75 µL ddH₂O), which was denatured for 10 min at 95°C and incubated at 4°C for 2 min. Then, 7.5 µL 10x dNTPs were added, whereby of dATP, dGTP and dTTP 20 µL of 100 mM were included, but only 10 µL of 100 mM dCTP, because additional 0.75 µL of the dye coupled to dCTP was added. The dye Cy3 was used for the precipitate-fraction and the dye Cy5 for the input-fraction. Addition of 0.75 µL exo-Klenow (1 U/µL), the large fragment of the *E. coli* DNA polymerase I and incubation for 12-16 h at 37°C lead to incorporation of the dyes into the amplified DNA. This reaction generally yielded 10-12 µg labelled genomic DNA fragments. After labeling, equal amounts of Cy3 labeled precipitate-DNA and Cy5-labeled input-DNA together with 180 µg of CotI-DNA, which blocks repetitive

sequences, were combined and precipitated. Precipitation of DNA was enabled by addition of 1/10 volume of 3 M NaAc (pH 5.2) and 2.5 volumes ethanol and incubation at -80°C for 15-20 min. The labeled DNA was precipitated by centrifugation at 13.000 rpm and 4°C for 30 min. Afterwards the DNA was dried and used for hybridization.

For hybridization the labelled DNA samples have to be solved in UltraHyb buffer, of which an aliquot was heated to 70°C whilst precipitation of the labeled DNA. 130 µL of UltraHyb buffer was added to each DNA-sample and vortexed. To resuspend the DNA pellets in UltraHyb buffer, the mixture was incubated for 30 min at 42°C and 500 rpm on a thermo-mixer and vortexed every 10 min. Meanwhile, the CGI-microarray was blocked for 15 min at 50°C in matrixCGH blocking buffer. The slides were washed twice in ddH₂O at RT and dried in 50-mL-Falcon tubes by centrifugation at 1000 rpm for 3 min. After complete resuspension of the DNA-pellet in UltraHyb buffer, the samples were denatured at 75°C for 10 min and re-annealing was allowed by cooling samples down to 42°C prior to hybridization. Subsequently, the DNA mixtures were applied to the microarrays mounted in a GeneTac Hybridization Station. The hybridizations were performed for a minimum of 36 h at 42°C in the hybridization station. Thereafter, the microarrays were automatically washed at 37°C with 2x SSC including 0.05 % (v/v) Tween-20 for 5 min, with Wash A buffer, which includes 50 % (v/v) formamide, 2x SSC and 0.1 % (v/v) Tween-20, for 5 min, and again with 2x SSC including 0.05 % (v/v) Tween-20 for 5 min. Immediately after the final washing step, the slides were unmounted and transferred to a glass cuvette containing 2x SSC at 43°C. Slides were incubated on a shaker at 80 rpm for 10 min under light-protection. Subsequently, slides were transferred to 1x PBS and incubated for 10 min at RT on a shaker at 20 rpm. Finally, slides were dried by centrifugation in 50-mL-Falcon tubes for 3 min at 1000 rpm and immediately scanned.

The microarrays were scanned with an Agilent DNA microarray scanner (Agilent Technologies, Palo Alto, USA) and images were analyzed using GenePix Pro 6.0 software (Axon Instruments, Burlingame, USA). Global normalization was performed in three steps. First, a normalization for print-order effects using the loess smoother, as described in (Smyth et al., 2003) was applied. In the second step, within-array normalization was done to normalize the log₂-ratios for each array separately by the use of print-tip loess (Smyth et al., 2003). Third, to normalize between arrays, quantile normalization was performed. All these normalization steps were carried out using zero weights for previously filtered spots and according to the instructions of Stefan Pfister as was previously described in (Pfister et al., 2007).

Data analysis by Ingenuity software

For analysis of 13q14-data, resulting from ChIP-on-chip experiments with the CGI-microarray, the web-based software Ingenuity (www.ingenuity.com) was used, which detects known pathways and networks the genes regulated by CGI-clones are taking part in. These pathways were used as quality control for ChIP-on-chip, to evaluate if their activation or inactivation, implied by H3K4me2 or macroH2A-enrichment, respectively, made sense in the hematopoietic samples. For genes with described functions the software can analyze molecular interactions to other entities, shows functional and disease relationships or the occurrence within a canonical pathway. Deregulated genes had to fulfill the following criteria to be analyzed by Ingenuity software: The probability to be differentially marked ("expressed" within the software) had to be at least 98 %. In addition, the deregulation had to be at least 1.5-fold. For the analysis RefSeq numbers were extracted from the microarray clone data base within the UCSC-browser (<http://data.microarrays.ca/cpg/index.htm>) for accordant clones. In case no RefSeq number was available, the GenBank number was used for identifying genes. Uploaded IDs are then mapped to an Entrez/Gene based database. This allows the relation of any uploaded ID to a gene/protein entity in the Knowledge Base (KB). Mapped IDs are then queried for molecular interactions to other entities within the KB.

In case of interactions to other entities, these networks are displayed in sketches. The generated pathways to which most IDs map are ranked by the software. For ChIP-on-chip experiments measuring an active chromatin mark, i.e. H3K4me2 enrichment, pathways associated with the hematopoietic system that were ranked among the top five pathways were considered as validation for the experiment.

To further validate ChIP-on-chip experiments there was no direct and simple control in hands, instead literature databases were searched for similar experiments with similar results. Seven CpG-island-sequences found to be H3K4me2-enriched in a study by Miao and Natarajan (2005) were also on the CpG-island (CGI)-microarray used in the present study. Consequently, the enrichment ratios at these seven sequences determined by Miao and Natarajan (Miao and Natarajan, 2005) were used as control ratios for H3K4me2-enrichment in ChIP-on-chip experiments.

2.19.3 Array-based profiling of reference-independent methylation status (aPRIMES)

The aPRIMES experiments presented here were carried out in collaboration with Verena Fleig. V. Fleig performed hybridization of CLL-tumor samples and B-cell control samples according to the protocol described below and acquired and analyzed all genome wide data. The data acquired at the same time for 13q14-CGI clones were independently analyzed by me and exclusively used in this work.

In the standard aPRIMES protocol by Pfister et al. (2007) 500 ng genomic DNA was restricted to completion with 10 U MseI for 3 h in a final volume of 10 mL in the buffer provided by the supplier (New England Biolabs, Beverly, USA). Subsequent heat inactivation was carried out at 65°C for 20 min. The MseI-fragments were then subjected to linker-mediated PCR as essentially described by Klein et al. (1999). Briefly, 1 mL each of 100 mM stock solution (MWG, Ebersberg, Germany) primer ddMse11 and primer Lib1 were annealed in 1 mL One-Phor-All-Buffer and 3 mL ddH₂O. The annealing was started at a temperature of 65°C and was shifted down to 15°C with a ramp of 18°C/min. At 15°C, 10 mL MseI fragments, 2 mL of ATP (10 mM) and 2 mL T4-DNA-Ligase (10 U, Roche) were added, and primers and DNA fragments were ligated overnight. Half of the resulting ligated MseI fragments were digested with the restriction enzyme McrBC (New England Biolabs, Beverly, MA, USA) for 8 h following the conditions recommended by the supplier. The other half of the MseI fragments was digested with two methylation-sensitive endonucleases, HpaII (recognition site CCGG, 3 h, 37°C) and BstUI (recognition site CGCG, 3 h, 60°C), according to the recommendations of the supplier. Digested DNA fragments were then treated with 1 mL Proteinase K (20mg/mL; Invitrogen, Karlsruhe) for 1 h at 37°C with subsequent heat inactivation at 80°C for 10 min. For the following amplification step, 10 mL consisting of 2 mL 10x Expand Long Template buffer 1 (Boehringer, Mannheim), 1 mL dNTPs (10 mM), 1 mL Lib1 primer, 1 mL expand long template DNA polymerase mixture (3.5 U/μL; Boehringer, Mannheim) and 5 mL ddH₂O were added to the 20 mL reaction volume. A MWG thermo cycler was programmed to 72°C for 3 min, followed by 20 cycles at 94°C for 30 s, 62°C for 30 s and 72°C for 90 s. Final elongation was carried out at 72°C for 10 min. The PCR products were recovered by ethanol precipitation. DNA was eluted in 30 mL 0.1x TE, pH 8 (Pfister et al., 2007).

Several internal controls are available for aPRIMES: First, spike CGIs from rice were used as positive controls for methylation. For this, ten rice CGIs were PCR amplified and printed onto the microarray. All sample DNAs for aPRIMES were spiked with ~10 pg of each *in vitro* methylated rice-CGI to control methylation and methylation-sensitive digestion. Mitochondrial CGI clones that were present in the original library were used as controls for unmethylated and allelically/partially methylated CGIs (Pfister et al., 2007).

All Sequence information for CGI clones was obtained from a publicly available database at <http://data.microarrays.ca/cpg/index.htm>. Chromosomal annotation of sequences is based on the University of California at Santa Cruz (UCSC) genome database (Freeze, May 2004).

2.20 Statistical analysis

Mean values, median values, standard deviations, Student's t tests, Kolmogorov-Smirnov tests and variance of statistical sample sets were calculated in the Excel software (Microsoft, USA) by standard formulas included in the software. The statistical analyses performed in this study comprised the calculation of p-values to detect significant differences between two sample sets. Normalized enrichment values of different sample groups were compared to each other by statistical means. All analyzed samples were assigned to distinct (sub-) groups, which are listed in Table 2-13. The values for distinct enrichment at several loci that belong to one functional element for all samples of one subgroup were analyzed together as one sample set in t tests.

Table 2-13: Sample (sub)-groups used for statistical analyses.

No.	Sample group	Included samples
1	CLL H3K4me2-enriched	CLL1; CLL2; CLL3; CLL4; CLL5
2	CLL macroH2A-enriched	CLL1; CLL2; CLL3; CLL4; CLL5
3	Wild type H3K4me2-enriched	P1; P2; P3; P4; P5; PB5; PB6; PB7; PB17; PB18; PB19
4	Wild type macroH2A-enriched	PL1; PU1; PU2; PB17; PB18
5	Young wild type H3K4me2-enriched (<45 years)	P1; P2; P3; P4; P5
6	Young wild type macroH2A-enriched (<45 years)	PL1; PU1; PU2;
7	Old wild type H3K4me2-enriched (>45 years)	PB5; PB6; PB7; PB17; PB18; PB19
8	Old wild type macroH2A-enriched (>45 years)	PB17; PB18
9	Wild type B-cells H3K4me2-enriched	PB9; PB12
10	Wild type T-cells H3K4me2-enriched	PB9; PB12

Table 2-14 depicts the groups that were compared to each other together with reason for comparison and finally, Table 2-15 depicts the analyzed functional elements and the included loci of these calculations.

Table 2-14: Sample groups compared by statistical analyses. Numbers of groups are as depicted in Table 2-14

Sample group #1	Sample group #2	Feature/reason
1 CLL	3 wild type	Compare H3K4me2 in CLL vs wild type
2 CLL	4 wild type	Compare macroH2A in CLL vs wild type
1 CLL	5 young wild type	Compare H3K4me2 in CLL vs young wild type
1 CLL	7 old wild type	Compare H3K4me2 in CLL vs old wild type
2 CLL	6 young wild type	Compare macroH2A in CLL vs young wild type
2 CLL	8 old wild type	Compare macroH2A in CLL vs old wild type
5 young wild type	7 old wild type	Compare H3K4me2 in young vs old wild type
6 young wild type	8 old wild type	Compare macroH2A in young vs old wild type

Significant differences ($p < 0.05$) between group 1 and group 2 were determined by Welch's Two sample t-test calculated with the R software package. By using R for calculating p-values, the software automatically calculated t-values. These t-values were further used to determine which sample group contained the higher enrichment. If $t > 0$, the first of the two compared groups includes higher values than the second, and if $t < 0$ the second group

contains higher values than the first; whereby the t-value itself correlates with the degree of value-differences between the two groups.

Table 2-15: Loci included in analyses of functional elements. e = exon; first number depicts number of exon of the gene; second number after "." depicts number of loci within exon designed for qPCR. Capitalized letter plus number = CGI and loci within CGI analyzed.

Functional elements	Included loci	No. of included loci
KPNA3	K3e6.1; K3e3.1; K3e1.1	3
C13ORF1	C13e4.1; C13e3.1; C13e2.1	3
RFP2	R2e1.1; R2e2.1; R2e3.1	3
BCMS	B1e1.1; B1e2.1; B1e3.1	3
BCMSUN/DLeu2	B2e2.2; B2e3.1; B2e3.2; B2e4.2	4
CGI A	A1; A2; A3	3
CGI B	B2; B3; B4	3
CGI C	C1; C2; C3	3
CGI D	D1; D2; D3; D4; D5	5
CGI E	E1; E2; E3; E4; E5; E6; E7; E12	8

2.20.1 Normal distribution of sample sets

To fulfill the criteria for statistical t-tests, normal distribution and equal variance of sample sets, all analyzed data sets were tested for their distribution and variance. Testing of normal distribution was either taken out by visual inspection of qq-plots in the R software or by Kolmogorov-Smirnov tests.

The Kolmogorov-Smirnov test (K-S test) is a nonparametric test of equality of one-dimensional probability distributions. It was used to compare two samples, because the K-S statistics quantify the distance between the empirical distribution functions of two samples. The null distribution of this statistic was calculated under the null hypothesis that the samples are drawn from the same distribution. These distance values were calculated by the K-S formula and compared to standard values. If the calculated value was smaller or equal to the standard value a normal distribution of the corresponding sample set was assumed. Calculations were carried out in Excel-datasheets for all analyzed combinations of sample sets.

The Kolmogorov-Smirnov statistic, which has the underlying empirical distribution function F_n for $n =$ independently and identically distributed observations X_i is defined as:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{X_i \leq x}$$

2.20.2 Variance of sample sets

The second criterion for selecting the appropriate t-test was the variance of sample sets. All combinations of sample sets were tested for their variance using this function:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

These calculations were carried out in and by the Excel software (Microsoft, USA). The formula depicted above describes the population variance of a finite population of the size N ,

whereby \bar{x} is the population mean. The variance of tested populations was unequal in all cases which is the reason why the Student's t-test was abolished. Instead, for normally distributed sample sets of unequal variance the Welch's Two sample t-test was applied to calculate p-values.

2.20.3 Welch Two sample t-test

In statistics, Welch's t-test is an adaptation of Student's t-test intended for use with two samples having possibly unequal variances (Welch, 1947). The Welch's t-test defines the statistic t by the formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

where \bar{X}_i is the i^{th} sample mean, s_i^2 is the i^{th} sample variance and N_i is the i^{th} sample size. Unlike in Student's t-test, the denominator is not based on a pooled variance estimate. The degrees of freedom ν associated with this variance estimate is approximated using the Welch-Satterthwaite equation, which was included in the calculations of the R software:

$$\nu = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{\frac{s_1^4}{N_1^2 \cdot \nu_1}}{+} \frac{\frac{s_2^4}{N_2^2 \cdot \nu_2}}$$

In this formula ν_i is N_{i-1} and the degrees of freedom are associated with the i^{th} variance estimate.

Using the R software the command "t-test", computed t and ν together with the t-distribution to test the null hypothesis that two population means were equal (using a two-tailed test). In particular, the test yielded a p-value which might or might not give evidence sufficient to reject the null hypothesis. The null hypothesis was rejected if $p < 0.05$, which corresponds to a tolerated error of 5 %.

2.20.4 Normalization of CHIP data

Values determined for *CDH12* –CpG-island-promoter sequences were used to normalize data generated by H3K4me2-ChIPs and values for *HK2*-CGI-promoter sequences were used for normalization of macroH2A-ChIPs. The normalized values were termed "x-fold enrichment versus control sequence". These control sequences were *CDH12* for all H3K4me2-ChIPs and *HK2* for all macroH2A-ChIPs. X-fold enrichment at all analyzed loci were calculated by the equation:

$$\text{x-fold enrichment vs control} = \frac{\text{fraction of input (loci)}}{\text{fraction of input (control)}}$$

for qPCR determined values, and by:

$$\text{x-fold enrichment vs control} = \frac{\text{ratio (13q-CGI-clone)}}{\text{ratio (control clone)}}$$

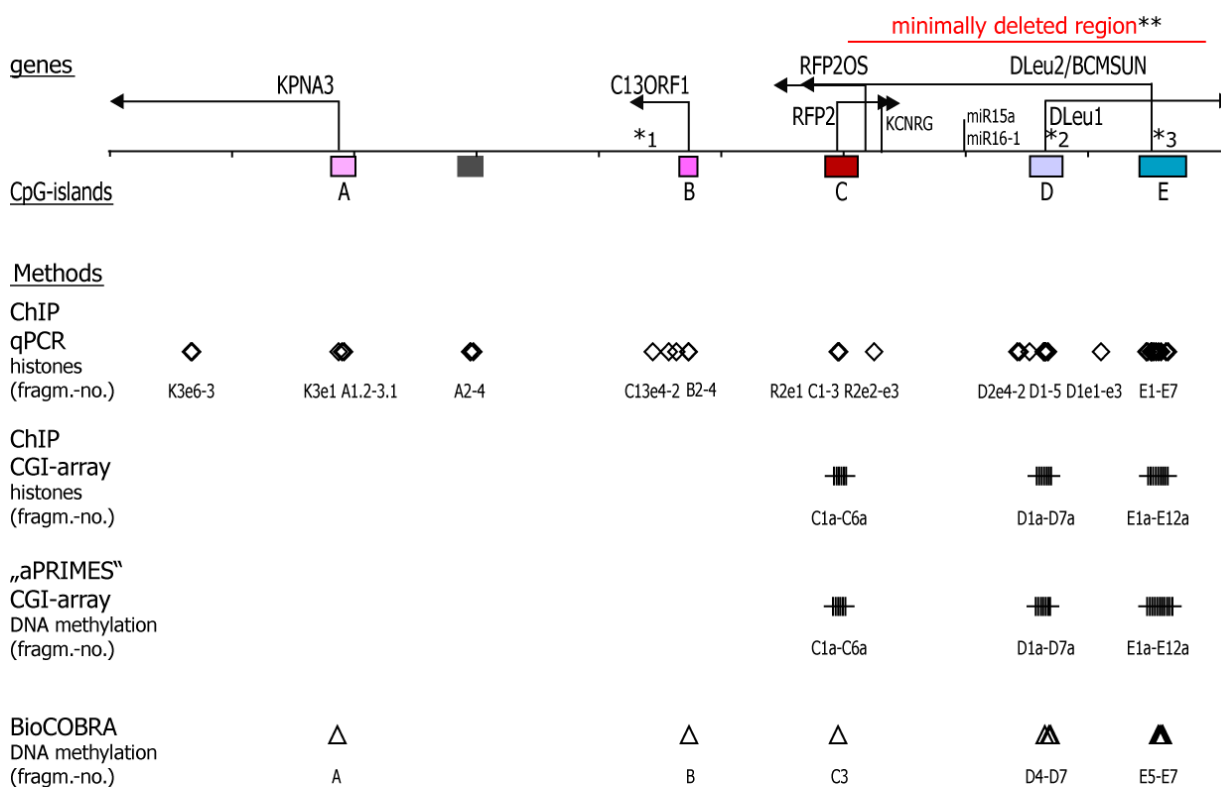
for ratios determined by microarray-hybridization of ChIP-DNA in ChIP-on-chip experiments.

2.21 Experimental set up

2.21.1 Analysis of chromosomal region 13q14.3 by several methods

Figure 2-3 shows exact localization of fragments analyzed by diverse methods for DNA-methylation (aPRIMES, BioCOBRA) and enrichment of two histone modifications (ChIP; ChIP-on-chip) in this work.

The most detailed analysis including several fragments analyzed in at least three exons of the five genes and at least three fragments analyzed in CpG-islands of the critical region in 13q14.3 was carried out by ChIP and qPCR analysis. Fragments analyzed by qPCR ranged in size from 50-250 bp. The fragments of CpG-islands C, D and E that are present on the CpG-island-array, span the respective CpG-island in 1 kb-tiling parts. The CpG-island-microarray was used for ChIP-on-chip and aPRIMES analyses. BioCOBRA analyses were performed by A. Philippen (see chapter 2.17) and she analyzed one to three fragments within each of the five CpG-islands of 13q14.3 and determined restricted DNA-methylation at these fragments. Sequence analysis was performed for the three indicated SNPs in Figure 2-3 (by *1-3) in heterozygous samples to analyze allele-specific enrichment of the two histone modifications (*2 and *3), or to detect monoallelic expression (*1; *C13ORF1*).



*1-3 = Single nucleotide polymorphisms (SNPs): *1= rs706617; *2=rs12100048; *3=rs9568354

** Stilgenbauer et al., 1998

Figure 2-3: Overview of localization of analyzed fragments in chromosomal region 13q14.3 by several methods.

3 Results

The aim of the present study was to identify the epigenetic regulatory mechanism that controls expression of candidate tumor suppressor genes from chromosomal region 13q14.3 and has lost its function in CLL cells. Part of this region is lost in over 50 % of patients (Stilgenbauer et al., 1998), pointing to the involvement of the critical region in 13q14.3 in the development of CLL. In order to identify epigenetic regulatory elements chromatin immunoprecipitation (ChIP) was established and specific epigenetic modifications at CpG-islands and genes of the region were quantified. Identification of the epigenetic code at 13q14.3 should elucidate the epigenetic regulatory mechanism of the region and its relevance for the pathogenesis of CLL.

3.1 Establishment and validation of chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) is a powerful tool to detect and quantify protein DNA interactions and chromatin modifications in different tissues.

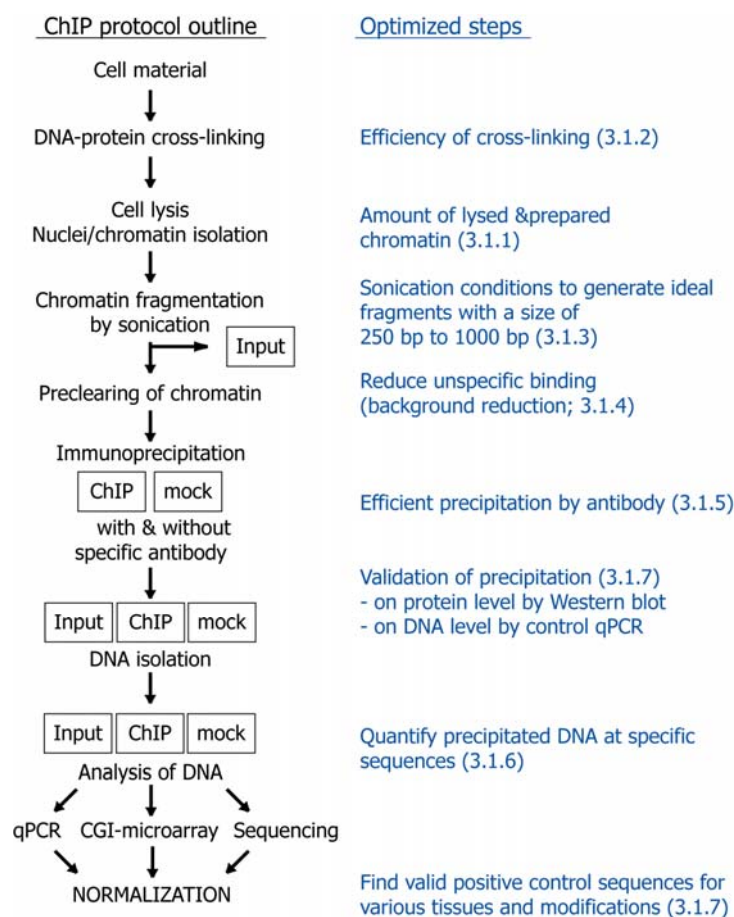


Figure 3-1: Overview of the outline and optimized steps of ChIP assays. Right side (blue): steps of the ChIP protocol that were optimized during this work. Numbers in brackets indicate chapters that describe optimization of respective steps.

In order to use ChIP to analyze chromatin modifications in CLL cells, the protocol needed to be adjusted to hematopoietic cells, which is described in this first chapter 3.1. A general outline of the ChIP protocol together with optimized steps is summarized in Figure 3-1.

3.1.1 Preparation of chromatin – nChIP versus xChIP

Two different protocols for ChIP were tested: “nChIP” (O'Neill and Turner, 2003) and “xChIP” (Orlando, 2000; Spencer et al., 2003; see chapter 2.14 for detailed protocols). The first step in the outline of the ChIP protocol (see Figure 3-1) had to be optimized in order to produce sufficient quality and quantity of chromatin from whole cell extracts for the following steps in the ChIP procedure. Using different buffers and different centrifugation steps, chromatin was prepared from whole cell extracts. A total of five approaches using different chromatin preparation conditions were tested, in order to find out the optimum condition for preparing chromatin with high quality and quantity. The resulting protein concentrations, served as a control for the efficiency of the respective approach (Table 3-1).

Table 3-1: Protein concentrations measured after chromatin preparation. a-e: Protein concentrations [$\mu\text{g}/\mu\text{L}$] determined after use of different buffers and protocols for chromatin preparation. Abbreviations: nd: not done, T: total chromatin, S: soluble chromatin, 10': digested chromatin, Uab: supernatant after precipitation with antibody (unbound); CPAb: precipitate of indicated antibody, Umock: mock supernatant, CPmock: mock precipitate. (see chapters 2.9.1-2.9.3 and chapter 2.14, for composition of buffers and detailed protocol).

a) Total chromatin preparation (nChIP-buffers/protocol)										
T	MN10'	Input	cleared	Input	Uab	CPAb	Umock	Cpmock	Tissue	Antibody
1.84	2.06	1.90	0.68	0.83	0	0.62	0.08	HL60	H3K9me3	
0.25	0.85	0.47	0.56	0.65	0.01	0.48	0	HL60	H3K4me3	
0.25	0.85	0.47	0.56	0.84	0.04	0.48	0	HL60	H3	
0.32	0.81	0.36	0.54	0.65	0.02	0.55	0.12	Jurkat	H3	
0.42	1.58	0.77	0.98	0.73	0.06	0.89	0.12	Jurkat	H3K4me3	
b) Without SDS-lysis-buffer preparation (nChIP-buffers+sonication-protocol)										
S	10'	Input	cleared	Input	Uab	CPAb	Umock	Cpmock	Tissue	Antibody
0.75	nd	0.12	0.12	0.40	0	0.17	0	HL60	H3K4me3	
1.27	nd	0.31	0.24	0.39	0	0.21	0	Jurkat	H3K4me3	
c) SDS-concentration down; Input diluted (xChIP-protocol with minimal SDS)										
S	10'	Input	cleared	Input	Uab	CPAb	Umock	Cpmock	Tissue	Antibody
0.33	1.18	0.53	0.64	0.76	0	0.65	0	HL60	H3K4me3	
0.34	1.65	0.76	0.79	0.83	0	0.75	0	Jurkat	H3K4me3	
d) Without SDS-lysis buffer (xChIP-protocol/buffers but without any SDS)										
S	10'	Input	cleared	Input	Uab	CPAb	Umock	Cpmock	Tissue	Antibody
nd	nd	0.58	0.69	0.69	1.33	0.65	1.49	HL60	H3K9me3	
nd	nd	1.68	1.43	1.39	0.32	1.14	0.41	Jurkat	H3	
e) Nuclei extraction prep (xChIP-protocol + nuclei extraction buffers)										
T	10'	Input	cleared	Input	Uab	CPAb	Umock	Cpmock	Tissue	Antibody
4.95	nd	2.67	2.73	2.17	1.18	2.08	0.98	HL60	H3	
2.32	nd	0.75	0.77	1.29	0.69	1.06	0.63	HL60	H3K9me3	

In the starting ChIP-fractions (i.e. "total", "soluble" and "MN10") protein concentrations ranging from 0.25 to 2 $\mu\text{g}/\mu\text{L}$ chromatin were measured. In the first three approaches, different buffers were used in order to increase chromatin amounts and quality (Table 3-1 a-c). Nearly half of the chromatin was lost during processing of "input chromatin" fractions (concentrations ranged from 0-0.05 $\mu\text{g}/\mu\text{L}$) and no protein could be precipitated. Testing the xChIP protocol without SDS in the lysis buffer, which is listed in Table 3-1d, led to unspecific background precipitation that was higher than precipitation by the indicated antibodies themselves (background = "CPmock": 1.48 $\mu\text{g}/\mu\text{L}$ versus precipitate = "CPab": 1.3 $\mu\text{g}/\mu\text{L}$ and CPmock: 0.41 $\mu\text{g}/\mu\text{L}$ versus CPab: 0.3 $\mu\text{g}/\mu\text{L}$; Table 3-1d). Fractions prepared with approach e, which combined the standard xChIP protocol and a nuclei extraction protocol, contained sufficient chromatin in input-fractions to subsequently precipitate protein by specific antibodies (0.6-1 $\mu\text{g}/\mu\text{L}$ in CPab; Table 3-1e).

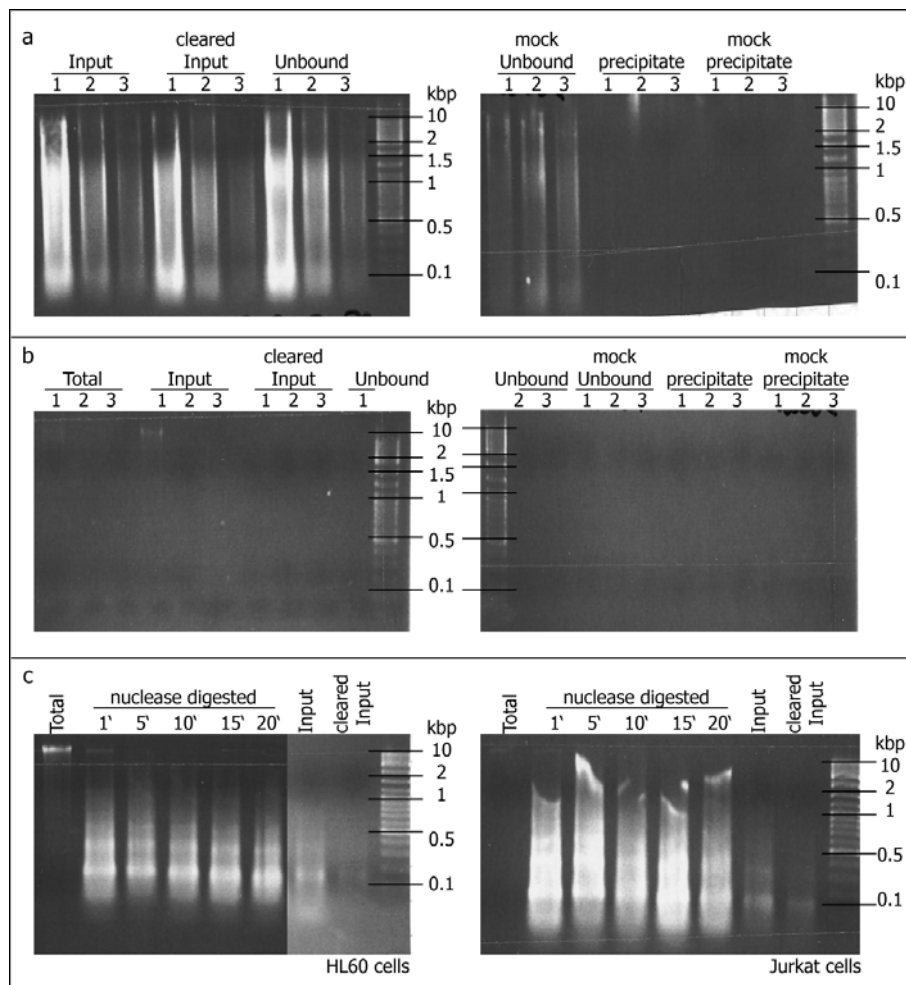


Figure 3-2: 1 %-agarose gel pictures of DNA isolated from ChIP-fractions analyzed to control chromatin preparation. a) DNA results/amounts by xChIP-protocol using 3 different amounts of HL60 cells, 1: 1×10^8 , 2: 1×10^7 & 3: 6×10^6 cells; b) same as (a) except for use of nChIP protocol; c) DNA amounts after nuclease digestion and input preparation without precipitation, left gel: HL60, right gel: Jurkat cells.

DNA content of ChIP-fractions needed to be controlled in parallel to protein levels as ChIP implies precipitation of both proteins and DNA. DNA was isolated from all ChIP-fractions (see Table 3-1) and analyzed on 1 %-agarose gels by electrophoresis (Figure 3-2). Using the xChIP standard protocol (Spencer et al., 2003), DNA could be isolated from all fractions

except from precipitates (Figure 3-2). The amounts of DNA varied between approaches 1, 2 and 3 with different numbers of starting cells being used. Variations of DNA amounts were in accordance with the various starting numbers of cells. Comparable amounts of DNA could be isolated in all fractions, except for precipitates and mock-precipitates (Figure 3-2a). The use of the standard nChIP-protocol (O'Neill and Turner, 2003) led to quite diverse results. Hardly any DNA could be detected in the different fractions, regardless of the amount of cells being used as starting material (Figure 3-2b). In nChIP-approaches, chromatin was lost during preparation of actual input-fractions as seen by testing length of enzymatic digestion (Figure 3-2c). DNA could be isolated from each digested fraction (Figure 3-2c) but no or not enough DNA was contained in input or cleared input fractions. Ineffectiveness of enzymatic digestion was visible in total ChIP-fraction in Figure 3-2c that contained whole undigested genomic DNA showing as one band at the same height as the 10 kb marker on the left. Preparation of chromatin by nChIP-protocol often failed to produce any input-chromatin for actual precipitation (Figure 3-2c). It was not possible to prepare chromatin by the nChIP protocol, as DNA levels were too low.

3.1.2 To crosslink or not to crosslink?

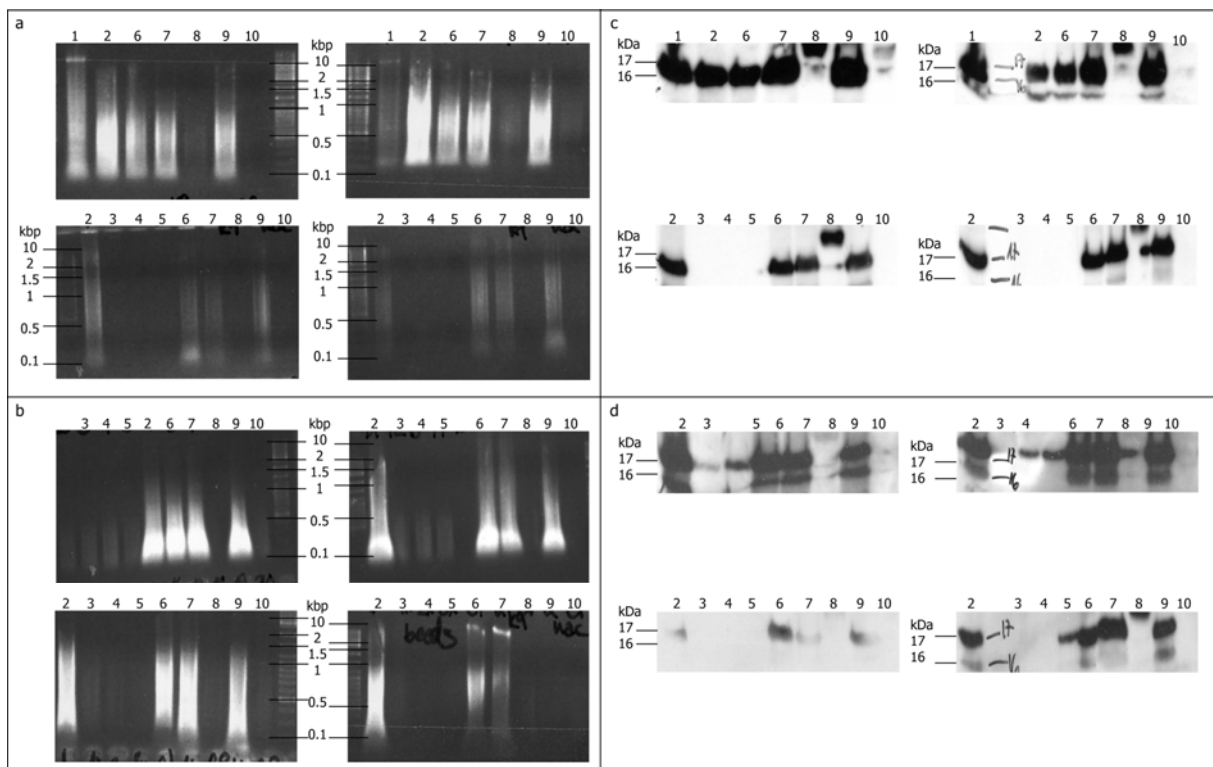


Figure 3-3: 1 %-agarose gel and Western blot pictures showing ChIP-fractions' content of DNA and protein. ChIP-fractions: 1: total; 2: input; 3-5: preclearing beads; 6: (pre)cleared input; 7: unbound; 8: precipitate; 9: mock unbound; 10: mock precipitate. a) DNA fractions after xChIP. Testing different crosslinking conditions: upper panel: 20 min at 4°C, lower panel: 10 min at 25°C. b) DNA fractions after xChIP without cross-linking; c) Western blot analysis of xChIP protein fractions analog to (a). d) Western blot analysis of xChIP- protein fractions analog to (b); precipitating antibody (a-d): α -H3K9me3. Detection of H3K9m3-amounts in (c) and (d) with α -H3 antibody (15 kDa); (a-d): left side: HL60 cells, right side: Jurkat cells.

The protocols for nChIP and xChIP differ in how the chromatin is prepared (chapter 3.1.2) and also in fixation of chromatin, which is only included in the xChIP protocol. Two different conditions for the formaldehyde cross-linking step of the xChIP protocol were tested and subsequently, DNA and protein content of all xChIP-fractions were analyzed (Figure 3-3). DNA content after cross-linking for 20 minutes at 4°C was high and almost identical in all fractions (Figure 3-3a, upper panel). DNA content after cross-linking for only 10 minutes at room temperature was strikingly lower (Figure 3-3a, lower panel). The variation in the size of DNA-fragments was higher with cross-linking as compared to no cross-linking (Figure 3-3a-b). To test the necessity of the cross-linking itself, the xChIP protocol was taken out according to the outline in Figure 3-1, but without the formaldehyde cross-linking step. Subsequently, DNA fragments generated by xChIP without cross-linking (Figure 3-3b) ranged in size only little between 100 and 400 bp, compared to a range of 100 bp to 1 kb generated by xChIP including cross-linking (Figure 3-3a). Moreover, the DNA amount was substantially higher in fractions prepared from xChIP approaches that included cross-linking (Figure 3-3a-b). The protein amounts were found to correlate with the DNA content (Figure 3-3), as high amounts of protein could be generated by xChIP-protocol including the cross-linking step (Figure 3-3c). Surprisingly, there was also considerably less unspecific binding of cross-linked chromatin to beads during preclearing (fractions 3 to 5 in Figure 3-3c-d) compared to chromatin without cross-links (Figure 3-3c-d). The benefit of formaldehyde cross-linking, a step included in the standard xChIP protocol is twofold, as it stabilizes chromatin and somehow reduces unspecific binding in ChIP procedures.

3.1.3 Chromatin fragmentation – Native nucleosomes versus sonicated chromatin fragments

Chromatin prepared according to the nChIP-protocol was fragmented by *Micrococcus* nuclease (MNase) digestion. In contrast to nChIP, chromatin prepared according to the xChIP protocol was stabilized by cross-linking it prior to fragmentation (chapter 3.1.3) and therefore could be fragmented into 100 to 1000 bp sized molecules by ultra-sonication. Figure 3-4a shows so-called "nucleosomal ladders" resulting from fragments of different length of nucleosomes (see chapter 2.14.1) of different nChIP control fractions. Length of nuclease digestion incubation time (2 to 10 minutes; Figure 3-4a left) had great impact on the amount of DNA later to be isolated. The DNA content of xChIP-fractions which were fragmented by ultra-sonication (Figure 3-4b) was higher in all fractions. High amounts of DNA fragments that ranged in size from 100 bp to 1 kb were isolated from xChIP, also in the precipitate xChIP fractions (insets in Figure 3-4b). The main difference of the two used ChIP-protocols was based on the fragmentation of chromatin in both procedures. In nChIP, native chromatin fragments according to nucleosomes should be generated by enzymatic digestion. The performance of the nuclease used to fragment chromatin is highly dependant on salt concentrations, amount of chromatin, amount and processivity of the nucleases, for instance. On the other hand, ultra-sonication that was performed in xChIP to fragment chromatin was a non-biased random process that fragments prior stabilized chromatin into reproducible sizes and size ranges once the conditions were optimized. In contrast, the enzymatic fragmentation that had the advantage of producing native chromatin-fragments according to nucleosomes was highly variable in its performance.

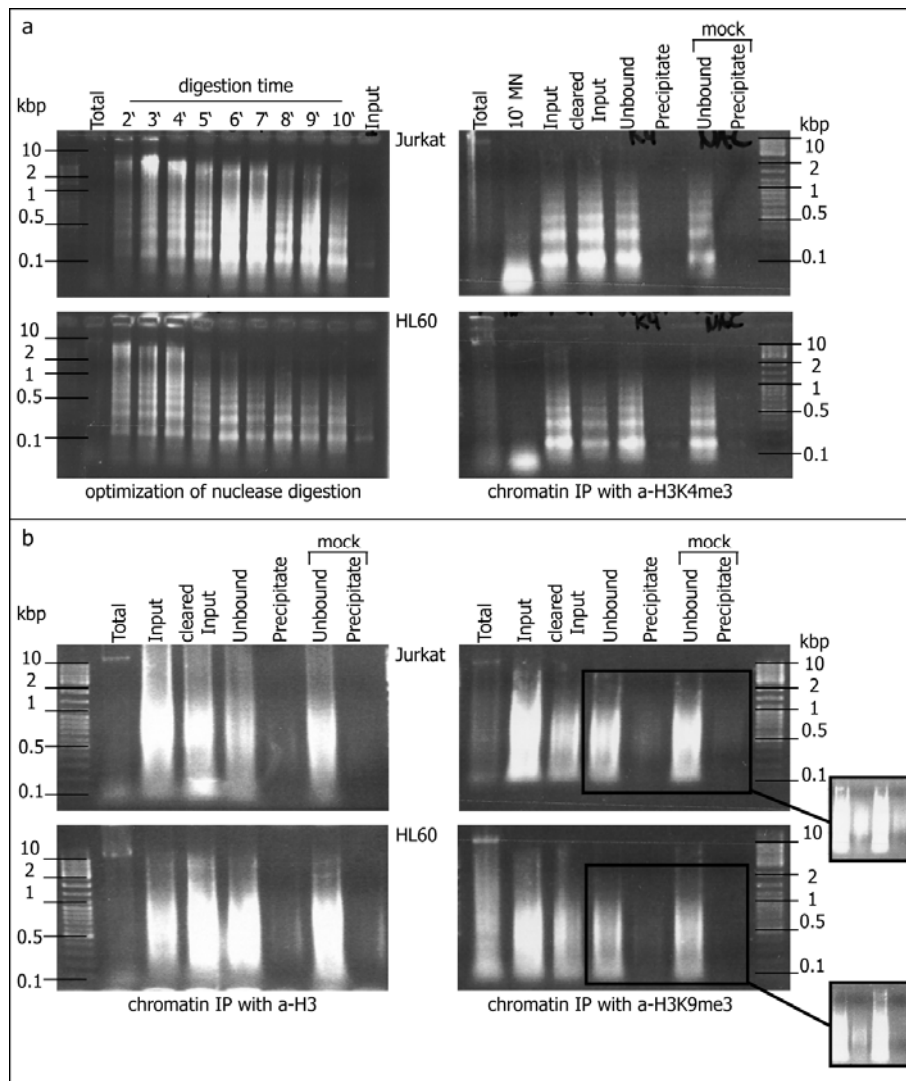


Figure 3-4: 1 %-agarose gel pictures to monitor DNA amounts isolated after chromatin fragmentation, testing different methods to chop chromatin. a) nucleosomal ladders of different nChIP-fractions after digestion with MNase. b) DNA fragments of xChIP-fractions after sonication. Insets: higher exposure of same DNA agarose gels.

3.1.4 Optimization of the precipitation procedure

For the precipitation of chromatin-antibody complexes by protein A/G coupled beads, five different approaches (bead-type and preclearing conditions) were tested. The aim was to achieve the lowest unspecific binding to beads and highest precipitation efficiency. Specific precipitation of antibody-chromatin complexes without co-precipitation of unspecific chromatin was achieved by use of protein A agarose/salmon sperm DNA pre-mix (Upstate, USA) and three times preclearing for 15 minutes at 4°C conditions (lowest line in Figure 3-5). In this approach, negative control fractions contained no chromatin (black boxes and dotted black boxes), while antibody-chromatin complexes were specifically precipitated (green box, Figure 3-5).

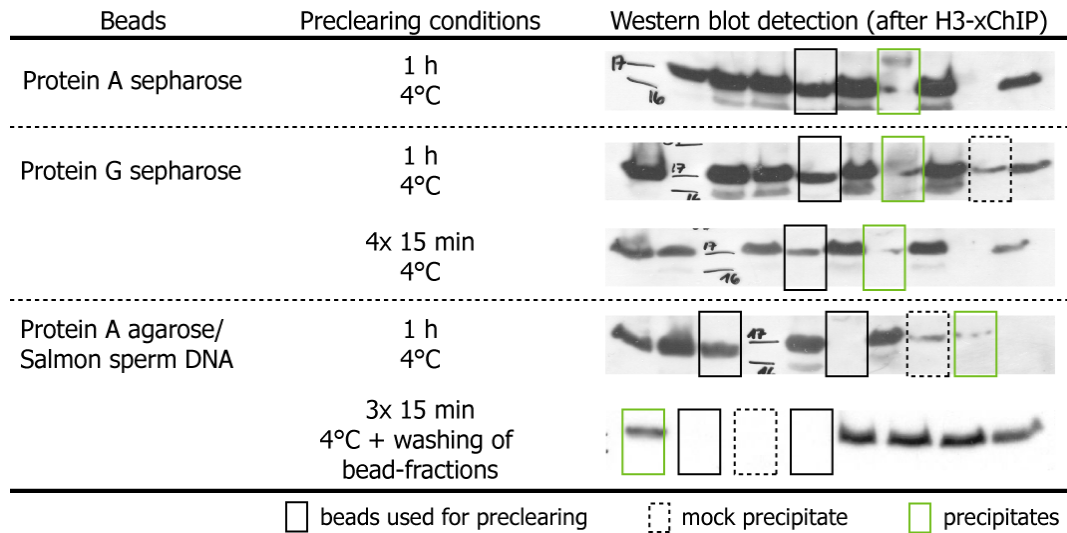


Figure 3-5: Testing of different beads and preclearing conditions to optimize precipitation efficiency and minimize background. "beads": different types of beads (sepharose/agarose) and coupled proteins (protein A/G) tested.

3.1.5 Performance of antibodies – Which antibodies precipitate chromatin?

Performance of antibodies is crucial for successful ChIP. It was desirable for antibodies to precipitate target chromatin efficiently and detect target protein in Western blot analyses. Five of six tested antibodies precipitated target proteins ("+"-bands in Figure 3-6a), except for anti-H4. Some antibodies precipitated protein in IPs but no chromatin in ChIP experiments, for example anti-H3K4me3 and anti-H3K9me3 (Figure 3-6a-b). Twelve antibodies were tested in ChIP for efficient precipitation of target chromatin (Figure 3-6b). Anti-Hp1 γ (Figure 3-6b) did not detect protein with the expected size in Western blot analyses. Anti-CTCF, anti-H3K9me3 and anti-H3K27me3 did not enrich any target chromatin in precipitates (see Figure 3-6b). Of the twelve antibodies tested only three precipitated their targets sufficiently and reproducibly, these were anti-H3K4me2, anti-macroH2A1.2 and anti-H3 (Figure 3-6b).

