UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1363

Prognostic markers and DNA methylation profiling in lymphoid malignancies

SUJATA BHOI

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017

ISSN 1651-6206
ISBN 978-91-513-0053-5
urn:nbn:se:uu:diva-328616

Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Dag Hammarskjölds väg 20, Uppsala, Thursday, 19 October 2017 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Professor H Denis Alexander (N Ireland Centre for Stratified Medicine, Ulster University, Altnagelvin Area Hospital, Derry/Londonderry).

Abstract

Bhoi, S. 2017. Prognostic markers and DNA methylation profiling in lymphoid malignancies. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1363. 77 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0053-5.

In recent years, great progress has been achieved towards identifying novel biomarkers in lymphoid malignancies, including chronic lymphocytic leukemia (CLL) and mantle cell lymphoma (MCL), at the genomic, transcriptomic and epigenomic level for accurate risk-stratification and prediction of treatment response. In **paper I**, we validated the prognostic relevance of a recently proposed RNA-based marker in CLL, *UGT2B17*, and analyzed its expression levels in 253 early-stage patients. Besides confirming its prognostic impact in multivariate analysis, we could identify 30% of IGHV-mutated CLL (M-CLL) cases with high expression and poor outcome, which otherwise lacked any other poor-prognostic marker. In **paper II**, we investigated the prognostic impact of a previously reported 5 CpG signature that divides CLL patients into three clinico-biological subgroups, namely naive B-cell-like CLL (n-CLL), memory B-cell-like CLL (m-CLL) and intermediate CLL (i-CLL), in 135 CLL patients using pyrosequencing. We validated the signature as an independent marker in multivariate analysis and further reported that subset #2 cases were predominantly classified as i-CLL, although displaying a similar outcome as n-CLL. In **paper III**, we investigated the methylation status and expression level of miR26A1 in both CLL (n=70) and MCL (n=65) cohorts. High miR26A1 methylation was associated with IGHV-unmutated (U-CLL) and shorter overall survival (OS) in CLL, while it was uniformly hypermethylated in MCL. Furthermore, overexpression of miR26A1 resulted in significant downregulation of EZH2 that in turn led to increased apoptosis. In **paper IV**, we performed DNA methylation profiling in 176 CLL cases assigned to one of 8 major stereotyped subsets (#1-8) in relation to non-subset CLL (n=325) and different normal B-cell subpopulations. Principal component analysis of subset vs. non-subset CLL revealed that U-CLL and M-CLL subsets generally clustered with n-CLL and m-CLL, respectively, indicating common cellular origins. In contrast, subset #2 emerged as the first defined member of the i-CLL subgroup, which in turn alludes to a distinct cellular origin for subset #2 and i-CLL patients. Altogether, this thesis confirms the prognostic significance of RNA and epigenetic-based markers in CLL, provides insight into the mechanism of miRNA deregulation in lymphoid malignancies and further unravels the DNA methylation landscape in stereotyped subsets of CLL.

Keywords: Lymphoid malignancies, CLL, MCL, prognostic markers, micro-RNA, methylation, stereotyped subsets

Sujata Bhoi, Department of Immunology, Genetics and Pathology, Experimental and Clinical Oncology, Rudbecklaboratoriet, Uppsala University, SE-751 85 Uppsala, Sweden.

ISSN 1651-6206
ISBN 978-91-513-0053-5
urn:nbn:se:uu:diva-328616 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-328616)

Dedicated to my beloved family

Supervisors and Examining Board

Main supervisor	**Richard Rosenquist Brandell**, Professor, MD Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala
Co-supervisor	**Larry Mansouri**, Associate Professor, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala
Faculty Opponent	**H Denis Alexander,** DPhil FRCPath Professor of Stratified Medicine N Ireland Centre for Stratified