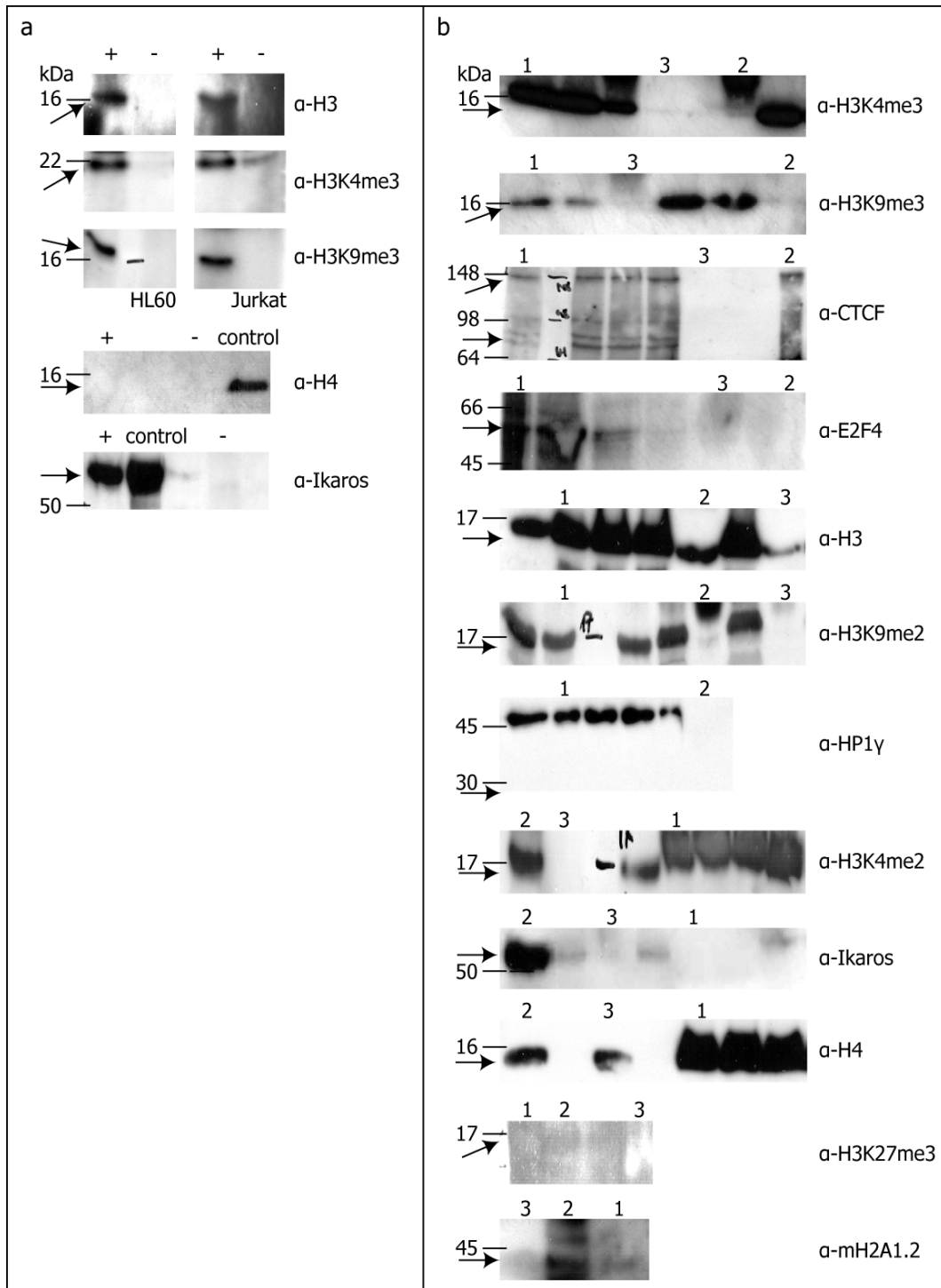


Figure 3-6: Western blot of IP and ChIP-fractions check of twelve different antibodies. a) IP; +: precipitated protein, -: mock, control: cell lysate; Western blot (WB) detection with same antibodies used for initial IPs; sizes of proteins in WB: H3: ~15 kDa, H3K4me3 & H3K9me3: ~17 kDa, H4: ~11 kDa and Ikaros: 50-60 kDa; b) ChIP; relevant Western blot bands were marked with numbers: 1 - input, 2 - precipitate, 3 - mock precipitate; sizes of proteins: see (a) and CTCF: 82+160 kDa, E2F4: 50-60 kDa, HP1 γ : 25 kDa; H3K4me2, H3K9me2 & H3K27me3: ~17 kDa and macroH2A1.2: ~42 kDa. HL60 cells were used (unless otherwise indicated); arrows indicate target proteins.

3.1.6 Optimization of ChIP – How to analyze precipitated DNA?

In order to determine the enrichment of different chromatin marks at the chromosomal region 13q14.3, precipitated ChIP-DNA was first analyzed by end-point PCR. As controls, PCRs were established to detect sequences known to be enriched for the respective histone modification and also for sequences known not to be enriched and controlled on genomic DNA (see "pos. control" in Figure 3-7). Products could be amplified from ChIP-fractions containing low amounts of DNA (preclearing beads, precipitate and mock precipitate; Figure 3-7). 300,000 cells per ChIP-fraction were used per PCR-reaction. However, for unknown reasons, no products were amplified in input-, cleared input- and unbound-fractions. Furthermore, no quantification of amplified DNA-products was possible by end-point PCR analyses. Subsequently, in order to quantify DNA from different ChIP-fractions, quantitative PCR (qPCR) was used to analyze ChIP-DNA (see chapter 3.1.9 and Figure 3-9).

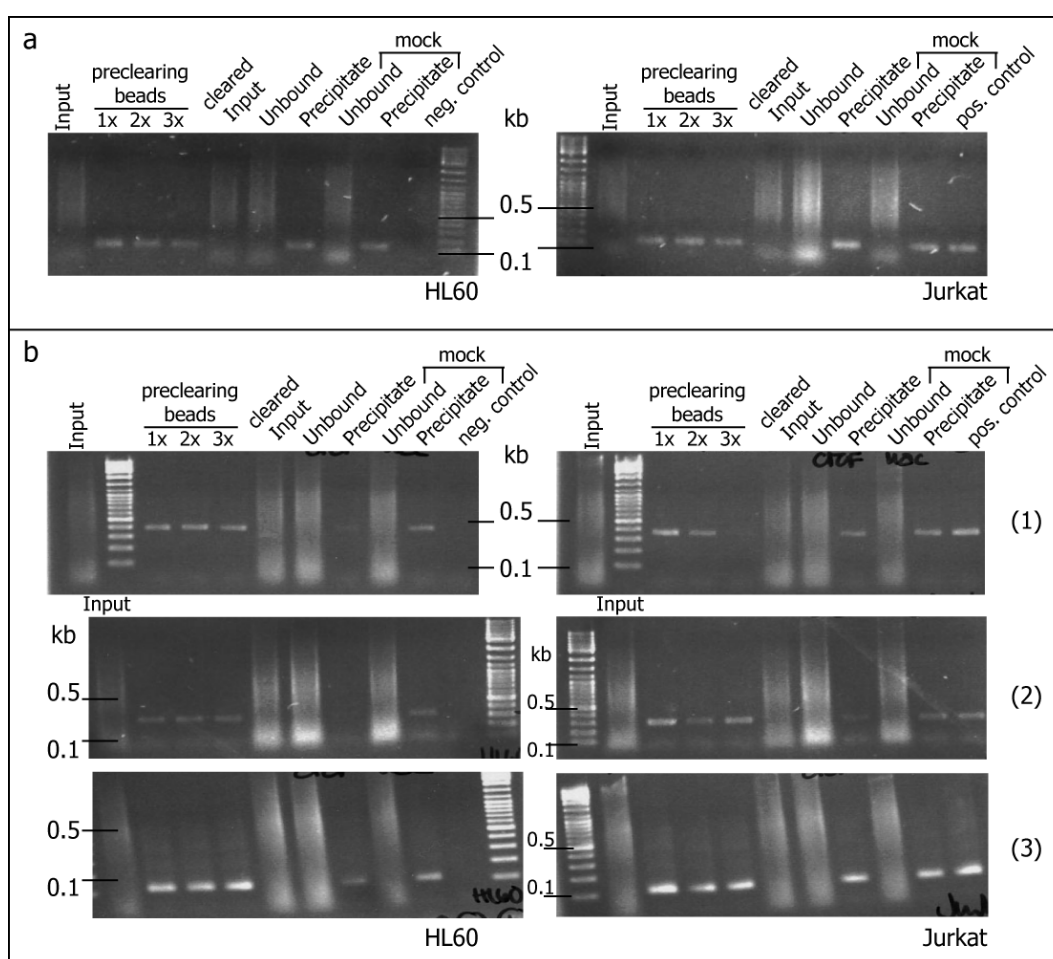


Figure 3-7: PCR analyses of ChIP-precipitates. a) PCR-analysis of DNA from anti-H3-xChIP, expected PCR product: MYH1 (110 bp; control); b) PCR-analysis of DNA from anti-CTCF-xChIP, expected PCR products: (1) H19-CTCF-binding site (385 bp; positive control), (2) cMyc-CTCF-binding site (284 bp; positive control), (3) MYH1 (110 bp; negative control); 15 cycles per PCR.

3.1.7 Validation of chromatin precipitation in H3K4me2- and macroH2A-ChIPs in CLL and non-malignant cells

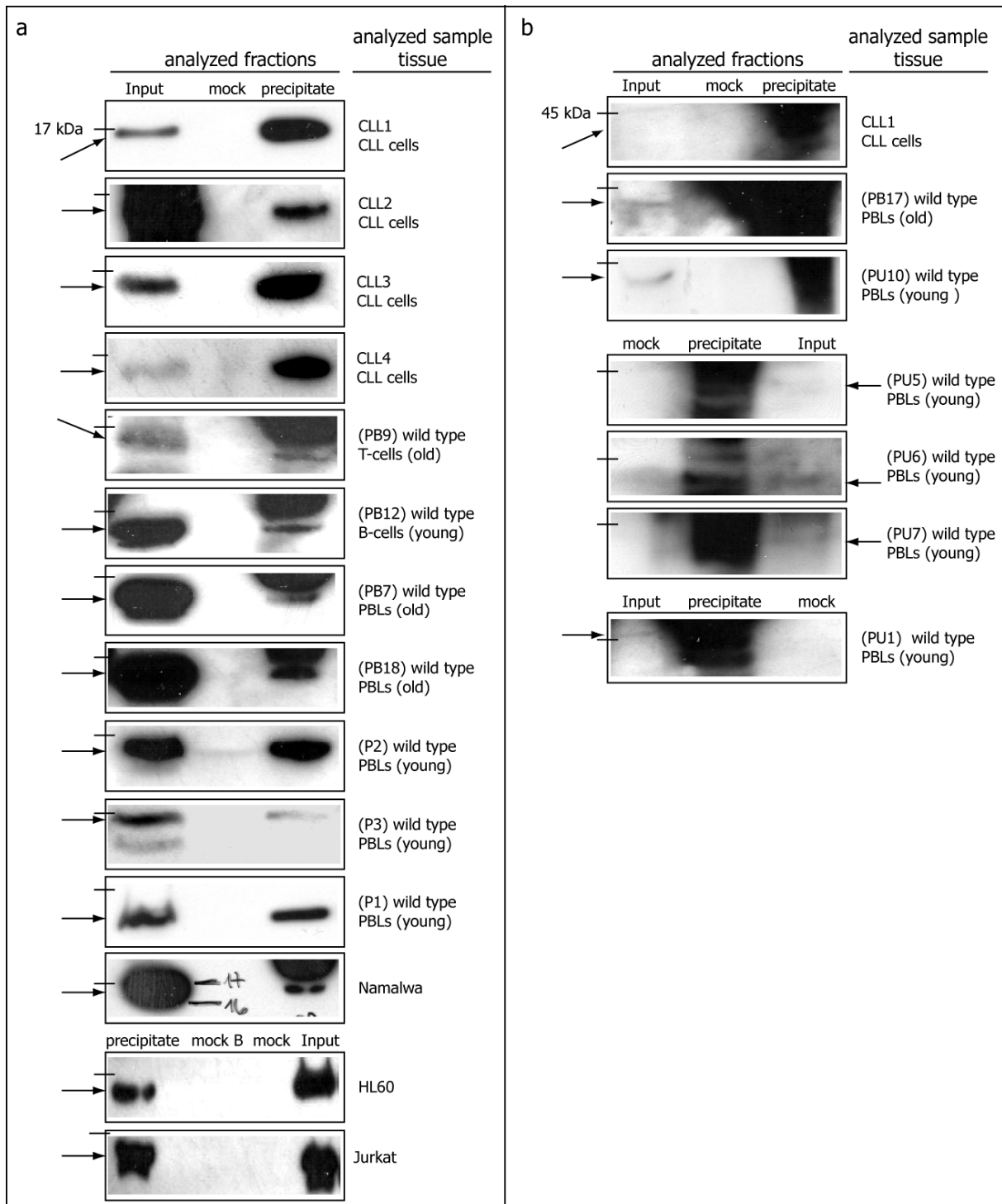


Figure 3-8: Western blot of H3K4me2- and macroH2A-precipitates. a) H3K4me2-precipitate-, mock- and input-fractions in different samples. MW H3K4me2 = ~17 kDa. b) macroH2A-precipitates analog to (a). MW macroH2A = ~42 kDa. Arrows point to target protein in precipitate fractions. Western blot detection with same antibody used for initial ChIP.

ChIP was used to quantify the presence of two epigenetic marks, H3K4me2 and macroH2A, at chromosomal region 13q14.3 in non-malignant and CLL cells (chapters 3.2 and 3.3). Specific precipitation of each ChIP-experiment was validated on two levels. First, it was tested whether target protein was specifically enriched in precipitate-fractions only, and not in mock-fractions (Figure 3-8). Detection by Western blot analyses was performed with the same antibodies used for ChIP-precipitation. H3K4me2 was specifically precipitated in different CLL patient samples (CLL1-CLL4; Figure 3-8a), in non-malignant B- and T-cells, in 13 samples of probands of two age groups ("young" and "old"; Figure 3-8a) and in the cell lines Namalwa, HL60 and Jurkat. Enrichment of macroH2A in ChIP-fractions was more difficult to detect by Western blot analysis, but successfully precipitated macroH2A could be detected in CLL patients (CLL1 in Figure 3-8b) and in healthy probands of the two age groups (Figure 3-8b).

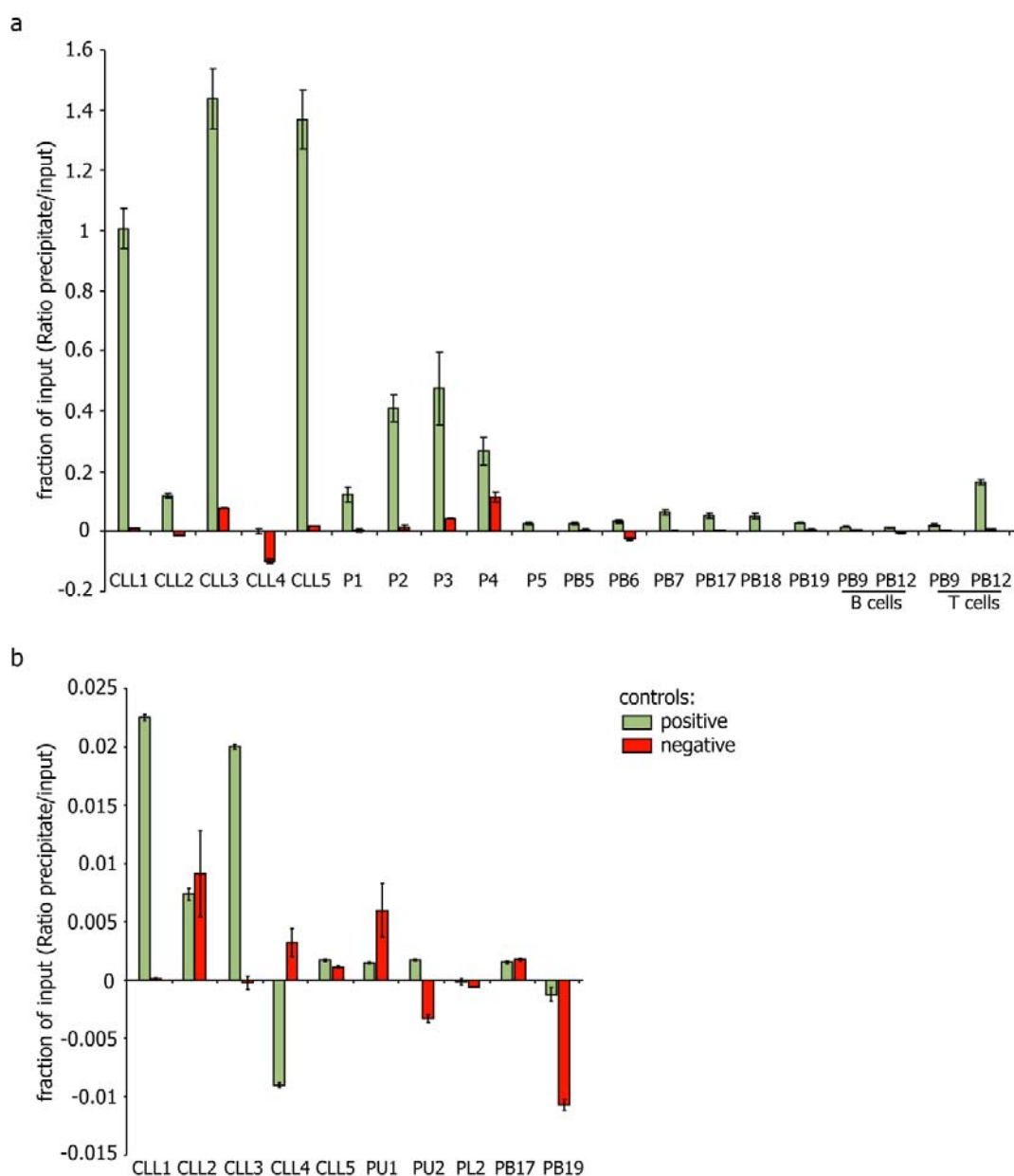


Figure 3-9: Fraction of input of H3K4me2- and macroH2A-ChIP measured by qPCR. a) H3K4me2: *GAPDH*-promoter (green bars; positive control), *MYH1*-promoter (red bars; negative control). b) macroH2A: *IL8*-promoter (green bars; positive control) and *MYH1*-promoter (red bars; negative control).

Second, the DNA-parts of precipitated chromatin were quantified by control-qPCR reactions on positive and negative (i.e. according to the enrichment of the analyzed chromatin mark) promoter sequences. ChIP was considered valid when the positive control promoter sequences were more strongly enriched than negative control promoter sequences and secondly, when precipitated fractions were enriched, while mock ChIP-fractions were not. Surprisingly, H3K4me-enrichment at the positive control promoter (*GAPDH*; green bars in Figure 3-9a) varied significantly between different tissues. Highest enrichment at *GAPDH* was found in CLL3 and CLL5 (ratio of 1.39 and 1.45), in wild type P3 (0.48) and in wild type T-cells PB12 (0.2; Figure 3-9a). In contrast, at promoter sequences used as negative control (*MYH1*; red bars in Figure 3-9a), enrichment was significantly lower than at sequences used as positive controls in all analyzed samples.

In general, enrichment of macroH2A was significantly lower in all tissues compared to H3K4me2-enrichment (Figure 3-9b). The highest ratio of macroH2A-enrichment was measured in the sample CLL1 (0.024; Figure 3-9b). Defining negative control promoters for macroH2A was slightly more difficult than for H3K4me2-enrichment. This could be seen by rather high enrichment of these negative controls measured in qPCR in several samples (CLL2, PU1 and PB17 in Figure 3-9b).

Table 3-2: Enrichment of control clones of H3K4me2-ChIP on chip experiments in HL60 used for validation. H3K4me2-enrichment ratios (Enr.) and percentiles (Per.) of ChIP-on-chip experiments for each sample. Columns: H3K4me2-enriched CpG-islands in THP1 cells (Miao and Natarajan, 2005); location of CGI to gene: up (upstream); within or down (downstream). Below the actual table a short description of gene functions of the control clones is given.

CpG-clone		29B09	115B10	63F01	33B01	15A04	5B06	13F12
Gene symbol		TNFRSF7	HOXA10	S100A10	BAT1	RHO6	VDR	DR1
Chromosome		12	7	1	6	12	12	1
location CGI to gene		3'	5'	3'	5'	3'	5'	within
HL60	Enr.	0.459	0.453	0.573	0.387	0.541	0.452	0.364
	Per.	10th	10th	10th	25th	10th	10th	10th
CLL1	Enr.	NA	0.428	0.053	NA	NA	NA	NA
	Per.	-	25th	50th	-	-	-	-
CLL2	Enr.	1.084	0.671	NA	0.126	NA	1.550	1.471
	Per.	10th	25th	-	50th	-	10th	10th
CLL3	Enr.	0.173	0.723	0.678	-0.093	NA	1.603	0.135
	Per.	50th	25th	10th	75th	-	10th	50th
CLL4	Enr.	0.813	NA	0.682	-0.387	NA	0.896	0.793
	Per.	10th	-	25th	75th	-	10th	10th
CLL5	Enr.	0.704	NA	0.3190	-0.152	NA	0.464	0.397
	Per.	10th	-	50th	75th	-	25th	25th
P4	Enr.	0.593	NA	0.158	0.376	0.004	0.470	0.152
	Per.	10th	-	75th	25th	50th	90th	75th
PB17	Enr.	-0.101	NA	0.428	0.036	-0.011	0.423	0.085
	Per.	50th	-	90th	50th	75th	90th	75th
Described gene function:		HLA-associated transcript	Rho family GTPase	TNF receptor superfamily member	Homeobox A10	Cytochrome P540, subunit	S100 calcium binding protein	TBP associated phosphor-protein

H3K4me2-ChIP-on-chip experiments were validated by enrichment of seven control promoter sequences chosen from the study by Miao and Natarajan (2005) that were also represented on the employed CpG-island microarray (CGI-array; see chapter 2.19.1 for details).

Accordingly, ChIP-on-chip experiments were validated if at least one of these seven clones had H3K4me2-enrichment ratios that were among the top ten of respective microarrays in one experiment. Subsequently, H3K4me2-ChIP-on-chip experiments in CLL and wild type cells detected similar enrichment-ratios for at least one of the seven control promoter clones (Table 3-2). In primary cells, enrichment of only one of the control clones was considered sufficient for validation of the ChIP-experiment.

Table 3-3: Enrichment of control CpG clones close to *CDH12* & *HK2*. Ratios and percentiles of control clones used for normalization of ChIP on chip-generated data. Values marked with asterisks (*) were used for normalization of arrays of which *CDH12*- or *HK2*-clones gave no signal. NA = no signal measured ("not available").

Sample No.	Sample type (tissue)	Precipitating antibody	Ratio CpG CDH12 H3K4me2	Percentile	Ratio CpG HK2 macroH2A	Percentile
CLL1	CLL	H3K4me2	NA	-	NA	-
CLL2	CLL	H3K4me2	0.225*	50th	0.244	50th
CLL3	CLL	H3K4me2	0.203	50th	-0.174	75th
CLL4	CLL	H3K4me2	NA	-	-0.141	75th
CLL5	CLL	H3K4me2	NA	-	-0.205	75th
PB17	Wt	H3K4me2	1	10th	-0.063	75th
P4	Wt	H3K4me2	0.949*	10th	1	10th
CLL2	CLL	macroH2A	0.103	50th	0.977*	10th
CLL5	CLL	macroH2A	NA	-	1	10th
PU1	Wt	macroH2A	NA	-	NA	-
PB17	Wt	macroH2A	NA	-	0.938*	10th

The last step of normalization of measured ChIP data in order to avoid generation of false-positives was normalization relative to control sequences (Mathieu et al., 2005). Normalization of raw ChIP data was included to achieve better comparability among different cell types analyzed and also among different platforms used for ChIP-DNA analysis (i.e. qPCR and CGI-array). For this purpose, two control promoters that were enriched in either H3K4me2 or macroH2A, but both in non-malignant cells and CLL cells, were identified from the ChIP-on-chip experiments (Table 3-3). The two clones selected as positive control sequences for H3K4me2- (*CDH12*) and macroH2A-ChIPs (*HK2*) fulfilled the criterion that they are enriched in both types of analyzed tissue in order to serve as control sequence, which was the case for the selected clones CDH12 and HK2 (Table 3-3).

The same sequences were measured and used as positive control sequences in subsequent qPCR analyses (Table A-1; Appendix A). The measured ratios of these positive control clone sequences were used for normalization of ChIP-on-chip experiments and analog the measured 'fraction of input'-values were used for normalization of qPCRs (see chapter 2.20.4 for details).

3.2 Chromatin modifications at 13q14.3 in non-malignant cells

Chromatin immunoprecipitation was established (chapter 3.1) to characterize epigenetic modifications at chromosomal region 13q14.3 in non-malignant and CLL cells. First, the normal chromatin status of the region in wild type cells is described in chapter 3.2.

3.2.1 Are there differences in enrichment of H3K4me2 between genes and promoters of the 13q14.3 region?

Enrichment of H3K4me2 at chromosomal region 13q14.3 was measured in eleven samples of peripheral blood lymphocytes (PBL) from healthy donors in order to identify genes and promoters that carry epigenetic marks associated with transcriptional activity. Considering the overall H3K4me2-enrichment at all measured loci of the region, samples could be assigned to three subgroups according to the level of enrichment of H3K4me2, but independently of other characteristics of the samples (Table 2-3). The first group consisted of highly H3K4me2-enriched samples (PB18, P3, PB5 and PB17), the second group included medium H3K4me2-enriched samples (P5, P2, PB7 and PB6) and the third group consisted of samples with only very low detectable enrichment of H3K4me2 (P1, P4 and PB19; Figure 3-10). However, the members assigned to each of the three groups were heterogeneous regarding their gender and age (see Table 2-3) and shared no other similarities.

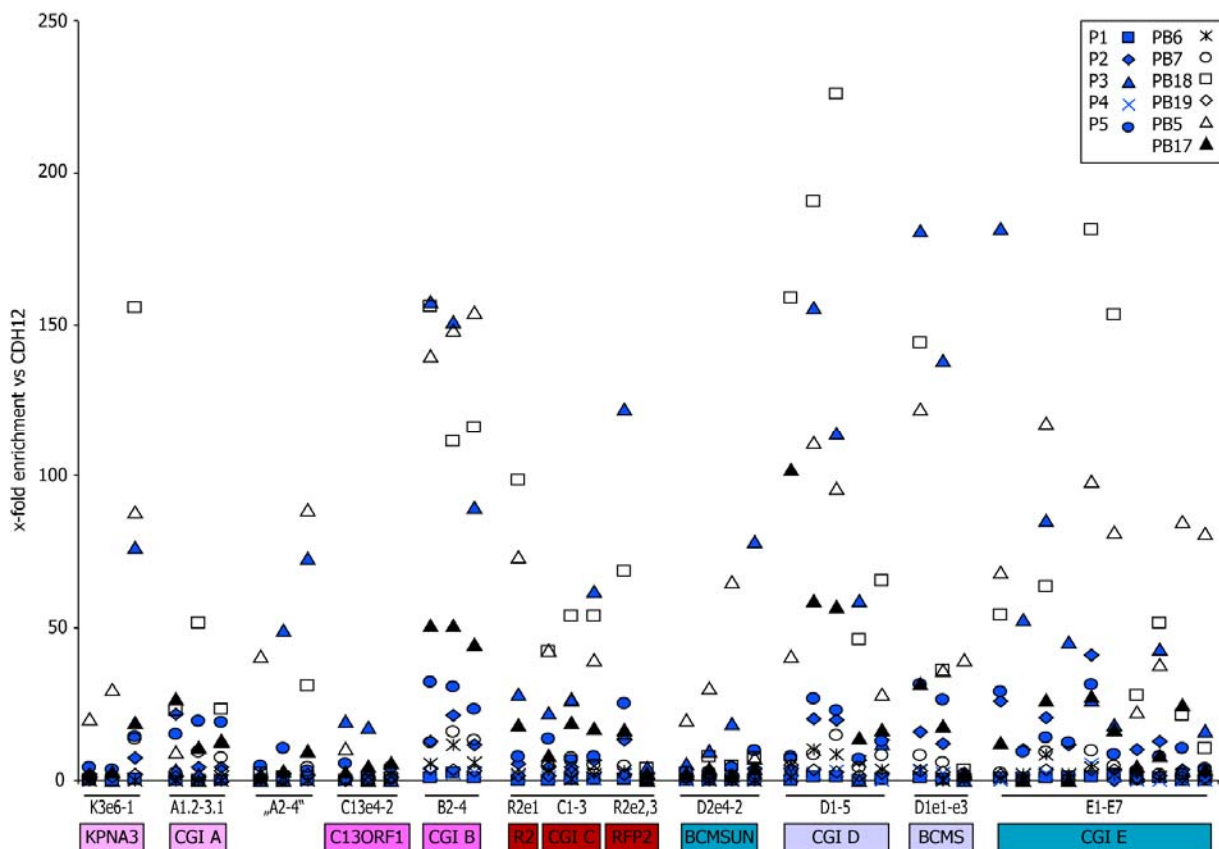


Figure 3-10: H3K4me2-enrichment at chromosomal region 13q14.3 genes (exons; i.e. e1-e4) and promoters (CpG-islands; CGI) in eleven wild type samples.

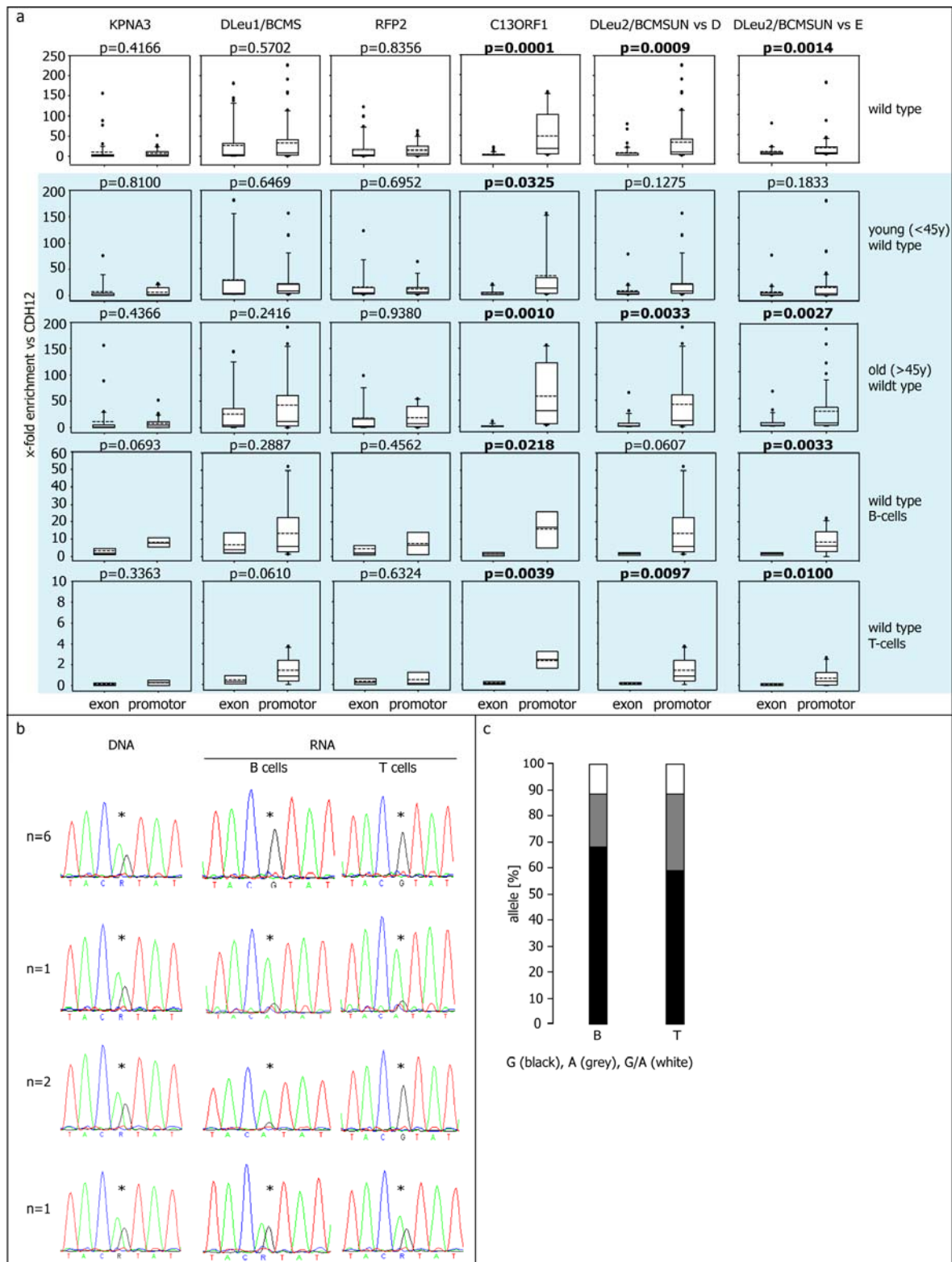
The H3K4me2-enrichment varied in different samples (i.e. individuals), but not at different genes and promoter CpG-islands of chromosomal region 13q14.3 (Figure 3-10). The pattern of H3K4me2-enrichment at distinct loci within promoters and genes was the same in all measured samples regardless of the level of enrichment (Figure 3-10). The highest median H3K4me2-enrichment in non-malignant cells was detected at CpG-islands B (11.7 to 16.1 fold enrichment versus *CDH12*), D (4.4 to 14.7 fold enrichment versus *CDH12*) and the centromeric part of CGI E (E1-E3; 3.3 to 9.5 fold enrichment versus *CDH12*; Table A-2, Appendix A). Highest H3K4me2-enrichment could be measured at CpG-island D (i.e. 225-fold enrichment versus *CDH12*; Figure 3-10). On the other hand, the lowest median H3K4me2-enrichment was measured at the third and sixth exon of *KPNA3* (0.1 and 0.6 fold enrichment versus *CDH12*), at all analyzed exons of *C13ORF1* (0.2 to 0.5 fold enrichment versus *CDH12*) and at *DLeu2/BCMSUN* exons (1.1 to 1.4 fold enrichment versus *CDH12*; Table A-2, Appendix A). Only one locus at 13q14.3 had no H3K4me2 in non-malignant cells: At the third exon of *KPNA3* no detectable enrichment could be measured in five of the eleven samples. In general, promoter CpG-islands and exonic regions had comparable H3K4me2-enrichment, except for the CpG-islands B and E that are located in the promoter regions of *C13ORF1* and *DLeu2/BCMSUN*, respectively (Figure 3-10; Table A-2, Appendix A). However, a general pattern of more H3K4me2 at promoter regions of 13q14.3 (CGIs) than at the exonic regions was detected (Figure 3-10).

In summary, medium levels of the active chromatin mark H3K4me2 were found at 13q14.3 in non-malignant cells. The pattern of H3K4me2-enrichment was the same in all samples; however, the levels of H3K4me2 differed interindividually. Remarkably, the exonic region of *KPNA3* showed almost no detectable H3K4me2, whereas at CGI D the highest measured H3K4me2 amount was measured.

3.2.2 Promoter-restricted H3K4me2-pattern correlates with monoallelic expression

Rougeulle et al. (2003) could show that monoallelically expressed genes are marked by promoter-restricted H3K4me2-enrichment. In order to test whether the monoallelically expressed genes at 13q14.3 would carry a similar pattern of histone modification; the enrichment of H3K4me2 was compared between gene exon and gene promoter regions. Significant differences ($p < 0.05$; chapter 2.20) between H3K4me2-enrichment at exons and the promoter of each gene could be detected, where promoters showed higher enrichment of H3K4me2 compared to the gene exons. Figure 3-11a shows box plots of enrichment of H3K4me2 at exons and corresponding promoters at five genes of the chromosomal region 13q14.3 analyzed in wild type samples. Promoter-restricted H3K4me2-enrichment was detected at *C13ORF1* ($p = 0.0001$) and at *DLeu2/BCMSUN* with both promoter-CpG-islands localized at transcription start sites of different splicing variants of *DLeu2* (Figure 1-11), CGI D ($p = 0.0009$) and CGI E ($p = 0.0014$). However, no such differential enrichment could be detected at *KPNA3* ($p = 0.4166$), *BCMS* ($p = 0.5702$) or *RFP2* ($p = 0.8356$; Figure 3-11a, yellow box). Monoallelic expression of *C13ORF1* in non-malignant hematopoietic tissue was confirmed by sequencing of a SNP (rs9568354) in the coding region of the gene. Sequencing of samples heterozygous for this SNP showed that only one allele was expressed in nine out of ten healthy probands (Figure 3-11b). Monoallelic expression could be determined for *C13ORF1* independently of the allelotype at the SNP (Figure 3-11c). Expression status of the other four genes has been determined in a previous study by Mertens et al. (2006).

3 Results



Briefly, *DLeu2/BCMSUN*, *RFP2* and splicing variants of *BCMS* were shown to be monoallelically expressed, while *KPNA3* was biallelically expressed. Four of the five genes in chromosomal region 13q14.3 were found to carry the histone mark that reflected their expression status, i.e. *KPNA3* and the short splicing variants of *BCMS* showed no restricted H3K4me2-enrichment in agreement with their biallelic expression while monoallelically expressed genes *C13ORF1* and *DLeu2/BCMSUN* showed promoter restricted H3K4me2-enrichment (Figure 3-11a). *RFP2* was shown to be monoallelically expressed in B- and T-cells (Mertens et al., 2006), but does not carry the respective promoter-restricted H3K4me2-enrichment.

All wild type samples were further subdivided into four groups related to age or tissue of the sample origins. In these subgroups, i.e. young and old wild type, B-cells and T-cells, more or less the same pattern as in the complete wild type group was detected with some exceptions. The subgroups T-cells and old wild type showed the same promoter restricted H3K4me2-enrichment at the same genes, *C13ORF1* and *DLeu2/BCMSUN*, as was detected for the complete wild type group. Lack of restricted enrichment was detected at the same genes, as well (*KPNA3*, *RFP2* and *BCMS*; Figure 3-11a, blue box). However, one remarkable difference was the lack of detection of the promoter-restricted enrichment of H3K4me2 – the mark for monoallelic expression – in probands below the age of 45 years at the *DLeu2/BCMSUN* gene, with neither of the two eligible promoters (CGI D or CGI E; Figure 3-11a, blue box). In the B-cell subgroup promoter restricted H3K4me2-enrichment at *DLeu2/BCMSUN* could only be detected with one of the two eligible promoters, namely CGI E ($p=0.0033$), and not with CGI D ($p=0.0607$; Figure 3-11a, blue box).

In no wild type subgroup a pattern of histone modifications specific for monoallelic expression was found for *RFP2*. Neither was promoter restricted H3K4me2-enrichment detected at the biallelically expressed genes *KPNA3* and *BCMS* (Figure 3-11a, blue box). The epigenetic pattern of monoallelic expression, i.e. promoter-restricted H3K4me2-enrichment, correlated with the monoallelic expression of the genes *DLeu2/BCMSUN* and *C13ORF1*, which was furthermore shown to be indeed monoallelically expressed. Surprisingly, the mark was not found at the monoallelically expressed *RFP2*, but was neither detected at the biallelically expressed genes *KPNA3* and *BCMS*.

3.2.3 MacroH2A-enrichment at 13q14.3 in non-malignant cells

Enrichment of macroH2A at chromosomal region 13q14.3 was characterized by ChIP and qPCR analyses in five PBL samples of healthy donors in order to identify epigenetically silenced loci. Overall, macroH2A was detected only in very low amounts at the region (Figure 3-12 and Table A-3, Appendix A). In two samples (PL2 and PB19) almost no macroH2A-enrichment at all could be detected with only minor exceptions at several distinct loci within exonic regions of *BCMS*, *KPNA3*, *C13ORF1* and *DLeu2/BCMSUN* (0.02 to 0.9 fold enrichment versus *HK2* and within distinct loci within CGI D and CGI E (0.6 and 0.7 fold enrichment versus *HK2*, respectively; Figure 3-12). As seen for H3K4me2-enrichment (chapter 3.2.1) the levels of macroH2A-enrichment varied between different samples/individuals, but the pattern of distribution was similar at genes and promoter CpG-islands of chromosomal region 13q14.3 (Figure 3-12). The highest median macroH2A-enrichment in wild type was measured at CpG-islands C and D and at exons of the gene *DLeu2/BCMSUN* (0.9 to 2.2 fold enrichment versus *HK2*; Table A-3, Appendix A). The largest part of the analyzed region 13q14.3 had no or only very little macroH2A-enrichment in non-malignant cells. Notably,

cell-samples that originated from donors of above 45 years of age (i.e. old wild type) had more macroH2A than the other samples.

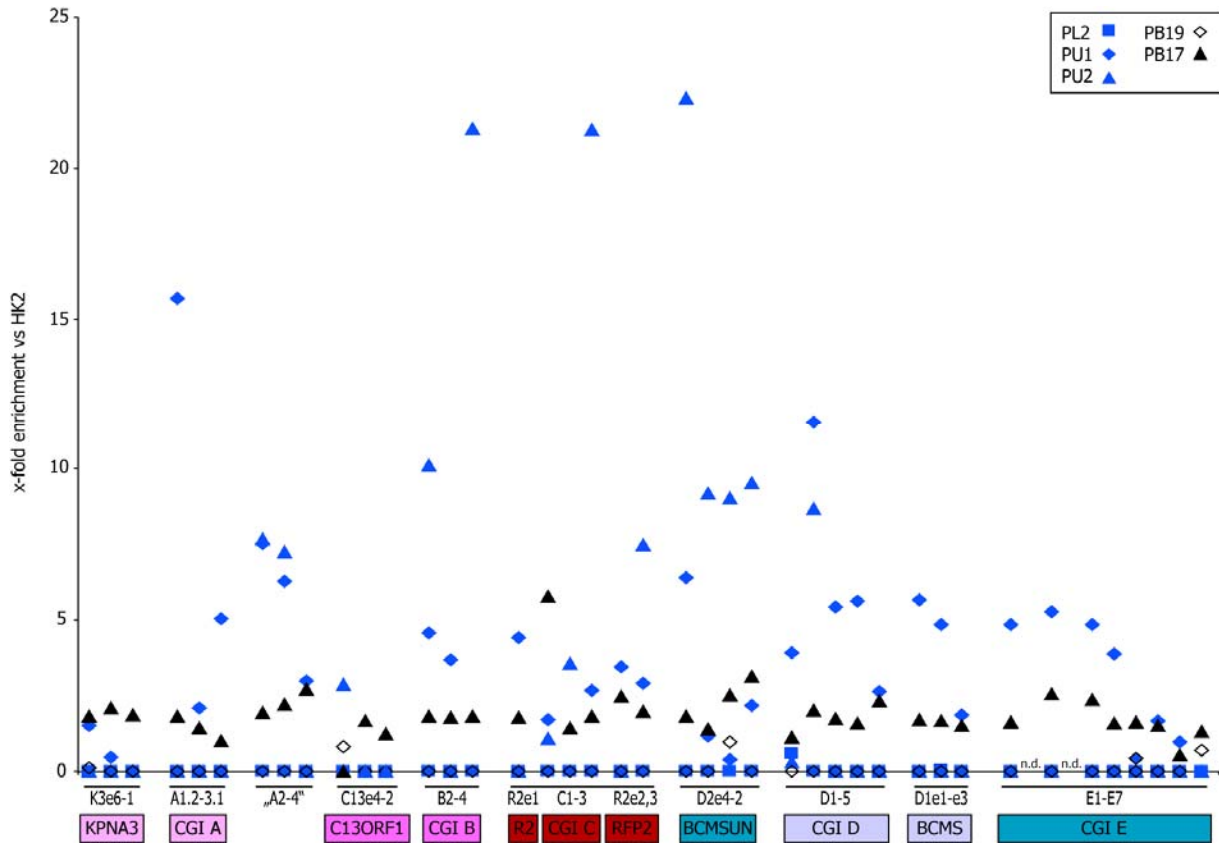


Figure 3-12: MacroH2A-enrichment at chromosomal region 13q14.3 genes (exons; i.e. e1-e4) and promoters (CpG-islands; CGI) in five wild type samples.

In general, macroH2A was higher enriched at exonic regions than at the CpG-islands in promoters of the region (Figure 3-12; Table A-3). Compared to H3K4me2-enrichment at chromosomal region 13q14.3, macroH2A was found to be much less enriched (reduced by a factor of ten; Figure 3-10 and Figure 3-12). Overall, enrichment of macroH2A at promoters and exons was reciprocal to H3K4me2-enrichment; i.e. while H3K4me2 was found at CpG-islands but not at exons, macroH2A was detected at exons but not at CpG-islands.

3.2.4 Comparison of histone code and DNA-methylation of chromosomal region 13q14.3

In order to determine the epigenetic code at the critical region in 13q14.3 in non-malignant cells, histone modifications were measured and, DNA-methylation was determined. The enrichment of the active chromatin mark H3K4me2 and of the repressive mark macroH2A in five to eleven non-malignant wild type samples at genes and CpG-islands in the region 13q14.3 is summarized in Figure 3-13a. The DNA-methylation at 13q14.3 was analyzed in nine non-malignant wild type samples by the aPRIMES method (chapter 2.19.3) in collaboration with Verena Fleig (Figure 3-13b). The epigenetic code of 13q14.3 was determined by correlating these three distinct epigenetic modifications to each other at genes and CpG-islands in the region (Figure 3-13).

KPNA3 and CGI A were neither remarkably enriched with H3K4me2 nor with macroH2A. Only one locus within the exons of *KPNA3* showed relatively high H3K4me2-enrichment (average

40 fold enrichment versus *CDH12*), whereas one locus in CGI A showed higher macroH2A-enrichment (average 6 fold enrichment versus *HK2*) compared to the other loci of *KPNA3* and CGI A (Figure 3-13a). DNA-methylation of CGIs A and B was determined by BioCOBRA analysis, which was performed by Angela Philippen. Thereby, 0 % DNA-methylation was measured at CGI A in non-malignant cells (data not shown). Thus, no distinct epigenetic modifications that could be specifically correlated to the transcriptional status of the gene were detected at *KPNA3* or CGI A.

At the exonic region of *C13ORF1* no remarkable enrichment, neither of H3K4me2 nor of macroH2A, was detected. In contrast, at CGI B located in the promoter region of the gene, comparatively high H3K4me2-enrichment (average 50 to 150 fold enrichment versus *CDH12*) and also relatively high macroH2A-enrichment was measured (average 5 to 12 fold enrichment versus *HK2*, Figure 3-13a). Also, DNA-methylation at CGI B varied from 0 % to 90 % in different samples, with an average of 40 % methylation of the CpG-island. The gene *C13ORF1* showed promoter-restricted H3K4me2-enrichment that correlates with its monoallelic expression status (chapter 3.2.2) and moreover, active and repressive chromatin marks were determined together at the promoter region in non-malignant cells.

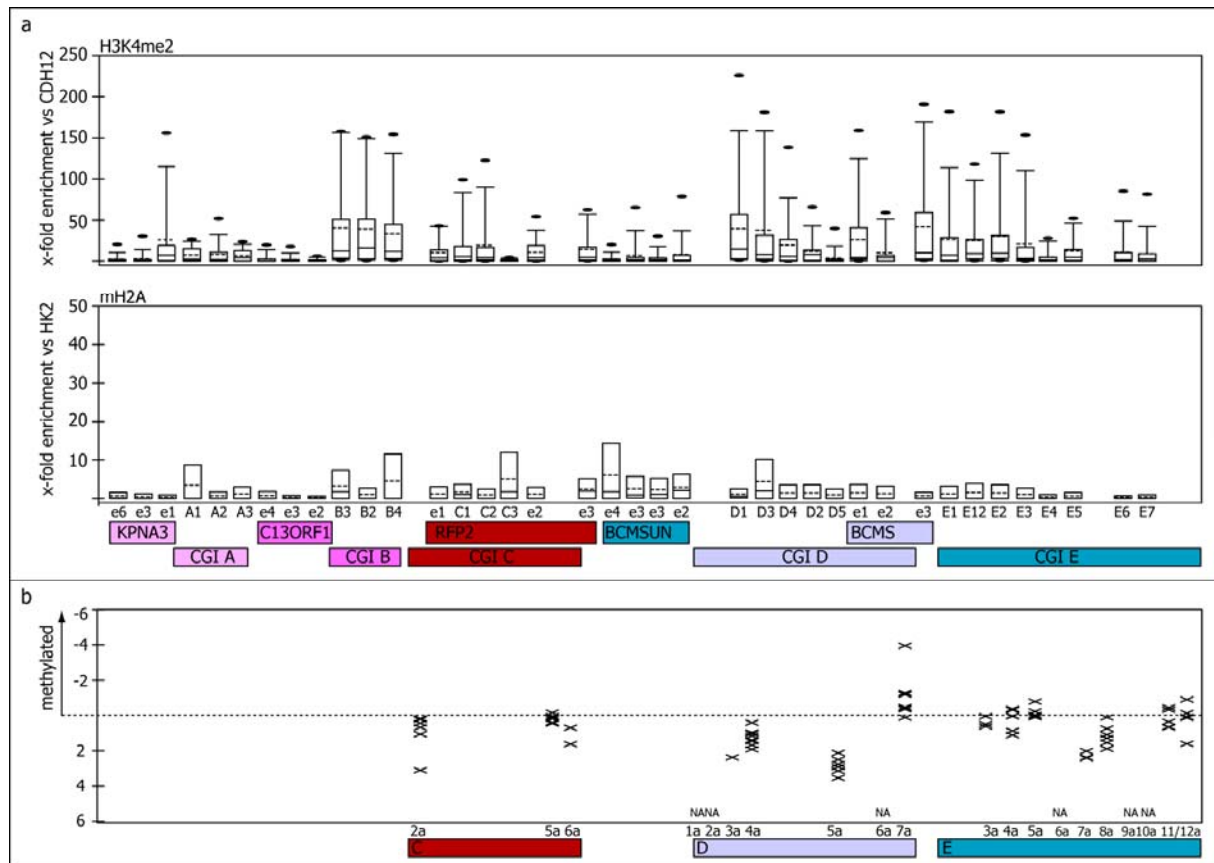


Figure 3-13: Wild type epigenetic modifications at 13q14.3. a) Box plots of histone modifications for each analyzed loci of 13q14.3 in wild type. b) Wild type DNA-methylation at 13q14.3 measured by aPRIMES. 13q14.3 not drawn to scale; relation of loci to each other (in a & b) is correct (compare overview of methods and analyzed fragments in Figure 2-3).

No distinct differences between gene and promoter sequences in enrichment of H3K4me2 and macroH2A were detected at *RFP2* and CGI C. Both histone modifications were found to be more or less homogeneously distributed in rather low amounts (i.e. H3K4me2: average

20 fold enrichment versus *CDH12*; macroH2A: average 5 fold enrichment versus *HK2*; Figure 3-13a). Moreover, no DNA-methylation was measured at CGI C (Figure 3-13b). Again, as seen at *KPNA3* and CGI A, no distinct epigenetic modifications that could be specifically correlated to the transcriptional status of the gene could be detected at *RFP2* or CGI A. However, in contrast to *KPNA3* and CGI A, both chromatin marks were found (in comparatively low amounts) at *RFP2* and CGI A.

CGI D and BCMS showed a rather heterogeneous enrichment of H3K4me₂, while macroH2A-enrichment was found to be homogeneously distributed in low amounts at the CpG-island and exonic region (Figure 3-13). H3K4me₂ was high at the first and third exon of *BCMS* (average 50 to 200 fold enrichment versus *CDH12*), but was significantly lower at the second exon (average 10 fold enrichment versus *CDH12*) in non-malignant cells. H3K4me₂ was high at the most centromeric locus within CpG-island D (average 50 to 230 fold enrichment versus *CDH12*) and became gradually less towards its most telomeric part (average of 5 to 50 fold enrichment versus *CDH12*, Figure 3-13a). In contrast, macroH2A was medium enriched at CGI D with an average of 2 fold enrichment versus *HK2* with one exception at locus D3 (average of 6 fold up to 12 fold enrichment versus *HK2*). The DNA-methylation determined at CGI D and BCMS divided the CpG-island into two parts: No DNA-methylation was found at the centromeric part, but DNA was methylated at the most telomeric part (D7a Figure 3-13b). Thus, H3K4me₂-enrichment and DNA-methylation were found to be significantly different at the most telomeric part of CGI D compared to the (larger) centromeric part (Figure 3-13). A similar division into two differentially methylated segments was also detected at CGI E.

The exonic region of the gene *DLeu2/BCMSUN* was comparable to the other exonic regions in 13q14.3; H3K4me₂-enrichment was low (average 20 fold enrichment versus *CDH12*), but macroH2A was high with an average of 8 fold enrichment versus *HK2* (Figure 3-13a). At the CpG-island located within the promoter region of the gene (i.e. CGI E), H3K4me₂ enrichment could be divided into two halves. The centromeric half of CGI E had high levels of H3K4me₂ (average 40 to 160 fold enrichment versus *CDH12*), whereas the telomeric part had significantly lower levels (average 20 fold enrichment versus *CDH12*, Figure 3-13a). Surprisingly, the centromeric half of CGI E showed low DNA-methylation in non-malignant cells, whereas the neighboring telomeric half did not (Figure 3-13b). An abrupt change of DNA-methylation was measured between the two halves at locus E6a. At locus E6a the average DNA-methylation was around 5 % to 20 %, and up to 80 % in one sample, whereas in the neighboring loci of E6a, 0 % DNA-methylation was measured in non-malignant cells. Analog to *C13ORF1* and CGI B, promoter-restricted H3K4me₂-enrichment was determined at *DLeu2/BCMSUN* and CGI E that correlates with monoallelic expression of *DLeu2/BCMSUN* in non-malignant cells (chapter 3.2.2). Furthermore, CGI E could be divided into two halves according to epigenetic marks determined and a region of interest that showed special chromatin marks at one distinct locus within CGI E was identified.

Remarkably, in both CpG-islands located in the 5'-regions of the two noncoding RNA genes of the critical region in 13q14.3, *BCMS* and *BCMSUN*, abrupt changes in DNA-methylation became evident that divide the CpG-islands into a methylated and an unmethylated part. The two distinct loci identified within CGIs D and E were further analyzed to determine further characteristics of these putative locus control elements.

3.2.5 Allelic modifications of CpG-islands D and E – Is there a regulatory element in CGI D or CGI E?

Locus control regions (LCR) are characterized among other features by differential allelic methylation of DNA and by differential allelic histone modifications (Li et al., 2002). ChIP experiments were carried out in non-malignant cells that were heterozygous for SNPs within the CpG-islands in order to detect differential allelic enrichment of H3K4me2 and macroH2A.

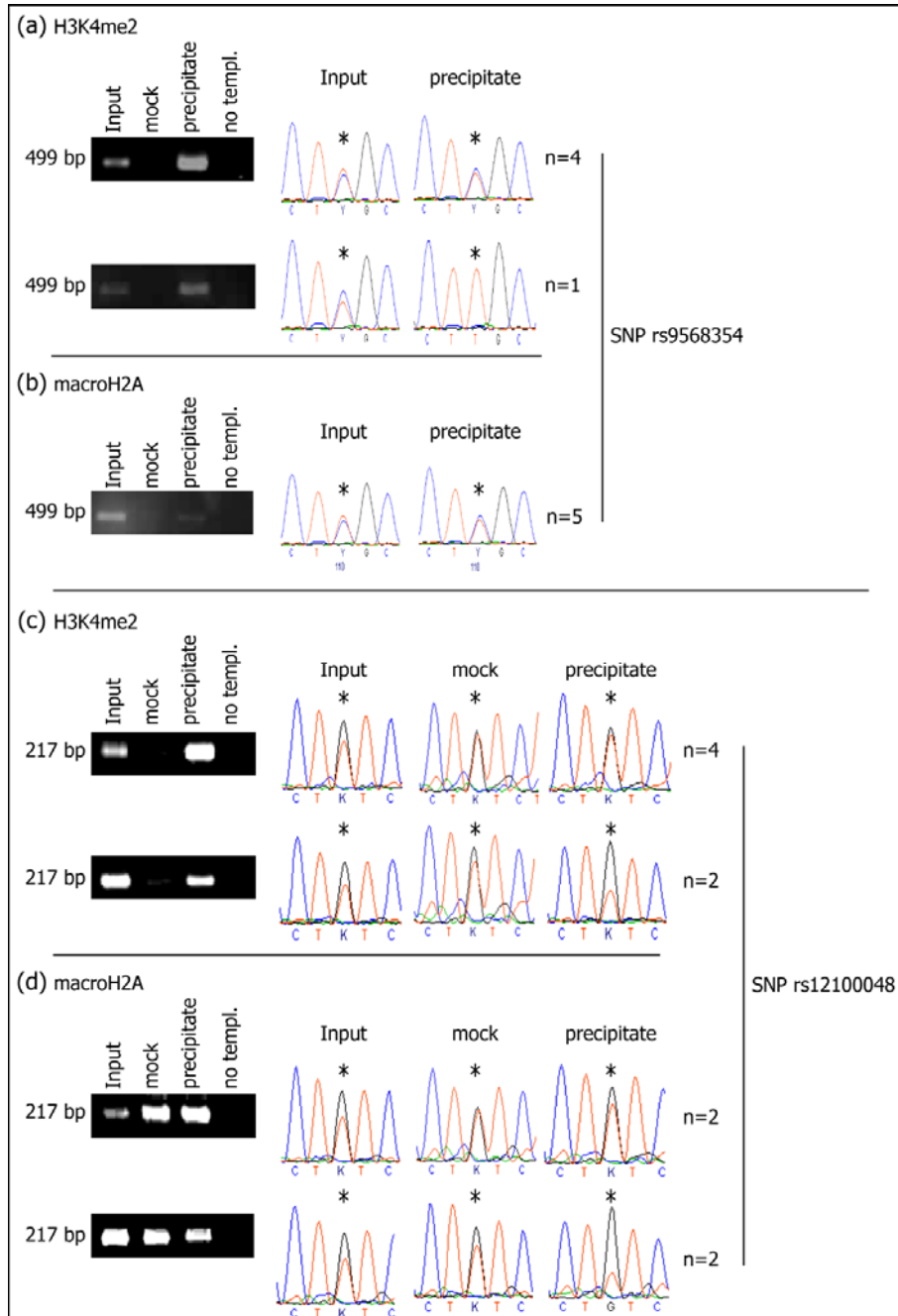


Figure 3-14: Allelic enrichment of H3K4me2 and macroH2A at distinct loci within CpG-islands D and E. a) allelic H3K4me2-enrichment at SNP rs9568354 within CGI E (E6); sequences of input DNA versus precipitated DNA, measured in n=5 probands. b) Allelic macroH2A-enrichment at CGI E (n=5). c) Allelic H3K4me2 at SNP rs12100048 within CGI D measured in n=6 probands; d) allelic macroH2A-enrichment at CGI D (n=4).

Especially interesting were the loci with abrupt changes of DNA-methylation in CGI D (SNP rs12100048) and CGI E (SNP rs9568354; chapter 3.2.4). Allelic H3K4me2-enrichment at CpG-island E (was detected in 1 out of 5 wild type samples, with only one allele detectable in precipitated sequences (Figure 3-14a). In contrast, no differential enrichment of macroH2A was detected at CGI E (Figure 3-14b). CpG-island D was neither significantly differentially enriched with H3K4me2 (Figure 3-14c) nor with macroH2A (Figure 3-14d) in any analyzed sample. Unfortunately, end-point-PCR reactions performed prior to sequencing in collaboration with Angela Philippen were unequally specific in individual samples (Figure 3-14c and d). Mostly, but not in all cases, the target could be amplified also from the unspecifically bound DNA in mock fractions as is exemplified in Figure 3-14d. Two of six (H3K4me2) and two of four (macroH2A) analyzed heterozygous probands depicted in Figure 3-14c and (d) show a preferred enrichment of the chromatin marks at only one of the two copies, which was evident in unequal sizes of curves determined at the SNP (*). The size of the sequencing curves is correlated with the amount of DNA from the respective allele present in the analyzed samples. Thus, in both putative LCR regions of 13q14.3, a tendency for monoallelic enrichment of the two histone marks was detected.

3.2.6 Histone modifications in different age groups – What influence has the age on epigenetic modifications at 13q14.3?

As CLL patients have a median age of 72 years at diagnosis (Dighiero and Hamblin, 2008), age-matched wild type samples were measured as appropriate control tissue and compared to non-age-matched wild type samples. For neither modification, H3K4me2- and macroH2A-enrichment, significant differences could be detected between young (<45 years) and old (>45 years) probands at genes and CpG-islands of chromosomal region 13q14.3 (Figure 3-15; Table 3-4). However, H3K4me2 tended to be stronger enriched in the old age group at all genes and CpG-islands analyzed. Calculated t-values were used to determine which sample group contained the higher enrichment (Table 3-4).

Table 3-4: T- and p-values calculated by Welch's two sample t-test; for comparison of young versus old subgroups' enrichment of H3K4me2 and macroH2A. T-values implicate which group contains greater values: negative t = young < old; positive t = young > old wild type.

Young vs old wild type	H3K4me2		macroH2A	
	p-value	t-value	p-value	t-value
KPNA3 exons	0.298	-1.062	0.141	-1.671
CGI A	0.214	-1.270	0.332	1.028
C13ORF1 exons	0.288	1.093	0.502	0.691
CGI B	0.274	-1.114	0.187	1.438
CGI C	0.234	-1.213	0.466	0.757
RFP2 exons	0.759	-0.310	0.351	0.972
CGI D	0.246	-1.184	0.110	1.688
BCMS exons	0.898	0.130	0.519	0.666
BCMSUN exons	0.953	0.060	0.081	1.901
CGI E	0.126	-1.547	0.914	0.109

Only at three loci the young age group showed a higher H3K4me2-enrichment than the old age group, these loci were *C13ORF1* ($t=1.1$), *BCMS* ($t=0.1$) and *DLeu2/BCMSUN* ($t=0.1$; Table 3-4). The old age group showed a much higher variance of enrichment ratios at distinct genes and CpG-islands than the young age group. This was observed by comparison of 95th percentiles in young and old box plots and taking into account the outliers of the box plots in each age group (Figure 3-15a). Enrichment of macroH2A at chromosomal region 13q14.3 was reciprocal to the H3K4me2-enrichment. Higher macroH2A-enrichment was detected in the young age group with only one exception at *KPNA3* ($t=-1.6$; Table 3-4 and Figure 3-15b). Here the young age group showed a greater variance of enrichment ratios at single loci (Figure 3-15b).

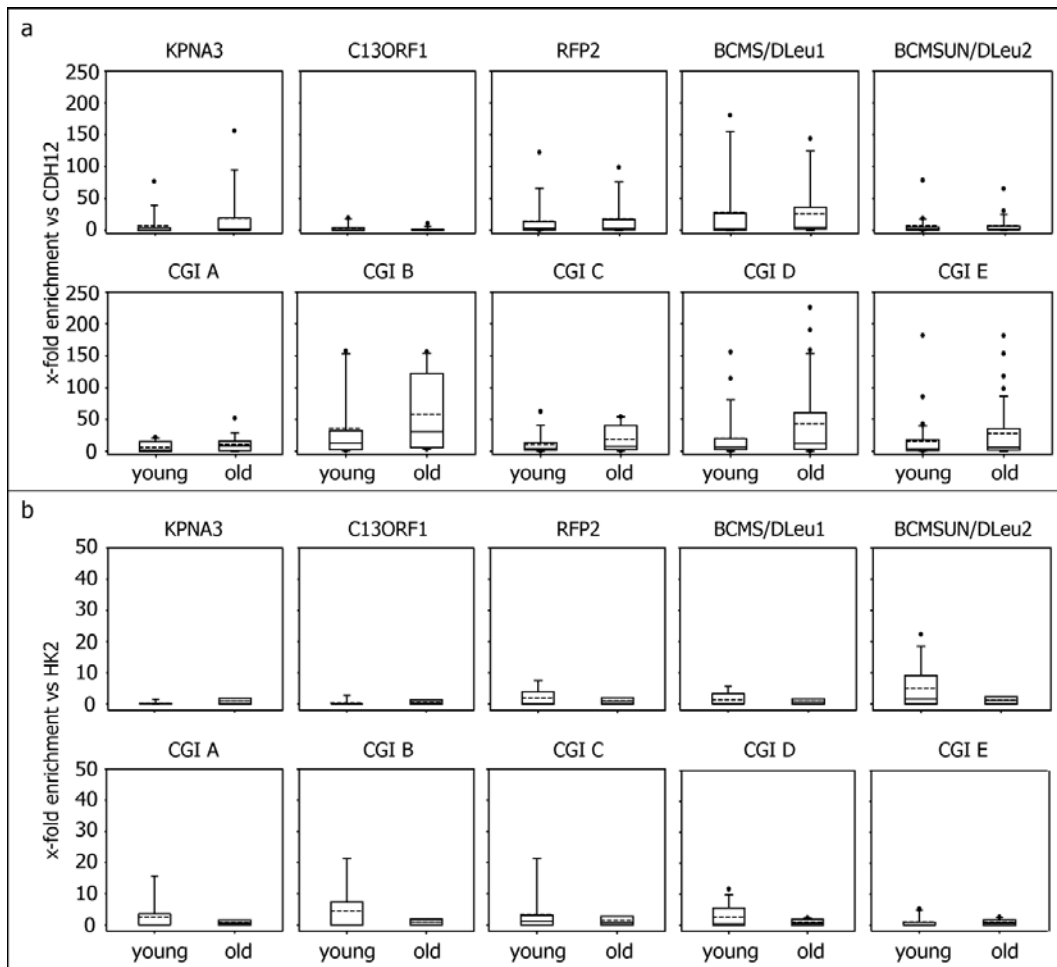


Figure 3-15: Histone modifications compared in two different age groups. a) Box plots comparing H3K4me2-enrichment in young and old wild type groups at five genes and CpG-islands of chromosomal region 13q14.3. b) macroH2A-enrichment as in (a). Dotted line within box plots = mean value.

The comparison of two age groups showed no significant differences regarding histone modification changes; only tendencies were observed. Old probands showed stronger enrichment of H3K4me2 (Figure 3-15a), while young probands showed higher macroH2A-enrichment (Figure 3-15b).

3.2.7 Differences in H3K4me2-enrichment between B- and T-cells

CLL is a B-cell disease. In the present study, epigenetic modifications of CLL cells were compared to that of non-malignant peripheral blood lymphocytes (PBL), a cell sample that includes B-cells, but also T-cells, Natural killer (NK) cells and macrophages. To ensure that epigenetic CLL-specific changes could be detected when comparing CLL cells to this mixed cell population, H3K4me2-enrichment was also separately determined in CD19⁺-sorted B-cells and in T-cells (Figure 3-16). Comparison of H3K4me2-enrichment in B- and in T-cells showed that B-cells had around 10 fold more H3K4me2 at 13q14.3 than T-cells (Figure 3-16). The highest median enrichment in both B- and T-cells was measured at CGI B, CGI D and the centromeric half of CGI E (4.1 to 28.3 fold enrichment versus *CDH12* in B-cells and 0.2 to 3.1 fold in T-cells; Table A-2, Appendix A), whereas lowest median H3K4me2-enrichment was measured at *KPNA3*, *C13ORF1* and *DLeu2/BCMSUN* (Figure 3-16 and Table A-2, Appendix A). Thus, the general pattern of H3K4me2-enrichment was the same in B- and T-cells. As previously determined also in PBL samples (chapter 3.2.1), the active chromatin mark was higher enriched at CpG-islands than at exonic regions of 13q14.3.

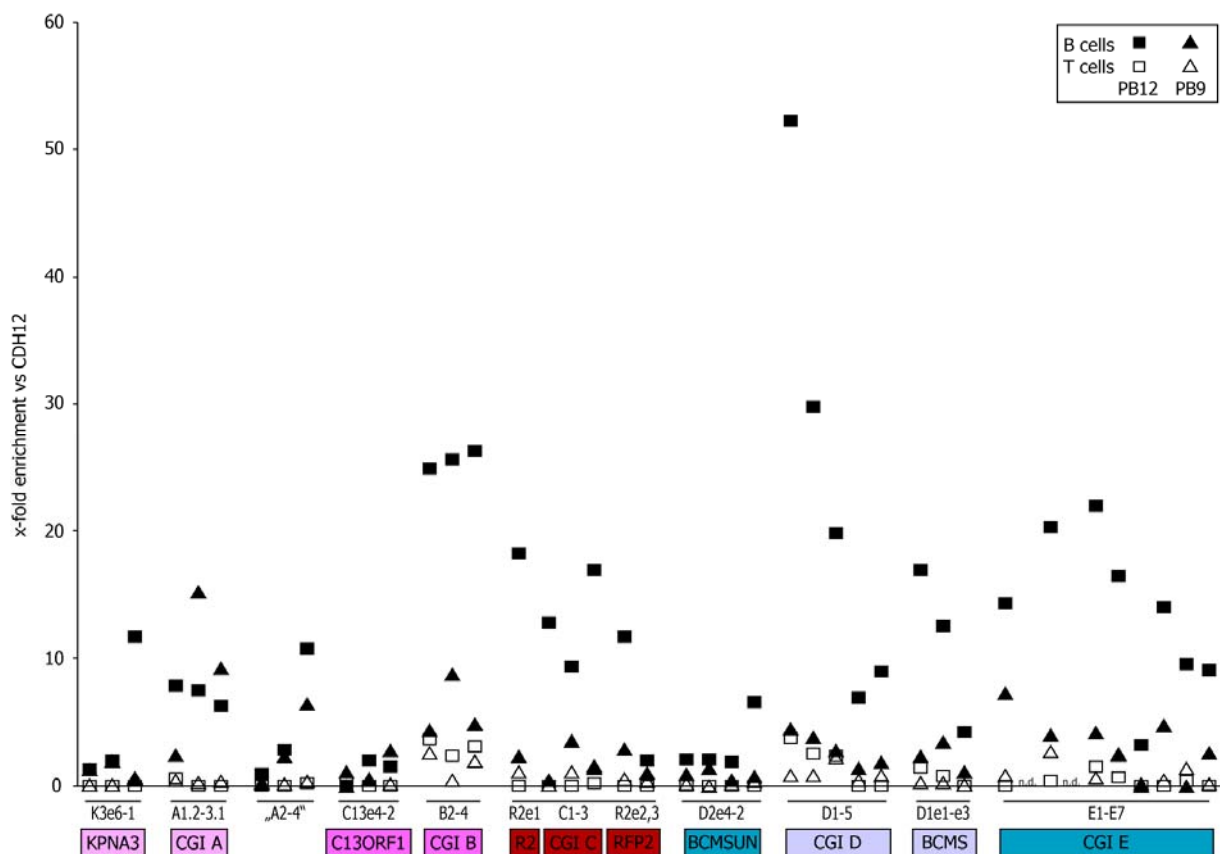


Figure 3-16: H3K4me2-enrichment at chromosomal region 13q14.3 in wild type B- (black symbols) and T-cells (white symbols).

In summary, no other differences than the general 10-fold higher enrichment of H3K4me2 at 13q14.3 in B-cells, was detected between B- and T-cells. The enrichment pattern of H3K4me2 with high enrichment at CGIs and low enrichment at gene exons was the same in B-cells, T-cells and PBL samples (Figure 3-10; Figure 3-16). Therefore, it seemed to be feasible to determine aberrations of the epigenetic code of 13q14.3 in CLL by comparison of epigenetic modifications of CLL cells to non-malignant PBL samples (chapter 3.3).

3.3 Aberrant chromatin of chromosomal region 13q14.3 in CLL

Epigenetic modifications were characterized at chromosomal region 13q14.3 in wild type (chapter 3.2) and in CLL samples. Chapter 3.3 describes characterization of two histone modifications, H3K4me2- and macroH2A-enrichment, and of DNA-methylation at 13q14.3 in CLL samples. Aberrant chromatin and epimutations in CLL were detected by comparing to chromatin of non-malignant cells in the region.

3.3.1 H3K4me2-enrichment at 13q14.3 in CLL

Enrichment of H3K4me2 at chromosomal region 13q14.3 was characterized in five CLL samples to identify epigenetic marks at genes and promoters that reflect the transcriptional status of the genes. Overall H3K4me2-enrichment at all measured loci of the region in CLL cells was significantly higher (10 to 100 fold) than in non-malignant cells (Figure 3-10) and compared to the control promoter *CDH12* (up to 1000 fold; Figure 3-17).

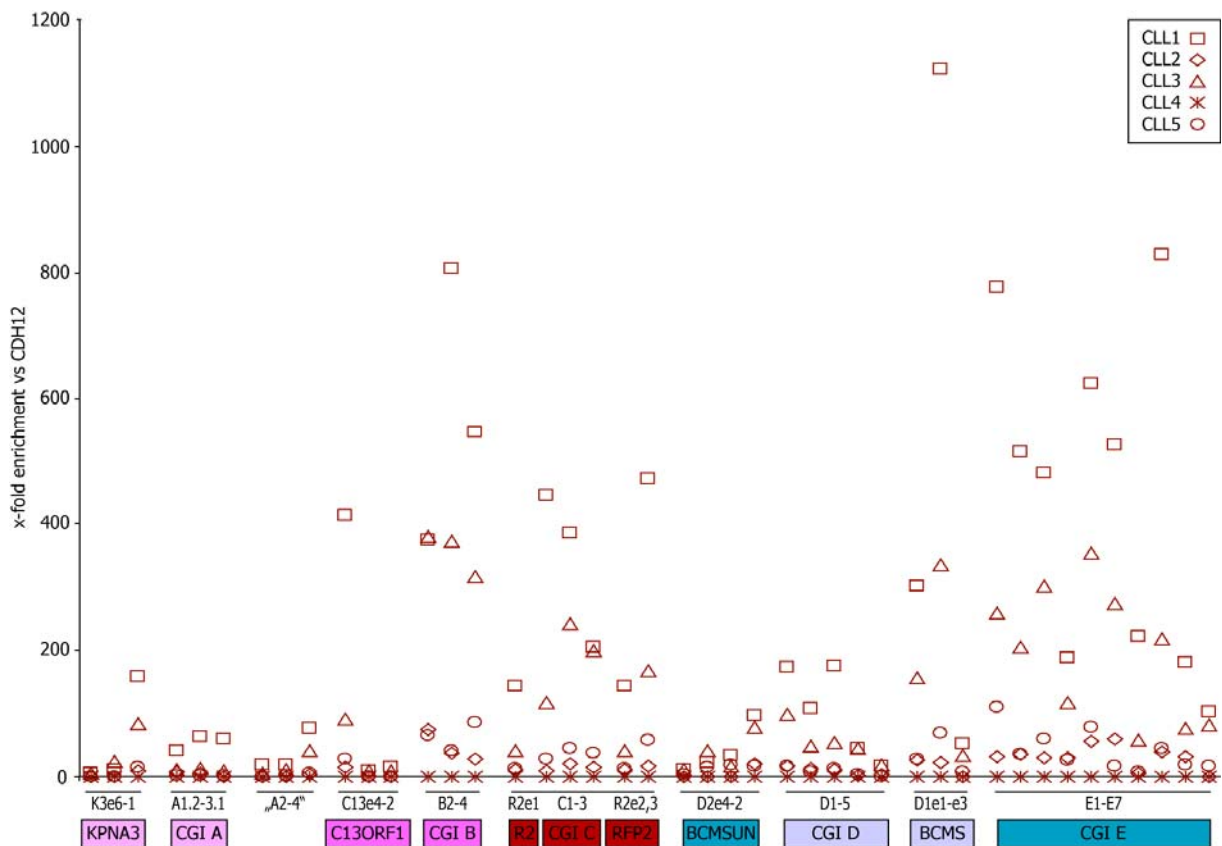


Figure 3-17: H3K4me2-enrichment at chromosomal region 13q14.3 in five CLL samples.

Highest median H3K4me2-enrichment was detected at CpG-islands B, C and the centromeric half of CGI E (27 to 109 fold enrichment versus *CDH12*). In contrast to non-malignant cells, CGI D had no significant high H3K4me2-enrichment in CLL cells (Figure 3-17) whereas high H3K4me2 was measured at the genes *RFP2* and at *BCMS* (58 and 28 to 68 fold enrichment versus *CDH12*, respectively; Table A-4, Appendix A). The overall pattern of H3K4me2-enrichment at 13q14.3 was similar in all CLL cell samples. However, when dividing the CLL

samples into subgroups according to genomic aberrations identified in the CLL patients (Table A-4, Appendix A), it was determined that del(13q) CLL samples had significantly lower H3K4me2-enrichment than CLL samples with retention of both copies (see chapter 3.3.6 below for details). The pattern of high H3K4me2-enrichment at CGIs and low enrichment at exonic regions was generally the same as in non-malignant cells. Interestingly, CGI D and CGI A had the lowest measured H3K4me2-levels of all five CGIs analyzed in CLL cells.

In summary, the most striking difference between non-malignant and CLL cells regarding enrichment of the active chromatin mark at the region 13q14.3, was quantitative, i.e. enrichment of H3K4me2 was at least 10 fold higher in CLL samples. High levels of H3K4me2 were measured at 13q14.3 in CLL cells and the pattern of H3K4me2-enrichment was the same in all samples. However, the levels of H3K4me2 varied largely between individual CLL samples analyzed (Figure 3-17).

3.3.2 Promoter-restricted H3K4me2-pattern correlates with monoallelic expression and is not changed in CLL

The histone pattern of monoallelically expressed genes (Rougeulle et al., 2003) of 13q14 was already characterized in non-malignant cells (chapter 3.2.2). In an analogous manner, promoter-restricted H3K4me2-enrichment at chromosomal region 13q14.3 was characterized in CLL cells. Figure 3-18 shows, analogous to Figure 3-11, box plots of ratios determined for gene exons and gene promoters of five genes of 13q14.3 analyzed in five CLL samples. Promoter-restricted H3K4me2-enrichment in CLL was detected at *C13ORF1* ($p=0.0238$) and with high significance at *DLeu2/BCMSUN* with the promoter CGI E ($p=0.0006$), but neither at *DLeu2/BCMSUN* with the second eligible promoter (Figure 2-3; CGI D; $p=0.2007$), nor at *KPNA3* ($p=0.6222$), *BCMS* ($p=0.1737$) and *RFP2* ($p=0.29$; Figure 3-18).

The epigenetic pattern of monoallelic enrichment was retained in CLL at all genes except for one splicing variant of *DLeu2/BCMSUN* in combination with CpG-island D. Thus, the epigenetic mark at one splicing variant of the gene *DLeu2/BCMSUN* is not detectable in CLL cells anymore.

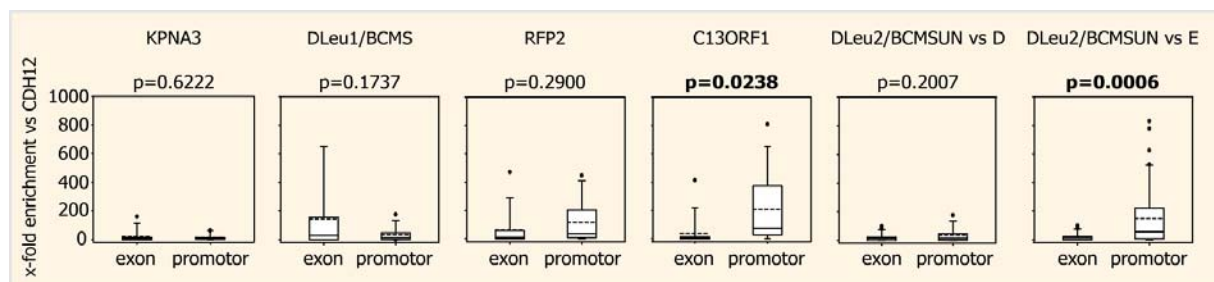


Figure 3-18: H3K4me2-enrichment at 13q14.3-genes in CLL at exons and at promoters. Dotted line within each box plots shows mean value; p-values calculated by Welch's two sample t-test; bold p values = significant ($p < 0.05$).

3.3.3 MacroH2A-enrichment at 13q14.3 in CLL

The enrichment of macroH2A at chromosomal region 13q14.3 was characterized by ChIP and qPCR analyses in five CLL samples in order to identify (aberrantly) silenced loci at chromosomal region 13q14.3. MacroH2A in CLL cells was high compared to non-malignant

cells (Figure 3-12; Figure 3-19). Medium levels of macroH2A (30 fold enrichment versus *HK2*) were distributed over the chromosomal region 13q14.3. Overall, macroH2A was detected in lower amounts than H3K4me2 at the 13q14 region in CLL cells (Figure 3-19 and Table A-5, Appendix A). The highest median macroH2A-enrichment at 13q14.3 in CLL cells was measured at exonic regions of the five genes analyzed (0.5 to 3.5 fold enrichment versus *HK2*; Table A-5, Appendix A). Surprisingly, also CpG-islands A, B, C and E had comparatively high macroH2A (0.1 to 3.5 fold enrichment versus *HK2*), whereas no macroH2A was detected at CGI D and at *KPNA3* in CLL cells (Figure 3-19; Table A-5, Appendix A). Interestingly, the del(13q) subgroup of CLL samples had more macroH2A at 13q14.3 than the subgroup of CLL samples that retained both copies of the critical region (Table A-5, Appendix A).

The change of macroH2A-enrichment in CLL cells compared to non-malignant cells was again quantitative, i.e. as seen for H3K4me2 enrichment, also macroH2A enrichment was measured in higher amounts at 13q14.3 in CLL cells. Furthermore, the pattern of macroH2A-enrichment was changed in CLL cells in such that it was found also in high amounts at CGIs and not exclusively at exonic regions in 13q14.3.

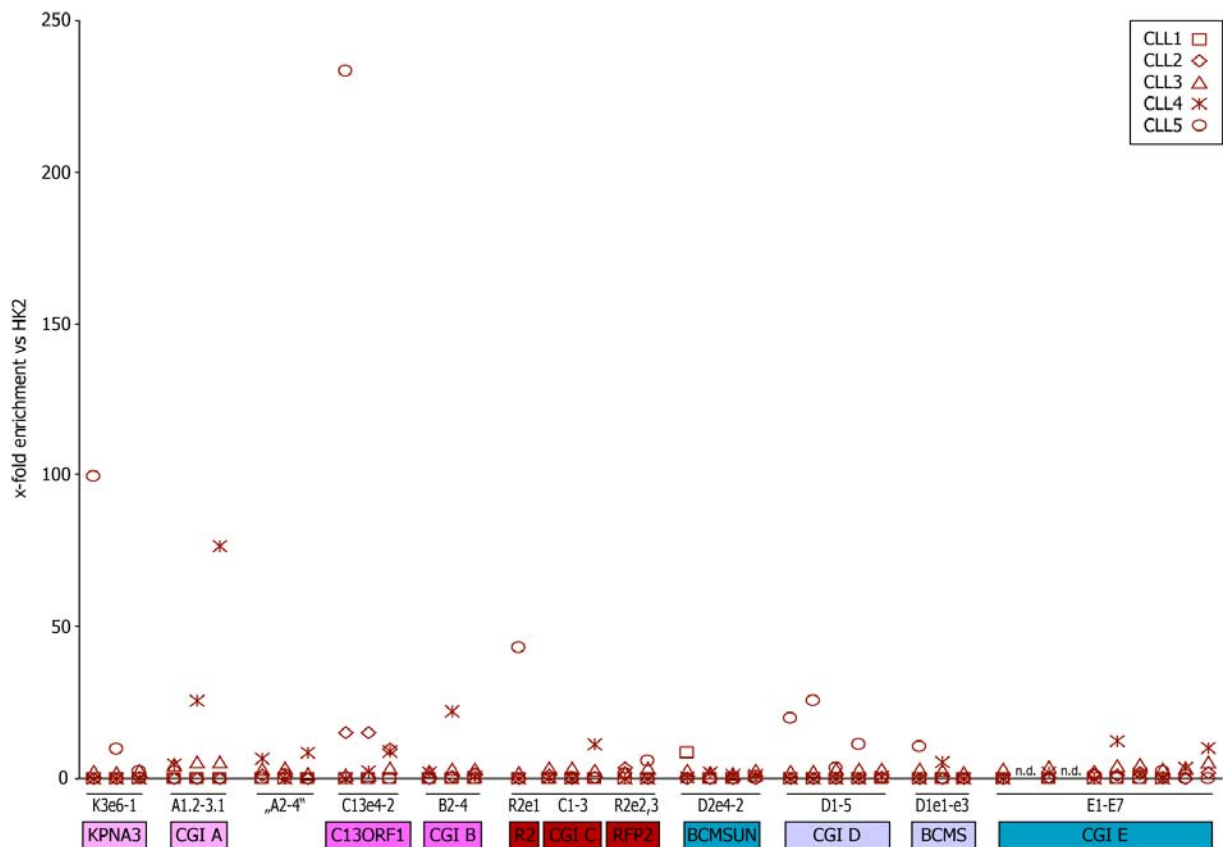


Figure 3-19: MacroH2A-enrichment at chromosomal region 13q14.3 in five CLL samples.

3.3.4 Histone code and DNA-methylation of chromosomal region 13q14.3 in CLL

In order to determine the aberrant epigenetic code at the chromosomal region 13q14.3 in CLL cells, histone modifications and DNA-methylation were determined at the five genes and

3 Results

CpG-islands in the region. Enrichment of H3K4me2 and macroH2A as well as DNA-methylation at 13q14.3 is summarized in Figure 3-20. The epigenetic code at 13q14.3 in CLL cells was determined by correlating the three marks to each other at genes and promoter regions. Subsequently the determined epigenetic code of CLL cells was analyzed for correlations to the epigenetic code determined in non-malignant cells (chapter 3.2.4).

KPNA3 and CGI A were neither remarkably enriched with H3K4me2 nor with macroH2A in CLL cells with only two exceptions. In the exonic region of *KPNA3* comparatively high amounts of H3K4me2 were measured (100 fold enrichment versus *CDH12*), whereas CGI A showed high macroH2A-enrichment (15 fold enrichment versus *HK2*; Figure 3-20a). In CLL cells 0 % DNA-methylation was detected at CGI A (A. Philippen, personal communication). Thus, no distinct epigenetic code that marks the transcriptional status of *KPNA3* was determined in CLL cells.

Except for comparatively high H3K4me2-enrichment at the fourth exon of *C13ORF1* (average 140 fold enrichment versus *CDH12*), neither H3K4me2- nor macroH2A-enrichment was measured at the exonic region of the gene (Figure 3-20a). H3K4me2-enrichment was found to be promoter-restricted in CLL cells, i.e. CGI B was strongly enriched, whereas *C13ORF1* was not. DNA-methylation was observed to be decreased in CLL cells at CGI B. In contrast to non-malignant cells where average DNA-methylation was 40 %, methylation was decreased to 20 % (range: 0 % to 60 %) in CLL cells.

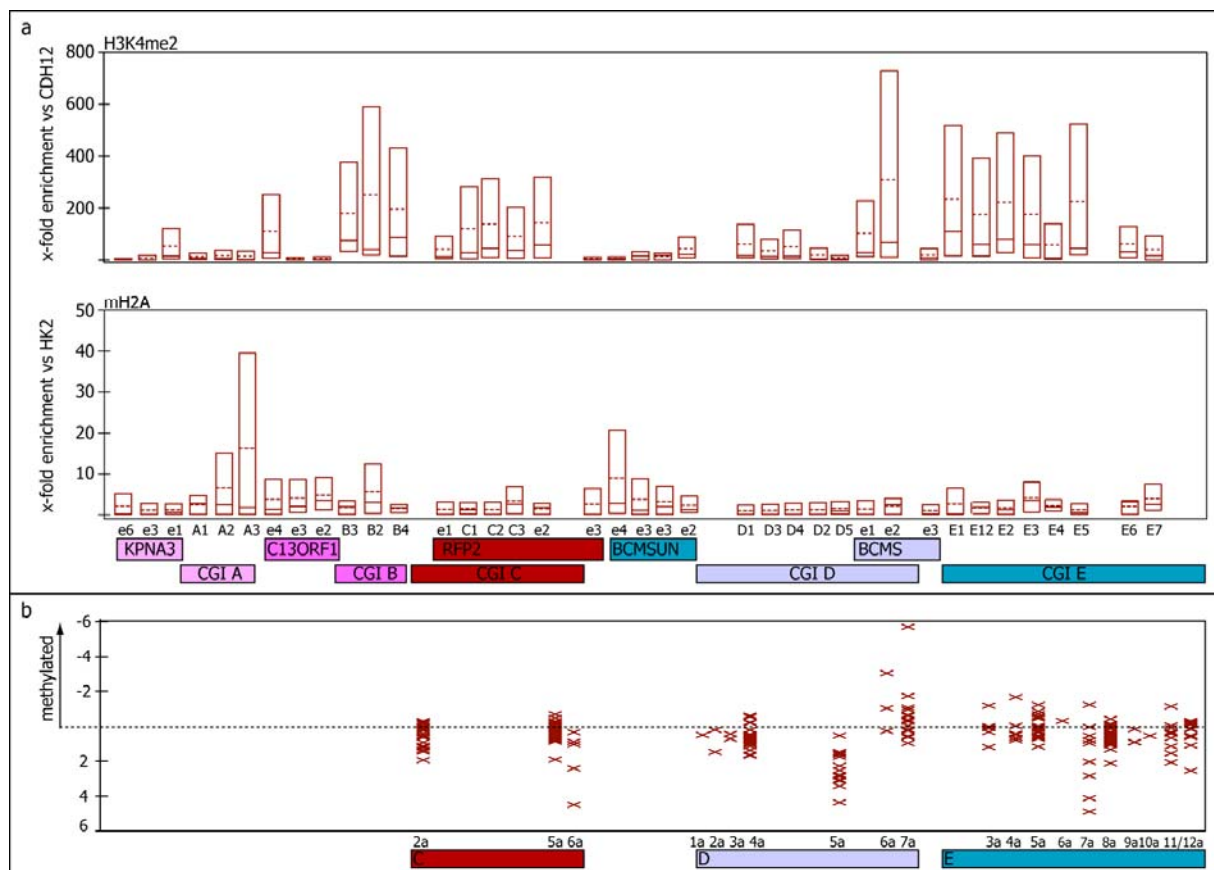


Figure 3-20: Epigenetic modifications at 13q14.3 in CLL samples. a) Box plots of histone modifications H3K4me and macroH2A for each fragment of 13q14.3 in five CLL samples. b) DNA-methylation at 13q14.3 measured by aPRIMES. 13q14.3 loci indicated are not drawn to scale, but their relation to each other in (a) and (b) is correct.

The overall epigenetic code of *C13ORF1* in CLL cells was merely the same as determined in non-malignant cells (chapter 3.2.4): promoter-restricted H3K4me2-enrichment at *C13ORF1* and CGI B was determined, together with presence of both activating (i.e. H3K4me2) and repressive (macroH2A and DNA-methylation) chromatin marks in CLL cells.

At *RFP2* and CGI C that are located within the critical region in 13q14.3, high amounts of H3K4me2 (average 200 fold enrichment versus *CDH12*) and medium macroH2A-enrichment was measured in CLL cells (Figure 3-20a). Furthermore, some of the analyzed samples showed DNA-methylation at CGI C (Figure 3-20b). Thus, epigenetic modifications in CLL cells differ from that in non-malignant cells, where neither DNA-methylation nor H3K4me2-enrichment was measured (chapter 3.2.4). However, no distinct epigenetic code specific for the transcriptional state of *RFP2* could be determined in CLL cells as both activating and repressive marks were measured at the exonic region and at CGI C (Figure 3-20).

At CGI D medium H3K4me2- and low macroH2A-enrichment was measured (average of 80 fold enrichment versus *CDH12* and average of 3 fold enrichment versus *HK2*, respectively) in CLL cells. However, at *BCMS*, which is located in the most telomeric part of the CpG-island D, H3K4me2-enrichment was significantly higher and more variable than measured at the centromeric CGI. Furthermore, CGI D showed no wide DNA-methylation in the centromeric part, only the most telomeric parts showed DNA-methylation in CLL cells (Figure 3-20b). DNA-methylation at the telomeric part of CGI D is less abundant in CLL cells than in non-malignant cells (chapter 3.2.4). A change of the methylation pattern between the two adjacent loci, D6a and D7a was evident (Figure 3-20b); at the first locus (D6a) 10 to 100 % DNA-methylation was measured, whereas the second locus (D7a) had less methylation varying between 0 % and 40 %. The division of CGI D into two different parts according to its epigenetic code was determined in CLL cells and in non-malignant cells (Figure 3-20; chapter 3.2.4). However, the most telomeric part including the candidate LCR carried repressive epigenetic marks in non-malignant cells (Figure 3-13), whereas it had mostly active marks in CLL cells (Figure 3-20).

At the exons of *DLeu2/BCMSUN* no H3K4me2- and low macroH2A-enrichment was measured. The centromeric half of CGI E was comparatively highly enriched with H3K4me2, while it was low at the telomeric half (Figure 3-20a). MacroH2A was homogeneously low enriched at CGI E with a slightly higher enrichment at the centromeric half (Figure 3-20a). Promoter-restricted H3K4me2-enrichment that reflects monoallelic expression of *DLeu2/BCMSUN* was detected in CLL cells. In non-malignant cells the locus E6a showed DNA-methylation, while the neighboring loci did not (Figure 3-13b). However, in CLL cells E6a and its neighboring loci showed DNA-methylation in the minority of CLL samples analyzed (Figure 3-20b). Thus, repressive DNA-methylation but activating histone modifications could be determined at the candidate LCR within CGI E in CLL cells, whereas a repressive epigenetic code was determined in non-malignant cells (chapter 3.2.4).

In summary, two candidate LCR loci within CGIs D and E that were determined to be differentially modified compared to the rest of the respective CGIs in non-malignant cells (chapters 3.2.6 and 3.2.7), also showed differential epigenetic modifications in CLL cells. Strikingly, the DNA-methylation status of these candidate LCRs is different in CLL compared to non-malignant cells (see also below, chapter 3.3.5). Whereas both candidate LCRs in non-malignant cells had predominantly repressive epigenetic modifications including distinct DNA-methylation (Figure 3-13), the candidate LCR in CGI D had activating epigenetic modifications and the one in CGI E had both active and repressive chromatin marks in CLL cells (Figure 3-20).

3.3.5 Aberrant chromatin at 13q14.3 in CLL?

In order to identify aberrant chromatin at 13q14.3 in CLL cells, ratios of H3K4me2- and macroH2A-enrichment were compared in CLL and wild type samples by Welch's Two Sample t-test (chapter 2.20.3). Figure 3-21 shows box plots of the ratios determined for analyzed CLL samples (CLL) and for analyzed wild type samples (wild type) at distinct loci of chromosomal region 13q14.3 (genes and CpG-islands). Significant differences of H3K4me2-enrichment between CLL and wild type were determined at *DLeu2/BCMSUN* ($p=0.039$), CpG-island B ($p=0.0175$), CpG-island C ($p=0.0155$) and CpG-island E ($p=0.0006$). The significant differences resulted in all cases from an aberrantly higher H3K4me2-enrichment in CLL samples (Figure 3-21a). Predominantly, aberrant enrichment of H3K4me2 was detected at CpG-islands located within gene promoter regions that are responsible for regulation of gene expression and not at the coding region of analyzed genes. *DLeu2/BCMSUN* was the only aberrantly H3K4me2-enriched gene of chromosomal region 13q14.3 (Figure 3-21a).

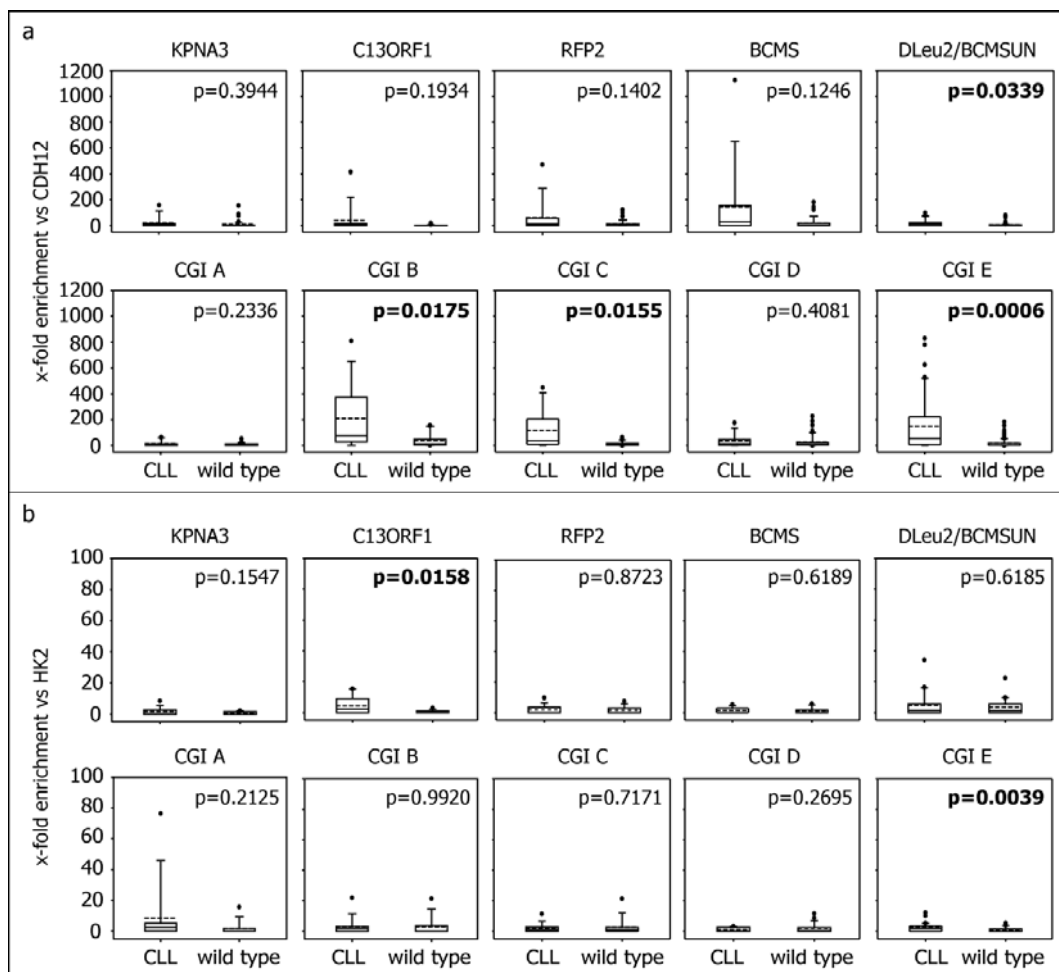


Figure 3-21: Differential enrichment of histone modifications at chromosomal region 13q14.3 in CLL compared to wild type. a) Box plots of differential H3K4me2-enrichment for five genes and five CpG-islands of the region, b) box plots analog to (a) showing macroH2A-enrichment; dotted lines within box plots = mean values; bold p values = significant ($p<0.05$).

Enrichment of macroH2A at chromosomal region 13q14.3 was merely the same in CLL and wild type as no significant differences could be determined. Only two exceptions were detected, *C13ORF1* ($p=0.0158$) and CpG-island E ($p=0.0039$) were significantly differentially

macroH2A-enriched in CLL compared to wild type (Figure 3-21b). While macroH2A was predominantly detected at the gene exons in non-malignant cells, it was significantly higher enriched at CGI E located within the promoter region of the ncRNA gene *DLeu2/BCMSUN*. Similar to H3K4me2-enrichment, macroH2A enrichment was generally higher in CLL samples than in wild type samples. However, at most analyzed loci the mean values determined for CLL and wild type samples did not differ significantly (Figure 3-21).

CLL mainly affects elderly adults, as can be seen from the median age of onset of 72 years. To this reason, wild type samples that were analyzed to determine aberrations in CLL were further subdivided into old and young wild type subgroups (see also chapter 3.2.6; Table 3-5). The enrichment of both modifications H3K4me2 and macroH2A, in these two wild type subgroups were compared to those of the CLL group. The only difference between the whole wild type group and the two subgroups, when comparing them to CLL, was determined at the gene *DLeu2/BCMSUN* (yellow fill in Table 3-5).

However, differential enrichment of H3K4me2 at *DLeu2/BCMSUN* was not significant between the young wild type subgroup and CLL and neither between the old subgroup and CLL ($p=0.0642$ and $p=0.089$, respectively). All other determined p-values had the same tendency as those calculated for the average of all wild type samples versus CLL. Significant differences were determined at CGI B, CGI C and CGI E for H3K4me2-enrichment in young and old wild type versus CLL and at *C13ORF1* and CGI E for macroH2A-enrichment (Table 3-5). At the same loci significant differences were previously determined for the comparison of wild type versus CLL (Figure 3-21). This completes the previous finding in non-malignant cells that no significant differences in enrichment of neither H3K4me2 nor macroH2A could be detected at 13q14.3 between the two wild type subgroups young and old (Figure 3-15).

Table 3-5: P values determined by Welch's two sample T test for comparison of the CLL group to wild type subgroups young and old. Bold p values = significant ($p<0.05$). Yellow fill: different from CLL versus complete wild type group

13q site	CLL vs old wild type		CLL vs young wild type	
	H3K4me2	mH2A	H3K4me2	mH2A
	p-value	p-value	p-value	p-value
KPNA3	0.8861	0.5196	0.2861	0.0608
CGI A	0.5615	0.1458	0.1673	0.2771
C13ORF1	0.1887	0.0211	0.2113	0.0139
CGI B	0.0344	0.1648	0.0174	0.6276
CGI C	0.0258	0.6614	0.0147	0.5927
RFP2	0.1878	0.3855	0.1615	0.8272
CGI D	0.5972	0.5144	0.2519	0.1718
BCMS	0.1410	0.2700	0.1508	0.9058
BCMSUN	0.0642	0.0878	0.0890	0.8824
CGI E	0.0003	0.0126	0.0001	0.0113

3.3.6 Correlation of macroH2A- and H3K4me2-enrichment at 13q14.3 with different karyotypes in CLL

Data before was shown as CLL median values calculated from single CLL samples of five different patients. A next interesting was to analyze and compare differences in enrichment of histone modifications at 13q14.3 in each single patient (Figure 3-22 and Figure 3-23).

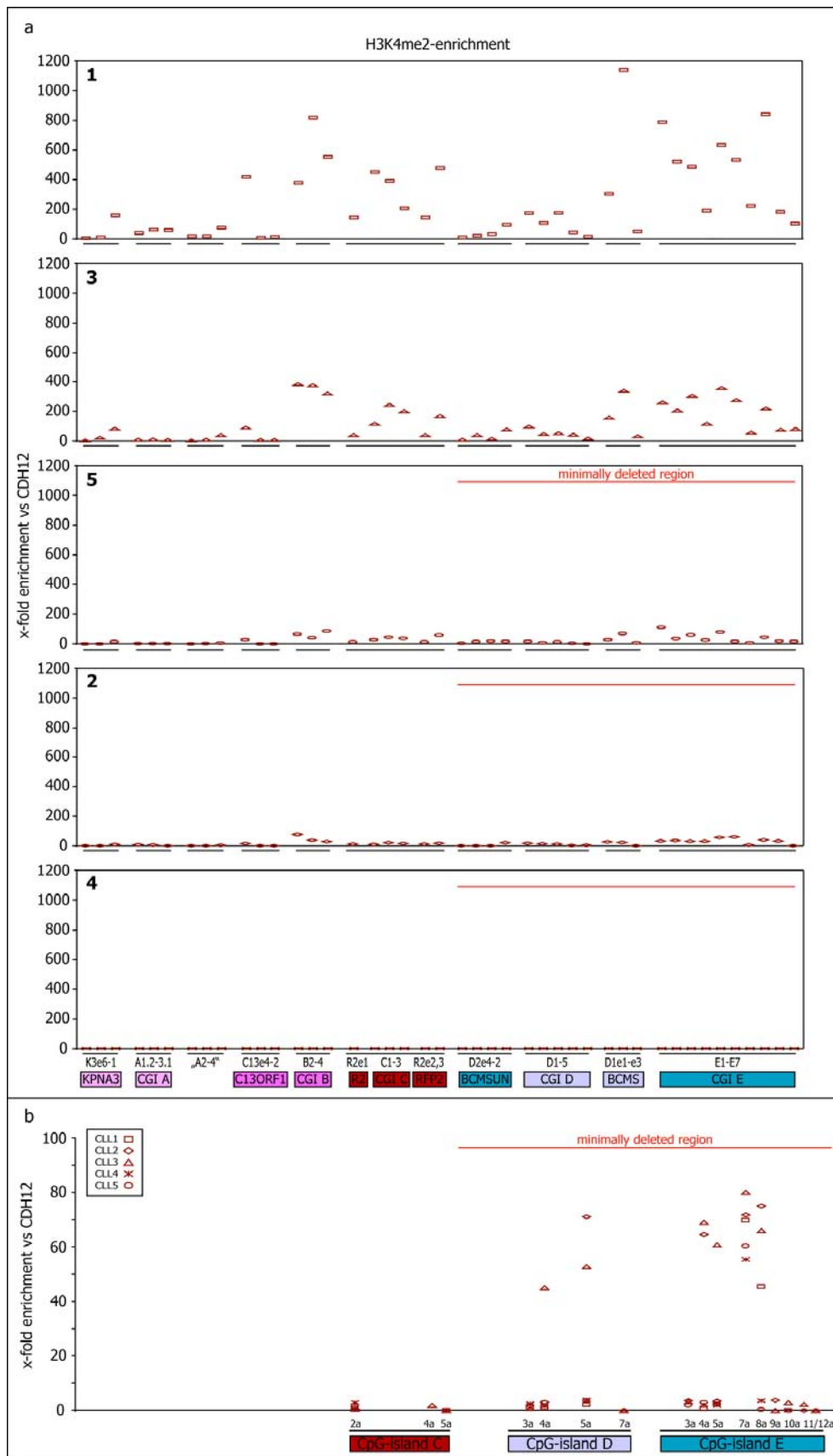


Figure 3-22: Differential enrichment of H3K4me2 in five CLL patients. a) qPCR analysis of CHIP in samples CLL1-5; b) ChIP-on-chip. Relative localization of fragments to each other in (a) and (b) are correct, but not scaled.

For the three patients showing a deletion of chromosomal material at 13q14.3 (del(13q); i.e. CLL5, CLL2 and CLL4), the minimally deleted region is depicted within the diagrams (red line within diagrams; Figure 3-22 and Figure 3-23).

No differences in the general pattern of distribution of neither H3K4me2 (Figure 3-22) nor macroH2A (Figure 3-23) were detected in the five analyzed CLL samples. However, the del(13q) CLL samples were distinct from the two CLL samples that derived from patients with other genomic aberrations. The del(13q) CLL samples showed overall low enrichment of H3K4me2 at 13q14.3. The sample CLL4 had no H3K4me2-enrichment at all, whereas very low overall H3K4me2-enrichment at 13q14.3 was determined in samples CLL2 and CLL5 (Figure 3-22a). In the same CLL samples macroH2A-enrichment at 13q14.3 was comparatively high (Figure 3-23) and thus, it was found to be reciprocal to H3K4me2. The retained copy of 13q14.3 in del(13q) CLL samples carried exclusively the repressive histone modification (macroH2A; Figure 3-23), whereas the analyzed active chromatin mark (H3K4me2) was almost completely lost (Figure 3-22).

In contrast to del(13q) CLL samples, two of the five analyzed CLL samples derived from patients that retained both copies of 13q14.3 and these had significantly different histone modifications at 13q14.3 compared to the del(13q) CLL samples. The first sample derived from a male CLL patient with a normal karyotype (CLL1) and showed the overall highest H3K4me2-enrichment ratios at 13q14.3. The highest H3K4me2-enrichment in sample CLL1 was detected at the candidate LCR locus within the telomeric part of CGI D (300 to 1122 fold enrichment versus *CDH12*; Figure 3-22a). MacroH2A-enrichment was detected only in very low amounts at 13q14.3 and was mostly found to be reciprocal to H3K4me2 in sample CLL1. Also in the second CLL sample with retention of both alleles of 13q14.3 (CLL3), H3K4me2-enrichment was significantly higher compared to the del(13q)-samples.

Sample CLL3 that derived from a male CLL patient with a 17p deletion, showed overall high H3K4me2-enrichment at 13q14.3, which was approximately 10 fold lower than in the CLL1 sample (Figure 3-22a). Analog to the sample CLL1, macroH2A-enrichment at 13q14.3 in sample CLL3 was reciprocal to H3K4me2-enrichment and detectable only in low amounts at the region. Thus, the retention of both copies of 13q14.3 in CLL cells was correlated with significantly higher abundance of the active chromatin mark (H3K4me2) and with loss of the repressive mark (macroH2A) at 13q14.3. Thus, the general patterns of enrichment at 13q14.3 of both analyzed chromatin marks were found to be essentially the same in all five patients. However, comparing the analyzed del(13q) CLL samples to those CLL samples with retention of both copies, the deletion of genomic material from the chromosomal region 13q14.3 could be correlated with loss of the active chromatin mark at 13q14.3. Not all data generated by qPCR analyses of single CLL samples were completely comparable to data obtained from ChIP-on-chip analyses of the same samples (Figure 3-22b; Figure 3-23b).

In general, less data was obtained in ChIP-on-chip analyses because the CpG-island microarray employed included only three of the five CpG-islands of 13q14.3 (CGIs C, D and E; see chapter 2.19.1). The most striking differences between both analyses were detected in the sample CLL4, which had detectable H3K4me2-enrichment at CGIs C, D and E determined by ChIP-on chip in contrast to analyses by qPCR that detected no H3K4me2-enrichment at all at 13q14.3 (Figure 3-22b). Similarly, macroH2A-enrichment at 13q14.3 in samples CLL2 and CLL5 was determined in slightly different amounts by ChIP-on-chip compared to qPCR analyses. Using ChIP-on-chip, a high enrichment of macroH2A was determined at CGIs C, D and E, which was not detected by qPCRs, whereas no macroH2A was measured in CLL5 by ChIP-on-chip, which was measured by qPCRs (Figure 3-23b).

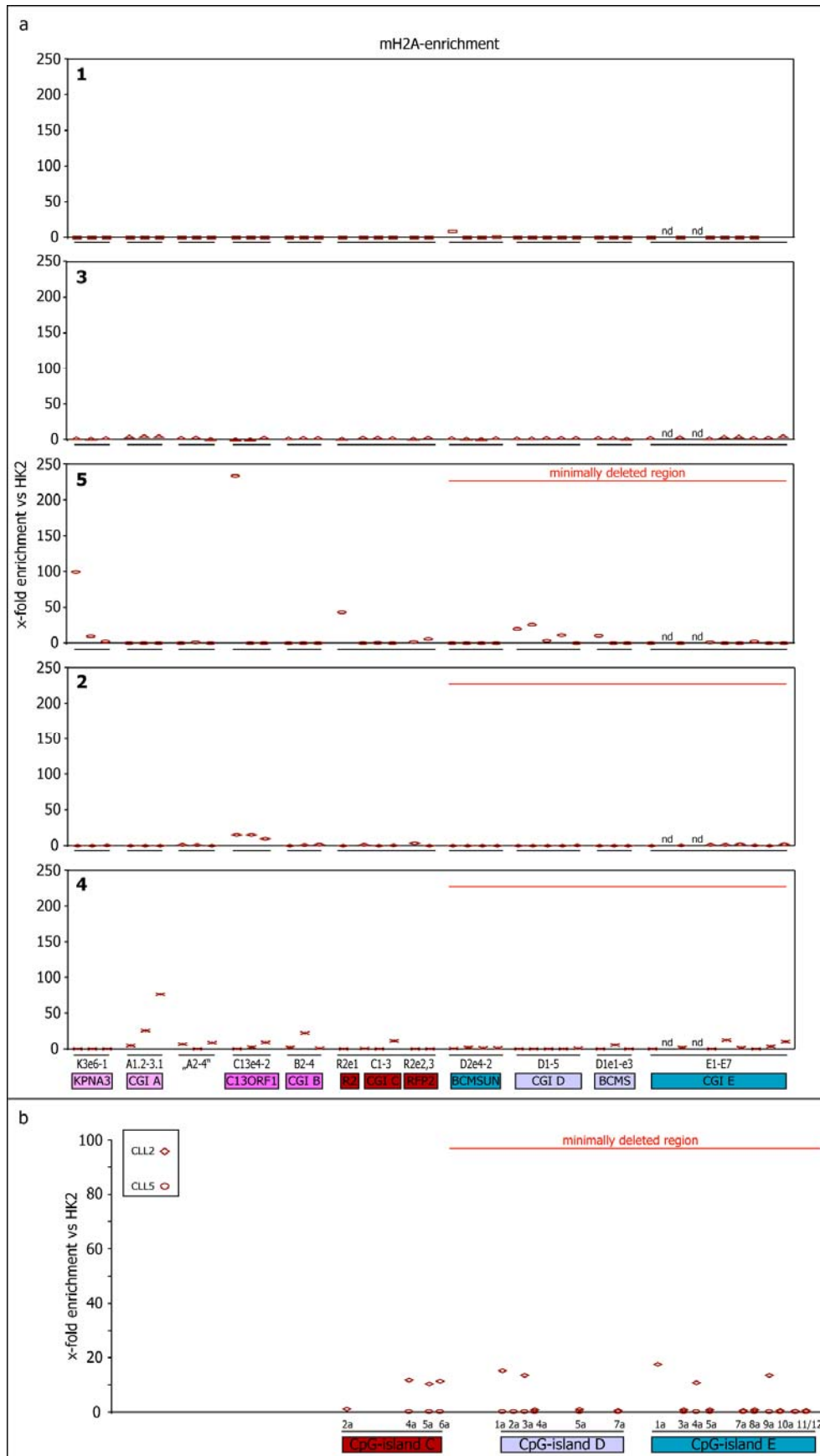


Figure 3-23: Differential enrichment of macroH2A in five CLL patients. a) qPCR analysis of CHIP in samples CLL1-5; b) ChIP-on-chip in samples CLL2 and CLL5. Relative localization of fragments to each other in (a) and (b) are correct, but not scaled.

Apart from these examples, the general results obtained from ChIP-on chip were comparable to results obtained by qPCR for both analyzed histone modifications in CLL cells, which were described above.

3.3.7 Identification of a locus control element in CGI D or E

Figure 3-24 and Figure 3-25 combine data from Figure 3-13 and Figure 3-20 for a better comparison of differential DNA-methylation and enrichment of histone modifications in CLL and wild type samples at distinct loci in chromosomal region 13q14.3.

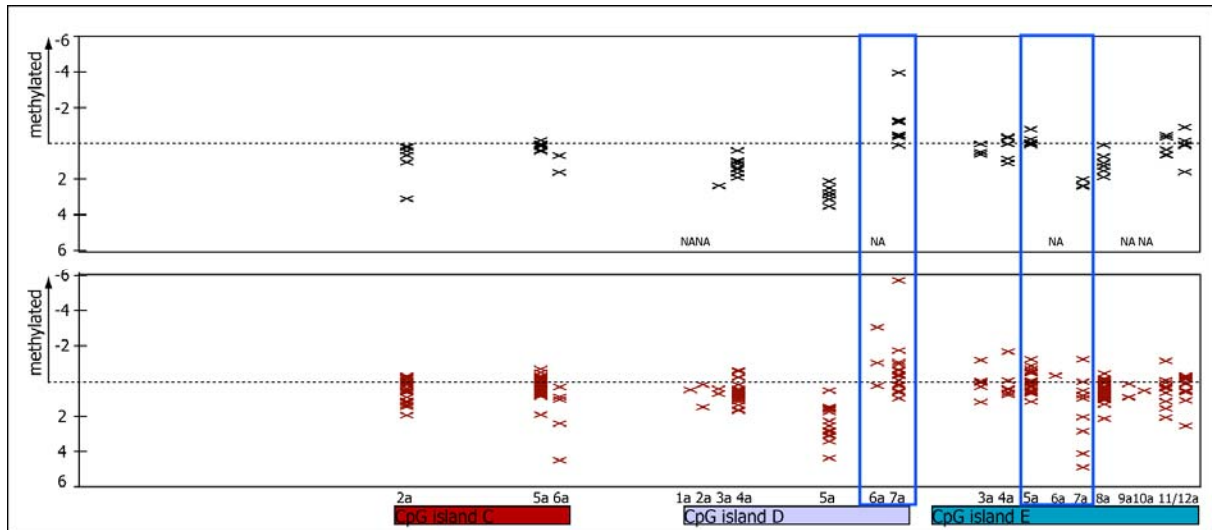


Figure 3-24: Summary of DNA-methylation patterns in wild type and CLL samples. DNA-methylation at 13q14 in wild type (black; n=9) and CLL (red; n=32) determined by aPRIMES. Blue boxes indicate loci with distinct epigenetic features within CGIs D and E.

Within CpG-islands D and E two small loci characterized as candidate LCRs in chapters 3.2.5 and 3.3.4, were analyzed. The epigenetic marks of these small regions were very distinct from the rest of the respective CpG-islands. These fragments had epigenetic characteristics that were not found elsewhere in the chromosomal region 13q14.3 (marked by blue frames in Figure 3-24 and Figure 3-25). In CGI D at loci D6a and D7a (Figure 3-24) remarkable differences in DNA-methylation between CLL and wild type samples were determined. At D6a two of three CLL samples were methylated. Unfortunately, no data were available for DNA-methylation at D6a in wild type samples, but at the adjacent locus D5a, no DNA-methylation was measured in non-malignant cells (Figure 3-24).

All analyzed wild type samples were methylated at the adjacent locus D7a, whereas some CLL samples were not, which reflects a loss of DNA-methylation at D7a in CLL compared to wild type. Consecutive BioCOBRA analysis by A. Philippen (see chapter 2.17) at D6a revealed hypomethylation of D6a in CLL samples compared to wild type samples (data not shown). The wild type samples showed DNA-methylation (70 % to 100 %), whereas CLL samples showed a highly variable but reduced DNA-methylation at the candidate LCR within CGI D (0 % and 100 %). However, no comparable differential methylation was detected at the adjacent locus D7a, where wild type and CLL samples both had a similar extent of DNA-methylation (0 % to 40 %). In summary, at the second most telomeric part of CGI D analyzed (D6a), which harbors one of two candidate LCRs of 13q14.3, DNA-methylation was determined in non-malignant cells, whereas this was found to be reduced in CLL cells (Figure

3 Results

3-24). In CpG-island E that harbors the second candidate LCR of 13q14.3, an abrupt change of DNA-methylation was detected at loci E5a to E7a. At E5a in both CLL and wild type samples, DNA-methylation was measured in some but not in all samples. At E6a no data could be obtained and E7a was not methylated in wild type, but in some of the CLL samples (Figure 3-24a). BioCOBRA analysis revealed hypermethylation of locus E6a in CLL in contrast to wild type (data not shown). At the locus E5a, DNA was methylated in CLL and wild type samples (60 % to 100 %), whereas at E6a DNA-methylation was more frequently measured in CLL cells. The adjacent locus E7a, however, showed no DNA-methylation, neither in CLL nor in wild type samples (data not shown).

Thus, DNA-methylation at the candidate LCR within CGI E (E6a) was evident in CLL cells but not in non-malignant cells (Figure 3-24). It was further analyzed if these differences in DNA-methylation between CLL and wild type samples at the two candidate LCRs within CpG-islands D and E were marked as well in the histone code (Figure 3-25).

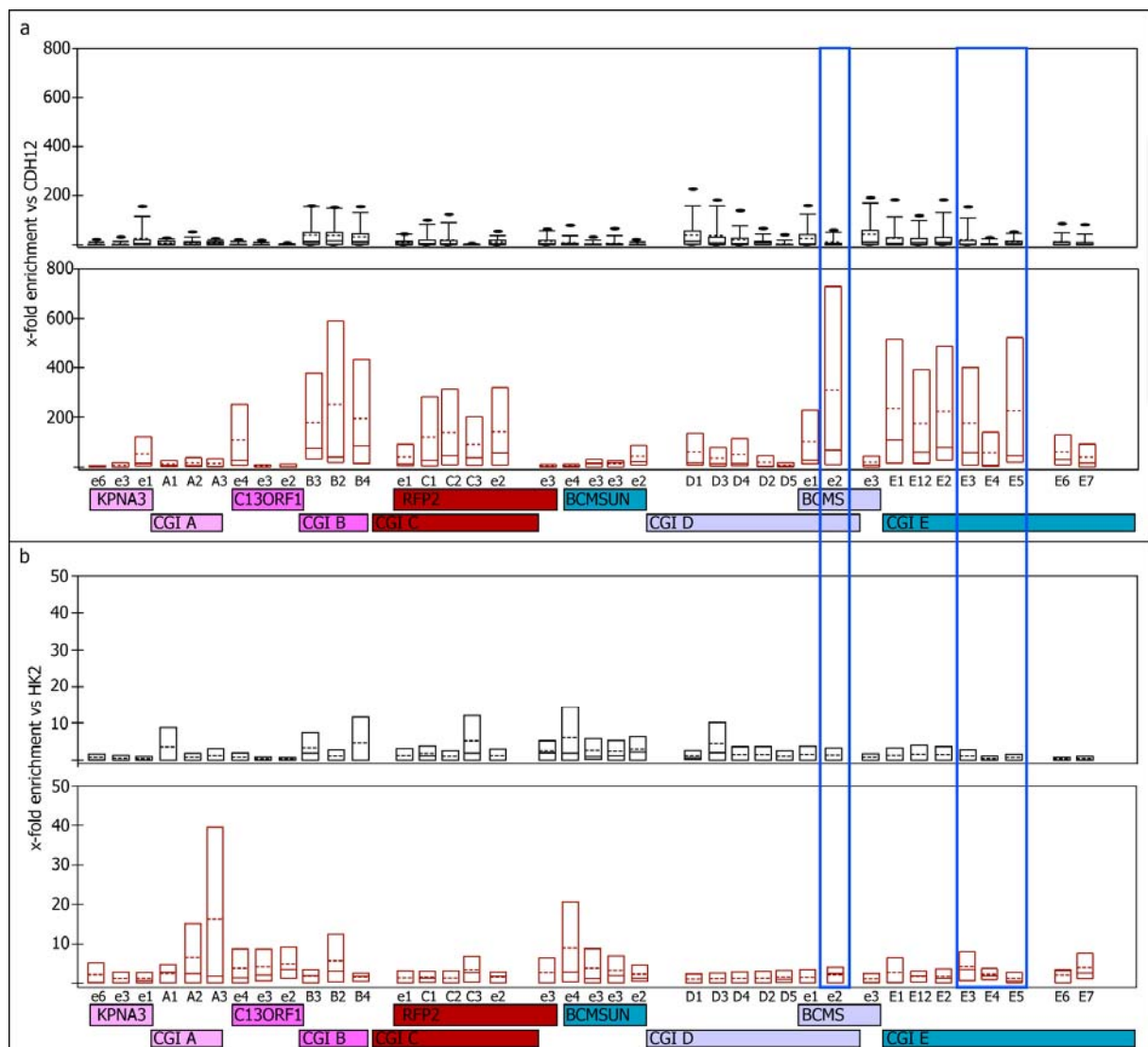


Figure 3-25: Summary H3K4me2- and macroH2A-enrichment in wildtype and CLL samples. a) H3K4me2-enrichment in wild type (black; n=11) and CLL (red; n=5) and d) macroH2A-enrichment in wild type (black; n=5) and CLL (red; n=5) at 13q14.3. Blue boxes indicate breakpoints within CGIs D and E.

Within CGI D, H3K4me2-enrichment could be shown to be significantly higher at the candidate LCR locus (D6a) and at its adjacent locus (D5a) in CLL samples compared to wild type samples (Figure 3-25). Thus, loss of DNA-methylation at D6a was detected together with a gain of an active chromatin mark at the candidate LCR within CpG-island D in CLL cells. Surprisingly, also within CpG-island E the candidate LCR locus and its two adjacent loci (E5a to E7a) were significantly higher H3K4me2-enriched in CLL samples compared to wild type samples (Figure 3-25).

Interestingly, the lowest H3K4me2-enrichment was measured at the candidate LCR locus itself (E6a) compared to the rest of the centromeric half of CGI E in CLL cells (Figure 3-25a). In contrast to the candidate LCR in CGI D, the second candidate LCR within CGI E had contrary epigenetic modifications in CLL cells, but not in non-malignant cells (Figure 3-24; Figure 3-25). Unlike H3K4me2-enrichment and DNA-methylation determined at the two candidate LCRs in 13q14.3, macroH2A-enrichment was homogeneous in comparatively low amounts at both CpG-islands. No differential macroH2A-enrichment could be determined, neither between CLL and wild type samples, nor between the candidate LCRs and the whole CpG-islands.

It was shown that most but not all genes in the region 13q14.3 showed an epigenetic code that reflected their transcriptional status in non-malignant cells. At two genes, *C13ORF1* and *DLeu2/BCMSUN*, promoter-restricted H3K4me2-enrichment was determined in non-malignant cells, which was in line with their monoallelic expression. The promoter-restricted H3K4me2-pattern specific for monoallelic expression was not changed in CLL cells.

In CLL cell samples derived from CLL patients with different karyotypes, a deletion of genomic material from 13q14.3 (del(13q)) and loss of active chromatin marks could be correlated.

Strikingly, two candidate locus control regions (LCRs) were identified by their special epigenetic features in non-malignant cells that could be shown to be differentially marked in CLL and non-malignant cells. The LCRs are located in the 5'-regions of the two large noncoding RNA genes, therefore, the LCRs might be involved in regulation of transcription of the ncRNAs. Differential epigenetic modification in non-malignant and CLL cells at these LCRs is of special interest as this could be an epimutation in 13q14.3 that has impact on the regulation of the region. Identification of candidate LCRs in 13q14.3 was a step towards elucidation of the regulatory mechanism in the critical region in 13q14.3.

Aberrant epigenetic modifications of 13q14.3 in CLL cells were most evident at the candidate LCRs, but furthermore a significant general gain of the active chromatin mark H3K4me2 at 13q14.3 was measured in CLL cells. The detailed analysis of the epigenetic code in non-malignant cells could help to determine the regulatory mechanism in 13q14.3 and furthermore, to identify epigenetic aberrations at 13q14.3 in CLL cells that might be relevant for the pathogenesis of CLL.

4 Discussion

CLL is the most common leukemia among adults in the Western world. Still, the molecular pathomechanism remains unknown. CLL is characterized by a defined set of well described genomic abnormalities, most of which have been shown to predict the survival of patients. The most frequent genomic aberration in CLL is a deletion of a critical region in 13q14.3 (del(13q)), which occurs in over 50 % of patients (Döhner et al., 2000). Apart from CLL tumors, the critical region in 13q14.3 is lost in a number of other tumors in high frequencies, for example around 60 % of patients with Mantle Cell Lymphoma (MCL) show a deletion of 13q14.3, which is the secondmost frequent genomic aberration in MCL (Bentz et al., 2000). In addition to hematological malignancies also other tumor entities show frequent loss of 13q14.3, including prostate carcinomas (Chen et al., 2001), squamous cell carcinomas (Ogawara et al., 1998) and leiomyosarcomas (Wang et al., 2003). The high frequency of deletion of 13q14 in CLL cells and in a large variety of other tumor cells suggests that this region harbors a central pathomechanism involved in malignant transformation.

Del(13q) in CLL has been correlated with a good prognosis unlike deletions in chromosomal regions like 11q22-23 and 17p13, which are associated with short (79 months) and very short (32 months) survival times, respectively (Döhner et al., 2000). Deletion of the first copy of 13q14 is always observed as first aberration and never as an acquired secondary aberration during clonal evolution in CLL (Stilgenbauer et al., 2007). However, biallelic loss of 13q14.3, i.e. additional loss of the second copy, is the second-most common additionally acquired aberration during clonal evolution in CLL after deletion of 17p13 (Stilgenbauer et al., 2007). The critical region in 13q14.3 seems to be of major importance in suppression of tumor development. Furthermore, it is likely that the regulation of expression of the candidate genes in the region is involved in the tumor suppressor mechanism of 13q14.3. The high frequency of 13q14 deletions in CLL patients and the observation that del(13q14) is possibly a causative aberration in CLL, point to a central role of the tumor suppressor mechanism in 13q14. Although the tumor suppressor mechanism in 13q14 is deleted in more than 50 % of CLL patients, it remains largely unclear how the mechanism functions and how a defective mechanism prevents tumorigenesis. To date, epigenetic silencing of one copy of 13q14.3 was shown indirectly by genomewide demethylation and deacetylation of the critical region in 13q14.3 in tumor cell lines (Mertens et al., 2006). Global demethylation as well as global deacetylation in tumor cell lines resulted in reactivation of expression of monoallelically silenced genes at 13q14.3. Strikingly, biallelic expression of a splicing variant of the ncRNA gene *BCMS* in the critical region in 13q14.3 could be reactivated by demethylation of CGI E. Therefore, it was important to gain insight into epigenetic regulation of the tumor suppressor mechanism in 13q14.3. The analysis and quantification of epigenetic modifications at the regulatory regions (CpG-islands; CGIs) and genes in 13q14.3 in non-malignant hematopoietic cells of healthy donors provides insight into the epigenetically regulated transcriptional state of the genes in the region. Identification of active and silenced genes and/or alleles in 13q14.3, i.e. of the distinct epigenetic code of 13q14.3 in non-malignant cells, can support the model of two differentially marked copies of 13q14.3 and the hypothesis of loss of the active copy in CLL.

4.1 Epigenetic regulation of 13q14.3?

The molecular functions of the genes within and in direct neighborhood of the critical region in 13q14.3 are mostly unclear. To date, the functions of 13q14.3 candidate proteins can only be inferred from homology of single protein domains of the protein coding genes in the region (*KPNA3*, *C13ORF1*, *RFP2* and *KCNRG*), which suggest potential functions of the proteins. Genes homologous to 13q14 genes are functionally involved in several pathways including NF κ B-signaling, repression of transcription, hematopoiesis and inhibition of proliferation (Tschuch, 2006). Previous studies showed that 13q14 candidate proteins interact with each other and with additional proteins that are involved in various cellular processes including regulation of translation and regulation of gene expression (Serra-Barrionuevo, 2008). These two studies indicate that the genes in the critical region in 13q14.3 might form a functional gene cluster, which is involved in cellular pathways that could contribute to CLL pathogenesis. Thus it is likely, that the candidate genes of 13q14.3 are involved in the pathomechanism of CLL.

The 13q14 candidate genes were shown to be downregulated in CLL cells (Mertens 2002), but despite extensive sequence analyses of the critical region in CLL no pathogenic point mutations could be identified that would lead to loss of function of a potential tumor suppressor gene in the region (Liu et al., 1997; Stilgenbauer et al., 1998). Additionally, the candidate genes are monoallelically expressed in non-malignant B- and T-cells. Together with the finding that 13q14.3 replicates asynchronously due to differential chromatin packaging of the two alleles (Mertens et al., 2006) these are strong indications for an epigenetic regulatory tumor suppressor mechanism in 13q14.3.

In this thesis, it was shown that promoter-restricted enrichment of H3K4me₂, which is a chromatin mark for monoallelically expressed genes (Rougeulle et al., 2003), was stably detected at the 13q14.3 genes *C13ORF1* and *DLeu/BCMSUN* in non-malignant cells (Figure 3-11) and the chromatin pattern at these two genes was also retained in CLL cells (Figure 3-18). The chromatin pattern of monoallelic expression was expected to be present at the gene *DLeu2/BCMSUN*, as it was previously shown to be monoallelically expressed in B- and T-cells of healthy donors (Mertens et al., 2006). Unexpectedly, we detected this chromatin pattern also at *C13ORF1* and subsequently, it was shown that the gene is indeed monoallelically expressed in B- and T-cells.

Thus, it was determined that the monoallelic expression status of the protein-coding gene *C13ORF1* and of the ncRNA gene *DLeu2/BCMSUN* is marked in the histone code (Figure 3-11). These findings point to further epigenetic regulation of the tumor suppressor mechanism and to an epimutation in 13q14.3 that leads to its inactivation in CLL cells. In order to elucidate the epigenetic regulatory pathomechanism of CLL, it was important to determine the epigenetic modifications at 13q14.3 in non-malignant and CLL cells.

4.2 Analysis of epigenetic modifications at 13q14.3

Chromatin remodeling, histone modifications and other chromatin-related processes play a crucial role in gene regulation. Chromatin immunoprecipitation (ChIP) has been recognized as useful technique to study these processes (Haring et al., 2007). In the last years, evidence has accumulated that suggests that the histone marks in cancer cells are similarly deregulated as has been shown for DNA-methylation in cancer cells (Fraga et al., 2005b).

Moreover, analogous to a pattern of global DNA-hypomethylation coupled to promoter hypermethylation that is found in cancer cells (Gargiulo and Minucci, 2008), specific histone modifications are thought to be changed in cancer cells into distinct patterns that are correlated to the tumor type (Fraga et al., 2005b).

ChIP-analyses of epigenetic modifications in non-malignant and CLL cells were performed in order to be able to identify an aberrant epigenetic code of 13q14.3 that can be correlated with the deregulated gene expression related to CLL pathogenesis. The strategy to identify the histone modifications at 13q14.3 was to quantify active and repressive chromatin marks at the critical region to determine the transcriptional status of the candidate genes. Using ChIP active and repressive chromatin marks were quantified at the CpG-islands (CGIs) and genes in the critical region in 13q14.3. However, quantification of three epigenetic marks determines only a part of the epigenetic code of 13q14.3. Nonetheless, repressive histone modifications together with DNA-methylation were considered as strong indication of transcriptional silencing of the measured loci. In order to quantify one activating (H3K4me2) and one repressive (macroH2A) histone modification at the critical region in 13q14.3, the chromatin immunoprecipitation (ChIP) technique was established.

4.2.1 Establishment of the ChIP protocol

Several steps had to be optimized during establishment of the ChIP protocol in order to use it with primary hematopoietic cells (Figure 3-1). Virtually every step of the procedure had to be optimized carefully and was modified for primary hematopoietic cells. Here, it is suggested to modify the common ChIP protocol (Orlando, 2000) for each application, each antibody and each cell type of interest as this was essential to achieve successful chromatin precipitations, which was also recently described by others (Papantonis and Lecanidou, 2008; Hanlon and Lieb, 2004; Clayton et al., 2006).

Large cell numbers are necessary for each ChIP experiment to compensate for i) loss of cell material after reversion of cross-links, ii) loss of cell material to surfaces throughout the procedure, and iii) impaired signal to noise ratio upon reduction of input material (Collas and Dahl, 2008). However, the primary material from tumors is very limited; therefore several cell lines were needed for ChIP optimization. As cell lines deriving from CLL tumors (MEC1 and MEC2) were not growing sufficiently well in culture, cell lines deriving from other hematological tumors that were rapidly growing into large cell numbers were used for establishment of ChIP. Three different cell lines were used to establish the ChIP procedures, which were HL60 (AML tumor of female origin), Jurkat (T-cell leukemia of male origin), and Namalwa (Burkitt's lymphoma of male origin). It remains uncertain whether aberrant epigenetic modifications determined by ChIP are specific for the original tumor or are rather derived from cultivational artifacts. Moreover, the cell lines used in the present study did not derive from CLL so the measured epigenetic modifications in HL60, Jurkat and Namalwa cells might serve as potential epigenetic profile in AML, in T-cell leukemia and in Burkitt's lymphoma, respectively and can therefore not be used to compare them to CLL profiles. All three cell lines showed reduced enrichment ratios of active and repressive chromatin marks at 13q14.3 loci compared to primary CLL cells and the distribution patterns of both marks differed from those determined in CLL cells (data not shown). One explanation for these observations is that the epigenetic profiles of 13q14.3 differ among various tumor entities, which would indicate that epigenetic profiles of 13q14.3 might be used to distinguish between different tumor entities. Another explanation for different profiles determined in primary CLL cells and cancer cell lines is that epigenetic modifications might change over

time in cultivated cells. Smiraglia and colleagues studied hypermethylation phenotypes of cancer cell lines in comparison to the primary malignancies (Smiraglia et al., 2001) and found that the cell lines exhibit significantly higher levels of CpG-island-hypermethylation than the primary malignancies they represent.

Thus, it becomes clear that cell lines might not be the ideal model system for the analysis of aberrant epigenetic modifications in cancer. However, further analyses are needed to answer the questions of whether the epigenetic profiles of the cell lines used here are indeed representative of the original tumor or not, and whether epigenetics of AML or Burkitt Lymphoma are different from CLL profiles. The focus of the current work was to analyze the epigenetic profile of 13q14.3 in CLL and non-malignant cells.

The first step of the ChIP protocol that had to be optimized in order to extract sufficient amounts of chromatin from hematopoietic cells is the chromatin preparation step (Figure 3-1). As Spencer et al. (2003) noted it was essential, which salt and/or detergents are included in ChIP-buffers, because each detergent/salt concentration can make a difference in the accuracy of determining the distribution of a protein along a specific DNA-sequence. Here, it was tested whether the SDS content of buffers was crucial for quality and quantity of prepared chromatin in ChIP. It was seen that the amount of input chromatin in nChIP-approaches (O'Neill and Turner, 2003) was significantly higher than that prepared in xChIP approaches. The two protocols differ in preparation of chromatin, in the fragmentation of chromatin and in the spectrum of analyzable chromatin modifications: Using the nChIP protocol, chromatin is prepared that can be fragmented by nuclease digestion into native chromatin fragments determined by nucleosomes. While native chromatin can be used to analyze histone modifications, chromatin prepared and fragmented by ultra-sonication using the xChIP approach can be used to analyze in addition to histone modifications also transcription factors or other chromatin-binding proteins that are fixed to the chromatin during the procedure. The main difference between the two protocols to this end was the buffers, because nChIP buffers were composed in such way that enzymatic digestion of chromatin is allowed. However, quality and/or quantity of nChIP-chromatin were not sufficient as it was not possible to precipitate targets efficiently (Table 3-1a-c). Moreover, a large disadvantage of the nChIP procedure is that only histone modifications can be analyzed with native chromatin. Analysis of transcription factors is impossible due to the lack of fixation to chromatin, which is included in the xChIP protocol. However, the broader application of ChIP is mainly interesting for future applications and so optimization of nChIP protocol was not immediately stopped. The standard xChIP protocol (Orlando, 2000) had the large disadvantage that the quantity of chromatin in fractions prepared from hematopoietic cells varied largely and no precipitation of chromatin could be achieved (Figure 3-2). It became clear that it was necessary to first enrich chromatin rather than use whole cell extracts for ChIP. Chromatin from whole cell extract thus was enriched by the use of nuclei extraction protocol originally developed for standard immunoprecipitation (IP).

The implementation of the nuclei extraction protocol prior to chromatin preparation enriched cell nuclei and depleted the cell debris, which was beneficial for the quality and the quantity of chromatin. The enrichment of nuclei stably produced sufficient and reproducible quality and quantity of chromatin, which could be used for subsequent precipitation in ChIP.

Formaldehyde cross-linking was assumed to be unnecessary when analyzing histone modifications, because chromatin is very stable and would need no further stabilization of histone-DNA-bonds. However, cross-linking ensures preservation of the chromatin structure

throughout isolation and ChIP procedures (Haring et al., 2007) and proved to be advantageous for the ChIP results. If cross-linking was included in xChIP procedures the chromatin was preserved as could be seen by the amount of DNA and protein that could be isolated afterwards (Figure 3-3). Moreover, a bias of DNA-fragment sizes was avoided (Figure 3-3a), compared to xChIP-protocol applied without cross-linking (Figure 3-3b), where an unexpected and unreproducible bias of DNA-fragments was produced. When comparing the two standard protocols nChIP and xChIP, it became evident that without cross-linking (i.e. as in nChIP) it is more likely that the DNA from chromatin fractions is lost and only proteins can be precipitated (Figure 3-3).

In this study, using the nChIP-protocol failed to precipitate chromatin without cross-linking (Figure 3-3). Therefore, the cross-linking step of the xChIP protocol proves to be essential for success. It was also noted by others that they systematically failed to obtain significant amounts of precipitate without cross-linking (Haring et al., 2007).

The third and last step that showed superiority of the xChIP over the nChIP protocol was the fragmentation of chromatin. In nChIP, fragmentation was carried out by nuclease digestion, which produced fragments that are multiples of nucleosomal DNA length of 146 bp in size (Figure 3-4a). The advantage of digesting chromatin is that the chromatin is digested into native fragments, i.e. the native chromatin structure provided by the nucleosomes is retained. On the other side, fragmentation by ultra-sonication is a completely random process that generates chromatin fragments with sizes varying from 100 bp to 1 kb (Figure 3-4b) with the bulk of fragments having a size between 500-900 bp. The fragments produced by ultra-sonication were more reproducible in size and quantities compared to enzymatic digestion, which resulted in higher variability of fragments. The nucleosomal ladders were not reproducible, which might be due to variation in the performance of the used enzymes. A second disadvantage of nuclease digestion of chromatin was that no cross-linking could be included as chromatin would not be recognized by the nuclease anymore. Ultra-sonication thus proved to be a robust tool that yielded reproducible and comparable amounts of chromatin fragments that could be used for precipitation in ChIP.

It was shown in this study that chromatin has to be stabilized by cross-linking with formaldehyde and that chromatin fragments are produced more robust by ultra-sonication than by enzymatic digestion. This led to the use of an optimized standard xChIP protocol for the analysis of hematopoietic cells because nChIP-protocol failed to include these essential steps. The chromatin prepared for ChIP was enriched from nuclei of whole cell extracts by the use of nuclei extraction procedures and buffers, which substitute the SDS-lysis buffer and preparation of chromatin procedures from the original xChIP protocol (Orlando, 2000).

Efficient precipitation of target chromatin is essentially dependent on the antibodies used. Nonetheless, also the choice of agarose beads and conditions for pre-clearing and precipitation is essential to minimize the binding of unspecific chromatin to agarose beads. Three sequential steps of pre-clearing with little amount of agarose beads proved to circumvent most efficiently the binding and co-precipitation of unspecific chromatin (Figure 3-5). The quality of the primary antibody used for ChIP is the most critical part for successful precipitations (Spencer et al., 2003). Not every antibody that worked well in standard immunoprecipitation reactions, worked as well in ChIP as was seen for several tested antibodies including α -H3K4me3, α -H3K9me3 and α -Ikaros (Figure 3-6a) and also noted by others (Taneyhill and Adams, 2008). This can partially be explained by the fact that epitopes to which antibodies bind may be masked by chromatin architecture and/or are altered

beyond antibody recognition after cross-linking. Nine of the twelve antibodies tested here had to be excluded from further analysis, because they could not be used for Western blot detection of precipitated chromatin (i.e. α -HP1 γ) and/or because they did not enrich target chromatin in precipitates (i.e. α -H3K9me3, α -CTCF or α -H3K27me3; Figure 3-6).

The remaining three of twelve tested antibodies, α -H3K4me2, α -macroH2A and α -H3 (Figure 3-6), could be used for quantification of histone modifications at chromosomal region 13q14.3 as only these three precipitated their respective targets efficiently and reproducibly. Antibodies that precipitate chromatin in ChIP were all polyclonal antibodies. The superiority of polyclonal over monoclonal antibodies can be explained by the characteristic of polyclonal antibodies, which recognize different epitopes of the same target. Thus, polyclonal antibodies are in principle able to precipitate more of the target chromatin even when one or several epitopes of the chromatin are masked following formaldehyde cross-linking. In contrast, monoclonal antibodies only can recognize one epitope and are therefore prone to be influenced by epitope masking.

In order to validate successful precipitation of target proteins two methods were applied. First, the proteins precipitated in ChIP were analyzed in Western blots for specific and efficient precipitation of the target. Second, precipitated DNA was analyzed by qPCR to ensure specific enrichment of control sequences. Western blot detection by the anti-macroH2A antibody was inefficient compared to anti-H3K4me2 and anti-H3 antibodies. The anti-macroH2A antibody only detected macroH2A in Western blot when it was highly enriched as in the ChIP-precipitates. Non-enriched amounts of macroH2A could not be detected with satisfying results (Figure 3-8b). As it was suggested that endogenous macroH2A is not detectable in cross-linked chromatin fractions by Western blots (A. Ladurner, personal communication) reversion of cross-links was performed prior to Western blot analysis. However, the reversion of cross-links also did not produce satisfying Western blot results (data not shown). Thus, only the precipitations with anti-H3K4me2 could be sufficiently validated by Western blots (Figure 3-8a).

The latter antibody also precipitated its target H3K4me2 more efficiently, as was seen from qPCR quantification of precipitated sequences (Figure 3-9). Resulting enrichment was calculated as "fraction of input", which was also measured by qPCR (Figure 3-9; see also chapter 2.16.10). The antibody anti-H3K4me2 precipitated its target by a factor of 10-100 more efficiently than did the anti-macroH2A antibody. This can be due to either a more efficient precipitation by the antibody used or to a higher abundance of H3K4me2 in the analyzed cells, whereby most likely both reasons apply here. MacroH2A1.2 is detected in lower quantities (1/10) in non-malignant hematopoietic cells and CLL cells than H3K4me2, and this is most likely due to lower endogenous amounts of macroH2A in general.

This observation suggests that active chromatin marks (H3K4me2) are widely established in the epigenome, whereas the exchange of histone variants for canonical histones (i.e. macroH2A is a variant of canonical H2A) is more uncommon as an inactive chromatin mark. This might be the reason why macroH2A had been detected more frequently at distinct sites only, e.g. at inactivated X-chromosomal genes (Changolkar and Pehrson, 2006). MacroH2A occurrence only at the inactivated X-chromosomal genes also points to a monoallelic enrichment of macroH2A. Also, macroH2A plays a role at imprinted loci, which are rare (Choo et al., 2006), and this would explain the generally lower abundance of the histone variant. However, macroH2A could be detected at the analyzed loci in this study, for example at the positive control sequence of the IL8 promoter region, which was used as positive

control (Figure 3-9b, green bars), and also at several loci of the critical region in 13q14.3, which will be discussed in more detail below.

As the ChIP-procedure is susceptible for generation of false-positive signals by unspecific precipitation, it is essential to normalize qPCR and microarray measured raw data. Three normalization steps were performed before interpretation of the ChIP-data: i) subtraction of background (i.e. subtraction of C_T -values of mock- from precipitate-fractions; (see also Mutskov and Felsenfeld, 2004)); ii) relation of precipitate to input (fraction of input; (Tariq et al., 2003)) and iii) normalization relative to control sequences (x-fold enrichment versus control; (Mathieu et al., 2005, see chapter 2.20.4 for details). These three steps were used in order to generate valid data for the quantification of epigenetic marks at distinct genomic regions. The normalization steps were included after every ChIP experiment to achieve better comparability among different cell samples and also among different platforms used for analysis of ChIP-DNA (i.e. ChIP-on chip-and ChIP and qPCR). Control sequences used for the third normalization step were selected from ChIP-on-chip experiments with the CGI-microarray. Two clones were selected as positive control sequences for H3K4me2- (*CDH12*) and macroH2A-ChIPs (*HK2*) and had to fulfill the criterion that they were enriched in both types of analyzed tissue, i.e. non-malignant and CLL cells, in order to serve as control sequence. The sequences found in promoter CpG-islands of the genes *CDH12* and *HK2* fulfilled this criterion and consecutively, these sequences were used as positive control for normalization. Only after (three-fold) normalization, statistical analyses were performed and resulting data was interpreted.

Normalization of ChIP-data is necessary, because it cannot be controlled by other means. For instance in expression analyses, normalization can be performed by testing expression levels of housekeeping genes in different tissues. However, levels of expression do not necessarily reflect changes in histone modifications (Haring et al., 2007). Thus, the described three steps of normalization of data were performed in order to link validated epigenetic modifications to transcriptional state of genes at 13q14.3.

The ChIP methodology was established to quantify epigenetic modifications at the critical region in 13q14.3 in non-malignant and CLL cells. It was noted also by others that it is essential to establish and optimize each step of the protocol to the cell system and to antibodies that are intended to be used. Although the ChIP methodology appears to be simple, it is indeed a complicated procedure. In the present as well as in other studies several factors had to be optimized in order to successfully precipitate chromatin of interest by ChIP in a given system (Spencer et al., 2003; Haring et al., 2007; Collas and Dahl, 2008). These factors include the quality of the primary antibody, the degree of sample ultra-sonication and formaldehyde cross-linking, the amount of protein A agarose, and the types of salt and/or detergent washes, which have been discussed above. Consequently, ChIP could be used as valuable tool for studying the chromatin modifications that are involved in the processing of DNA from the critical region in 13q14.3.

4.2.2 Why of all histone modifications quantify H3K4me2 and macroH2A at 13q14.3?

ChIP was used to characterize two histone modifications at 13q14.3, which are at first glance more uncommon: enrichment of H3K4me2 and macroH2A. However, they were chosen for several specific reasons.

The activating mark H3K4me2 is distributed specifically at active genes, but more broadly as the more commonly analyzed mark H3K4me3, which is enriched only at the transcription start sites of active genes (Wysocka et al., 2006). It was repeatedly shown that H3K4me3 correlates with the onset of transcription (Bernstein et al., 2005; Pokholok et al., 2005; Santos-Rosa et al., 2002), while H3K4me2 resides elsewhere in the vicinity of active genes (Schubeler et al., 2004). As the main interest in this work was to characterize CpG-islands (CGIs), noncoding sequence and exonic regions in 13q14.3, quantification of H3K4me2 was chosen as it was described to be also enriched at these regions of active genes (Mellor et al., 2008). Thus, the chance of detecting the active chromatin mark at the small analyzed 13q14 loci was maximized. Support for this choice came from recent studies that analyzed the genome-wide distribution of up to 20 histone modifications in parallel by a method called "ChIP-Seq" that combines ChIP with next-generation sequencing (Barski et al., 2007; Mellor et al., 2008). ChIP-Seq allows genome-wide screening of the epigenetic code at functional elements, i.e at transcription start sites (TSSs), at enhancers, at promoters and at intergenic and exonic regions. The screen by Barski et al. (2007) included genome-wide quantification of H3K4me2-enrichment and 18 other histone modifications. In the screen, H3K4me2 was found to be enriched at TSSs and also at intergenic regions in a distinct pattern similar to H3K4me3, but with a much broader distribution than H3K4me3.

Therefore, H3K4me2 enrichment can be considered as sufficiently significant mark of active genes as it was found together with H3K4me3, which is significantly correlated to transcriptional activity (Barski et al., 2007; Wysocka et al., 2006).

Analysis of macroH2A-enrichment was chosen to detect inactive chromatin sites at 13q14, because it was shown to be allelically enriched at imprinted loci (Choo et al., 2006) in addition to its abundance at silenced loci. MacroH2A is found for example at the inactive X chromosome in female mammals (Costanzi and Pehrson, 1998). The allelic deposition of macroH2A is of special interest as the current model for the tumor suppressor mechanism in 13q14 involves differential chromatin packaging of the two copies of 13q14.3 (Mertens et al., 2006), which results in one active and one inactive copy. Although 13q14 is not imprinted, it has several features of imprinted loci that includes monoallelic expression of two large noncoding RNA genes that span the critical region, and that promoter-restricted H3K4me2-enrichment is detected at monoallelically expressed genes (Figure 3-11; Figure 3-18), a pattern that was first recognized at imprinted and monoallelically expressed genes (Rougeulle et al., 2003). A fourth characteristic of imprinted regions in the critical region in 13q14.3 is, supposedly, the localization of a locus control region (LCR), analogous to imprinting control regions (ICEs) in imprinted loci, in one of the two CGIs in the 5' regions of the two large noncoding RNA genes.

On the basis of these similarities to imprinted loci it was feasible to assume that macroH2A can be detected at 13q14. As the model of epigenetic regulation of the critical region in 13q14.3 suggests differential chromatin packaging of the two copies, macroH2A- as well as H3K4me2-enrichment was analyzed for monoallelic enrichment at 13q14.3 (Figure 3-16). Detailed characterization of H3K4me2- and macroH2A-enrichment at the genes and CpG-islands in the critical region in 13q14.3 was used to determine active and inactive chromatin marks of the region in non-malignant cells and epigenetic aberrations in CLL cells.

4.3 Epigenetic modifications at 13q14.3 in non-malignant hematopoietic cells

Quantification of the two histone modifications H3K4me2 and macroH2A together with DNA-methylation at the critical region in 13q14.3 in non-malignant cells was analyzed in several aspects. The following pivotal questions will be discussed in the following sub-chapters:

- i) What is the epigenetic state of the candidate genes in the critical region in 13q14.3? (Chapter 4.3.1).
- ii) Is the monoallelic expression of candidate genes reflected in the histone code? Is this true for all candidate genes? What are the exceptions? (Chapter 4.3.2).
- iii) Epigenetic modifications are changed during cellular aging. Do we observe epigenetic differences between cells derived from young healthy donors and older ones? What are the differences? (Chapter 4.3.3).
- iv) Epigenetic modifications are tissue-dependent and differ between different cell types. In the present study, peripheral blood lymphocyte (PBL) samples were analyzed as non-malignant control cell material even though they are not primarily composed of B-cells. Does it make sense to compare these non-malignant PBLs to CLL cells that are derived from B-cells? Are there different epigenetic profiles detectable in non-malignant B- and T-cells? (Chapter 4.3.4).
- v) The current model of the epigenetically regulated tumor suppressor mechanism in 13q14.3 includes monoallelic expression of the candidate genes, differential chromatin packaging and asynchronous replication timing of the two alleles (Mertens et al., 2006). Do chromatin marks also show preferential allelic enrichment at 13q14.3 that would support the model of one active and one silent 13q14.3 copy? Furthermore, is there a locus control element/region (LCR) located in one of the CpG-islands? (Chapters 4.3.5 and 4.4.4).

4.3.1 The epigenetic code of 13q14.3 in non-malignant hematopoietic cells

It was proposed that each cell type within every tissue carries a distinct epigenome (Esteller, 2007; Fraga and Esteller, 2005), which is specifically altered during malignant transformation. Thus, in order to analyze the epigenetic changes in CLL cells, first the wild type epigenetic code of 13q14.3 was determined in B-cells and in a mixture of hematopoietic cells. Two histone modifications (i.e. H3K4me2 and macroH2A) were analyzed in detail at 36 loci within 13q14.3 (Figure 2-3) in non-malignant cells from eleven healthy donors.

The genome wide distribution of H3K4me2 is strongly correlated with the distribution of H3K4me3, which is enriched specifically at active genes (Barski et al., 2007). Here, H3K4me2-enrichment was analyzed in order to identify CpG-islands (CGIs) in the 5'-regions of 13q14.3 candidate genes that are actively marked. Astonishingly, the overall pattern of H3K4me2-enrichment at all single loci in the critical region in 13q14.3 is the same or nearly the same in all analyzed cells (Figure 3-10). Considering H3K4me2 at 13q14.3 alone, the observation of an identical distribution pattern of the active histone mark in cells derived from eleven donors supports the existence of the above described tissue-specific epigenome. When considering only one histone modification, it is questionable whether the

transcriptional state of a gene can be determined. However, the active chromatin mark is found most frequently at three CGIs in 13q14.3: CGIs B, D and E showed highest enrichment of the active chromatin mark (Figure 3-10). CGIs A and C in 13q14.3 only showed reduced levels of the active mark. Moreover, several analyzed loci in 13q14.3 are located in the exons of candidate genes (Figure 2-3) and these were shown to generally contain fewer active marks than the 5' regions of the genes (i.e. the CGIs; Figure 3-10). Thus, a first analysis of the distribution of the active chromatin mark at 13q14.3 suggests that the three CGIs B, D and E are actively marked whereas CGIs A and C are not. Whether A and C are silenced can only be decided after analysis of more than one epigenetic mark. Therefore, macroH2A-enrichment was analyzed as inactive chromatin mark at the same loci in 13q14.3. MacroH2A is higher enriched at CGIs A and C than at CGIs B, D and E, which is in accordance with H3K4me2-enrichment. High enrichment of macroH2A together with lack of H3K4me2 at CGIs A and C suggests that both CGIs might be marked as inactive. However, when considering also the DNA-methylation at these two CGIs, silencing is not supported or not (yet) further stabilized by DNA-methylation. CGI A and CGI C do not show any (or only very little) DNA-methylation in non-malignant cells (Figure 3-13 and A. Philippen, personal communication). The two histone modifications both suggest inactivity of the CGIs, but DNA-methylation does not support silencing of the CGIs. Moreover, the genes *KPNA3* and *RFP2* are both expressed in non-malignant cells (Mertens et al., 2006), which excludes at least complete silencing of the two CGIs in the promoter region of the genes. The enrichment of macroH2A at 13q14 was generally lower (by factor 10-100) than that of H3K4me2. This can be due to the finding that macroH2A-enrichment can be allele-specific at imprinted loci (Choo et al., 2006) and suggests that it might indeed be enriched at only one copy of the monoallelically expressed 13q14 region. Furthermore, it was detected that macroH2A is especially lower enriched in non-malignant cells derived from probands that are above 45 years old. This suggests that enrichment of macroH2A might decrease during aging or is correlated with the aging process (see chapter 4.3.3).

Quantification of two histone marks and of DNA-methylation at the CpG-islands of 13q14.3 showed an unequal occurrence of activating and repressive chromatin marks at the five CGIs and also at the two copies of 13q14.3. This unequal distribution of chromatin marks suggests a distinct epigenetic code at 13q14.3 that regulates expression of the candidate genes.

Epigenetic marks that constitute the epigenome of a given cell (Fraga and Esteller, 2005), specifically mark the transcription status of genes at the genes promoter region. Therefore all three epigenetic modifications analyzed here, were considered if they constitute a distinct transcriptional state of the genes in the critical region in 13q14.3 in the chromatin. No distinct epigenetic mark was detected for *KPNA3*. Only medium and non-significant enrichment of the active and silent histone marks was detected along with a lack of DNA-methylation at CGI A in non-malignant cells (Figure 3-13). *KPNA3* and CGI A are located outside of the critical region in 13q14.3 and lack a distinct epigenetic mark, which suggest that *KPNA3* is also outside of the epigenetic regulation in 13q14.3.

In contrast, at the *C13ORF1* locus, a gene located within the critical region, promoter-restricted enrichment of H3K4me2 was detected (Figure 3-11) that is correlated with monoallelic expression of the gene (see chapter 4.3.2). Furthermore, a comparatively high macroH2A-enrichment is detected at CGI B. This may point to a bivalent mark of the *C13ORF1* promoter, which might be the reason for selective monoallelic expression of the gene in most but not all healthy probands and of most but not all splice variants of *C13ORF1*. The DNA-methylation at CGI B also reflects a potential bivalent epigenetic code at

the sequences of the promoter and the gene, as DNA-methylation was found to be highly variable ranging from 0 % to 80 % (A. Philippen, personal communication). These observed interindividual differences might reflect allelic differences in DNA-methylation, but also age differences or tissue differences as samples containing diverse lymphocytes (PBL) derived from healthy probands with an age between 23 and 54 years were analyzed.

The epigenetic code of CGI B and *C13ORF1* points to an important role of the candidate gene in the pathomechanism of CLL. Its protein function is not clear yet, but it has been shown to be involved in NF κ B-signaling, in NOTCH-signaling and in AKT-signaling, which are pathways that regulate B-cell development, hematopoiesis and proliferation, respectively (Tschuch, 2006). Deregulation of these signaling pathways by *C13ORF1* can directly be linked to the pathogenesis of CLL, which points to major importance of *C13ORF1* for the CLL-pathomechanism.

The first and second exons of *RFP2*, of which the sequence overlaps with the sequence of CGI C (Figure 1-11), show minor DNA-methylation (Figure 3-13). However, this is contrasted by distribution of the histone modifications. H3K4me2 is equally distributed along the sequence of *RFP2* exons and CGI C, and so is macroH2A. Only one locus within CGI C (Figure 3-13) shows the highest detected enrichment of macroH2A in this region. The candidate gene *RFP2* behaves very peculiar when considering the analyzed epigenetic modifications together with the known transcriptional state of the gene. *RFP2* is monoallelically expressed in non-malignant B- and T-cells and also in CLL cells (Mertens et al., 2002). However, in contrast to *C13ORF1*, the monoallelic expression is not reflected in the histone code, as no significant promoter-restricted H3K4me2-enrichment was detected (see below, and Figure 3-11). Neither was a distinct epigenetic code determined for *RFP2* and CGI C. All three analyzed epigenetic marks are found at *RFP2* and CGI C in medium levels, but none of them points to a distinct transcriptional state of the gene (i.e. *RFP2* is neither marked as active nor as inactive). The lack of an unambiguous epigenetic mark at *RFP2* and CGI C on the other hand supports the model of two epigenetically distinct copies of 13q14.3. As the abundance of active and inactive marks in medium amounts, may reflect that each mark is enriched at only one of the two copies. If one copy carries only active marks and the other only inactive, the performed analysis that cannot discriminate between the two alleles but measures the marks of both copies together would produce the observed results. *RFP2* could therefore be an interesting candidate to be analyzed in more detail with respect to differential chromatin packaging of the two 13q14.3 copies. It will be interesting to learn how the two alleles are differentially marked at the *RFP2* locus and how monoallelic expression of the gene is epigenetically regulated.

Interestingly, the two ncRNA genes and the CGIs D and E that are located in the 5'-regions of the ncRNA genes show a special distribution of epigenetic marks that was not observed anywhere else in the critical region in 13q14.3 (Figure 3-13). Regarding the epigenetic modifications at CGI D, the CpG-island is divided into two parts. Most of CGI D is actively marked (Figure 3-13). In the centromeric part we find high enrichment of H3K4me2, low macroH2A-enrichment and no DNA-methylation. This abruptly changes at the most telomeric part of the CGI, which overlaps with the first and second exons of *BCMS*. Here significant silencing was measured, as H3K4me2-enrichment is comparatively low and DNA is heavily methylated (Figure 3-13). The silencing is apparently not stabilized via macroH2A-enrichment, which is very low all over the sequence of CGI D (Figure 3-13a).

CGI E again shows a mix of active and silent segments, but this pattern is not as clear concerning the distinct modifications as was described for CGI D. H3K4me2-enrichment is high all over CGI E, except at two loci located in the middle of CGI E where it is significantly

lower (Figure 3-13a). Starting at locus "E4a" DNA-methylation in non-malignant cells is observed in a stretch within the CGI, which is about 3000 bp in length (Figure 3-13b; A. Philippen, personal communication). No significant macroH2A-enrichment could be detected at the whole CGI E (Figure 3-13a, lower panel).

In summary, CGIs D and E show a mixture of activating and inactively marked parts in non-malignant cells. CGI D is marked with chromatin modifications associated with active genes except for the most telomeric part of CGI D, which overlaps with exons of *BCMS*. All three analyzed marks, H3K4me₂-, macroH2A-enrichment and DNA-methylation, are in concordance with active euchromatin at the centromeric larger part of CGI D. However, the most telomeric part, within exon two of *BCMS*, shows marks corresponding to a silent state. As was seen for the centromeric larger part of CGI D, all analyzed chromatin marks are in concordance with each other and mark the most telomeric part of CGI D as silenced. CGI E can be considered as divided into two halves: The centromeric half, which is marked as active, and the telomeric half, which is less active and shows some marks for silencing. The special epigenetic features of these two distinct loci within CGIs D and E suggest the localization of LCRs within these loci, which will be discussed in more detail below (chapter 4.3.5).

A model for epigenetic regulation of the gene expression from the critical region in 13q14.3 is summarized in Figure 4-1 at the end of chapter 4.3. The epigenetic code of the critical region in 13q14.3 determined by the analyzed chromatin marks in the present study is summarized in Table 4-1. At the five analyzed genes and CpG-islands in their promoter regions in 13q14.3, no exclusive marks for transcriptional activity or silencing were detected. However, when considering that only three epigenetic modifications of many more that constitute the epigenetic code (Figure 1-5), were analyzed here, it is possible that some specificities may be missed.

Table 4-1: Summary of epigenetic marks at the genes and CGIs located in the chromosomal region 13q14.3. +: respective mark was found; -: no mark; /: not measured.

Gene/CGI in 13q14.3	Active mark (H3K4me ₂)	Inactive mark (macroH2A)	DNA-methylation	Sum of epigenetic marks
KPNA3 CGI A	+ +	+ +	/ -	active?/inactive? active?/inactive?
C13ORF1 CGI B	- +++	++ +	/ -(+)	"monoallelic active"
RFP2 CGI C	- +	++ +	- +	inactive active?/inactive?
BCMS CGI D LCR1	- ++ (+)	++ + ++	+ - +	inactive active LCR1: inactive
BCMSUN CGI E LCR2	- ++ (+)	++ + +	/ - +	"monoallelic active" LCR2: inactive

In summary, the three marks together point to active chromatin marks at the CGIs within the 5' regions of the genes *KPNA3*, *C13ORF1*, *RFP2* and *BCMS* in non-malignant cells (Table 4-1). The active marks at these genes correlate with their transcriptional activity. Moreover, the occurrence of chromatin marks with contrary transcriptional impact at the same positions could point to allelic differences of the two 13q14 copies. The procedure applied here did not allow differentiation between the two copies; therefore, differential packaging of the alleles

can only be concluded from the observation that supposedly opposing epigenetic modifications are detected when analyzing both alleles together. This conclusion is supported by earlier findings that suggested differential chromatin packaging of the two 13q14-copies that leads to asynchronous replication timing of the two alleles (Mertens et al., 2006).

Accordingly, the genes in the critical region are monoallelically expressed from the active copy of 13q14.3, whereas the second copy is silent in non-malignant cells.

4.3.2 Histone code of monoallelic expression at 13q14.3

Most genes carry chromatin marks reflecting their expression status. Rougeulle et al. (2003) showed that imprinted and monoallelically expressed genes have promoter-restricted H3K4me2-enrichment that correlates with monoallelic expression of the genes. In order to test whether the pattern could be detected at the monoallelically expressed genes at 13q14.3, H3K4me2-enrichment was determined at genes and promoter regions. The chromatin pattern of monoallelic expression as described by (Rougeulle et al., 2003) was found at the genes *DLeu2/BCMSUN* and *C13ORF1* in the critical region in 13q14.3. *DLeu2/BCMSUN* was earlier described to be indeed monoallelically expressed (Mertens et al., 2006). However, the analysis of the chromatin pattern for monoallelic expression led to identification of another gene in the region that carries this mark, i.e. *C13ORF1*. Here we could show that *C13ORF1* is another gene in the critical region that is monoallelically expressed in non-malignant cells (Figure 3-11). Thus, it became clear that almost all genes in the critical region in 13q14 are monoallelically expressed in non-malignant cells. Monoallelic expression is determined for the genes *C13ORF1*, *DLeu2/BCMSUN*, *RFP2* and splicing variants of *BCMS* (see Figure 3-11b and Mertens et al., 2006)), while *KPNA3* is biallelically expressed. All of these genes but *RFP2* were found to carry the histone mark that reflects their expression status. Except for the biallelically expressed *KPNA3* and the short splicing variant of *BCMS* that showed no restricted H3K4me2-enrichment, monoallelically expressed genes like *C13ORF1* and *DLeu2/BCMSUN* showed promoter-restricted H3K4me2-enrichment (Figure 3-11). *RFP2* was shown to be monoallelically expressed in B- and T-cells (Mertens et al., 2006), but promoter restricted H3K4me2-enrichment was not detected.

The chromatin modification pattern that was correlated with monoallelic expression is remarkably stable regarding different tissue types and different ages of analyzed cells. It was found to be exactly the same in almost all subgroups of healthy probands, i.e. in probands below and above the age of 45 years (young and old wild type subgroup) and also in B- and T-cells. Nonetheless, two exceptions were detected. The first exception is observed at *DLeu2/BCMSUN* with both eligible promoters (CGI D and E), where promoter-restricted H3K4me2 could not be detected in probands below the age of 45 years despite the fact that the gene is monoallelically expressed (Figure 3-12). The second exception is found in the B-cell subgroup, where the promoter-restricted H3K4me2-enrichment at *DLeu2/BCMSUN* could only be detected when taking CGI E into account and not with CGI D. These observations suggest that promoter-restricted enrichment of H3K4me2 might be dynamic. The chromatin mark for monoallelic expression might be set up at *DLeu2/BCMSUN* at a later point in lifetime (after the age of 45 years) and not early in development as has been described for other chromatin marks and for the epigenetic code of imprinted genomic regions (Choo et al., 2006; Wood and Oakey, 2006; Reik et al., 2001; Santos et al., 2002). The epigenetic code for monoallelic expression that is found at *DLeu2/BCMSUN* with the promoter consisting of CGI D might on the other hand also be tissue-specific. The pattern for monoallelic expression is found at a potential TSS of a splicing variant of *DLeu2/BCMSUN* in T-cells only. The finding

suggests specific monoallelic expression of this variant in T-cells, as it was not found in the B-cell subgroup. It remains to be determined whether the “loss” of the chromatin mark for monoallelic expression in B-cells is due to tissue-specificity of monoallelic expression of the *DLEU2/BCMSUN* splicing variant.

The promoter-restricted H3K4me2-enrichment of monoallelically expressed genes in 13q14.3 is not changed in CLL cells (Figure 3-18). The chromatin mark was detected at the monoallelically expressed genes *C13ORF1* and *DLeu2/BCMSUN* also in CLL cells and it was not detected at any biallelically expressed gene of 13q14. The chromatin pattern that reflects the expression status of the genes therefore seems not to be affected by an epimutation or aberrant epigenetic changes in CLL.

4.3.3 Are epigenetic marks at 13q14.3 changed during aging?

Some epigenetic changes can be assigned to environmental and stochastic changes that occur during the lifetime of an organism (Fraga et al., 2005a). Changes in the epigenetic landscape that have been associated with aging include DNA-methylation changes and changes in histone modifications H4K16ac and H4K20me3 (Fraga and Esteller, 2007).

Therefore, it was essential to determine epigenetic modifications in age-matched non-malignant cell samples in order to be able to later distinguish epigenetic aberrations related to CLL from epigenetic changes that occur during aging. Healthy probands were stratified according to age with a threshold set at 45 years. The “young” subgroup included five samples of probands below the age of 45 years. The mean age of the young subgroup was 32 years with a range of 23 to 43 years. The “old” subgroup contained six samples of probands with an age of 45 years and above. The mean age of the old subgroup was 50 years with a range of 48 to 54 years. Unfortunately, an ideal CLL age-matched group of probands with a mean age of 65 years was not available. The differences of the age of probands subjected to the two subgroups may not be sufficient to detect significant differences between the subgroups. Neither H3K4me2- nor macroH2A-enrichment at 13q14.3 was significantly changed in young or old subgroups compared to each other (Figure 3-14;

Table 3-4). No significant differences could be measured between the subgroups; however, some tendencies were noted. First, H3K4me2-enrichment tended to be higher in the old subgroup at most 13q14 genes and CGIs, except for the genes *C13ORF1*, *BCMS* and *DLeu2/BCMSUN*. This indicates epigenetic changes at 13q14.3 that occur during aging. At the exonic regions of three genes aging-related epigenetic modifications were correlated with a modest gain of active chromatin marks. Nevertheless, the differences in enrichment were not significant and thus they can also be related to interindividual differences of epigenetic modifications. Second, macroH2A-enrichment was higher enriched in the young subgroup. MacroH2A was higher at four of the genes in 13q14.3 and at all five CGIs, the only exception was *KPNA3*, which was higher macroH2A-enriched in the old subgroup. Thus, the inactive chromatin mark seems to get lost during aging.

No significant p-values to describe differences between the subgroups could be calculated (Table 3-4); this might, in addition to other reasons, be due to the similarity of the two age groups. The majority of the eleven analyzed healthy probands was between 42 to 50 years old. The small variation of the age of the probands led to the cut-off threshold for the young subgroup and the starting age of the old subgroup of 45 years. Thus, the selection of the threshold was limited by the available sample variation. However, these probands may be

too similar regarding their aging-related epigenetic modifications, to allow detection of differences between the two subgroups. Another explanation might be that the threshold of 45 years was either not high enough or too high. The epigenetic changes during aging are maybe occurring rarely and more slowly so that no changes are detectable in 10 years time period.

In order to detect epigenetic changes during aging in hematopoietic cells or B-cells, it will be necessary to analyze more samples in each age group and maybe to redefine the threshold used to stratify the groups. Another possibility to gain insight into the aging related epigenetic changes might be to analyze a healthy "old" group with a median age of 65 years (i.e. with the same median age as CLL patients). However, this age group is usually not amenable to sampling, especially not via blood donation. It would further be interesting to analyze several healthy young age groups with a median age below 20 years and below 10 years, for instance. Again, the problem could be sampling, especially with the very young probands. However, if samples could be collected, it would be possible to detect significant differences of epigenetic modifications in non-malignant cells that are related to the aging process. In addition, with this kind of set-up it would be possible to determine at which age the first epigenetic changes appear in lifetime.

It is not clear to which extent the observed changes in histone modifications at 13q14 have an impact on altered expression of the genes. It is well possible that these epigenetic marks are only accumulating during the cellular aging process, without direct effects on transcriptional regulation, unless they reach a specific threshold. However, it is highly likely that the epigenetic differences observed in old cells when compared to young, predispose these aged cells to the tumorigenic transformation process as was already described (Fraga et al., 2007). H3K4me2 enrichment is higher in older cells and therefore it could be easier to transform these cells towards CLL, where a significantly higher H3K4me2-enrichment was detected (Figure 3-17; chapter 4.4). The aberrant H3K4me2-enrichment in CLL cells leads most likely to up-regulation of expression of the noncoding RNA, which in turn might induce aberrant regulation of the critical region in 13q14. Meanwhile, a repressive mark is lost during the aging process, which might contribute to aberrant activation of distinct promoters of the critical region in 13q14.3 observed in CLL cells.

In the present study, evidence was found that the epigenetic changes of non-malignant cells during aging predispose aged cells to tumorigenic transformation. In hematopoietic cells, these changes include a modest gain of an active chromatin mark at the CpG-islands in 13q14.3 along with modest loss of an inactive mark at the critical region. With the experimental set-up chosen here, no significant differences between young and aged non-malignant cells could be detected. However, both changes of epigenetic modifications are determined as substantial aberrations in transformed CLL cells (chapter 4.4.1).

4.3.4 Tissue-specificity of epigenetic marks at 13q14.3. Are there differences between B- and T-cells?

Cell types of different tissues hold specific epigenomes, which is the key through which multicellular organisms can exist where different cell types harbor an identical genome. The components of epigenomes are DNA-methylation, post-translational modifications of histones (PTM) and high order organization of chromatin in cell nuclei (Gargiulo and Minucci, 2008). CLL is a B-cell disease, i.e. aberrant CLL cells are derived from B-cells. However, the genes located in the critical region in 13q14.3 were shown to be monoallelically expressed in both

B- and T-cells derived from healthy probands (Mertens et al., 2006). As the expression pattern of 13q14.3 genes is essentially the same in B- and T-cells the epigenetic mark that regulates the expression has to be set early in development. The epigenetic regulation of monoallelic expression of 13q14.3-genes is set most likely at the stage of a lymphocyte progenitor cell state, and maintained during further differentiation into B- and T-lymphocytes (Figure 1-6). Therefore, with respect to monoallelic expression of the 13q14.3 candidate genes, it was interesting to also determine epigenetic modifications in both B- and T-cells. Additionally, the fact that malignant CLL cells are derived from B-cells poses the questions whether samples consisting of unselected peripheral blood lymphocytes (PBLs), which were used as non-malignant control, are the right control tissue to be able to detect epigenetic differences or aberrations in CLL cells. In order to address these questions, two samples of CD19⁺-sorted B-cells and of T-cells derived from two healthy donors were analyzed for H3K4me2-enrichment at 13q14. The detected B-cell profile of H3K4me2-enrichment of 13q14 is similar to the PBL profiles determined earlier (Figure 3-10; Figure 3-15). When comparing B- to T-cell profiles (Figure 3-15), a generally up to ten-fold higher H3K4me2-enrichment was determined in B-cells at 13q14.3. Interestingly, the greatest enrichment fold-changes are evident at CGI B, D and E where the differences reach a factor of ten. In contrast, the overall pattern of H3K4me2-distribution at 13q14.3 is the same in B- and T-cells. However, the H3K4me2-enrichment in T-cells never reached a value of 10 fold enrichment versus *CDH12*. These results suggest that the T-cell influence in PBL control samples for CLL is negligible as it was found to be generally below ten-fold enrichment versus *CDH12*. Moreover, H3K4me2-enrichment was found to be significantly lower in T-cells when compared to B-cells. Interestingly, the lower enrichment of the active chromatin mark H3K4me2 at 13q14.3 was in line with the previously described lower expression of the candidate genes of 13q14 in T-cells (Mertens et al., 2002). It will be interesting to further analyze if epigenetic modifications are different in B- and T-cells and if differential modifications have an impact on the differential gene expression in these cell types. It can be concluded that the T-cell-specific influence in PBL samples is negligible and that PBL samples can thus be used as control tissue for CLL cell samples. It seemed thus feasible to compare the PBL profiles of epigenetic modifications at 13q14.3 to CLL profiles, with the purpose to identify epigenetic aberrations in CLL cells.

4.3.5 Differential epigenetic marks at the two 13q14 copies? Model of the epigenetically regulated tumor suppressor mechanism in the critical region in 13q14.3.

Analysis of allelic variation in gene expression and of allelic variation in epigenetic modifications helps to identify regulatory single nucleotide polymorphisms (SNPs) and regulatory genomic elements (Kadota et al., 2007). The most characteristic features of locus control elements (LCRs) that regulate transcription of gene clusters in mammalian genomes, are differential epigenetic modifications of the two alleles (Li et al., 2002). Differential epigenetic modifications include DNA-methylation, but also histone modifications. LCRs are described further to be differentially methylated (and probably modified) with respect to the region they are localized.

In this thesis, two candidate LCRs in the critical region in 13q14.3 could be identified, the first within a telomeric 2 kb part of CGI D and the second within a 2 kb part in the middle of CGI E (chapter 4.3.1; Figure 3-13). Next, healthy probands were recruited that are heterozygous for a SNP located nearby the candidate LCR locus. Allelic enrichment of both

macroH2A and H3K4me2 was determined by ChIP combined with subsequent sequencing of the identified SNPs (LCRs). Thus, potential differential allelic enrichment of the two chromatin marks could be detected at the LCRs in samples derived from heterozygous probands. Analysis of these heterozygous samples allowed discrimination of the two copies and thus sequencing of ChIP-DNA precipitates could be used to detect differential enrichment at the two copies. Clear allelic differences were only detected at the candidate LCR within CGI E (Figure 3-16). At the locus in CGI E H3K4me2 was indeed enriched at only one copy of 13q14.3 in one of five probands (Figure 3-16a). A tendency for monoallelic enrichment of H3K4me2 could also be detected at the second potential LCR within CGI D, because only one of the two copies was found to preferentially carry the active chromatin mark in two of six analyzed probands (Figure 3-16c). Surprisingly, macroH2A-enrichment showed no clear differential enrichment between the two copies of 13q14.3 at CGI E (Figure 3-16b) and a preference for monoallelic enrichment in two of four probands at CGI D (Figure 3-16d). The allelic differential enrichment of H3K4me2 and macroH2A point to involvement of these specifically marked loci in the regulation of the tumor suppressor mechanism. However, allelic enrichment of H3K4me2 could only be detected in one out of five probands, and a preference for monoallelic enrichment of H3K4me2 in two out of six probands, which might be due to the location of the analyzed SNP. The SNPs are probably located near but not exactly in the LCRs. Unfortunately, allelic enrichment cannot be further analyzed because no other SNPs are available in the determined candidate LCR regions.

In addition to characterization of epigenetic regulation by candidate LCRs in the critical region in 13q14.3, the candidate genes in the region are mostly monoallelically expressed, except for *KPNA3* and *BCMS* (Figure 3-11 and (Mertens et al., 2006)). Both candidate LCRs are found within CGIs that are located in the 5' region of the two large noncoding RNA genes *BCMS* and *BCMSUN*. Several large noncoding RNA (ncRNA) species have been described to control gene expression (Zaratiegui et al., 2007; Andersen and Panning, 2003; Shamovsky and Nudler, 2006) with one of the best described examples for a ncRNA that regulates expression of an imprinted gene cluster being the ncRNA *Air* (Rougeulle and Heard, 2002). Moreover, the majority of large ncRNAs have been described to be involved in regulation of expression of other genes in *cis* (Mattick, 2003). Imprinted genes have been shown to be organized as linkage groups in gene clusters (Lafon-Hughes et al., 2008), which might be true also for the candidate genes in the critical region 13q14.3 as these share several similarities with imprinted loci, although the region is not imprinted (Mertens et al., 2006). Furthermore, it has been shown that transcription of regulatory ncRNAs located in imprinted gene clusters, are regulated by so-called imprinting control elements (ICEs), which belong to the family of LCRs (Pauler et al., 2007; Li et al., 2002). Methylation of CpG-dinucleotides in imprinted regions create differentially methylated regions (DMRs) either in the active or the inactive copy (Lafon-Hughes et al., 2008).

Thus, a potential functional gene cluster in the critical region in 13q14.3 is postulated, where the transcription of the genes is interdependent and/or depending on epigenetic regulation. The postulation of a functional gene cluster in 13q14.3 comes from several findings within our group. The identification of candidate LCRs in the critical region in 13q14.3 that regulate the transcription of the large ncRNA genes indicates a regulatory role of these ncRNAs. If the two large ncRNAs are involved in the regulation of the candidate genes, the genes are most likely regulated together as a gene cluster. Moreover, (Tschuch, 2006) found the genes to be involved in the same signaling pathways in a functional study of the candidate genes. Recently, (Serra-Barrionuevo, 2008) also showed interaction of the proteins from 13q14.3 in

a yeast-two hybrid screen and furthermore, cooperation of these proteins in large protein complexes and in the same signaling pathways. The findings of cooperative involvement of the proteins in the same signaling pathways and/or in the same protein complexes point to their functional relationship with each other.

Considering the epigenetic modifications detected in the present study at the gene cluster and at functional elements in the critical region, a model of epigenetic regulation at 13q14 is suggested that consists of three features:

- i) 13q14 is present in one active copy and one silent copy in non-malignant cells,
- ii) The large ncRNAs *BCMS* and *BCMSUN* most likely are involved in the regulation (in *cis*) of expression of the other genes in 13q14, and
- iii) Differential epigenetic modifications of a candidate LCR region controls transcription of the large ncRNA genes, and via regulation of the ncRNAs the LCR controls transcription of the gene cluster in the critical region (Figure 4-1).

Evidence for this model, especially for the existence of one active and one silent 13q14.3 copy comes on the one hand from the monoallelic expression of the genes, which was shown for all genes but *KPNA3* in the region (Figure 3-11; and (Mertens et al., 2006)). On the other hand the determined allelic variation in the enrichment of histone modifications at the candidate LCRs in CpG-islands D and E are strong hints that support the model for epigenetic regulation of the tumor suppressor mechanism in 13q14.3.

4.3.6 Consequences of the epigenetic code at 13q14.3 for the tumor suppressor mechanism in the critical region 13q14.3

The epigenetic marks determined in this study are not distinct on a gene-level (Table 4-1), except for the promoter-restricted H3K4me2-enrichment at CGI B and *C13ORF1*, which is a distinct chromatin mark of monoallelic expression of *C13ORF1*. At most other CpG-islands in the critical region in 13q14.3, a bivalent epigenetic pattern is determined, as both marks for active and inactive transcription were detected (Table 4-1; summarized in Figure 4-1). The existence of bivalent promoters was suggested primarily in embryonic stem cells, where bivalent chromatin domains carrying both repressive and activating histone modifications mark promoters of developmental regulator genes (Ku et al., 2008). The finding of 'bivalent' marks at CpG-islands located within promoter regions in 13q14.3 in differentiated non-malignant cells, however, rather point to differential modifications at the two copies of 13q14. Allelic variation of enrichment of active and inactive chromatin marks was shown for a small locus in CGI E (chapter 4.3.5; Figure 3-16). The observation of activating and repressive chromatin marks at the same gene locus identified two types of loci; the first were imprinted loci and abundance of contrary marks is explained by allelic differences. The second includes loci that were known or suspected to be regulated by antisense interfering transcripts that arise from CpG-islands at the 3'-region or by LCR-controlled ncRNAs (Mellor et al., 2008).

Whether these modifications are mutually exclusive or both present on the same allele awaits further analysis of SNPs in the region. On the other hand, the 'bivalent'-like pattern in promoter regions (CGIs) of the genes in the critical region may also point to incomplete epigenetic silencing/activation of the respective gene locus. Nonetheless, the expression

status of the genes in the critical region in 13q14 is reflected in the epigenetic modifications detected at the CGIs in their 5'-regions (Figure 4-1a).

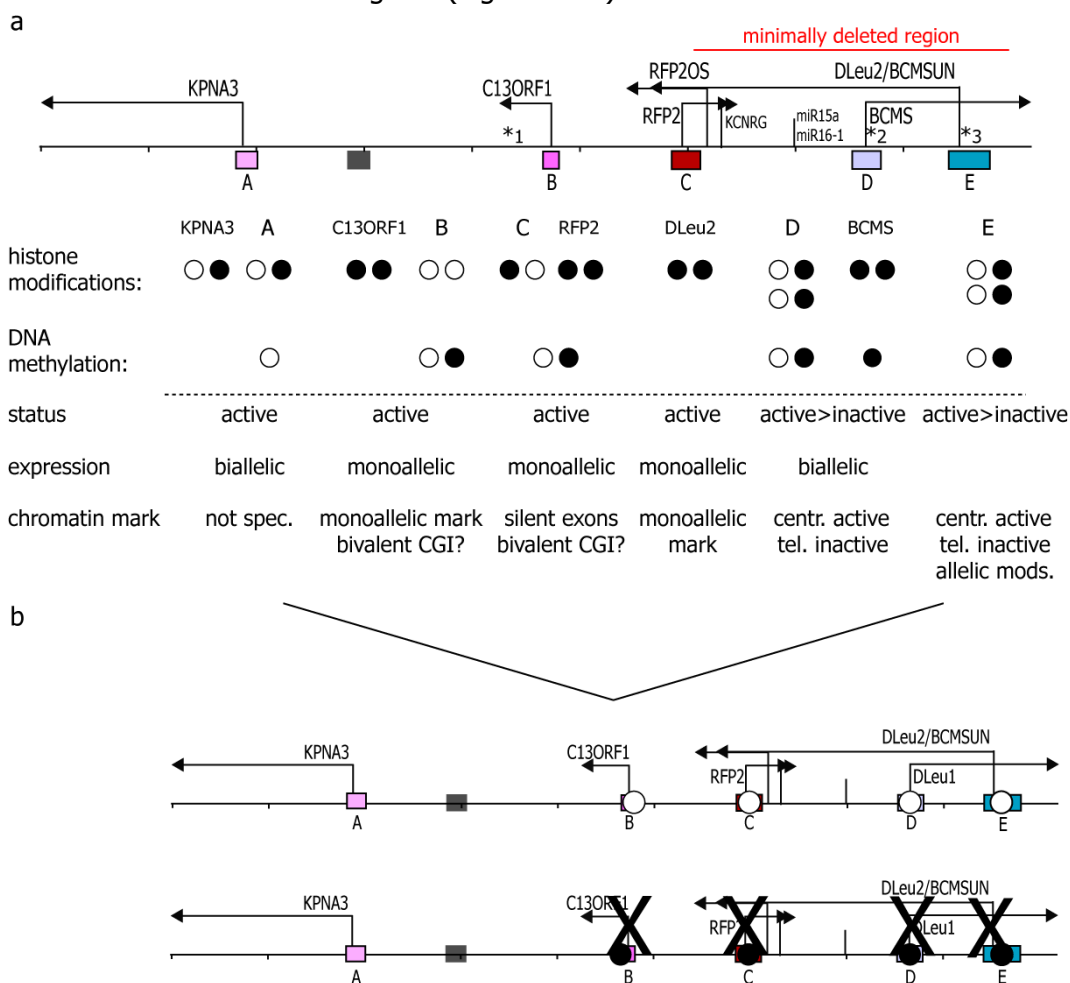


Figure 4-1: Epigenetic modifications at 13q14.3 in non-malignant cells. Model of the epigenetic regulatory mechanism in 13q14.3; white circles = active chromatin mark; black circles = inactive chromatin mark. (a) Schematic summary of epigenetic modifications at genes and CGIs of 13q. (b) Proposed model of resulting gene expression from two copies of 13q14. The regulatory impact of the ncRNAs is omitted in the figure.

The epigenetic pattern for monoallelic expression (Rougeulle et al., 2003) was found at the monoallelically expressed genes *C13ORF1* and *BCMSUN*, but not at the biallelically expressed genes *KPNA3* and *BCMS*. *RFP2* was previously described to be monoallelically expressed (Mertens et al., 2006); still, promoter-restricted H3K4me2-enrichment at *RFP2* and CGI C could not be detected in the present study (Figure 3-11).

The epigenetic code of the two large ncRNA genes and the two CGIs in their 5'-regions that span the critical region in 13q14.3 show peculiarities that are not observed anywhere else in the region. Both CGIs are divided into a centromeric active part and a telomeric inactive part (i.e. half in the case of CGI E; Figure 4-1a; Figure 3-13). Moreover, evidence for differential allelic enrichment at breakpoints within the CGIs was detected (Figure 3-16). The abrupt change in chromatin marks within the CGIs suggests the existence of candidate LCRs that could be involved in regulation of transcription in 13q14.3. The transcription of the ncRNAs seems to depend on the epigenetic modification status of the LCRs that determine whether it is active or inactive. Thus, the epigenetic status of the LCRs could determine the

transcriptional status of the ncRNAs, which in turn regulate expression of the other genes in the critical region of 13q14.3. This model of regulation is analogous to models described for imprinted loci that are regulated by transcriptional regulation of large noncoding RNAs (Rougeulle and Avner, 2004). Further investigations are important to determine exact functions and regulation of the tumor suppressor mechanism in 13q14.3. Overexpression of the ncRNAs in tumor cell lines has no effect on the expression of the other genes in 13q14.3 (D. Mertens, personal communication). Therefore, in order to determine the regulatory function of the ncRNAs, it will be necessary to transplant the whole genomic region into B-cell progenitor cells, and to analyze the resulting epigenetic modifications and the expression of the other genes in the critical region. For linking differential chromatin modifications at the potential LCRs in 13q14.3 to transcription of the ncRNA genes, further studies implying modulation of epigenetic modifications at the LCRs in cancer cell lines are needed. The proposed studies would need to demonstrate significant correlation of the epigenetic status of the LCR with the transcription rate of the ncRNA genes.

Finally, the epigenetically regulated tumor suppressor mechanism in 13q14.3 is not an imprinting mechanism, but results in monoallelic expression of the genes in the critical region. The regulatory impact of the epigenetically controlled transcription of the ncRNA genes in the critical region in 13q14.3 has not yet been fully understood. However, the findings of this study suggest that in non-malignant B- and T-lymphocytes derived from healthy probands the two 13q14.3-alleles differ, which results in one active and one silent copy. Thus, differential chromatin packaging of the two copies of 13q14.3 leads to the observed monoallelic expression of the candidate genes in the region (Figure 4-1).

4.4 The aberrant epigenetic code of 13q14.3 in CLL

The development of cancer in humans is the result of a multi-step process that involves activation of oncogenes and/or inactivation of tumor suppressor genes (TSGs) (Hanahan and Weinberg, 2000). These two steps arise not only due to mutations, but also as the result of a translocation or an altered transcription rate (Spannhoff et al., 2008). Histone modifications and DNA-methylation have been recognized as important features and markers of tumorigenic transformation of cancer cells (Esteller, 2006). To date, several studies provided important insights into epigenetic alterations of cancer cells and into the contribution of epigenetic alterations in the pathogenesis of tumors (Gargiulo and Minucci, 2008).

Several genomic aberrations have been well described in CLL and could be linked to prognosis and survival of CLL patients (Döhner et al., 2000). The high frequency of deletion of a critical region of chromosomal band 13q14.3 in the majority of CLL patients points to a tumor suppressor mechanism in the region that is causally involved in the pathogenesis of CLL (chapter 4.3.6). The genetic aberrations in CLL have been characterized in detail in the last decades by many groups (Liu et al., 1993; Döhner et al., 1997; Liu et al., 1997; Kapanadze et al., 1998; Stilgenbauer et al., 1998; Döhner et al., 1999; Hamblin et al., 1999; Döhner et al., 2000; Kapanadze et al., 2000). However, these extensive studies failed to detect a pathogenic point mutation of a putative tumor suppressor gene in 13q14 in CLL cells. The absence of pathogenic mutations within candidate genes from analysis of patients with monoallelic loss at 13q14 suggests either haploinsufficiency or epigenetic suppression of the retained allele. Recent data support the latter suggestion as favoured mechanism (Dickinson et al., 2006; Haslinger et al., 2004; Kienle et al., 2005; Mertens et al., 2006). An epigenetic tumor suppressor mechanism in 13q14 was postulated (Mertens et al., 2006) that results in monoallelic expression of the gene cluster in the critical region and allelic silencing of 13q14 in non-malignant cells (chapter 4.3). This tumor suppressor mechanism would explain the almost complete downregulation of candidate genes and microRNAs observed in CLL (Calin et al., 2002; Cotter and Auer, 2007; Mertens et al., 2002). To date, none of the described targets of the microRNAs including *BCL2* (Cimmino et al., 2005) and *CCND1* (Chen et al., 2008), was shown to be specifically down-regulated by the miRNAs in CLL cells. Calin et al. (2002) showed that absence of miR15/miR16 in CLL is associated with mutated *IgV_H*-genes and with deletions from the critical region in 13q14.3. Normally, miR15/miR16 are involved in the control of gene expression of their proposed target *BCL2* (Cimmino et al., 2005). Thus, their absence in CLL could play a major role in the prevention of apoptosis that is observed in CLL cells. However, the absence of miR15/miR16 was not confirmed in another study focusing on these two microRNAs in CLL: Fulci et al. (2007) found low expression of the miRNAs in only 12 % of patients analyzed, despite 54 % having a deletion of 13q14.3. All biallelically deleted del(13q) patients showed striking downregulation of miR15. However, this pronounced downregulation of both miRNAs was not paralleled by any significant increase in *BCL2* expression (Fulci et al., 2007).

Therefore, the link of the microRNAs and CLL pathogenesis is considered to be a part of inactivation of the tumor suppressor mechanism in 13q14.3 rather than the cause for inactivation in CLL. The microRNAs are most likely co-regulated by the large ncRNAs in the critical region as part of the functional gene cluster. The postulated model for the tumor suppressor mechanism involves co-regulation of candidate genes including the protein coding and microRNA genes, by the two large noncoding RNAs, *BCMS* and *BCMSUN* that span the critical region.

In the current study, epigenetic modifications at 13q14.3, including H3K4me2- and macroH2A-enrichment as well as DNA-methylation, were analyzed in CLL cells to identify an epimutation that can lead to deregulation and subsequent inactivation of the postulated tumor suppressor mechanism in 13q14.3 in CLL patients. The aging process could possibly account for epigenetic changes occurring later in life time. Also, tissue specific differences between B- and T-cells might account for interindividual differences in the epigenetic code detected in non-malignant cells, but no significant differences could be determined when comparing B- and T-cells of healthy probands.

Hence, the determined epigenetic features of the critical region in non-malignant cells (Figure 4-1) together with the proposed model of epigenetic regulation of 13q14.3 (chapter 4.3.6) can now be used to identify epigenetic aberrations at 13q14.3 in CLL cells. Elucidation of the epigenetic regulatory mechanism was used to determine aberrations in CLL that might lead to deregulation or inactivation of the epigenetic tumor suppressor mechanism in 13q14. By comparison of modification patterns in CLL and in non-malignant cells, aberrations of 13q14 in CLL cells could be identified. These aberrations will be discussed in the following chapter and possible effects on the function and regulation of the tumor suppressor mechanism will be considered.

4.4.1 Aberrant chromatin modifications at 13q14.3 in CLL

Global DNA-hypomethylation coupled with promoter hypermethylation has been accepted as supplementary hallmark of cancer (Gargiulo and Minucci, 2008; Esteller, 2006). Additionally, it has been suggested that histone marks are also specifically deregulated in cancer cells (Fraga et al., 2005b) and all kinds of epigenetic modifications constitute a tumor epigenome that distinguishes cancer cells from non-malignant cells (Esteller, 2007). Thus, alterations of the epigenetic code are directly linked to an aberrant gene expression in tumor cells. To date, epigenetic aberrations detected in CLL include global DNA hypomethylation (Wahlfors et al., 1992), promoter hypermethylation and consecutive silencing of the death-associated protein kinase 1 (*DAPK1*) gene (Raval et al., 2007), of the transcription factor TWIST2 (Raval et al., 2005), and also genome wide promoter hypermethylation (Raval et al., 2006). Furthermore, the epigenetic alterations described in CLL lead to deregulation of the genes *ID4* (Yu et al., 2005), secreted frizzled-related protein 4 (*SFRP4*) (Liu et al., 2006) and of various other genes (Smiraglia and Plass, 2002).

Some of the above mentioned genes as well as the microRNAs could be linked to CLL pathogenesis. The first example is *SFRP4*, which was shown to be involved in Wnt-signaling (Liu et al., 2006). Wnt-signaling controls apoptosis and is required for normal B-cell development, whereas aberrant activation of this pathway has been observed in CLL (Rosenwald et al., 2001; Lu et al., 2004). Liu et al. (2006) showed aberrant DNA-methylation and silencing of *SFRP4* as well as of additional *SFRP* family members in primary CLL samples. The family member *SFRP1* was hypermethylated and down-regulated in analyzed CLL samples suggesting this epigenetic event as critical step during leukemogenesis. Silencing of *SFRPs* in CLL by CpG-island-hypermethylation is one possible mechanism contributing to aberrant activation of the Wnt signaling pathway in CLL that results in defective apoptosis of CLL cells. Two hallmarks in the pathogenesis of CLL are the escape of CLL cells from apoptosis and the low proliferative activity of CLL cells due to their escape from the cell cycle arrest in G₀/G₁ phase. The result of the defective apoptosis *in vivo* is accumulation of CLL cells in the peripheral blood of patients (Dighiero and Hamblin, 2008).

Regarding the proposed functions and interactions of the candidate genes of the critical region in 13q14.3 (see chapter 1.3.2), their CpG-island-hypermethylation and epigenetic silencing (due to histone modifications) might lead analog to aberrant activation of Wnt-signaling pathways by SFRPs, to disturbed NF κ B-signaling by 13q14.3 candidate genes (Tschuch, 2006), which constitutes another critical step during the pathogenesis of CLL.

Here, two histone modifications and DNA-methylation at 13q14.3 were analyzed in CLL, to detect epigenetic aberrations in CLL cells that could lead to identification of an epimutation responsible for the inactivation of the tumor suppressor mechanism in CLL.

The most striking epigenetic aberration detected in CLL cells is the significantly higher enrichment of H3K4me2 at all analyzed loci in 13q14.3. Surprisingly, a significant and drastic gain of the active mark was determined at the tumor suppressor locus 13q14.3 in CLL. This was rather unexpected because with the definition of the epigenetic regulation of the tumor suppressor mechanism in 13q14.3, whereby one copy is active and the second is inactive, it was assumed that the active copy along with active epigenetic marks would be lost in CLL cells. As a loss of function of the tumor suppressive mechanism is evident in CLL cells, it was expected that the inactive copy would be retained in CLL patients. However, the significant enrichment of chromatin marks associated with transcriptional activity at 13q14 was determined as median value from all CLL samples analyzed, which derived from three del(13q) patients and two patients that retained both copies. Individual differences between these patients will be discussed below (chapter 4.4.3).

The general gain of activating chromatin marks at 13q14.3 was analyzed in more detail and showed aberrant patterns in CLL at CGI B, CGI C and CGI E. These three CGIs showed significantly higher H3K4me2-enrichment in CLL compared to non-malignant cells (Figure 3-23), which was not detected at CGI A and CGI D (Figure 3-17; Figure 3-23). In relation to the other analyzed CpG-islands in 13q14.3, at CGI D a lower enrichment of the active mark was detected in CLL cells compared to non-malignant cells (Figure 3-17). Furthermore, a slight increase of inactive chromatin marks at CGI D could be measured in CLL, which can lead to inactivation of one of the two large ncRNA genes, most likely of *BCMS*, as CGI D is located in the 5'-region of the *BCMS*. Also, a gain of active chromatin at CGI C was detected in CLL cells, which can lead to aberrant activation of the gene *RFP2*. The significant higher enrichment of H3K4me2 detected at CGIs B and E, which are located in the 5'-regions of the monoallelically expressed genes *C13ORF1* and *DLeu2/BCMSUN*, respectively, may lead to upregulation of the genes.

In summary, a general 10-1000-fold higher enrichment of H3K4me2 and a loss of the distinct distribution of H3K4me2 observed in non-malignant cells could be detected at 13q14 in CLL (Figure 3-17). Moreover, loss of active chromatin marks at CGI D was detected, which might be linked to aberrant silencing of CGI D in CLL cells. The last finding indicates a loss of activity of one regulatory ncRNA that might be subject to specific epigenetic silencing in CLL. In contrast to non-malignant cells, macroH2A is enriched at CGI D in CLL, which supports the aberrant silencing of CGI D in CLL. Indeed, one of the splicing variants of this ncRNA gene was earlier shown to be significantly down-regulated in CLL patients (Mertens et al., 2002). Loss of activity of the ncRNA might have great impact on the regulation of the critical region in 13q14.3, because the ncRNA was proposed to be involved in the epigenetic regulation of the tumor suppressor mechanism in 13q14.3 (chapter 4.3.5). The exact regulatory function of the ncRNA still has to be determined, but in the present study and in previous studies (Mertens et al., 2002; Tschuch, 2006; Serra-Barrionuevo, 2008; Wolf et al.,

2001) the importance of this ncRNA gene for the pathomechanism of CLL was shown directly and indirectly.

As a marker for repressive chromatin, enrichment of macroH2A was measured at 13q14.3. Similar to findings in non-malignant cells (chapter 4.3.1), the enrichment of macroH2A at 13q14.3 was found to be significantly lower than that of H3K4me2 also in CLL cells. In CLL a distinct pattern of macroH2A enrichment differing from that seen in non-malignant cells becomes apparent: MacroH2A is highly enriched at exonic regions and generally lower at CGIs (Figure 3-19). This was not detected in non-malignant cells, where macroH2A is spread more or less homogenously over the whole region 13q14.3 at low levels (Figure 3-12). A modest gain of the repressive mark in CLL cells, however, points to potential aberrant allelic silencing of 13q14.3 in CLL cells. As seen already in non-malignant cells, enrichment of macroH2A at promoters and genes is reciprocal to H3K4me2-enrichment. While H3K4me2 can be found in high amounts at the promoter regions and not at gene exons, macroH2A was detected in high amounts at exons and not, or only in lower amounts, at the promoters. The abundance of macroH2A at exonic regions points to a possible silencing role of macroH2A itself. MacroH2A might be involved in recruitment of the chromatin to heterochromatic regions, in recruitment of silencing factors and in the prevention of binding of transcriptional activators, which would lead to repression of transcription of underlying genes.

Some evidence is available that suggests DNA-methylation in cancer cells is not involved in the silencing of active genes, but occurs at genes that are already silent in order to silence them permanently (Lafon-Hughes et al., 2008). This observation might explain the discordance of DNA-methylation and histone modification patterns detected at some parts of the region 13q14.3 in CLL cells. The CGIs A and C are not methylated in CLL and non-malignant cells. DNA-methylation at CGI B is found in a range between 0 and 90 % in non-malignant cells, but is slightly reduced in CLL cells to a range between 0 % and 70 % (Figure 3-20). Differences in DNA-methylation at 13q14.3 between non-malignant cells and CLL cells are found at CGIs D and E. In non-malignant cells CGI D shows DNA-methylation at its most telomeric parts, whereas in CLL cells methylation at these loci is partially lost. Some CLL patients have completely lost the methylation at this locus, whereas others still maintain the same degree of methylation that was observed in non-malignant cells (70-100 %).

The same phenomenon was observed in CGI E. DNA-methylation is lost in CLL at locus "E5a" (Figure 3-23), which is methylated in all analyzed non-malignant cell samples, but only in half of the CLL samples. Thus, analysis of DNA-methylation (Figure 3-23) indicates potential silencing of the candidate LCR in CLL cells. The DNA-methylation patterns determined at the candidate LCR in CGI E are not distinct, nonetheless, all findings both in non-malignant and CLL cells at candidate LCRs within CGIs D and E, point to a unique function of these loci that are potentially involved in regulatory control of gene expression from the critical region in 13q14.3 and are thus candidate LCRs.

When considering the epigenetic code of 13q14.3 that is constituted by the three analyzed chromatin marks, differences between non-malignant and CLL cells can be detected that lead to identification of epimutations in CLL (Table 4-2).

In CLL cells the chromatin at CGI A in the 5' region of the gene *KPNA3* in 13q14 is not clearly marked. Figure 3-20 shows abundance of the active chromatin mark (H3K4me2-enrichment) together with comparatively high enrichment of the inactive mark macroH2A. However, no

DNA-methylation was detected at CGI A (data not shown). Altogether the three different epigenetic marks do not fit the current models of a clear epigenetic code at CGI A, which could lead to impaired expression of *KPNA3* in CLL. In non-malignant cells, at CGI A no significant level of any of the three epigenetic marks was detected (Table 4-1).

Table 4-2: Summary of epigenetic marks at the genes and CGIs located in the chromosomal region 13q14.3 in CLL cells. +: respective mark was found; -: no mark; /: not analyzed.

Gene/CGI in 13q14.3	Active mark (H3K4me2)	Inactive mark (macroH2A)	DNA-methylation	Sum of epigenetic marks	Aberrations compared to non-malignant cells (epimutations)
KPNA3 CGI A	(+) (+)	++ ++	/ -	(active?)/inactive? (active?)/inactive?	more macroH2A less H3K4me2
C13ORF1 CGI B	- +++	++ +	/ -(+)	"monoallelic active" active	none
RFP2 CGI C	(+) +++	(+) (+)	- +	active active	more H3K4me2 less macroH2A
BCMS CGI D LCR1	(+) (+) +++	(+) (+) (+)	+ - -	inactive inactive? LCR1: active	silenced localized activation
BCMSUN CGI E LCR2	- +++ ++	++ + (+)	/ + +	"monoallelic active" active? LCR2: inactive	none localized activation

In CLL cells, however, H3K4me2 is detected in comparably low amounts along with a significantly high enrichment of macroH2A. MacroH2A enrichment suggests aberrant silencing of the promoter region, but enrichment of macroH2A does not coincide with DNA-methylation. However, enrichment of macroH2A could explain the aberrant silencing of *KPNA3* in CLL (Mertens et al., 2002). Silencing of *KPNA3* is interesting as it is one of the genes in 13q14 linked to NFκB-signaling. It has been shown to be a potential carrier protein that shuttles NFκB into the nucleus (Tschuch, 2006; Fagerlund et al., 2005; Fagerlund et al., 2008). This points to a potential selective advantage of CLL cell clones that have partially lost *KPNA3* and are therefore less efficient in translocating NFκB into the nucleus, where NFκB regulates expression of its target genes. This would explain how NFκB-function might be lost in CLL cells.

The epigenetic marks at CGI B, *C13ORF1*, CGI C, *RFP2* and *BCMSUN* remain unchanged in CLL cells compared to non-malignant cells, except for the overall 10 to 100 fold higher enrichment of histone modifications described earlier. The chromatin mark of monoallelic expression at *C13ORF1* (and CGI B) is unaffectedly stable in CLL cells. *RFP2* carries active marks and no inactive mark (no macroH2A and no DNA-methylation) and *BCMSUN* exons are marked as silent with no H3K4me2- and high macroH2A-enrichment (Figure 3-20).

The overall epigenetic code of candidate genes in 13q14.3 in and outside of the critical region is practically the same in non-malignant and CLL cells. Nonetheless, also aberrant chromatin modifications at 13q14.3 were detected when comparing CLL cells to non-malignant cells. A general higher H3K4me2- and a higher macroH2A-enrichment were detected in CLL cells. Significant differences of enrichment of chromatin marks between non-malignant and CLL cells, are detected at CGIs B, C and E and at the exonic region of *BCMSUN* (Figure 3-23). Finally, distinct epimutations in CLL cells are detected at candidate LCRs located within CGIs D and E in the 5'-regions of the ncRNA genes in the critical region.

Thus, the aberrant modifications in CLL are most evident at the promoter regions in 13q14 and at the sequence of ncRNA gene itself.

Aberrations affecting promoter or regulatory regions in diseases are quite common, as these directly control gene expression. However, activation of genomic loci by epigenetic marks is usually not achieved by alteration of only one histone mark. The histone code seems to be far more complex (Jenuwein and Allis, 2001; Jost and Galm, 2007; Strahl and Allis, 2000). Nonetheless, the significantly stronger enrichment of H3K4me2 at 13q14 in CLL is quite striking. It is highly likely that this has an effect on gene expression, but the effect appears not to be direct, as it is not measurable by upregulated expression of the candidate genes in CLL. On the other hand, also significant higher enrichment of macroH2A at CGI E and *C13ORF1* is detected in CLL cells, which could result in aberrant silencing (Figure 3-23). *C13ORF1* shows significant levels of macroH2A enrichment only in CLL patients. Nonetheless, the gene is monoallelically expressed in non-malignant and CLL cells and furthermore, the chromatin mark for monoallelic expression remains unchanged in CLL.

Considering the model of the epigenetic regulatory mechanism in 13q14.3 (chapter 4.3.6), pronounced de-regulation of the ncRNAs might have an indirect but drastic effect on the gene expression of the other genes in the critical region. The proper expression of the ncRNAs could be crucial for maintenance of the tumor suppressor mechanism in 13q14. Importantly, the expression of the ncRNA might be changed in CLL by an epimutation within their promoter region, which might have impact on the deregulation of the whole tumor suppressor mechanism in 13q14 in CLL and it might further lead to inactivation of the tumor suppressive function of 13q14. Loss of this function might be a hallmark in the process of malignant transformation to CLL.

In summary, the generally higher enrichment of both activating and repressive chromatin marks presents as an epigenetic aberration in CLL cells. They might affect the ncRNA genes as they include loss of the monoallelic chromatin mark at *DLeu2/BCMSUN* in CLL and aberrant macroH2A-enrichment at CGI E, which might lead to silencing of the ncRNA. Whether the loss of chromatin marks associated with monoallelic expression is the cause or the consequence for aberrant macroH2A at the second promoter region including CGI E is not known. Furthermore, the distinct aberrant epigenetic activation of the two candidate LCRs in 13q14.3 points to an epimutation in CLL that affects most likely the transcription of one or both ncRNA genes that span the critical region. Finally, deregulated transcription of the ncRNA(s) would suffice for deregulation and subsequent inactivation of the postulated tumor suppressor mechanism in the critical region in 13q14.3.

4.4.2 Epimutations or aging-related changes at 13q14.3?

Several changes of the epigenome of a particular cell can be associated with cellular senescence and the aging process itself (Fraga et al., 2005a; Issa, 2003). Moreover connections between aging and cancer have been made, whereby the link of these processes is the epigenetic language (Fraga et al., 2007). CLL is a disease that mainly affects elderly people with a median age of 72 years (Dighiero and Hamblin, 2008). To control for epigenetic changes that occur during aging and are, thus, unrelated to the disease itself, H3K4me2- and macroH2A-enrichment in non-malignant cells ("old wild type") that derived from an "old" organism with a median age of 51 years were analyzed and epigenetic profile of this group was compared to the CLL profile (Table 3-5). Only modest epigenetic

alterations of these "old" non-malignant cells were detected that can be related to the aging process (chapter 4.3.3). The epigenetic alterations that seem to accumulate during aging of non-malignant cells are significantly different from aberrations that are observed in CLL cells (Table 3-5). Nonetheless, the epigenetic modifications observed in "old" non-malignant cells are similar to the detected epigenetic changes in CLL cells. A modest higher enrichment of both activating and repressive marks was detected when comparing "old" to "young" non-malignant cells (Table 3-4), but differences were not significant. Moreover, when comparing these young and old wild type subgroups to CLL (Table 3-5), no significant differences can be detected that are related to the age of the subgroups (Figure 3-21).

So, indeed, an aging-related epigenetic code that is similar to the epigenetic aberrations characteristic of CLL cells, was detected in non-malignant cells. Thus, the altered epigenome determined in "old" non-malignant cells might present a predisposition of these cells to turn into CLL cells.

4.4.3 Correlation of epigenetic aberrations at 13q14.3 and genomic aberrations in five CLL patients

Genomic aberrations of CLL patients can be correlated to prognosis and survival in CLL (Döhner et al., 2000). Here, epigenetic modifications at 13q14.3 in CLL cells of five patients were analyzed. Even though statistical significance could not be reached when comparing this limited number of CLL samples, observations and tendencies inferred from these five cases will be discussed in this chapter. Some of these observations might provide a basis for further investigations and possible limitations of future experiments.

The five patients can be grouped according to their karyotypes into CLL with only one copy of 13q14 ((del)13q), and CLL patients that retain both copies, whereof one patient had a deletion of 17p13 ((del)17p) and the other one had a no genomic aberrations (norm). Considering these two CLL-patient subgroups, it was possible to analyze whether loss of genetic material from the critical region in 13q14.3 is linked to the observed epigenetic aberrations in CLL. The question arises whether epigenetic modifications at 13q14.3 would correlate with the deletion status of the region. The proposed model of epigenetic regulation of 13q14.3 comprises the differentiation of the two copies into one active and one silent copy. Depending on whether the active or the silent allele is lost in del(13q) patients, quantification of the active mark in these patients is expected to result in either equal amounts, or in no or lower detected amounts as determined in the first group. Accordingly, all three analyzed del(13q) patients (CLL2, CLL4 and CLL5; Figure 3-22), supposedly lost the active copy as these have only very low H3K4me2-enrichment at 13q14.3. Surprisingly, those two patients that retained both 13q14 copies do not have amounts of the active mark that are comparable to that of non-malignant cells. Instead, they show H3K4me2-enrichment that is around 10 to 1000-fold higher than in non-malignant cells and also 3 to 100 fold higher than in (del)13q patients (Figure 3-22). The high enrichment of H3K4me2 at 13q14 in cells that retain both copies points to further established differences between the two copies in these CLL patients.

In principle, this finding supports the model of an active and an inactive copy of 13q14.3, which was already found in non-malignant cells. The inactive copy in CLL cells, however, is marked by a lesser amount of active chromatin marks and additionally by higher amounts of the inactive chromatin mark, which was not detected in non-malignant cells. Moreover,

compared to amounts in non-malignant cells, the CLL cells with retention of both copies also gained active chromatin marks.

The analyzed del(13q) patients seem to have lost the active copy, which would explain the strong reduction of H3K4me2-enrichment detected at 13q14. Support for this finding comes from the analyzed repressive chromatin mark, macroH2A-enrichment at 13q14 (Figure 3-23): The three del(13q) patients show macroH2A enrichment at 13q14. At some analyzed loci in 13q14.3 macroH2A was found in high amounts, which resembles a selected gain of repressive marks at 13q14.3 compared to non-malignant cells. In contrast, the normal and del(17p) patients have almost no detectable macroH2A enrichment at the critical region in 13q14.3.

In conclusion, CLL cells with both copies of the critical region show a gain of the activating chromatin mark that is coupled to reduction of the repressive mark and all analyzed del(13q) patients lost the active copy. In summary, the findings suggest that in patients with retention of both copies, 13q14.3 is strongly activated by strong H3K4me2 enrichment and reduction of macroH2A.

However, the active chromatin mark in 13q14 is not equally distributed at the region neither in non-malignant cells, nor in the CLL patients analyzed. It is not only differentially enriched when comparing promoter regions to exonic sequences, but differs also within particular CpG-islands analyzed (Figure 3-10; Figure 3-22). Only one distinct locus is heavily activated in CLL cells, which is the candidate LCR within CGI D (Figure 3-22). Furthermore, the second candidate LCR located in CGI E shows also significantly higher enrichment of the active chromatin mark. In contrast to the distinct activation of the LCR locus in CGI D, the activation at the second candidate LCR affects the complete centromeric half of CGI E (Figure 3-22), pointing to CGI D harbouring the effective LCR in the region. The activation of the candidate LCR that controls expression of the regulatory ncRNA in the critical region in 13q14.3, can lead to deregulation of the ncRNA.

Considering the model for the epigenetic regulatory mechanism in the critical region in 13q14.3 (chapter 4.3.6), aberrant upregulation of the regulatory ncRNA might lead to complete silencing of other genes in the region, which are initially only downregulated upon loss of one copy. Complete inactivation of the tumor suppressor mechanism in 13q14.3 can be either caused by epigenetic deregulation of the retained copy described above or by biallelic deletions of 13q14.3 during clonal evolution.

These findings suggest an epimutation in 13q14 being responsible for the inactivation of the epigenetic regulatory mechanism of 13q14 in CLL. Inactivation of the candidate genes in the critical region likely plays a role for the pathogenesis of CLL, as the genes could be linked to essential pathways that support and control normal B-cell development (Mertens et al., 2006; Serra-Barrionuevo, 2008; Tschuch et al., 2008; Chen et al., 2008). Further, the findings in the present study suggest that the regulatory impact of the ncRNA, the LCR or another part of the critical region in 13q14 might be affected by this epimutation and thus lead to deregulation and inactivation of the tumor suppressive function of 13q14 and to development of CLL.

4.4.4 Identification of a potential epimutation in the postulated LCR of the critical region in 13q14.3

Generally, locus control regions (LCR) are described to regulate expression of gene clusters within imprinted regions (Li et al., 2002). Commonly, these LCRs are differentially DNA-methylated on the two copies and also have differential chromatin marks on the two copies. The organization of the critical region in 13q14.3 with the two large ncRNA genes and evident monoallelic expression of the genes is very similar to that of imprinted loci. However, the region was shown not to be imprinted (Mertens et al., 2006). Still, the topological organization suggests the existence of an LCR that (co-) regulates the transcription rate of the region (chapter 4.3.5).

In the present study, both promoter regions of the two large ncRNA genes were shown to contain distinct loci that have differential allelic epigenetic marks (Figure 3-16). Moreover, specific aberrant activation of the candidate LCR within CGI D was detected and also activation of the candidate LCR along with aberrant activation of the centromeric half of CGI E in CLL cells.

For instance, CGI D carries no significant amount of the repressive chromatin mark macroH2A; most of the CpG-island neither shows H3K4me2-enrichment and/or DNA-methylation. Only at the most telomeric loci analyzed in CGI D (Table 4-2; Figure 3-20), H3K4me2-enrichment in 13q14.3 was detected. Also DNA-methylation was detected in CLL cells at the most telomeric part of CGI D. However, DNA-methylation was less abundant in CLL cells than in non-malignant cells, which might point to loss of differential DNA-methylation at this locus. The loss of DNA-methylation reflects a loss of silencing and together with the gain of the active chromatin mark; this reflects an activation of the most telomeric part of CGI D, i.e. of the candidate LCR in CLL. The aberrant activation of the candidate LCR in the promoter region of the ncRNA genes would lead to aberrant transcription of the ncRNA(s), i.e. either up- or down-regulation of their transcriptional rates. Aberrant transcription of the regulatory ncRNA in the critical region would lead to deregulated expression of the other genes in 13q14.3, regardless of whether the ncRNA is activated or repressed. Measurable effects of this deregulation of the tumor suppressor mechanism might include the observed significant down-regulation of *RFP2*. Analogous to the epigenetic aberrations detected at the candidate LCR in CGI D, the second potential LCR within CGI E shows aberrant epigenetic modifications (Figure 3-20) that might lead to deregulation of an ncRNA in the critical region and to subsequent inactivation of the tumor suppressor mechanism.

In summary, two candidate LCRs in the 5'-regions of both ncRNA genes spanning the critical region in 13q14.3 could be determined. Still, functionality of these LCRs has to be further determined. However, epimutations in CLL cells were detected that specifically affect only these distinct candidate LCR loci, which points to their functional involvement in controlling the transcription of other regulatory genes, most likely the ncRNA genes, in 13q14.3. Thus, it can be explained how the detected epimutation in 13q14.3 can lead to deregulation and inactivation of the tumor suppressor mechanism. Finally, as these epimutations do not affect the complete CpG-islands or other already described functional elements in the region, but rather reside within the noncoding sequence, it becomes clear why the epimutations were not detected earlier.

4.4.5 Model of epigenetic deregulation of the tumor suppressor mechanism in 13q14.3 in CLL

Unlike other leukemias, CLL is a heterogeneous disease without a common disease-causing translocation, deletion or mutation (D'Arena et al., 2003). There are a few well characterized genomic aberrations that could be linked to prognosis and survival of CLL patients (Döhner et al., 2000). However, the pathomechanism of CLL is still unclear. Recently, epigenetic modifications have been recognized as hallmarks of tumorigenesis (Esteller, 2008). Unlike other hallmarks of tumorigenic transformation in cancer cells, epimutations are in principle reversible and thus, they are potential targets for effective new treatments in cancer. Some common epigenetic alterations have been described that specifically affect CLL cells. Several studies showed differential aberrant DNA-methylation in CLL compared to non-malignant cells (Raval et al., 2007; Plass et al., 2007; Raval et al., 2006; Yu, 2006; Rush et al., 2004; Raval et al., 2005). However, DNA-methylation is only one component of epigenetic modulation in chromatin remodeling and transcriptional regulation. Important additional components include expression of histone methyltransferases, methyl-CpG-binding proteins and histone modifications that potentially direct interaction of histone methyltransferases (HMTs) with chromatin and chromatin associated proteins and histone acetylases and deacetylases (Yu, 2006). Recently, it has been recognized that there are connections between histone methylation patterns and cancer progression and therefore, specific histone methyltransferases present as promising targets for future cancer treatment (Spannhoff et al., 2008). *SETDB2* encodes a putative HMT that is located just centromeric to the chromosomal region 13q14.3 (Serra-Barrionuevo, 2008; Yu, 2006). The substrate specificity of this putative HMT has not yet been identified. However, immunofluorescent staining of manipulated cells that over-express *SETDB2* showed predominant localization of the protein in the cytoplasm (Tschuch, 2006). This makes a role of *SETDB2* for trimethylating histone H3 at lysine 9 (H3K9me3) not impossible but unlikely (Tschuch, 2006).

Nonetheless, the HMT-activity of *SETDB2* presents a potential link between the aberrant histone methylation determined in CLL and molecular function of a candidate gene from the genomic region of 13q14.3. If *SETDB2* or another HMT with specificity for dimethylation of H3K4 would be envisaged as target for future cancer treatment, first the substrate specificity of the HMT activity and its potential involvement in alterations of the histone modifications at 13q14.3, have to be determined. Their identification represents a promising target for future treatment of CLL, as their elimination or activation could be used to restore the tumor suppressor mechanism at 13q14.3.

The involvement of aberrant histone modifications in the progression of CLL have not been fully addressed here. However, putative deregulatory effects on the tumor suppressor mechanism in 13q14.3 caused by epimutations of candidate LCRs within the region were determined in CLL cells (chapter 4.4.3 and 4.4.5). The tumor suppressor mechanism in the critical region in 13q14.3 is epigenetically regulated, which is suggested by differential chromatin packaging of the two alleles (chapter 4.3.5) and results in the detected monoallelic expression of the genes in non-malignant cells (chapters 3.2.2 and 3.2.7, Mertens et al., 2006). Loss of the single active copy of 13q14.3 in CLL patients would suffice for complete loss of function of the tumor suppressor mechanism in the critical region (Figure 4-1) and results in the observed downregulation of candidate genes and miRNA genes in CLL cells (Calin et al., 2002; Mertens et al., 2002). Here, it was determined that loss of function of the tumor suppressor mechanism can be caused either by the genomic deletion of the active copy or by epigenetic deregulation of the region caused by an

epimutation in candidate LCRs in the critical region (chapter 4.4.4). The epimutation in the LCR most probably leads to deregulated transcription of a regulatory ncRNA that is suggested to control the expression of the other genes in the critical region in 13q14.3 (chapter 4.3.6). Consequently, the loss of regulated expression of the genes in 13q14.3 would lead to inactivation of the tumor suppressive function in CLL cells. Deregulation and subsequent inactivation of the tumor suppressor mechanism by epimutations in the critical region in 13q14.3 would explain the downregulation of the candidate genes and microRNA genes determined in del(13q) CLL patients and may be causative for the pathogenesis of CLL.

In conclusion, a specific epigenetic code at 13q14.3 was identified in this thesis that is changed in CLL and could therefore regulate the tumor suppressor mechanism in the critical region in 13q14.3. Epigenetic marks were found that are specific for the previously uncovered monoallelic expression of 13q14.3 candidate genes in non-malignant cells. Evidence is presented for a model of epigenetic regulation that implies differential chromatin packaging of the two copies of 13q14.3. Moreover, the elimination of technical hurdles in the analysis of epigenetic modifications in primary hematopoietic and CLL cells is a step forward towards complete understanding of how cellular functions are regulated by the epigenome in health and disease. The established ChIP-method can be used in future for detailed and complete characterization of the epigenetic features at 13q14.3 and will help to fully elucidate the complex epigenetic regulatory network that controls the tumor suppressor mechanism in 13q14.3. Next steps should be to fully understand the pathogenicity of deregulation of the 13q14.3 tumor suppressor mechanism in CLL and thereby provide a basis for development of more efficient and specific therapies for CLL in the future.

Summary

B-cell chronic lymphocytic leukemia (CLL) is the most common leukemia among adults in the Western world with a median age of 72 years of CLL patients at diagnosis. The most common genomic aberration in CLL is the deletion of a critical region in chromosomal band 13q14.3, which is deleted in more than 50 % of patients. The high frequency of deletions of 13q14.3 in CLL and other tumors points to a tumor suppressor mechanism localized in the critical region. The candidate genes localized in the critical region in 13q14.3 have been shown to be monoallelically expressed. Towards elucidation of the tumor suppressor mechanism in 13q14.3, the epigenetic modifications of genes and CpG-islands in the region were analyzed in this thesis. Monoallelic expression of almost all genes in the critical region in 13q14.3 in non-malignant cells was detected to be marked by epigenetic modifications at promoter and exonic regions of these genes. A chromatin pattern specific for monoallelically expressed and imprinted genes, which is promoter-restricted enrichment of dimethylated lysine 4 of histone H3 (H3K4me2), was detected at the genes *DLeu2/BCMSUN* and *C13ORF1*. Detection of this specific chromatin pattern led to identification of a novel monoallelically expressed gene in the region: *C13ORF1*. Here, *C13ORF1* was in addition to *RFP2* and *DLeu2/BCMSUN*, shown to be monoallelically expressed in B- and T-cells from healthy donors.

The chromatin immunoprecipitation (ChIP) methodology was established in order to quantify two chromatin modifications at the critical region. ChIP was technically established and optimized to analyze enrichment of the histone mark H3K4me2 that is correlated with an active transcription state, and of the histone variant macroH2A, which was shown to be allelically enriched at monoallelically silenced loci, at 13q14.3. In addition, also DNA-methylation of the CpG-islands was analyzed. Thus, an epigenetic code of the candidate genes that correlates with their transcription status was determined in non-malignant hematopoietic and CLL cells. In non-malignant cells, the epigenetic code identifies one active and one inactive copy of the critical region in 13q14.3. Thereby, the genes *DLeu2/BCMSUN* and *C13ORF1* are monoallelically expressed from the epigenetically active allele and their expression status is marked by a distinct chromatin pattern. In contrast, the distinct chromatin pattern could not be detected at the gene *RFP2* that was earlier shown to be monoallelically expressed. The quantification of H3K4me2 and macroH2A occurrence at *RFP2* points to allelic differences of enrichment at the gene locus, although this could not be further defined by the chosen methods. However, allelic enrichment of chromatin marks could be detected at two distinct loci in the 5'-regions of the large noncoding RNA genes, *BCMS* and *DLeu2/BCMSUN*, in samples derived from healthy probands. Thereby, in this study, two candidate locus control regions (LCRs) were identified that are located in the CpG-islands D and E in the promoter region of the two large ncRNAs and showed differential DNA-methylation and histone modifications. The differential epigenetic code and the localization of the candidate LCRs suggest their involvement in the regulation of expression of the two large ncRNA genes that span the critical region in 13q14.3. In summary these findings suggest a model of the epigenetically regulated tumor suppressor mechanism in 13q14.3 that leads to monoallelic expression of the candidate genes in the region. This model consists of the features that i) 13q14.3 is present in one active and one silent chromosome copy in non-malignant healthy cells, ii) the large ncRNAs *BCMS* and *BCMSUN* are likely involved in the regulation (in *cis*) of expression of the other genes in 13q14.3, and iii) differential epigenetic modifications of candidate LCRs control transcription of the large

ncRNA genes. Accordingly, the epigenetic status of the LCRs is critical for transcription of the ncRNAs that in turn regulate the expression of the candidate genes in the critical region. Surprisingly, the most evident epigenetic aberration at 13q14.3 was the generally higher enrichment of the active chromatin mark at the region in CLL cells. However, also the inactive chromatin mark was increased at 13q14.3 in CLL cells. Significant differences of enrichment of H3K4me2 and macroH2A between non-malignant and CLL cells, were detected at CpG-islands B, C and E and at the exonic region of *DLeu2/BCMSUN*, which clearly showed that the aberrant modifications in CLL are most evident at the promoter regions of the genes *C13ORF1*, *RFP2* and *DLeu2/BCMSUN*, respectively. The chromatin pattern that marks monoallelic expression at *C13ORF1* and *DLeu2/BCMSUN* remained unchanged in the analyzed CLL cells. However, in CLL cells derived from patients with different genomic aberrations, the deletion of genomic material from 13q14.3 and loss of active chromatin marks could be correlated. This suggests that in del(13q) CLL patients, indeed the active copy of 13q14.3 is lost, which would explain downregulation of expression of the genes in the critical region observed in these patients. In line with the proposed model of the tumor suppressor mechanism in 13q14.3, the most evident aberrant epigenetic modifications of 13q14.3 in CLL cells were detected at the two candidate LCRs, where distinct epimutations were detected. Thus, an epimutation in 13q14.3 was shown that could be responsible for the inactivation of the epigenetic regulatory mechanism of 13q14.3 in CLL cells. Further, these findings suggest that the regulatory impact of the ncRNAs, the LCR or another yet unidentified part of the critical region in 13q14.3 is affected by this epimutation, which leads to deregulation and inactivation of the tumor suppressive function of 13q14.3.

In conclusion, a specific epigenetic code at the genes in the critical region in 13q14.3 was determined that is correlated with their transcriptional state. Evidence was obtained for a model of epigenetic regulation that implies differential chromatin packaging of the two copies of 13q14.3 and results in monoallelic expression of the candidate genes in healthy B- and T-cells. Furthermore, two candidate LCRs were identified in the critical region in 13q14.3 in non-malignant cells that showed distinct epimutations in CLL cells. The localization and the detected epimutations of the LCRs in CLL point to their role in regulation of the ncRNA genes that most likely regulate expression of the other genes in the critical region. On the basis of these findings, further characterization of epigenetic features at 13q14.3 will help to fully elucidate the complex epigenetic regulatory network that controls the tumor suppressor mechanism in 13q14.3. The findings presented here will help to understand the pathogenicity of (epigenetic) inactivation of the tumor suppressor mechanism in CLL and thereby provide the basis for development of more efficient and specific therapies for CLL in the future.

Zusammenfassung

Die chronisch lymphatische Leukämie des B-Zell Typs (B-CLL) ist in den westlichen Ländern die am häufigsten auftretende Leukämie bei Erwachsenen. Durchschnittlich sind die Patienten zum Zeitpunkt der Diagnose 72 Jahre alt. Die häufigste genomische Veränderung bei Patienten ist die Deletion einer kritischen Region in der chromosomalen Bande 13q14.3, die bei über 50 % aller B-CLL Patienten auftritt. Die Häufigkeit des Verlusts der kritischen Region in 13q14.3 in B-CLL und anderen Tumoren deutet auf einen wichtigen Tumorsuppressormechanismus hin, der in der kritischen Region lokalisiert ist. Außerdem konnte bisher für fast alle Kandidatengene der Region 13q14.3 monoallelische Expression nachgewiesen werden.

In der vorliegenden Arbeit wurden epigenetische Modifikationen an Genen und CpG-Inseln der chromosomalen Region untersucht, um zur Aufklärung des Tumorsuppressormechanismus beizutragen. Die monoallelische Genexpression der kritischen Region in 13q14.3 ist durch epigenetische Modifikationen des Chromatins der Promotoren und Genregionen markiert. Ein Chromatinmodifikationsmuster, das spezifische monoallelisch exprimierte und auch geprägte Gene markiert, ist die Anreicherung von dimethyliertem Lysin 4 des Histons H3 (H3K4me₂), die eng auf den Promoterbereich begrenzt ist. Dieses spezifische Muster wurde an den Genen *C13ORF1* und *DLeu2/BCMSUN* entdeckt und führte zur Identifizierung eines weiteren monoallelisch exprimierten Gens der Region 13q14.3. Zusätzlich zu den bereits gezeigten Genen *RFP2* und *DLeu2/BCMSUN*, konnte auch für *C13ORF1* eine monoallelische Expression in gesunden B- und T-Zellen gezeigt werden.

Die Chromatin Immunopräzipitation (ChIP) Methode wurde etabliert, um zwei Chromatinmodifikationen an der chromosomalen Bande 13q14.3 zu quantifizieren. Hierfür musste die ChIP-Methode zunächst technisch etabliert und optimiert werden, um die Anreicherung von H3K4me₂, was mit einem aktiven Transkriptionsstatus korreliert, und von der Histonvariante macroH2A, welche mit einem reprimiertem Transkriptionsstatus und allelisch unterschiedlicher Anreicherung korreliert, analysieren zu können. Zusätzlich wurde auch die DNA-Methylierung der Region 13q14.3 untersucht. Auf diese Weise wurde der epigenetische Code der Kandidatengene, der mit deren Transkriptionsstatus korreliert ist, in gesunden hämatopoetischen Zellen und in B-CLL Zellen identifiziert. In gesunden hämatopoetischen Zellen markieren die epigenetischen Modifikationen eine aktive und eine inaktive Kopie der kritischen Region in der chromosomalen Bande 13q14.3. Die Gene *C13ORF1* und *DLeu2/BCMSUN* werden allein von der aktiven Kopie exprimiert und ihr monoallelicher Transkriptionsstatus ist im Chromatin markiert. Für das bereits als monoallelisch exprimiert beschriebene Gen *RFP2* konnte dieses spezifische Chromatinmuster in der vorliegenden Arbeit nicht gefunden werden. Allerdings deutet die Quantifizierung von H3K4me₂ und macroH2A darauf hin, dass es auch an diesem Genlocus allelische Unterschiede gibt, welche mit der angewendeten Methode jedoch nicht näher definiert werden konnten. Im Gegensatz zum *RFP2* Locus, konnte eine allelische Anreicherung der untersuchten Chromatinmodifikationen an zwei definierten Loci in der 5'-Region der nichtkodierenden RNA (ncRNA) Gene, *BCMS* und *DLeu2/BCMSUN*, in gesunden Zellen gezeigt werden. Mittels der in der vorliegenden Studie definierten, allelisch unterschiedlichen Anreicherung der Histonmodifikationen und auch der DNA-Methylierung, wurden zwei Kandidaten-Locus-Kontroll-Regionen (LCR) identifiziert, die in den CpG-Inseln D und E in den Promoterregionen der beiden großen ncRNA Gene lokalisiert sind. Ihr differentieller epigenetischer Code sowie die Lokalisierung der Kandidaten-LCRs spricht für ihre Beteiligung an der Expressionsregulation der beiden ncRNA Gene, die die gesamte kritische Region überspannen.

Zusammenfassend führen diese Ergebnisse zu einem Modell des epigenetisch regulierten Tumorsuppressormechanismus in 13q14.3, der zur monoallelischen Expression der Kandidatengene führt. Das Modell besteht bisher aus drei ermittelten Merkmalen: i) die chromosomale Region 13q14.3 liegt in einer aktiven und einer inaktiven Kopie in gesunden hämatopoietischen Zellen vor, ii) die großen ncRNA Gene, *BCMS* und *DLeu2/BCMSUN*, sind wahrscheinlich an der Expressionsregulation der anderen Gene der Region (in *cis*) beteiligt und iii) die differentiellen epigenetischen Modifikationen der Kandidaten-LCRs kontrollieren die Transkription der großen ncRNA Gene. Somit wäre der epigenetische Status der LCRs entscheidend für die Transkription der ncRNA Gene, welche wiederum die Expression der Kandidatengene regulieren.

In B-CLL Zellen war die auffälligste epigenetische Aberration in 13q14.3 ein allgemein starker Anstieg der Anreicherung der aktivierenden aber auch der repressiven Chromatinmarkierung. Signifikante Unterschiede in der Anreicherung von H3K4me2 und macroH2A zwischen gesunden und B-CLL Zellen wurden an den CpG-Inseln B, C und D, sowie an der exonischen Region des Gens *DLeu2/BCMSUN* gefunden. Dies zeigte klar, dass aberrante Modifikationen in B-CLL am deutlichsten an den Promoterregionen der Gene *C13ORF1*, *RFP2* und *DLeu2/BCMSUN* auftreten. Das Chromatinmodifikationsmuster, das die monoallelisch exprimierten Gene *C13ORF1* und *DLeu2/BCMSUN* markiert, blieb in den analysierten B-CLL Zellen unverändert bestehen. Darüber hinaus konnte in B-CLL Zellen von Patienten mit unterschiedlichen genomischen Aberrationen die Deletion von genomischem Material der chromosomalen Region 13q14.3 mit dem Verlust der aktivierenden Chromatinmarkierungen korreliert werden. Dies deutet darauf hin, dass tatsächlich die aktive Kopie von 13q14.3 in monoallelisch deletierten Patienten mit B-CLL verloren gegangen ist, was auch die starke Herunterregulierung der Genexpression in der kritischen Region erklärt, die in diesen Patienten beobachtet wurde. In Übereinstimmung mit dem Modell des Tumorsuppressormechanismus in 13q14.3 wurde die auffälligste epigenetische Aberration an 13q14.3 in B-CLL Zellen an den Kandidaten-LCRs gefunden. Somit wurde eine Epimutation in 13q14.3 entdeckt, die für die Inaktivierung des Tumorsuppressormechanismus in 13q14.3 in B-CLL Zellen verantwortlich sein könnte. Die vorliegenden Ergebnisse zeigen weiterhin, dass der regulatorische Einfluss der ncRNAs, der LCRs oder einem anderen noch nicht näher identifizierten Bereichs der kritischen Region in 13q14.3 von dieser Epimutation betroffen sein könnte, und diese so zur Deregulierung und Inaktivierung der Tumorsuppressorfunktion in 13q14.3 führt.

Zusammenfassend konnte ein spezifischer epigenetischer Code der Gene in der kritischen Region in 13q14.3 identifiziert werden, der mit deren Transkriptionsstatus korreliert. Die hier gefundenen Ergebnisse unterstützen das Modell der epigenetischen Regulation, das differentielle Chromatinverpackung der beiden Kopien von 13q14.3 einschließt und zur monoallelischen Expression der Kandidatengene in gesunden B- und T-Zellen führt. Zusätzlich konnten zwei Kandidaten-LCRs in den Promoterregionen der beiden großen ncRNA Gene in der kritischen Region identifiziert werden, die distinkte Epimutationen in B-CLL Zellen aufweisen. Die Lokalisierung sowie die gefundenen Epimutationen der LCRs in B-CLL deuten auf eine wichtige Rolle der LCRs bei der Transkriptionsregulation der großen ncRNA Gene hin, die wahrscheinlich die Genexpression in der kritischen Region regulieren. Auf Basis der vorliegenden Ergebnisse ist nun eine weitere Charakterisierung des komplexen epigenetischen Netzwerks möglich, das den Tumorsuppressormechanismus in 13q14.3 kontrolliert. Die Ergebnisse der vorliegenden Studie tragen dazu bei, die Pathogenität der epigenetischen Inaktivierung des Tumorsuppressormechanismus in B-CLL zu verstehen und können somit eine Basis für die zukünftige Entwicklung von effizienteren und spezifischeren Therapien der B-CLL bilden.

Bibliography

- Ahmad, K., and Henikoff, S. (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* 9:1191-1200.
- Ahmed, F.E. (1980) The combined action of chemical carcinogens on DNA repair in human cells. *Radiation and environmental biophysics* 18:57-64.
- Andersen, A.A., and Panning, B. (2003) Epigenetic gene regulation by noncoding RNAs. *Curr Opin Cell Biol* 15:281-289.
- Ando, S., Yang, H., Nozaki, N., Okazaki, T., and Yoda, K. (2002) CENP-A, -B, and -C chromatin complex that contains the I-type alpha-satellite array constitutes the prekinetochore in HeLa cells. *Mol Cell Biol* 22:2229-2241.
- Avery, S., Nadal, E., Marin, D., Olavarria, E., Kaeda, J., Vulliamy, T., Brito Babapulle, F., Goldman, J.M., and Apperley, J.F. (2004) Lymphoid transformation in a CML patient in complete cytogenetic remission following treatment with imatinib. *Leukemia research* 28 Suppl 1:S75-77.
- Ballestar, E., Paz, M.F., Valle, L., Wei, S., Fraga, M.F., Espada, J., Cigudosa, J.C., Huang, T.H., and Esteller, M. (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. *Embo J* 22:6335-6345.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410:120-124.
- Baranova, A., Hammarsund, M., Ivanov, D., Skoblov, M., Sangfelt, O., Corcoran, M., Borodina, T., Makeeva, N., Pestova, A., Tyazhelova, T., Nazarenko, S., Gorreta, F., Alsheddi, T., Schlauch, K., Nikitin, E., Kapanadze, B., Shagin, D., Poltarau, A., Ivanovich Vorobiev, A., Zabarovsky, E., Lukianov, S., Chandhoke, V., Ibbotson, R., Oscier, D., Einhorn, S., Grander, D., and Yankovsky, N. (2003) Distinct organization of the candidate tumor suppressor gene RFP2 in human and mouse: multiple mRNA isoforms in both species- and human-specific antisense transcript RFP2OS. *Gene* 321:103-112.
- Baranova, A.V., Ivanov, D.V., Tiazhelova, T.V., and Iankovskii, N.K. (2004) [Structural-functional characteristics of the 13q14 region of the human genome in the search for potential tumor suppressor genes]. *Molekuliarnaia biologii* 38:203-212.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129:823-837.
- Bassing, C.H., Swat, W., and Alt, F.W. (2002) The mechanism and regulation of chromosomal V(D)J recombination. *Cell* 109 Suppl:S45-55.
- Begleiter, A., Mowat, M., Israels, L.G., and Johnston, J.B. (1996) Chlorambucil in chronic lymphocytic leukemia: mechanism of action. *Leukemia & lymphoma* 23:187-201.
- Bentz, M., Plesch, A., Bullinger, L., Stilgenbauer, S., Ott, G., Müller-Hermelink, H.K., Baudis, M., Barth, T.F., Möller, P., Lichter, P., and Döhner, H. (2000) t(11;14)-positive mantle cell lymphomas exhibit complex karyotypes and share similarities with B-cell chronic lymphocytic leukemia. *Genes, chromosomes & cancer* 27:285-294.
- Bergman, Y., Fisher, A., and Cedar, H. (2003) Epigenetic mechanisms that regulate antigen receptor gene expression. *Curr Opin Immunol* 15:176-181.
- Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., Schreiber, S.L., and Lander, E.S. (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120:169-181.

- Bernstein, B.E., Meissner, A., and Lander, E.S. (2007) The mammalian epigenome. *Cell* 128:669-681.
- Bhattacharya, S.K., Ramchandani, S., Cervoni, N., and Szyf, M. (1999) A mammalian protein with specific demethylase activity for mCpG DNA. *Nature* 397:579-583.
- Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* 16:6-21.
- Blanco-Aparicio, C., Renner, O., Leal, J.F., and Carnero, A. (2007) PTEN, more than the AKT pathway. *Carcinogenesis* 28:1379-1386.
- Blayney, L.M., Zissimopoulos, S., Ralph, E., Abbot, E., Matthews, L., and Lai, F.A. (2004) Ryanodine receptor oligomeric interaction: identification of a putative binding region. *J Biol Chem* 279:14639-14648.
- Blelloch, R., Wang, Z., Meissner, A., Pollard, S., Smith, A., and Jaenisch, R. (2006) Reprogramming efficiency following somatic cell nuclear transfer is influenced by the differentiation and methylation state of the donor nucleus. *Stem Cells* 24:2007-2013.
- Boyiadzis, M., Foon, K.A., and Pavletic, S. (2007) Hematopoietic stem cell transplantation for chronic lymphocytic leukemia: potential cure for an incurable disease. *Expert opinion on biological therapy* 7:1789-1797.
- Brena, R.M., Auer, H., Kornacker, K., and Plass, C. (2006) Quantification of DNA methylation in electrofluidics chips (Bio-COBRA). *Nat Protoc* 1:52-58.
- Burnett, A.K. (2005) The treatment of AML: current status and novel approaches. *Hematology (Amsterdam, Netherlands)* 10 Suppl 1:50-53.
- Burnette, W.N. (1981) "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analytical biochemistry* 112:195-203.
- Byrd, J.C., Lin, T.S., and Grever, M.R. (2006) Treatment of relapsed chronic lymphocytic leukemia: old and new therapies. *Seminars in oncology* 33:210-219.
- Cabelof, D.C., Ikeno, Y., Nyska, A., Busuttill, R.A., Anyangwe, N., Vijg, J., Matherly, L.H., Tucker, J.D., Wilson, S.H., Richardson, A., and Heydari, A.R. (2006) Haploinsufficiency in DNA polymerase beta increases cancer risk with age and alters mortality rate. *Cancer research* 66:7460-7465.
- Calin, G.A., Cimmino, A., Fabbri, M., Ferracin, M., Wojcik, S.E., Shimizu, M., Taccioli, C., Zanesi, N., Garzon, R., Aqeilan, R.I., Alder, H., Volinia, S., Rassenti, L., Liu, X., Liu, C.G., Kipps, T.J., Negrini, M., and Croce, C.M. (2008) MiR-15a and miR-16-1 cluster functions in human leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 105:5166-5171.
- Calin, G.A., and Croce, C.M. (2006) MicroRNA signatures in human cancers. *Nat Rev Cancer* 6:857-866.
- Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Alder, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C.M. (2002) Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 99:15524-15529.
- Calin, G.A., Liu, C.G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C.D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., Dell'Aquila, M.L., Alder, H., Rassenti, L., Kipps, T.J., Bullrich, F., Negrini, M., and Croce, C.M. (2004) MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proceedings of the National Academy of Sciences of the United States of America* 101:11755-11760.
- Calin, G.A., Pekarsky, Y., and Croce, C.M. (2007) The role of microRNA and other non-coding RNA in the pathogenesis of chronic lymphocytic leukemia. *Best Pract Res Clin Haematol* 20:425-437.
- Calin, G.A., Trapasso, F., Shimizu, M., Dumitru, C.D., Yendamuri, S., Godwin, A.K., Ferracin, M., Bernardi, G., Chatterjee, D., Baldassarre, G., Rattan, S., Alder, H., Mabuchi, H., Shiraishi, T., Hansen, L.L., Overgaard, J., Herlea, V., Mauro, F.R., Dighiero, G.,

- Movsas, B., Rassenti, L., Kipps, T., Baffa, R., Fusco, A., Mori, M., Russo, G., Liu, C.G., Neubergh, D., Bullrich, F., Negrini, M., and Croce, C.M. (2005) Familial cancer associated with a polymorphism in ARLTS1. *The New England journal of medicine* 352:1667-1676.
- Casillas, M.A., Jr., Lopatina, N., Andrews, L.G., and Tollefsbol, T.O. (2003) Transcriptional control of the DNA methyltransferases is altered in aging and neoplastically-transformed human fibroblasts. *Molecular and cellular biochemistry* 252:33-43.
- Catovsky, D. (1982) Symposium: classification of leukemia. 1. The classification of acute leukemia. *Pathology* 14:277-281.
- Cedar, H., and Bergman, Y. (2008) Choreography of Ig allelic exclusion. *Curr Opin Immunol* 20:308-317.
- Cervoni, N., Bhattacharya, S., and Szyf, M. (1999) DNA demethylase is a processive enzyme. *J Biol Chem* 274:8363-8366.
- Changolkar, L.N., and Pehrson, J.R. (2006) macroH2A1 histone variants are depleted on active genes but concentrated on the inactive X chromosome. *Mol Cell Biol* 26:4410-4420.
- Chen, C., Frierson, H.F., Jr., Haggerty, P.F., Theodorescu, D., Gregory, C.W., and Dong, J.T. (2001) An 800-kb region of deletion at 13q14 in human prostate and other carcinomas. *Genomics* 77:135-144.
- Chen, C.Z. (2005) MicroRNAs as oncogenes and tumor suppressors. *The New England journal of medicine* 353:1768-1771.
- Chen, R.W., Bemis, L.T., Amato, C.M., Myint, H., Tran, H., Birks, D.K., Eckhardt, S.G., and Robinson, W.A. (2008) Truncation in CCND1 mRNA alters miR-16-1 regulation in mantle cell lymphoma. *Blood* 112:822-829.
- Chena, C., Avalos, J.S., Bezares, R.F., Arrossagaray, G., Turdo, K., Bistmans, A., and Slavutsky, I. (2008) Biallelic deletion 13q14.3 in patients with chronic lymphocytic leukemia. Cytogenetic, FISH and clinical studies. *European journal of haematology*.
- Cheson, B.D. (2006) Monoclonal antibody therapy of chronic lymphocytic leukemia. *Cancer Immunol Immunother* 55:188-196.
- Choo, J.H., Kim, J.D., Chung, J.H., Stubbs, L., and Kim, J. (2006) Allele-specific deposition of macroH2A1 in imprinting control regions. *Hum Mol Genet* 15:717-724.
- Cimmino, A., Calin, G.A., Fabbri, M., Iorio, M.V., Ferracin, M., Shimizu, M., Wojcik, S.E., Aqeilan, R.I., Zupo, S., Dono, M., Rassenti, L., Alder, H., Volinia, S., Liu, C.G., Kipps, T.J., Negrini, M., and Croce, C.M. (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* 102:13944-13949.
- Clayton, A.L., Hazzalin, C.A., and Mahadevan, L.C. (2006) Enhanced histone acetylation and transcription: a dynamic perspective. *Mol Cell* 23:289-296.
- Collas, P., and Dahl, J.A. (2008) Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. *Front Biosci* 13:929-943.
- Collins, S.J., Gallo, R.C., and Gallagher, R.E. (1977) Continuous growth and differentiation of human myeloid leukaemic cells in suspension culture. *Nature* 270:347-349.
- Corcoran, M.M., Hammarsund, M., Zhu, C., Lerner, M., Kapanadze, B., Wilson, B., Larsson, C., Forsberg, L., Ibbotson, R.E., Einhorn, S., Oscier, D.G., Grandt, D., and Sangfelt, O. (2004) DLEU2 encodes an antisense RNA for the putative bicistronic RFP2/LEU5 gene in humans and mouse. *Genes, chromosomes & cancer* 40:285-297.
- Costanzi, C., and Pehrson, J.R. (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* 393:599-601.
- Cotter, F.E., and Auer, R.L. (2007) Genetic alteration associated with chronic lymphocytic leukemia. *Cytogenetic and genome research* 118:310-319.
- Cross, S.H., Charlton, J.A., Nan, X., and Bird, A.P. (1994) Purification of CpG islands using a methylated DNA binding column. *Nat Genet* 6:236-244.

- Culhane, J.C., and Cole, P.A. (2007) LSD1 and the chemistry of histone demethylation. *Curr Opin Chem Biol* 11:561-568.
- D'Arena, G., Di Renzo, N., Brugiattelli, M., Vigliotti, M.L., and Keating, M.J. (2003) Biological and clinical heterogeneity of B-cell chronic lymphocytic leukemia. *Leukemia & lymphoma* 44:223-228.
- Damle, R.N., Wasil, T., Fais, F., Ghiotto, F., Valetto, A., Allen, S.L., Buchbinder, A., Budman, D., Dittmar, K., Kowitz, J., Lichtman, S.M., Schulman, P., Vinciguerra, V.P., Rai, K.R., Ferrarini, M., and Chiorazzi, N. (1999) Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* 94:1840-1847.
- Del Principe, M.I., Del Poeta, G., Buccisano, F., Maurillo, L., Venditti, A., Zucchetto, A., Marini, R., Niscola, P., Consalvo, M.A., Mazzone, C., Ottaviani, L., Panetta, P., Bruno, A., Bomben, R., Suppo, G., Degan, M., Gattei, V., de Fabritiis, P., Cantonetti, M., Lo Coco, F., Del Principe, D., and Amadori, S. (2006) Clinical significance of ZAP-70 protein expression in B-cell chronic lymphocytic leukemia. *Blood* 108:853-861.
- Dickinson, J.D., Joshi, A., Iqbal, J., Sanger, W., Bierman, P.J., and Joshi, S.S. (2006) Genomic abnormalities in chronic lymphocytic leukemia influence gene expression by a gene dosage effect. *International journal of molecular medicine* 17:769-778.
- Dighiero, G., and Hamblin, T.J. (2008) Chronic lymphocytic leukaemia. *Lancet* 371:1017-1029.
- Döhner, H., Stilgenbauer, S., Benner, A., Leupolt, E., Kröber, A., Bullinger, L., Döhner, K., Bentz, M., and Lichter, P. (2000) Genomic aberrations and survival in chronic lymphocytic leukemia. *The New England journal of medicine* 343:1910-1916.
- Döhner, H., Stilgenbauer, S., Döhner, K., Bentz, M., and Lichter, P. (1999) Chromosome aberrations in B-cell chronic lymphocytic leukemia: reassessment based on molecular cytogenetic analysis. *J Mol Med* 77:266-281.
- Döhner, H., Stilgenbauer, S., Fischer, K., Bentz, M., and Lichter, P. (1997) Cytogenetic and molecular cytogenetic analysis of B cell chronic lymphocytic leukemia: specific chromosome aberrations identify prognostic subgroups of patients and point to loci of candidate genes. *Leukemia* 11 Suppl 2:S19-24.
- Edry, E., and Melamed, D. (2004) Receptor editing in positive and negative selection of B lymphopoiesis. *J Immunol* 173:4265-4271.
- Esteller, M. (2003) Relevance of DNA methylation in the management of cancer. *The lancet oncology* 4:351-358.
- Esteller, M. (2005) Aberrant DNA methylation as a cancer-inducing mechanism. *Annu Rev Pharmacol Toxicol* 45:629-656.
- Esteller, M. (2006) Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. *Br J Cancer* 94:179-183.
- Esteller, M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature reviews* 8:286-298.
- Esteller, M. (2008) Epigenetics in cancer. *The New England journal of medicine* 358:1148-1159.
- Fagerlund, R., Kinnunen, L., Kohler, M., Julkunen, I., and Melen, K. (2005) NF- κ B is transported into the nucleus by importin α 3 and importin α 4. *J Biol Chem* 280:15942-15951.
- Fagerlund, R., Melen, K., Cao, X., and Julkunen, I. (2008) NF- κ B p52, RelB and c-Rel are transported into the nucleus via a subset of importin α molecules. *Cellular signalling* 20:1442-1451.
- Fahrner, J.A., Eguchi, S., Herman, J.G., and Baylin, S.B. (2002) Dependence of histone modifications and gene expression on DNA hypermethylation in cancer. *Cancer research* 62:7213-7218.
- Feinberg, A.P., and Tycko, B. (2004) The history of cancer epigenetics. *Nat Rev Cancer* 4:143-153.

- Fraga, M.F., Agrelo, R., and Esteller, M. (2007) Cross-talk between aging and cancer: the epigenetic language. *Ann N Y Acad Sci* 1100:60-74.
- Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T.D., Wu, Y.Z., Plass, C., and Esteller, M. (2005a) Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America* 102:10604-10609.
- Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., Iyer, N.G., Perez-Rosado, A., Calvo, E., Lopez, J.A., Cano, A., Calasanz, M.J., Colomer, D., Piris, M.A., Ahn, N., Imhof, A., Caldas, C., Jenuwein, T., and Esteller, M. (2005b) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 37:391-400.
- Fraga, M.F., and Esteller, M. (2005) Towards the human cancer epigenome: a first draft of histone modifications. *Cell Cycle* 4:1377-1381.
- Fraga, M.F., and Esteller, M. (2007) Epigenetics and aging: the targets and the marks. *Trends Genet* 23:413-418.
- Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986) A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323:643-646.
- Fulci, V., Chiaretti, S., Goldoni, M., Azzalin, G., Carucci, N., Tavoraro, S., Castellano, L., Magrelli, A., Citarella, F., Messina, M., Maggio, R., Peragine, N., Santangelo, S., Mauro, F.R., Landgraf, P., Tuschl, T., Weir, D.B., Chien, M., Russo, J.J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., Guarini, A., Foa, R., and Macino, G. (2007) Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood* 109:4944-4951.
- Gao, C., Furge, K., Koeman, J., Dykema, K., Su, Y., Cutler, M.L., Werts, A., Haak, P., and Vande Woude, G.F. (2007) Chromosome instability, chromosome transcriptome, and clonal evolution of tumor cell populations. *Proceedings of the National Academy of Sciences of the United States of America* 104:8995-9000.
- Gargiulo, G., and Minucci, S. (2008) Epigenomic profiling of cancer cells. *Int J Biochem Cell Biol*.
- Gauld, S.B., Dal Porto, J.M., and Cambier, J.C. (2002) B cell antigen receptor signaling: roles in cell development and disease. *Science* 296:1641-1642.
- Ghia, P., and Caligaris-Cappio, F. (2006) The origin of B-cell chronic lymphocytic leukemia. *Seminars in oncology* 33:150-156.
- Gimelbrant, A., Hutchinson, J.N., Thompson, B.R., and Chess, A. (2007) Widespread monoallelic expression on human autosomes. *Science* 318:1136-1140.
- Gine, E., Bosch, F., Villamor, N., Rozman, M., Colomer, D., Lopez-Guillermo, A., Campo, E., and Montserrat, E. (2002) Simultaneous diagnosis of hairy cell leukemia and chronic lymphocytic leukemia/small lymphocytic lymphoma: a frequent association? *Leukemia* 16:1454-1459.
- Gine, E., Moreno, C., Esteve, J., and Montserrat, E. (2007) The role of stem-cell transplantation in chronic lymphocytic leukemia risk-adapted therapy. *Best Pract Res Clin Haematol* 20:529-543.
- Goi, K., Takagi, M., Iwata, S., Delia, D., Asada, M., Donghi, R., Tsunematsu, Y., Nakazawa, S., Yamamoto, H., Yokota, J., Tamura, K., Saeki, Y., Utsunomiya, J., Takahashi, T., Ueda, R., Ishioka, C., Eguchi, M., Kamata, N., and Mizutani, S. (1997) DNA damage-associated dysregulation of the cell cycle and apoptosis control in cells with germ-line p53 mutation. *Cancer research* 57:1895-1902.
- Goldmit, M., and Bergman, Y. (2004) Monoallelic gene expression: a repertoire of recurrent themes. *Immunol Rev* 200:197-214.

- Goll, M.G., and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. *Annual review of biochemistry* 74:481-514.
- Gonzalez-Zulueta, M., Bender, C.M., Yang, A.S., Nguyen, T., Beart, R.W., Van Tornout, J.M., and Jones, P.A. (1995) Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing. *Cancer research* 55:4531-4535.
- Greger, V., Passarge, E., Hopping, W., Messmer, E., and Horsthemke, B. (1989) Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Human genetics* 83:155-158.
- Grundy, R.G., Aledo, R., and Cowell, J.K. (1998) Characterization of the breakpoints in unbalanced t(5;11)(p15;p15) constitutional chromosome translocations in two patients with beckwith-wiedemann syndrome using fluorescence in situ hybridisation. *International journal of molecular medicine* 1:801-808.
- Guillouf, C., Rosselli, F., Krishnaraju, K., Moustacchi, E., Hoffman, B., and Liebermann, D.A. (1995) p53 involvement in control of G2 exit of the cell cycle: role in DNA damage-induced apoptosis. *Oncogene* 10:2263-2270.
- Hamblin, T.J., Davis, Z., Gardiner, A., Oscier, D.G., and Stevenson, F.K. (1999) Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* 94:1848-1854.
- Hamblin, T.J., Orchard, J.A., Ibbotson, R.E., Davis, Z., Thomas, P.W., Stevenson, F.K., and Oscier, D.G. (2002) CD38 expression and immunoglobulin variable region mutations are independent prognostic variables in chronic lymphocytic leukemia, but CD38 expression may vary during the course of the disease. *Blood* 99:1023-1029.
- Hammarsund, M., Corcoran, M.M., Wilson, W., Zhu, C., Einhorn, S., Sangfelt, O., and Grander, D. (2004) Characterization of a novel B-CLL candidate gene--DLEU7--located in the 13q14 tumor suppressor locus. *FEBS letters* 556:75-80.
- Hanahan, D., and Weinberg, R.A. (2000) The hallmarks of cancer. *Cell* 100:57-70.
- Hanlon, S.E., and Lieb, J.D. (2004) Progress and challenges in profiling the dynamics of chromatin and transcription factor binding with DNA microarrays. *Curr Opin Genet Dev* 14:697-705.
- Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C., and Stam, M. (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods* 3:11.
- Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405:486-489.
- Haslinger, C., Schweifer, N., Stilgenbauer, S., Dohner, H., Lichter, P., Kraut, N., Stratowa, C., and Abseher, R. (2004) Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. *J Clin Oncol* 22:3937-3949.
- Hattori, M., Sakamoto, H., and Yamamoto, T. (2001) DNA demethylase expression correlates with lung resistance protein expression in common epithelial ovarian cancers. *The Journal of international medical research* 29:204-213.
- Hattori, N., Nishino, K., Ko, Y.G., Hattori, N., Ohgane, J., Tanaka, S., and Shiota, K. (2004) Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. *J Biol Chem* 279:17063-17069.
- Henikoff, S., Furuyama, T., and Ahmad, K. (2004) Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet* 20:320-326.
- Herman, J.G., and Baylin, S.B. (2003) Gene silencing in cancer in association with promoter hypermethylation. *The New England journal of medicine* 349:2042-2054.
- Herman, J.G., Latif, F., Weng, Y., Lerman, M.I., Zbar, B., Liu, S., Samid, D., Duan, D.S., Gnarr, J.R., Linehan, W.M., and et al. (1994) Silencing of the VHL tumor-suppressor

- gene by DNA methylation in renal carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 91:9700-9704.
- Herman, J.G., Merlo, A., Mao, L., Lapidus, R.G., Issa, J.P., Davidson, N.E., Sidransky, D., and Baylin, S.B. (1995) Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer research* 55:4525-4530.
- Hirose, Y., Kiyoi, H., Iwai, M., Yokozawa, T., Ito, M., and Naoe, T. (2002) Successful treatment with imatinib mesylate of a CML patient in megakaryoblastic crisis with severe fibrosis. *International journal of hematology* 76:349-353.
- Holm, T.M., Jackson-Grusby, L., Brambrink, T., Yamada, Y., Rideout, W.M., 3rd, and Jaenisch, R. (2005) Global loss of imprinting leads to widespread tumorigenesis in adult mice. *Cancer cell* 8:275-285.
- Illingworth, R., Kerr, A., Desousa, D., Jorgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., Humphray, S., Cox, T., Langford, C., and Bird, A. (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS biology* 6:e22.
- Isakov, N., Wange, R.L., Burgess, W.H., Watts, J.D., Aebersold, R., and Samelson, L.E. (1995) ZAP-70 binding specificity to T cell receptor tyrosine-based activation motifs: the tandem SH2 domains of ZAP-70 bind distinct tyrosine-based activation motifs with varying affinity. *The Journal of experimental medicine* 181:375-380.
- Issa, J.P. (2003) Age-related epigenetic changes and the immune system. *Clinical immunology (Orlando, Fla)* 109:103-108.
- Ivanov, D.V., Tyazhelova, T.V., Lemonnier, L., Kononenko, N., Pestova, A.A., Nikitin, E.A., Prevarskaya, N., Skryma, R., Panchin, Y.V., Yankovsky, N.K., and Baranova, A.V. (2003) A new human gene KCNKG encoding potassium channel regulating protein is a cancer suppressor gene candidate located in 13q14.3. *FEBS letters* 539:156-160.
- Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., and Jaenisch, R. (2001) Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat Genet* 27:31-39.
- Jadayel, D.M., Lukas, J., Nacheva, E., Bartkova, J., Stranks, G., De Schouwer, P.J., Lens, D., Bartek, J., Dyer, M.J., Kruger, A.R., and Catovsky, D. (1997) Potential role for concurrent abnormalities of the cyclin D1, p16CDKN2 and p15CDKN2B genes in certain B cell non-Hodgkin's lymphomas. *Functional studies in a cell line (Granta 519)*. *Leukemia* 11:64-72.
- Janeway, C.A. (2002) *Immunologie*. Spektrum Akademischer Verlag GmbH, Heidelberg.
- Jansen, J., den Ottolander, G.J., Holdrinet, R.S., Tricot, G.J., and Hermans, J. (1984) Prognosis and therapy in hairy cell leukemia. *Seminars in oncology* 11:472-478.
- Jemal, A., Murray, T., Ward, E., Samuels, A., Tiwari, R.C., Ghafoor, A., Feuer, E.J., and Thun, M.J. (2005) Cancer statistics, 2005. *CA Cancer J Clin* 55:10-30.
- Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., and Thun, M.J. (2008) Cancer statistics, 2008. *CA Cancer J Clin* 58:71-96.
- Jenuwein, T., and Allis, C.D. (2001) Translating the histone code. *Science* 293:1074-1080.
- Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998) SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci* 54:80-93.
- Jimenez, G., Gale, K.B., and Enver, T. (1992) The mouse beta-globin locus control region: hypersensitive sites 3 and 4. *Nucleic acids research* 20:5797-5803.
- Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005) RAS is regulated by the let-7 microRNA family. *Cell* 120:635-647.
- Jones, P.A., and Baylin, S.B. (2002) The fundamental role of epigenetic events in cancer. *Nature reviews* 3:415-428.

- Jost, E., and Galm, O. (2007) EHA scientific workshop report: the role of epigenetics in hematological malignancies. *Epigenetics* 2:71-79.
- Kadota, M., Yang, H.H., Hu, N., Wang, C., Hu, Y., Taylor, P.R., Buetow, K.H., and Lee, M.P. (2007) Allele-specific chromatin immunoprecipitation studies show genetic influence on chromatin state in human genome. *PLoS Genet* 3:e81.
- Kapanadze, B., Kashuba, V., Baranova, A., Rasool, O., van Everdink, W., Liu, Y., Syomov, A., Corcoran, M., Poltarau, A., Brodyansky, V., Syomova, N., Kazakov, A., Ibbotson, R., van den Berg, A., Gizatullin, R., Fedorova, L., Sulimova, G., Zelenin, A., Deaven, L., Lehrach, H., Grander, D., Buys, C., Oscier, D., Zabarovsky, E.R., Einhorn, S., and Yankovsky, N. (1998) A cosmid and cDNA fine physical map of a human chromosome 13q14 region frequently lost in B-cell chronic lymphocytic leukemia and identification of a new putative tumor suppressor gene, *Leu5*. *FEBS letters* 426:266-270.
- Kapanadze, B., Makeeva, N., Corcoran, M., Jareborg, N., Hammarsund, M., Baranova, A., Zabarovsky, E., Vorontsova, O., Merup, M., Gahrton, G., Jansson, M., Yankovsky, N., Einhorn, S., Oscier, D., Grander, D., and Sangfelt, O. (2000) Comparative sequence analysis of a region on human chromosome 13q14, frequently deleted in B-cell chronic lymphocytic leukemia, and its homologous region on mouse chromosome 14. *Genomics* 70:327-334.
- Kay, N.E., Rai, K.R., and O'Brien, S. (2006) Chronic lymphocytic leukemia: current and emerging treatment approaches. *Clin Adv Hematol Oncol* 4:1-10; quiz 11-12.
- Kienle, D.L., Korz, C., Hosch, B., Benner, A., Mertens, D., Habermann, A., Kröber, A., Jäger, U., Lichter, P., Döhner, H., and Stilgenbauer, S. (2005) Evidence for distinct pathomechanisms in genetic subgroups of chronic lymphocytic leukemia revealed by quantitative expression analysis of cell cycle, activation, and apoptosis-associated genes. *J Clin Oncol* 23:3780-3792.
- Kitada, S., Andersen, J., Akar, S., Zapata, J.M., Takayama, S., Krajewski, S., Wang, H.G., Zhang, X., Bullrich, F., Croce, C.M., Rai, K., Hines, J., and Reed, J.C. (1998) Expression of apoptosis-regulating proteins in chronic lymphocytic leukemia: correlations with *In vitro* and *In vivo* chemoresponses. *Blood* 91:3379-3389.
- Klein, C.A., Schmidt-Kittler, O., Schardt, J.A., Pantel, K., Speicher, M.R., and Riethmuller, G. (1999) Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proceedings of the National Academy of Sciences of the United States of America* 96:4494-4499.
- Klein, G., Imreh, S., and Zabarovsky, E.R. (2007) Why do we not all die of cancer at an early age? *Adv Cancer Res* 98:1-16.
- Knudson, A.G., Jr. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* 68:820-823.
- Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* 128:693-705.
- Kröber, A., Bloehdorn, J., Hafner, S., Buhler, A., Seiler, T., Kienle, D., Winkler, D., Bangerter, M., Schlenk, R.F., Benner, A., Lichter, P., Döhner, H., and Stilgenbauer, S. (2006) Additional genetic high-risk features such as 11q deletion, 17p deletion, and V3-21 usage characterize discordance of ZAP-70 and VH mutation status in chronic lymphocytic leukemia. *J Clin Oncol* 24:969-975.
- Kröber, A., Seiler, T., Benner, A., Bullinger, L., Brückle, E., Lichter, P., Döhner, H., and Stilgenbauer, S. (2002) V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood* 100:1410-1416.
- Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., Adli, M., Kasif, S., Ptaszek, L.M., Cowan, C.A., Lander, E.S., Koseki, H., and Bernstein, B.E. (2008) Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* 4:e1000242.

- Küppers, R. (2005) Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* 5:251-262.
- Küppers, R., Klein, U., Hansmann, M.L., and Rajewsky, K. (1999) Cellular origin of human B-cell lymphomas. *The New England journal of medicine* 341:1520-1529.
- Küppers, R., Sousa, A.B., Baur, A.S., Strickler, J.G., Rajewsky, K., and Hansmann, M.L. (2001) Common germinal-center B-cell origin of the malignant cells in two composite lymphomas, involving classical Hodgkin's disease and either follicular lymphoma or B-CLL. *Mol Med* 7:285-292.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410:116-120.
- Lafon-Hughes, L., Di Tomaso, M.V., Mendez-Acuna, L., and Martinez-Lopez, W. (2008) Chromatin-remodelling mechanisms in cancer. *Mutation research* 658:191-214.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science* 294:853-858.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., de Jong, P., Catanese, J.J.,

- Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y.J. (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.
- Le Toriellec, E., Despouy, G., Pierron, G., Gaye, N., Joiner, M., Bellanger, D., Vincent-Salomon, A., and Stern, M.H. (2008) Haploinsufficiency of CDKN1B contributes to leukemogenesis in T-cell prolymphocytic leukemia. *Blood* 111:2321-2328.
- Lerner, M., Corcoran, M., Cepeda, D., Nielsen, M.L., Zubarev, R., Ponten, F., Uhlen, M., Hober, S., Grander, D., and Sangfelt, O. (2007) The RBCC gene RFP2 (Leu5) encodes a novel transmembrane E3 ubiquitin ligase involved in ERAD. *Molecular biology of the cell* 18:1670-1682.
- Li, B., Carey, M., and Workman, J.L. (2007) The role of chromatin during transcription. *Cell* 128:707-719.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69:915-926.
- Li, Q., Peterson, K.R., Fang, X., and Stamatoyannopoulos, G. (2002) Locus control regions. *Blood* 100:3077-3086.
- Lichter, P., and Ried, T. (1994) Molecular analysis of chromosome aberrations. In situ hybridization. *Methods Mol Biol* 29:449-478.
- Liu, T.H., Raval, A., Chen, S.S., Matkovic, J.J., Byrd, J.C., and Plass, C. (2006) CpG island methylation and expression of the secreted frizzled-related protein gene family in chronic lymphocytic leukemia. *Cancer research* 66:653-658.
- Liu, Y., Corcoran, M., Rasool, O., Ivanova, G., Ibbotson, R., Grander, D., Iyengar, A., Baranova, A., Kashuba, V., Merup, M., Wu, X., Gardiner, A., Mullenbach, R., Poltarau, A., Hultstrom, A.L., Juliusson, G., Chapman, R., Tiller, M., Cotter, F., Gahrton, G., Yankovsky, N., Zabarovsky, E., Einhorn, S., and Oscier, D. (1997) Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia. *Oncogene* 15:2463-2473.
- Liu, Y., Szekely, L., Grander, D., Soderhall, S., Juliusson, G., Gahrton, G., Linder, S., and Einhorn, S. (1993) Chronic lymphocytic leukemia cells with allelic deletions at 13q14 commonly have one intact RB1 gene: evidence for a role of an adjacent locus. *Proceedings of the National Academy of Sciences of the United States of America* 90:8697-8701.
- Lodish, H.F. (2000) *Molecular Cell Biology*. W.H. Freeman and company, New York.
- Lu, D., Zhao, Y., Tawatao, R., Cottam, H.B., Sen, M., Leoni, L.M., Kipps, T.J., Corr, M., and Carson, D.A. (2004) Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 101:3118-3123.
- Lujambio, A., Ropero, S., Ballestar, E., Fraga, M.F., Cerrato, C., Setien, F., Casado, S., Suarez-Gauthier, A., Sanchez-Cespedes, M., Git, A., Spiteri, I., Das, P.P., Caldas, C., Miska, E., and Esteller, M. (2007) Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer research* 67:1424-1429.
- Mabuchi, H., Fujii, H., Calin, G., Alder, H., Negrini, M., Rassenti, L., Kipps, T.J., Bullrich, F., and Croce, C.M. (2001) Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. *Cancer research* 61:2870-2877.
- Madrenas, J., Wange, R.L., Wang, J.L., Isakov, N., Samelson, L.E., and Germain, R.N. (1995) Zeta phosphorylation without ZAP-70 activation induced by TCR antagonists or partial agonists. *Science* 267:515-518.
- Maher, V.M., and McCormick, J.J. (1986) Role of DNA lesions and DNA repair in mutagenesis by carcinogens in diploid human fibroblasts. *Progress in clinical and biological research* 209A:245-253.
- Martin, G.S. (1970) Rous sarcoma virus: a function required for the maintenance of the transformed state. *Nature* 227:1021-1023.

- Mathieu, O., Probst, A.V., and Paszkowski, J. (2005) Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *EMBO J* 24:2783-2791.
- Matsuda, F., and Honjo, T. (1996) Organization of the human immunoglobulin heavy-chain locus. *Advances in immunology* 62:1-29.
- Mattick, J.S. (2003) Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays* 25:930-939.
- Mellor, J., Dudek, P., and Clynes, D. (2008) A glimpse into the epigenetic landscape of gene regulation. *Curr Opin Genet Dev* 18:116-122.
- Merlo, A., Herman, J.G., Mao, L., Lee, D.J., Gabrielson, E., Burger, P.C., Baylin, S.B., and Sidransky, D. (1995) 5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nature medicine* 1:686-692.
- Mertens, D., Wolf, S., Schroeter, P., Schaffner, C., Döhner, H., Stilgenbauer, S., and Lichter, P. (2002) Down-regulation of candidate tumor suppressor genes within chromosome band 13q14.3 is independent of the DNA methylation pattern in B-cell chronic lymphocytic leukemia. *Blood* 99:4116-4121.
- Mertens, D., Wolf, S., Tschuch, C., Mund, C., Kienle, D., Ohl, S., Schroeter, P., Lyko, F., Döhner, H., Stilgenbauer, S., and Lichter, P. (2006) Allelic silencing at the tumor-suppressor locus 13q14.3 suggests an epigenetic tumor-suppressor mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 103:7741-7746.
- Miao, F., and Natarajan, R. (2005) Mapping global histone methylation patterns in the coding regions of human genes. *Mol Cell Biol* 25:4650-4661.
- Migliazza, A., Bosch, F., Komatsu, H., Cayanis, E., Martinotti, S., Toniato, E., Guccione, E., Qu, X., Chien, M., Murty, V.V., Gaidano, G., Inghirami, G., Zhang, P., Fischer, S., Kalachikov, S.M., Russo, J., Edelman, I., Efstratiadis, A., and Dalla-Favera, R. (2001) Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. *Blood* 97:2098-2104.
- Moon, A.M., and Ley, T.J. (1990) Conservation of the primary structure, organization, and function of the human and mouse beta-globin locus-activating regions. *Proceedings of the National Academy of Sciences of the United States of America* 87:7693-7697.
- Moore, T., and Haig, D. (1991) Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* 7:45-49.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 51 Pt 1:263-273.
- Mutskov, V., and Felsenfeld, G. (2004) Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *EMBO J* 23:138-149.
- Nadkarni, J.S., Nadkarni, J.J., Clifford, P., Manolov, G., Fenyo, E.M., and Klein, E. (1969) Characteristics of new cell lines derived from Burkitt lymphomas. *Cancer* 23:64-79.
- Neuberger, M.S., Ehrenstein, M.R., Klix, N., Jolly, C.J., Yelamos, J., Rada, C., and Milstein, C. (1998) Monitoring and interpreting the intrinsic features of somatic hypermutation. *Immunol Rev* 162:107-116.
- Nowell, P.C. (1976) The clonal evolution of tumor cell populations. *Science* 194:23-28.
- O'Dwyer, M.E., Mauro, M.J., Kurilik, G., Mori, M., Balleisen, S., Olson, S., Magenis, E., Capdeville, R., and Druker, B.J. (2002) The impact of clonal evolution on response to imatinib mesylate (STI571) in accelerated phase CML. *Blood* 100:1628-1633.
- O'Neill, L.P., and Turner, B.M. (2003) Immunoprecipitation of native chromatin: NChIP. *Methods* 31:76-82.

- Oakes, C.C., Smiraglia, D.J., Plass, C., Trasler, J.M., and Robaire, B. (2003) Aging results in hypermethylation of ribosomal DNA in sperm and liver of male rats. *Proceedings of the National Academy of Sciences of the United States of America* 100:1775-1780.
- Ogawara, K., Miyakawa, A., Shiba, M., Uzawa, K., Watanabe, T., Wang, X.L., Sato, T., Kubosawa, H., Kondo, Y., and Tanzawa, H. (1998) Allelic loss of chromosome 13q14.3 in human oral cancer: correlation with lymph node metastasis. *International journal of cancer* 79:312-317.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99:247-257.
- Orlando, V. (2000) Mapping chromosomal proteins in vivo by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem Sci* 25:99-104.
- Orlando, V., Jane, E.P., Chinwalla, V., Harte, P.J., and Paro, R. (1998) Binding of trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early *Drosophila* embryogenesis. *Embo J* 17:5141-5150.
- Ozdog, H., Teschendorff, A.E., Ahmed, A.A., Hyland, S.J., Blenkiron, C., Bobrow, L., Veerakumarasivam, A., Burtt, G., Subkhankulova, T., Arends, M.J., Collins, V.P., Bowtell, D., Kouzarides, T., Brenton, J.D., and Caldas, C. (2006) Differential expression of selected histone modifier genes in human solid cancers. *BMC genomics* 7:90.
- Papantonis, A., and Lecanidou, R. (2008) A modified chromatin-immunoprecipitation protocol for silkworm ovarian follicular cells reveals C/EBP and GATA binding modes on an early chorion gene promoter. *Molecular biology reports*.
- Pauler, F.M., Koerner, M.V., and Barlow, D.P. (2007) Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet* 23:284-292.
- Pehrson, J.R., and Fried, V.A. (1992) MacroH2A, a core histone containing a large nonhistone region. *Science* 257:1398-1400.
- Pfister, S., Schlaeger, C., Mendrzyk, F., Wittmann, A., Benner, A., Kulozik, A., Scheurlen, W., Radlwimmer, B., and Lichter, P. (2007) Array-based profiling of reference-independent methylation status (aPRIMES) identifies frequent promoter methylation and consecutive downregulation of ZIC2 in pediatric medulloblastoma. *Nucleic acids research* 35:e51.
- Plass, C., Byrd, J.C., Raval, A., Tanner, S.M., and de la Chapelle, A. (2007) Molecular profiling of chronic lymphocytic leukaemia: genetics meets epigenetics to identify predisposing genes. *British journal of haematology* 139:744-752.
- Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D.K., and Young, R.A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 122:517-527.
- Rassenti, L.Z., Huynh, L., Toy, T.L., Chen, L., Keating, M.J., Gribben, J.G., Neuberg, D.S., Flinn, I.W., Rai, K.R., Byrd, J.C., Kay, N.E., Greaves, A., Weiss, A., and Kipps, T.J. (2004) ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *The New England journal of medicine* 351:893-901.
- Raval, A., Byrd, J.C., and Plass, C. (2006) Epigenetics in chronic lymphocytic leukemia. *Seminars in oncology* 33:157-166.
- Raval, A., Lucas, D.M., Matkovic, J.J., Bennett, K.L., Liyanarachchi, S., Young, D.C., Rassenti, L., Kipps, T.J., Grever, M.R., Byrd, J.C., and Plass, C. (2005) TWIST2 demonstrates differential methylation in immunoglobulin variable heavy chain mutated and unmutated chronic lymphocytic leukemia. *J Clin Oncol* 23:3877-3885.
- Raval, A., Tanner, S.M., Byrd, J.C., Angerman, E.B., Perko, J.D., Chen, S.S., Hackanson, B., Grever, M.R., Lucas, D.M., Matkovic, J.J., Lin, T.S., Kipps, T.J., Murray, F., Weisenburger, D., Sanger, W., Lynch, J., Watson, P., Jansen, M., Yoshinaga, Y.,

- Rosenquist, R., de Jong, P.J., Coggill, P., Beck, S., Lynch, H., de la Chapelle, A., and Plass, C. (2007) Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. *Cell* 129:879-890.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., and Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406:593-599.
- Rehm, H. (2002) *Der Experimentator: Proteinbiochemie / Proteomics*. 4. Auflage. Spektrum Heidelberg: Akad. Verlag 2002.
- Reik, W., Dean, W., and Walter, J. (2001) Epigenetic reprogramming in mammalian development. *Science* 293:1089-1093.
- Renart, J., Reiser, J., and Stark, G.R. (1979) Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proceedings of the National Academy of Sciences of the United States of America* 76:3116-3120.
- Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S., Guffanti, A., Minucci, S., Pelicci, P.G., and Ballabio, A. (2001) The tripartite motif family identifies cell compartments. *Embo J* 20:2140-2151.
- Rippe, K. (2007) Dynamic organization of the cell nucleus. *Curr Opin Genet Dev* 17:373-380.
- Robert-Koch-Institut (2008) Krebs in Deutschland 2003 – 2004. Häufigkeiten und Trends. In: Bundes, G.d. (ed). Robert-Koch-Institut, pp. 1-114.
- Rosenblum, B.B., Lee, L.G., Spurgeon, S.L., Khan, S.H., Menchen, S.M., Heiner, C.R., and Chen, S.M. (1997) New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic acids research* 25:4500-4504.
- Rosenwald, A., Alizadeh, A.A., Widhopf, G., Simon, R., Davis, R.E., Yu, X., Yang, L., Pickeral, O.K., Rassenti, L.Z., Powell, J., Botstein, D., Byrd, J.C., Grever, M.R., Cheson, B.D., Chiorazzi, N., Wilson, W.H., Kipps, T.J., Brown, P.O., and Staudt, L.M. (2001) Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *The Journal of experimental medicine* 194:1639-1647.
- Rougeulle, C., and Avner, P. (2004) The role of antisense transcription in the regulation of X-inactivation. *Curr Top Dev Biol* 63:61-89.
- Rougeulle, C., and Heard, E. (2002) Antisense RNA in imprinting: spreading silence through Air. *Trends Genet* 18:434-437.
- Rougeulle, C., Navarro, P., and Avner, P. (2003) Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression. *Hum Mol Genet* 12:3343-3348.
- Ruppel, M. (2003) Charakterisierung der Gene der chromosomalen Bande 13q14.3 hinsichtlich ihrer apoptosinduzierenden Wirkung. Fachbereich Biologie. Technische Universität Darmstadt, Darmstadt, p. 93.
- Rush, L.J., Raval, A., Funchain, P., Johnson, A.J., Smith, L., Lucas, D.M., Bembea, M., Liu, T.H., Heerema, N.A., Rassenti, L., Liyanarachchi, S., Davuluri, R., Byrd, J.C., and Plass, C. (2004) Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. *Cancer research* 64:2424-2433.
- Saito, Y., Liang, G., Egger, G., Friedman, J.M., Chuang, J.C., Coetzee, G.A., and Jones, P.A. (2006) Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer cell* 9:435-443.
- Sakai, T., Toguchida, J., Ohtani, N., Yandell, D.W., Rapaport, J.M., and Dryja, T.P. (1991) Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet* 48:880-888.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74:5463-5467.

- Santarosa, M., and Ashworth, A. (2004) Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. *Biochimica et biophysica acta* 1654:105-122.
- Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* 419:407-411.
- Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental biology* 241:172-182.
- Schagger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical biochemistry* 166:368-379.
- Schneider, U., Schwenk, H.U., and Bornkamm, G. (1977) Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. *International journal of cancer* 19:621-626.
- Schubeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., Bell, S.P., and Groudine, M. (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 18:1263-1271.
- Seiler, T., Döhner, H., and Stilgenbauer, S. (2006) Risk stratification in chronic lymphocytic leukemia. *Seminars in oncology* 33:186-194.
- Seligson, D.B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M., and Kurdستاني, S.K. (2005) Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* 435:1262-1266.
- Serra-Barrionuevo, L. (2008) Functional characterization of candidate tumor suppressor genes localized in the critical chromosomal region 13q14. Faculty of Natural Sciences. University of Ulm, Ulm, p. 152.
- Shamovsky, I., and Nudler, E. (2006) Gene control by large noncoding RNAs. *Sci STKE* 2006:pe40.
- Skoblov, M., Shakhbazov, K., Oshchepkov, D., Ivanov, D., Guskova, A., Ivanov, D., Rubtsov, P., Prasolov, V., Yankovsky, N., and Baranova, A. (2006) Human RFP2 gene promoter: unique structure and unusual strength. *Biochem Biophys Res Commun* 342:859-866.
- Smiraglia, D.J., and Plass, C. (2002) The study of aberrant methylation in cancer via restriction landmark genomic scanning. *Oncogene* 21:5414-5426.
- Smiraglia, D.J., Rush, L.J., Fruhwald, M.C., Dai, Z., Held, W.A., Costello, J.F., Lang, J.C., Eng, C., Li, B., Wright, F.A., Caligiuri, M.A., and Plass, C. (2001) Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum Mol Genet* 10:1413-1419.
- Smyth, G.K., Yang, Y.H., and Speed, T. (2003) Statistical issues in cDNA microarray data analysis. *Methods Mol Biol* 224:111-136.
- Spadari, S., Sutherland, B.M., Pedrali-Noy, G., Focher, F., Chiesa, M.T., and Ciarrocchi, G. (1987) Alteration of DNA tertiary structure by physical and chemical carcinogens: involvement in DNA repair processes. *Toxicologic pathology* 15:82-87.
- Spannhoff, A., Sippl, W., and Jung, M. (2008) Cancer treatment of the future: Inhibitors of histone methyltransferases. *Int J Biochem Cell Biol*.
- Spencer, V.A., Sun, J.M., Li, L., and Davie, J.R. (2003) Chromatin immunoprecipitation: a tool for studying histone acetylation and transcription factor binding. *Methods* 31:67-75.
- Statistisches-Bundesamt (2008) <http://www.destatis.de>. Statistisches Bundesamt Deutschland.
- Stehelin, D., Guntaka, R.V., Varmus, H.E., and Bishop, J.M. (1976a) Purification of DNA complementary to nucleotide sequences required for neoplastic transformation of fibroblasts by avian sarcoma viruses. *Journal of molecular biology* 101:349-365.

- Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976b) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260:170-173.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksoz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H., and Wanker, E.E. (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122:957-968.
- Stilgenbauer, S., Bullinger, L., Lichter, P., and Döhner, H. (2002) Genetics of chronic lymphocytic leukemia: genomic aberrations and V(H) gene mutation status in pathogenesis and clinical course. *Leukemia* 16:993-1007.
- Stilgenbauer, S., Döhner, H., Bulgay-Morschel, M., Weitz, S., Bentz, M., and Lichter, P. (1993) High frequency of monoallelic retinoblastoma gene deletion in B-cell chronic lymphoid leukemia shown by interphase cytogenetics. *Blood* 81:2118-2124.
- Stilgenbauer, S., Nickolenko, J., Wilhelm, J., Wolf, S., Weitz, S., Döhner, K., Boehm, T., Döhner, H., and Lichter, P. (1998) Expressed sequences as candidates for a novel tumor suppressor gene at band 13q14 in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. *Oncogene*, pp. 1891-1897.
- Stilgenbauer, S., Sander, S., Bullinger, L., Benner, A., Leupolt, E., Winkler, D., Kröber, A., Kienle, D., Lichter, P., and Döhner, H. (2007) Clonal evolution in chronic lymphocytic leukemia: acquisition of high-risk genomic aberrations associated with unmutated VH, resistance to therapy, and short survival. *Haematologica* 92:1242-1245.
- Strahl, B.D., and Allis, C.D. (2000) The language of covalent histone modifications. *Nature* 403:41-45.
- Suzuki, M.M., and Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nature reviews* 9:465-476.
- Szyf, M., and Bhattacharya, S.K. (2002) Extracting DNA demethylase activity from mammalian cells. *Methods Mol Biol* 200:163-176.
- Takeda, S., Fujiwara, T., Shimizu, F., Kawai, A., Shinomiya, K., Okuno, S., Ozaki, K., Katagiri, T., Shimada, Y., Nagata, M., Watanabe, T., Takaichi, A., Kuga, Y., Suzuki, M., Hishigaki, H., Takahashi, E., Shin, S., Nakamura, Y., and Hirai, Y. (1997) Isolation and mapping of karyopherin alpha 3 (KPNA3), a human gene that is highly homologous to genes encoding *Xenopus* importin, yeast SRP1 and human RCH1. *Cytogenetics and cell genetics* 76:87-93.
- Talcott, B., and Moore, M.S. (2000) The nuclear import of RCC1 requires a specific nuclear localization sequence receptor, karyopherin alpha3/Qip. *J Biol Chem* 275:10099-10104.
- Tam, C.S., and Keating, M.J. (2007) Chemoimmunotherapy of chronic lymphocytic leukemia. *Best Pract Res Clin Haematol* 20:479-498.
- Taneyhill, L.A., and Adams, M.S. (2008) Chapter 19 Investigating Regulatory Factors and Their DNA Binding Affinities Through Real Time Quantitative PCR (RT-QPCR) and Chromatin Immunoprecipitation (ChIP) Assays. *Methods Cell Biol* 87:367-389.
- Tariq, M., Saze, H., Probst, A.V., Lichota, J., Habu, Y., and Paszkowski, J. (2003) Erasure of CpG methylation in *Arabidopsis* alters patterns of histone H3 methylation in heterochromatin. *Proceedings of the National Academy of Sciences of the United States of America* 100:8823-8827.
- Tschuch, C. (2006) Functional characterization of candidate genes in the pathogenesis of B-cell chronic lymphocytic leukemia (B-CLL). Medical Faculty. Ruperto-Carola-University, Heidelberg, p. 132.
- Tschuch, C., Schulz, A., Pscherer, A., Werft, W., Benner, A., Hotz-Wagenblatt, A., Barrionuevo, L.S., Lichter, P., and Mertens, D. (2008) Off-target effects of siRNA specific for GFP. *BMC molecular biology* 9:60.

- Tycko, B., and Morison, I.M. (2002) Physiological functions of imprinted genes. *J Cell Physiol* 192:245-258.
- van Everdink, W.J., Baranova, A., Lummen, C., Tyazhelova, T., Looman, M.W., Ivanov, D., Verlind, E., Pestova, A., Faber, H., van der Veen, A.Y., Yankovsky, N., Vellenga, E., and Buys, C.H. (2003) RFP2, c13ORF1, and FAM10A4 are the most likely tumor suppressor gene candidates for B-cell chronic lymphocytic leukemia. *Cancer genetics and cytogenetics* 146:48-57.
- van Lohuizen, M. (1999) The trithorax-group and polycomb-group chromatin modifiers: implications for disease. *Curr Opin Genet Dev* 9:355-361.
- Verlaan, D.J., Ge, B., Grundberg, E., Hoberman, R., Lam, K., Koka, V., Dias, J., Gurd, S., Martin, N., Mallmin, H., Nilsson, O., Harmsen, E., Kwan, T., and Pastinen, T.M. (2008) Targeted screening of cis-regulatory variation in human haplotypes. *Genome Res.*
- Vogelstein, B., Lane, D., and Levine, A.J. (2000) Surfing the p53 network. *Nature* 408:307-310.
- Wahlfors, J., Hiltunen, H., Heinonen, K., Hamalainen, E., Alhonen, L., and Janne, J. (1992) Genomic hypomethylation in human chronic lymphocytic leukemia. *Blood* 80:2074-2080.
- Wang, R., Titley, J.C., Lu, Y.J., Summersgill, B.M., Bridge, J.A., Fisher, C., and Shipley, J. (2003) Loss of 13q14-q21 and gain of 5p14-pter in the progression of leiomyosarcoma. *Mod Pathol* 16:778-785.
- Weinberg, R.A. (2007) *The Biology of Cancer*. Garland Science, New York City.
- Welch, B.L. (1947) The generalization of "student's" problem when several different population variances are involved. *Biometrika* 34:28-35.
- Wirbelauer, C., Bell, O., and Schubeler, D. (2005) Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes Dev* 19:1761-1766.
- Wolf, S., Mertens, D., Schaffner, C., Korz, C., Dohner, H., Stilgenbauer, S., and Lichter, P. (2001) B-cell neoplasia associated gene with multiple splicing (BCMS): the candidate B-CLL gene on 13q14 comprises more than 560 kb covering all critical regions. *Hum Mol Genet* 10:1275-1285.
- Wolffe, A.P., Jones, P.L., and Wade, P.A. (1999) DNA demethylation. *Proceedings of the National Academy of Sciences of the United States of America* 96:5894-5896.
- Wood, A.J., and Oakey, R.J. (2006) Genomic imprinting in mammals: emerging themes and established theories. *PLoS Genet* 2:e147.
- Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., Wu, C., and Allis, C.D. (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442:86-90.
- Xiao, B., Wilson, J.R., and Gamblin, S.J. (2003) SET domains and histone methylation. *Current opinion in structural biology* 13:699-705.
- Yu, L., Liu, C., Vandeusen, J., Becknell, B., Dai, Z., Wu, Y.Z., Raval, A., Liu, T.H., Ding, W., Mao, C., Liu, S., Smith, L.T., Lee, S., Rassenti, L., Marcucci, G., Byrd, J., Caligiuri, M.A., and Plass, C. (2005) Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. *Nat Genet* 37:265-274.
- Yu, M.K. (2006) Epigenetics and chronic lymphocytic leukemia. *American journal of hematology* 81:864-869.
- Zaratiegui, M., Irvine, D.V., and Martienssen, R.A. (2007) Noncoding RNAs and gene silencing. *Cell* 128:763-776.
- Zenz, T., Mertens, D., Dohner, H., and Stilgenbauer, S. (2008) Molecular diagnostics in chronic lymphocytic leukemia - pathogenetic and clinical implications. *Leukemia & lymphoma* 49:864-873.

- Zeschnigk, M., Schmitz, B., Dittrich, B., Buiting, K., Horsthemke, B., and Doerfler, W. (1997) Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. *Hum Mol Genet* 6:387-395.
- Zhu, J., and Emerson, S.G. (2002) Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene* 21:3295-3313.
- Zippin, C., Cutler, S.J., Reeves, W.J., Jr., and Lum, D. (1971) Variation in survival among patients with acute lymphocytic leukemia. *Blood* 37:59-72.

Appendix A – Additional tables

Table A-1: ChIP and qPCR measured fraction of input-values of control CpG-island-promoters *CDH12* & *HK2* used for normalization. ** used for normalization of qPCR data. * calculated from duplicate qPCR reactions.

Sample No.	precipitating antibody	control promoter	fraction of input value**	standard deviation*
CLL1	H3K4me2	CDH12	0.0016	0.0001
CLL2	H3K4me2	CDH12	0.0022	0.0002
CLL3	H3K4me2	CDH12	0.0088	0.0017
CLL4	H3K4me2	CDH12	0.0048	0.0006
CLL5	H3K4me2	CDH12	0.0301	0.0059
P1	H3K4me2	CDH12	0.0436	0.0263
P2	H3K4me2	CDH12	0.0096	0.0072
P3	H3K4me2	CDH12	0.0036	0.0028
P4	H3K4me2	CDH12	0.0390	0.0068
P5	H3K4me2	CDH12	0.0049	0.0007
PB5	H3K4me2	CDH12	0.0001	0.0000
PB6	H3K4me2	CDH12	0.0030	0.0005
PB7	H3K4me2	CDH12	0.0051	0.0021
PB17	H3K4me2	CDH12	0.0012	0.0001
PB18	H3K4me2	CDH12	0.0004	0.0000
PB19	H3K4me2	CDH12	0.0059	0.0039
PB9 – B-cells	H3K4me2	CDH12	0.0016	0.0003
PB12 – B-cells	H3K4me2	CDH12	0.0052	0.0014
PB9 – T-cells	H3K4me2	CDH12	0.0452	0.0018
PB12 – T-cells	H3K4me2	CDH12	0.0022	0.0001
CLL1	macroH2A1.2	HK2	0.0006	0.0003
CLL2	macroH2A1.2	HK2	0.0027	0.0005
CLL3	macroH2A1.2	HK2	0.0005	0.0002
CLL4	macroH2A1.2	HK2	0.0015	0.0002
CLL5	macroH2A1.2	HK2	0.0000	0.0001
PU1	macroH2A1.2	HK2	0.0011	0.0001
PU2	macroH2A1.2	HK2	0.0008	0.0001
PL2	macroH2A1.2	HK2	0.0015	0.0003
PB17	macroH2A1.2	HK2	0.0011	0.0001
PB19	macroH2A1.2	HK2	0.0025	0.0006

Table A-2: Median H3K4me2-enrichment in wild type and several wild type subgroups. *Median values = x-fold enrichment vs. *CDH12*, wt = wild type, **CGI = CpG-island.

amplicon #	Median*					gene/CGI**
	wild type	young wt	old wt	B-cells	T-cells	
KPNA3e6.1	0.645	0.645	0.873	1.313	0.053	KPNA3
KPNA3e3.1	0.121	0.000	1.051	1.951	0.028	
KPNA3e1.1	7.294	7.294	16.371	6.120	0.325	
A1.2	2.414	3.457	5.676	5.136	0.536	CGI A
A2.1	1.869	1.869	5.347	11.303	0.118	
A3.1	4.266	1.003	9.982	7.757	0.189	
A3	0.553	0.553	1.707	0.609	0.111	non-related CGI
A4	1.240	1.984	0.853	2.557	0.110	
A2	2.931	1.376	7.077	8.556	0.271	
C13orf1e4.1	0.196	0.590	0.116	0.570	0.098	C13ORF1
C13orf1e3.1	0.484	0.603	0.057	1.216	0.279	
C13orf1e2.1	0.545	0.545	0.596	2.177	0.106	
B3	12.304	12.629	31.495	14.580	3.113	CGI B
B2	16.102	21.415	33.479	17.163	1.404	
B4	11.744	11.744	28.968	15.551	2.495	
RFP2e1.1	5.586	5.586	12.917	10.223	0.578	RFP2
C1	4.086	4.086	6.309	6.605	0.029	CGI C
C2	4.325	4.325	13.127	6.428	0.546	
C3	4.860	6.461	10.947	9.129	0.893	
RFP2e2.1	4.574	13.267	10.111	7.281	0.294	RFP2
RFP2e3.1	0.831	0.831	0.695	1.469	0.198	
BCMSUNe4.2	1.241	1.545	1.697	1.508	0.080	BCMSUN/ DLeu2
BCMSUNe3.1	1.413	1.477	2.288	1.754	0.000	
BCMSUNe3.2	1.411	1.411	1.587	1.171	0.101	
BCMSUNe2.2	1.141	3.984	2.590	3.630	0.170	
D1	4.458	3.320	23.793	28.262	2.269	CGI D
D3	10.175	20.125	34.658	16.750	1.697	
D4	14.742	19.659	35.972	11.296	2.289	
D2	4.408	6.624	3.015	4.164	0.232	
D5	8.238	12.062	12.482	5.425	0.408	
BCMSe1.1	8.074	16.102	19.865	9.603	0.817	BCMSe/ DLeu1
BCMSe2.1	5.863	12.048	11.887	7.966	0.519	
BCMSe3.1	1.123	0.325	1.737	2.675	0.053	
E1	7.227	25.928	7.389	10.731	0.402	CGI E
E1.1	nd	9.368	nd	nd	nd	
E12	9.154	14.138	17.853	12.142	1.551	
E1.2	nd	11.471	nd	nd	nd	
E2	9.557	26.213	18.537	13.056	1.031	
E3	3.338	2.040	10.801	9.404	1.550	
E4	1.225	1.644	2.843	1.604	0.079	
E5	4.738	7.641	5.039	9.338	0.215	
E6	1.809	3.453	11.663	4.730	0.887	
E7	2.577	2.121	3.246	5.802	0.105	

Table A-3: Median values of macroH2A-enrichment in wild type and two wild type-subgroups.
 *Median [x-fold enrichment vs. *HK2*], wt = wild type, **CGI = CpG-island.

amplicon #	Median*			gene/CGI**
	wild type	young wt	old wt	
KPNA3e6.1	0.121	0.000	0.966	KPNA3
KPNA3e3.1	0.000	0.000	1.047	
KPNA3e1.1	0.000	0.000	0.941	
A1.2	0.000	0.000	0.903	CGI A
A2.1	0.000	0.000	0.725	
A3.1	0.000	0.000	0.511	
A3	1.956	7.518	0.978	non-related CGI
A4	2.227	6.305	1.113	
A2	0.000	0.000	1.353	
C13orf1e4.1	0.000	0.000	0.407	C13ORF1
C13orf1e3.1	0.000	0.000	0.836	
C13orf1e2.1	0.000	0.000	0.617	
B3	1.828	4.581	0.914	CGI B
B2	0.000	0.000	0.887	
B4	0.000	0.000	0.917	
RFP2e1.1	0.000	0.000	0.886	RFP2
C1	1.082	1.082	2.894	CGI C
C2	0.000	0.000	0.711	
C3	1.816	2.677	0.908	
RFP2e2.1	0.000	0.000	1.242	RFP2
RFP2e3.1	1.963	2.897	0.982	
BCMSUNe4.2	1.836	6.400	0.918	BCMSUN/ DLeu2
BCMSUNe3.1	1.156	1.156	0.697	
BCMSUNe3.2	0.962	0.402	1.743	
BCMSUNe2.2	2.183	2.183	1.563	
D1	0.572	0.572	0.554	CGI D
D3	2.011	8.679	1.006	
D4	0.000	0.000	0.880	
D2	0.000	0.000	0.801	
D5	0.000	0.000	1.164	
BCMSe1.1	0.000	0.000	0.857	BCMS/ DLeu1
BCMse2.1	0.024	0.024	0.829	
BCMSe3.1	0.000	0.000	0.766	
E1	0.000	0.000	0.824	CGI E
E1.1	nd	nd	nd	
E12	0.000	0.000	1.284	
E1.2	nd	nd	nd	
E2	0.000	0.000	1.181	
E3	0.000	0.000	0.801	
E4	0.000	0.000	0.813	
E5	0.000	0.000	0.752	
E6	0.000	0.000	0.267	
E7	0.000	0.000	1.012	

Table A-4: Median H3K4me2-enrichment in CLL and two CLL-subgroups. *Median [x-fold enrichment vs *CDH12*], **17pdel = deletion of 17p, NK = normal karyotype +CGI = CpG-island.

amplicon #	Median*			gene/CGI ⁺
	CLL (all)	Del(13q) CLL	Del(17p)/NK** CLL	
KPNA3e6.1	0.526	0.000	5.666	KPNA3
KPNA3e3.1	0.288	0.000	17.507	
KPNA3e1.1	14.699	9.484	121.026	
A1.2	7.448	2.756	26.323	CGI A
A2.1	6.336	2.692	37.715	
A3.1	2.022	0.316	34.539	
A3	0.560	0.000	12.129	non-related
A4	1.116	0.000	14.783	
A2	5.920	4.581	58.832	CGI
C13orf1e4.1	27.467	14.720	252.176	C13ORF1
C13orf1e3.1	0.537	0.000	9.741	
C13orf1e2.1	0.832	0.000	11.622	
B3	75.399	64.997	377.080	CGI B
B2	40.604	37.430	590.605	
B4	86.261	28.272	431.714	
RFP2e1.1	12.364	11.132	91.714	RFP2
C1	27.309	9.603	282.237	CGI C
C2	45.326	20.092	313.145	
C3	37.123	14.179	202.285	
RFP2e2.1	58.018	17.481	319.237	RFP2
RFP2e3.1	1.472	0.000	11.014	
BCMSUNe4.2	3.864	0.000	11.655	BCMSUN/ DLeu2
BCMSUNe3.1	17.547	0.000	25.181	
BCMSUNe3.2	15.254	0.000	31.823	
BCMSUNe2.2	20.734	17.430	88.166	
D1	16.749	16.043	136.715	CGI D
D3	12.810	8.333	78.578	
D4	13.971	11.052	114.536	
D2	3.865	1.705	45.351	
D5	4.101	0.154	17.669	
BCMSe1.1	28.398	26.149	228.556	BCMS/ DLeu1
BCMse2.1	68.226	21.488	729.070	
BCMSe3.1	8.194	0.000	43.449	
E1	109.017	32.291	517.415	CGI E
E1.1	34.841	34.710	359.928	
E12	59.811	30.251	391.235	
E1.2	29.248	25.842	152.242	
E2	78.688	56.109	489.433	
E3	58.847	17.115	400.422	
E4	7.750	5.353	139.709	
E5	44.850	38.590	523.384	
E6	30.839	19.547	128.860	
E7	17.062	0.308	92.650	

Table A-5: Median values of macroH2A-enrichment in CLL and two CLL-subgroups (subdivided by karyotype). *Median [x-fold enrichment vs *HK2*], **17pdel = patient with deletion of chromosomal arm 17p, NK = patient with normal karyotype; ⁺CGI = CpG-island.

amplicon #	Median*			gene/CGI ⁺
	CLL (all)	Del(13q) CLL	Del(17p)/NK** CLL	
KPNA3e6.1	0.091	2.160	0.046	KPNA3
KPNA3e3.1	0.000	1.804	0.000	
KPNA3e1.1	0.540	2.575	0.270	
A1.2	2.724	4.616	0.000	CGI A
A2.1	2.438	5.463	0.000	
A3.1	1.780	5.265	0.000	
A3	3.285	4.435	0.645	non-related CGI
A4	0.899	3.307	0.450	
A2	1.571	3.005	0.000	
C13orf1e4.1	1.265	1.265	7.652	C13ORF1
C13orf1e3.1	2.018	2.018	7.666	
C13orf1e2.1	3.527	3.527	4.867	
B3	1.924	2.395	0.000	CGI B
B2	3.067	3.494	0.296	
B4	1.730	1.739	0.865	
RFP2e1.1	0.000	1.908	0.000	RFP2
C1	1.178	2.708	0.589	
C2	0.000	2.835	0.000	
C3	2.593	2.702	0.170	CGI C
RFP2e2.1	1.761	1.761	1.791	
RFP2e3.1	0.000	3.454	0.000	
BCMSUNe4.2	2.785	2.785	4.222	BCMSUN/DLeu2
BCMSUNe3.1	1.955	2.058	0.000	
BCMSUNe3.2	1.080	1.574	0.000	
BCMSUNe2.2	1.213	2.734	0.469	CGI D
D1	0.000	2.142	0.000	
D3	0.000	2.362	0.000	
D4	0.000	2.686	0.000	
D2	0.000	2.924	0.000	
D5	0.874	3.165	0.100	
BCMSe1.1	0.000	3.014	0.000	BCMS/DLeu1
BCMse2.1	2.409	2.778	0.000	
BCMSe3.1	0.082	1.908	0.041	
E1	0.121	3.080	0.000	CGI E
E1.1	nd	nd	nd	
E12	1.847	2.109	0.200	
E1.2	nd	nd	nd	
E2	1.207	2.462	0.604	
E3	3.518	4.147	0.544	
E4	1.844	2.877	0.922	
E5	0.452	2.217	0.226	
E6	3.174	3.391	0.000	
E7	2.502	5.313	1.020	

Appendix B – Primer sequences

Gene CpG-island	Primer forward Primer reverse	5'-3'- forward sequence 5'-3'-reverse sequence
CpG-island A	A1.2fwd A1.2rev	gcggaagcactcacttc agaaccccagcttgagaac
CpG-island A	A2.1fwd A2.1rev	gacctggcgactttactgc ggcattctcctcgttctcac
CpG-island A	A3.1fwd A3.1rev	gtgtcgcccagtttcttcat tgctgctgaaggctaagat
KPNA3	KPNA3e1.1_fwd KPNA3e1.1_rev	GCAGTAAAAGTCGCCAGGTCC GGGCTATACGTGCCCGTTG
KPNA3	KPNA3e3.1_fwd KPNA3e3.1_rev	AACAAAAGAGATGAACACTTATTGAAAAA AACATCTGAATCTTCTAGACTTTCTTCTTG
KPNA3	KPNA3e6.1_fwd KPNA3e6.1_rev	AAAAGTGTATCCAGTGACAGAAATCC TTGACTAGAATTGGTAAAATCCCAGAT
CpG-island B	B2fwd B2rev	CTTGAACACAGAACCGGGCCCA ACGCTGCGCCGACACCATTT
CpG-island B	B3fwd B3rev	CGCCCATTCGCGATGACAGACA TCGCAGATCTCTCCAGAAGCCG
CpG-island B	B4fwd B4rev	CGGCATCTCCTTCAGAGGGATGTG ATGGCCACCTCGGTGTTGTGCT
C13ORF1	C13ORF1e2.1_fwd C13ORF1e2.1_rev	GTTGTTATTGTAAAGAATGGAAGAAGAATATGT TGTAAGGTGCGCTGGCTAAA
C13ORF1	C13ORF1e3.1_fwd C13ORF1e3.1_rev	AATCAGATTCTCTTGGCCGA GCTCCATCATTCTCATCACCAG
C13ORF1	C13ORF1e4.1_fwd C13ORF1e4.1_rev	TGACCATGTGGAATTAATGTATACTTGA CCCTCGTATACCTGATGCTGGA
C13ORF1	C13ORF1e5.1_fwd C13ORF1e5.1_rev	TCATGACCCTGAGTCTTTGCCT GGGCGTAACTAAAGGACAGTGATT
CpG-island C	C1fwd C1rev	CGGAGCCGCGAGTCCATTTT TTCGAGGACCACCCCGCTTT
CpG-island C	C2fwd C2rev	CAGAAACCAGCGGGGCACTGTCAT CGGACGGAGCAGGTTTTCTGGA
CpG-island C	C3fwd C3rev	GCTACCAGCGTCTCCACATCCCCTA ATGCAACCAAACGCTGGCGG
CpG-island C	C1a_fwd C1a_rev	TAGAATGGGAGGCAGGTTTG ccgttgactaataatgcttagcc
CpG-island C	C2a_fwd C2a_rev	ggctaagcattattagcaacgg cgttcctccggcgctctc
CpG-island C	C3a_fwd C3a_rev	gagacgccggaaggaacg cctacaagctgatgtattggca
CpG-island C	C4a_fwd C4a_rev	tgccaaatacatcagctttagg gatgggattccacatattggc
CpG-island C	C5a_fwd C5a_rev	gaccaatattggtgaaatcccatc taggtttaacaggctctctagtagg
CpG-island C	C6a_fwd C6a_rev	cttcactagagagcctgttaaacta ataatcacatcagcaagaaatcac

Gene CpG-island	Primer forward Primer reverse	5'-3'- forward sequence 5'-3'-reverse sequence
RFP2	RFP2e1.1fwd RFP2e1.1rev	GTACCGTGCGGTCCCTGTAGTTGGA CTAGTGGGAAAGGCCGCGCG
RFP2	RFP2e1.2fwd RFP2e1.2rev	TACCGTGCGGTCCCTGTAGTTGGA TCGCTGAAGGTTTCGAGGACCAC
RFP2	RFP2e2.1fwd RFP2e2.1rev	TCGCTGAAGGTTTCGAGGACCAC GCTGATGTATTTGGCAGGGT
RFP2	RFP2e3.1fwd RFP2e3.1rev	TGCTTGAAGAAGATCTCACATGCCC GGAATTCCGCACACTCCCTTCTAAG
RFP2	RFP2e3.2fwd RFP2e3.2rev	TGGAGACCAGCTCCATTCAAGTGTC TGCATACTGGCATTTCGGAGAGAT
CpG-island D	D1fwd D1rev	GCAGCCAGGGCTTGGAAACTCTT CCGCCTTTTCTCGCCGTTTT
CpG-island D	D2fwd D2rev	ACAGGTTATCCTGTCTCTCCCGCT CGGGGTTGGCTCTAACGAATTT
CpG-island D	D3fwd D3rev	CAGACGCCCAAGTTGCCCTAA GGCGCGGGTCTACTCTCACTTAAT
CpG-island D	D4fwd D4rev	TCCCGAGTCCCGCTCTGCTACTTCT TACTTGGAGCAAAGGGCAGTCGGC
CpG-island D	D5fwd D5rev	GGCTGCCTCCACAGCTGTCAATACC GGTTGGAGTTTGCGCATGCGTA
CpG-island D	D1a_fwd D1a_rev	aaattagttgctggcatacaatg cttcaaatttgagaatacactagcg
CpG-island D	D2a_fwd D2a_rev	cgtagtgatttctcaaattgaag tgtaattcagtaactgcagtgaac
CpG-island D	D3a_fwd D3a_rev	gttcactgcagttactgaattagca ctgggtggttctatcaattcttactc
CpG-island D	D4a_fwd D4a_rev	gagtaagaattgatagaaaccaccag GGTCTACTCTCACTTAATCGCGA
CpG-island D	D5a_fwd D5a_rev	TCGCGATTAAGTGAGAGTAGACC GGTTGAATCGTACAGCGTGGT
CpG-island D	D6a_fwd D6a_rev	ACCACGCTGTACGATTCAACC TGCTACCCTTCTCCCACTTC
CpG-island D	D7a_fwd D7a_rev	GTGCCTGAAGTACTAATAAGGAAGTG taacatccttattgcctaacagg
BCMS	BCMSe1.1fwd BCMSe1.1rev	TTTTGCAAAGCCGCGGAGGT GGTAGCTATAAGACGACCCCTCGGC
BCMS	BCMSe1.2fwd BCMSe1.2rev	GCCCACAGGCATTTAGTCTACGTTG TTTCTTTTCTCCTAAGCAGGACCCG
BCMS	BCMSe2.1fwd BCMSe2.1rev	ATGAGGACACCTGAGGTTTCTAGA TTCTAAGACTTTGGGGCAGA
BCMS	BCMSe3.1fwd BCMSe3.1rev	CCTTCAGGAATTGAGTCACAATGCA AAAAGGGAAAGAATGGCTGGCAA
BCMS	BCMSe3.2fwd BCMSe3.2rev	TCTGGCAATAATAAAGTTGGGCAGC GGAAAGGGCCAAGAACTGCTGT
CpG-island E	E1fwd E1rev	TACCTCCTGGATTTACAACTGGG TCTTCAACAGTGTATAAATGCTACACAAAC
CpG-island E	E1.1fwd E1.1rev	CTTTCAGTTGTTGCCTCCAAACGG CGCGGTGAGCGCTATTTATAGGTAG

Gene CpG-island	Primer forward Primer reverse	5'-3'- forward sequence 5'-3'-reverse sequence
CpG-island E	E12fwd E12rev	GCGCCTCGTTCTATTTTTCT GGTCAGTCCGTACCGCCC
CpG-island E	E1.2fwd E1.2rev	GAACCCCCCGGCTCGATTTT TCCGCCTTCTCCTTTTCGCAA
CpG-island E	E2fwd E2rev	GGGAGTTTGCAGACCGGAG TCGTCATAAGGCTTTGAAGGAAAG
CpG-island E	E3fwd E3rev	CAGCACTCCTAAGACGGCG TGCAGCTTGCGCGA
CpG-island E	E4fwd E4rev	AATCTGGTCAGCTTAAAAGTGTCTC TGGACTCATTTAGGATCACCAGAT
CpG-island E	E5fwd E5rev	GTTTCAGACTTCAGTACCTAAAACGTTATG GACTCTAAATCGAGGTTTCTGATGCT
CpG-island E	E6fwd E6rev	CGCTAGTCTAGCCCAGCGT TCCTAACTCTTTGGGTCCCTGTAA
CpG-island E	E7fwd E7rev	CTGCTACAGCAGGCTTCGC AGTGCTACAACTCTTTCTCCATTTTCT
CpG-island E	E8fwd E8rev	ACTCCAGTGGTCTGCAAGCC GGCAAGCGATTGATTCCG
CpG-island E	E9fwd E9rev	GTGCATCCTGGTTAGAATTGGG GGCAAGCGATTGATTCCG
CpG-island E	E10fwd E10rev	AGGGAATTCTGTATTCTAGTTGTGGATCT TTGAACGACCTTGGGCCA
CpG-island E	E11fwd E11rev	TGTTAACTCAAGGCCGATTACCTT ttcctccacactcttctaca
CpG-island E	E1a_fwd E1a_rev	tctctcttaacatctacctctggat GACGCGTCCCATTGTGTGA
CpG-island E	E2a_fwd E2a_rev	TCACACAATGGGACGCGTC GACAGGTCAGTCCGTACCGC
CpG-island E	E3a_fwd E3a_rev	GCGGTACGGACTGACCTGTC TCTCACATCTAACTTCCAGTAACTTCTC
CpG-island E	E4a_fwd E4a_rev	GAGAAGTTACTGGAAGTTAGATGTGAGA CGCCGTCTTAGGAGTGCTG
CpG-island E	E5a_fwd E5a_rev	CAGCACTCCTAAGACGGCG TGGACTCATTTAGGATCACCAGAT
CpG-island E	E6a_fwd E6a_rev	atctggtgatcctaatgagtcca taggtactgaagtctgaacctagcc
CpG-island E	E7a_fwd E7a_rev	ggctaggttcagacttcagtaccta AGACGTGCAATAGCCACCG
CpG-island E	E8a_fwd E8a_rev	CGGTGGCTATTGCACGTCT ACGCTGGGCTAGACTAGCG
CpG-island E	E9a_fwd E9a_rev	CGTAGTCTAGCCCAGCGT GCGAAGCCTGCTGTAGCAG
CpG-island E	E10a_fwd E10a_rev	CTGCTACAGCAGGCTTCGC GGCAAGCGATTGATTCCG
CpG-island E	E11a_fwd E11a_rev	CGGAATCAATCGCTTGCC ttcctccacactcttctaca
CpG-island E	E12a_fwd E12a_rev	CGGAATCAATCGCTTGCC caaggatactgaagttctgtaagtagc

Gene CpG-island	Primer forward Primer reverse	5'-3'- forward sequence 5'-3'-reverse sequence
BCMSUN	BCMSUNe2.2fwd BCMSUNe2.2rev	TCTCATGTCATGTTCCAATGCA GGAGAACAGCCTCACTTCTTTGA
BCMSUN	BCMSUNe3.1fwd BCMSUNe3.1rev	TTGTAAAAACGGATGGGTGCA TCGATGCTGCTTGTGAGCTG
BCMSUN	BCMSUNe3.2fwd BCMSUNe3.2rev	CTTTGGCCACCATTATTGCATATT ATCGATGCTGCTTGTGAGCTGt
BCMSUN	BCMSUNe4.2fwd BCMSUNe4.2rev	GACCAGAAAGCAACTATATGATTACCTTT AAAGTTTGACATTTGCAATAGCATTC
GAPDH	GAPDHprom_fwd GAPDHprom_rev	CCGGGATTGTCTGCCCTAAT GCACGGAAGGTCACGATGT
MYH1	MYH1prom_fwd MYH1prom_rev	ACCACGATTTTCAGCAAGAATG GAGCAGGAAAATCGCAGAAC
CDH12	CDH12a96F04_fwd CDH12a96F04_rev	tgccatgtggagatgagaag acccaacattccaccaaagt
CDH12	CDH12b96F04_fwd CDH12b96F04_rev	gaccactgaaaagaaaggcatta acccaacattccaccaaagt
HK2	HK2c4F11_fwd HK2c4F11_rev	tctcgttccatgccttctct aatccaagcattcctgacg
HK2	HK2d4F11_fwd HK2d4F11_rev	tttgtccctcacctttcctg gagagaaggcatggaacgag
HPRT	HPRTprom1_fwd HPRTprom1_rev	CCATTTGCTGACTGTACTGTCCTAAGT ATTAGAGGAGTATAGGTATGCGTGTGTGT
HPRT	HPRTprom2_fwd HPRTprom2_rev	GATGACTGGAACCCGAAGAGC CCTTTCCCTCCCAACTCAGTCT
HPRT	HPRTe2.1_fwd HPRTe2.1_rev	ACCTAATCATTATGCTGAGGATTTGG TGCCATAATTAGTCCATGAGGAATAAAC
HPRT	HPRTe3.1_fwd HPRTe3.1_rev	TTGTAGCCCTCTGTGTGCTCAA CCAGCAGGTCAGCAAAGAATTTAT
HPRT	HPRTe4.1_fwd HPRTe4.1_rev	GGGACATAAAAAGTAATTGGTGGAGAT TTTTCTTCCCTTTCAAGATACATACCTT
HPRT	HPRTe8.1_fwd HPRTe8.1_rev	TTTGTAATGCCCTGTAGTCTCTGTAT GTCTGGCTTATATCCAACACTTCGT
cMYC	cMYCprom1_fwd cMYCprom1_rev	CGCGCTCTCCAAGTATACGTG GGAAAACGATGCCTAGAATGATTAAC
cMYC	cMYCprom2_fwd cMYCprom2_rev	GGGCAGCTGTTCCGCC TCTCCCTTTCTCTGCTGCTCC
cMYC	cMYCprom3_fwd cMYCprom3_rev	CCCAAAGCAGAGGGCGT AGCCGTCCAGACCCTCG
cMYC	cMYCe1.1_fwd cMYCe1.1_rev	GGCTTGGCGGGAAAAAGA CGAGTTAGATAAAGCCCCGAAAA
cMYC	cMYCe2.1_fwd cMYCe2.1_rev	CACCAACAGGAACTATGACCTCG GCTCGAATTTCTTCCAGATATCCTC
cMYC	cMYCe2.2_fwd cMYCe2.2_rev	TGGTCTTCCCCTACCCTCTCA ACTCCGTGAGGAGAGCAGA
cMYC	cMYCe3.1_fwd cMYCe3.1_rev	GCCTCCCTCCACTCGGAA GGCTGGTGCATTTTCGGTT
cMYC	cMYCe3.2_fwd cMYCe3.2_rev	TGGACTTTGGGCATAAAAAGAACTT CATTTAATGGCAATATTTACAGAGAAACAT

Gene CpG-island	Primer forward Primer reverse	5'-3'- forward sequence 5'-3'-reverse sequence
TUBB	TUBBprom1_fwd TUBBprom1_rev	CACTCAATTTTTCCCCATCAAT GCAACTCTGCAAGAGACCCAGTAT
TUBB	TUBBprom2_fwd TUBBprom2_rev	AAGTTCATGCCTTGGGTCACC GCTCCCAGCTGCAGTACCTCT
TUBB	TUBBprom3_fwd TUBBprom3_rev	GCCAATGGACGCCTGGTAC GGAGACAAAGCCTCATCGAGC
TUBB	TUBBe1.1_fwd TUBBe1.1_rev	CTTATTTTCTTGCCCCATACATACCTT GCACGATTTCCCTCATGGTTAA
TUBB	TUBBe2.1_fwd TUBBe2.1_rev	GACCCGCTGCACATATCCA GCCACACCCTTCCCCTAGA
TUBB	TUBBe3.1_fwd TUBBe3.1_rev	GGGAGGTGATCAGTGATGAACAT CATTGTAGTACACAGAGATGCGGTC
TUBB	TUBBe4.1_fwd TUBBe4.1_rev	AGCCATCCAGGAGCTCTTCAA TCCATGCCCTCGCCTGT
TUBB	TUBBe4.2_fwd TUBBe4.2_rev	CCCATCTCAGCTTCAAGGGAG AGGAGCTGAGTGAGGGAGGTAGA
ACTB	ACTBprom5_fwd ACTBprom5_rev	GAAGTGGCCAGGGCGG GCTGCGAGAATAGCCGGG
ACTB	ACTBe2.1_fwd ACTBe2.1_rev	CCAACTGGGACGACATGGAG GCAGCTCATTGTAGAAGGTGTGG
ACTB	ACTBe3.1_fwd ACTBe3.1_rev	ACCTTCAACACCCCAGCCA GAGGCGTACAGGGATAGCACAG
ACTB	ACTBe4.1_fwd ACTBe4.1_rev	AGGCATGGAGTCTGTGGC CACTTCATGATGGAGTTGAAGGTAGTT
ACTB	ACTBe5.1_fwd ACTBe5.1_rev	CACTGACTTGAGACCAGTTGAATAAAAAG TGGCCTCATTTTTAAGGTGTGC
UBE3A	UBE3Aprom1_fwd UBE3Aprom1_rev	AAGTTGCAAGTTGTAACACTATGAATTGT GGTACCAGGGTCACCATTTTAAAA
UBE3A	UBE3Aprom4_fwd UBE3Aprom4_rev	GGGTCACAGACAGCAGAAACTAAAA GCGACTGGACAGAAAGGGTAATC
UBE3A	UBE3Ae1.1_fwd UBE3Ae1.1_rev	AGATCCGTGTGTCTCCCAAG CCCCGTCGTCTCCTGTAGT
UBE3A	UBE3Ae2.1_fwd UBE3Ae2.1_rev	GAGCCTGATTTTAGAATTCACCA ACCATATTTGCCCCAACTTCTG
UBE3A	UBE3Ae3.1_fwd UBE3Ae3.1_rev	GCCTACGCTCAGATCAAGGT TTTTACAAGCTGTGGCCATTC
UBE3A	UBE3Ae7.1_fwd UBE3Ae7.1_rev	AACTGAGGGCTGTGGAAATG GGAGGGATGAGGATCACAGA
UBE3A	UBE3Ae8.1_fwd UBE3Ae8.1_rev	GTGACTTACTTAACAGAAGAGAAGGTATATGAAAT TGTACCAATGCCTCAGCACTAGAA
PEG3	PEG3prom1_fwd PEG3prom1_rev	AATTTCTTGCCTACCTTTCCAACAG GGTAAGTGTCTTCTCATTATTTGTTTTTAAAG
PEG3	PEG3prom4_fwd PEG3prom4_rev	ACGCAGTGGAAAGGTCTGTATTGA TCACACACTACTGAGTTGTTTTCCCA
PEG3	PEG3e4.1_fwd PEG3prome4.1_rev	GGCCCCAAATCTGTATGAGCTAG GACATCCGGCTCCTTAGTCAAGT
PEG3	PEG3e9.2_fwd PEG3prome9.2_rev	GCCACTCATCAAGATCCAAGAGA TGCTCCTGCTTACCTCGAC

Gene CpG-island	Primer forward Primer reverse	5'-3'- forward sequence 5'-3'-reverse sequence
Xist	Xistprom1_fwd Xistprom1_rev	TGGCTTTTATATGAGCTTTTTCTTCCT CATATGACACAGTAGTGAAGTTGACCCT
Xist	Xistprom3_fwd Xistprom3_rev	CCCTTTTTGTCCTTACTGGGTA TGAGGGCGGAGAGAGCATAA
Xist	Xiste1.1_fwd Xiste1.1_rev	GTTTCAGATTGTGGAGGAAAAGTGG CAAACGACTAGCCCTAAGCCG
Xist	Xiste1.3_fwd Xiste1.3_rev	TTCTTCATCTGGAGCACCTGC AGCACCTCTGCATATGTTCCCT
Xist	Xiste3.1_fwd Xiste3.1_rev	ACAGCTATAAGAGGCTCCAAATTAATCA CCAGATTCTCAAAGGGAAAGATATGA
Xist	Xiste4.1_fwd Xiste4.1_rev	CAGTGTTAGTGATCCATTCCCTTTG CCTCATGCCCCATCTCCAC
H19	H19CTCFbs_fwd H19CTCFbs_rev	CTCCTTCGGTCTCACCGCCTGGAT CCTTAGACGGAGTCGGAGCTG
H19	H19CTCFbs2_fwd H19CTCFbs2_rev	GAGGTGAATTTGCCACAGG CCCAAAGGAAATACTCCGGAA
HOXA9	HOXA9posK27_fwd HOXA9posK27_rev	CTCAGGAGCCTCGTGTCTTT GTGACCAGGTGGAGGTGTGT
HOXD12	HOXD12posK27_fwd HOXD12posK27_rev	GAACCTGCAGGCAAAGTTTC AGAGACTGCGCTCACACATC
PAX6	PAX6e5posK27_fwd PAX6e5posK27_rev	TCTTCTCTTCTTTTCCCCTTTC AGACACCACCGAGCTGATTC
CRHR2	CRHR2posK27_fwd CRHR2posK27_rev	CTTCAACGGCGTCAAGTACA GGGTCCGGAATCCTCTTTAC
ZMND1	ZMND1neg1K27_fwd ZMND1negK27_rev	TGCACGTACGTAGCCTTCAC TCCGAAATTTGTTGGAGCAT
PCK1	PCK1neg1K27_fwd PCK1negK27_rev	CCCACTGGGAACACAACTT CCTTTCTTCTCTTTGGATGATCT
APRT	APRTneg1K27_fwd APRTnegK27_rev	GCCTTGACTCGCACTTTTGT TAGGCGCCATCGATTTTAAG
MXD1	MXD1neg1K27_fwd MXD1negK27_rev	GAGATGGTGAGTCGGGGATA AACCACGCTCGACAAGAGAG
IL8	IL8mH2Apos1_fwd IL8mH2Apos1_rev	TTATAATTTATACCATGTAGC GGGTATCTTTAAACCAGACAGG
IL8	IL8mH2Apos2_fwd IL8mH2Apos2_rev	ATTTGATAAGGAACAAATAGGAAGTG GTGGCTTTTTATATCATCACCC
IL8	IL8mH2Aneg1_fwd IL8mH2Aneg1_rev	TCTTCTTCTGATAGACCAAACCTTTAAGG GATGGTGAAGATAAGCCAGCC
IL8	IL8mH2Aneg2_fwd IL8mH2Aneg2_rev	AAGCTTCTAGGACAAGAGCCAGG GGTCAGAAAGATGTGCTTACC

Publications

Articles based on this thesis:

Daniel Mertens, Angela Philippen, **Melanie Ruppel**, Danilo Allegra, Nupur Bhattacharya, Cordula Tschuch, Stephan Wolf, Irina Idler, Thorsten Zenz, Stephan Stilgenbauer. CLL and 13q14: miRs & more. Review. *Leukemia & Lymphoma*. 2009. In press.

Ruppel M., Philippen A., Fleig V., Döhner H., Stilgenbauer S., Lichter P. and Mertens D. The epigenetic code of the tumor suppressor mechanism in chromosomal band 13q14.3 in CLL. In preparation

Posters based on this thesis:

Melanie Ruppel, Daniel Mertens and Peter Lichter. Fine-mapping the histone code at 13q14.3, a critical region in B-cell Chronic Lymphocytic Leukemia. Presented at the European Hematology Association, Amsterdam, Netherlands, June 2006; the Genomes and Cancer Conference, Heidelberg, Germany, September 2006; and the Chromatin mediated biological decisions Conference, Marburg, Germany, October 2006.

Melanie Ruppel, Angela Philippen, Verena Fleig, Dirk Kienle, Hartmut Döhner, Peter Lichter and Daniel Mertens. A complex epigenetic tumor suppressor mechanism at 13q14 involves large ncRNA genes, monoallelic expression and miRNAs. Evaluation of the Research program Functional and Structural Genomics, Heidelberg October 2008.

Curriculum vitae

Particulars

Name: Melanie Ruppel
Title: Dipl. Biol.
Date of Birth: 13.11.1977
Place of Birth: Vreden, Germany
Family status: unmarried
Father: Manfred Ruppel
Mother: Monika Ruppel, geb. Hiepler

Education

Since July 2003 PhD student at the German Cancer Research Center in Heidelberg; Division of Molecular Genetics
Head: Prof. Dr. Peter Lichter

June 2003 Degree of Dipl. Biol. ('sehr gut'),
Technische Universität Darmstadt

October 1997 – July 2003 Study of Biology at the Technische Universität Darmstadt
Majors: Genetics, cell biology and animal physiology

June 1997 Abitur (A-levels), Score: 2.9

1988 – 1997 Gymnasium Georgianum, Vreden

1984 – 1988 Primary school, Vreden

Acknowledgments

My sincere thanks go to Prof. Dr. Peter Lichter for the opportunity to do my PhD thesis in his lab, for the motivating discussions and his thoughtful support over the years.

I sincerely thank Prof. Dr. H. Ulrich Göringer for his valuable support in finishing my thesis and for the opportunity to defend it at the Technische Universität Darmstadt.

I also thank Prof. Dr. Markus Löbrich for being the co-referee of my thesis and for his interest in this work.

I thank Prof. Dr. Gerhard Thiel and Prof. Dr. Boris Schmidt taking part in the examination committee in my dissertation defense.

I thank my supervisor Dr. Daniel Mertens for his support. The valuable comments and proofreading of this thesis have been of great help.

I thank the team at the Innere Medizin III at the Universitätsklinik Ulm for the cooperation.

I thank Angela Philippen for her collaboration and her contribution of BioCOBRA data to this thesis.

I thank Verena Fleig for her collaboration with aPRIMES DNA-methylation analyses.

I appreciate the valuable technical support of Dr. Bernhard Radlwimmer, Dr. Stefan Pfister and Dr. Björn Fritz with CGI-microarrays, hybridizations, analyses and the pipetting robot. I especially thank Stefanie Hofmann and Andrea Wittmann for their help and always friendly and patient support in the lab.

In the last years, Dr. Armin Pscherer and Dr. Stefanie Heck let me benefit from their extensive knowledge and advice. I thank both for the great atmosphere in the office and for all the delicious espressi.

Many thanks go to Natalia Becker for her help with statistical analyses in 'R'.

For the support and helpful advices whenever needed during my work in the lab, I thank Magdalena Schlotter, Frauke Devens, Petra Schröter, Stefanie Hofmann, Andrea Wittmann, Heidi Kramer, Sybille Ohl, Dr. Cordula Tschuch, Dr. Letitia Serra-Barrionuevo and Dr. Stephan Wolf.

For great technical support during his internship, I thank Christian Kern.

The "Frauenabende" have always been great fun. Thanks to Frauke Devens, Angela Schulz, Dr. Cordula Tschuch, Suzin Choi, Dr. Letitia Serra-Barrionuevo and Alexandra Farfsing.

For great lunch times despite the food, I thank Dr. Cordula Tschuch, Dr. Stefanie Heck, Dr. Armin Pscherer, Suzin Choi and Alexandra Farfsing. For the great food and the 'Spaghetti Club' experiences on fridays, I thank Dr. Cordula Tschuch, Kathrin Wildenberger, Frauke Devens, Suzin Choi, Natalia Becker, Angela Schulz, Jan Meier and Dr. Armin Pscherer.

Acknowledgments

I thank my colleagues of the Lichter-lab and of the Mertens-group for the friendly atmosphere within the groups and the smiles on the way.

I especially thank Dr. Cordula Tschuch and Dr. Letitia Serra-Barrionuevo for sharing with me the ups and downs in the lab during the PhD, and for the motivating discussions about science, life and everything else in and outside of the lab.

Rolf, I hope I will not have to pay for the dinner! ☺

I am most thankful for the unconditional support of my parents and my sisters, Steffie and Chrissie, who always believe in me and who still try to understand what I am actually doing.

Last, but not least I deeply thank Stephan Partsch for his thorough proofreading of this thesis, but most of all for his patience, his love and his caring support over the last years.

A scientist in his laboratory is not a mere technician: he is also a child confronting natural phenomena that impress him as though they were fairy tales.
Marie Curie

The process of scientific discovery is, in effect, a continual flight from wonder.
Albert Einstein

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

Heidelberg, 05.01.2009

Melanie Ruppel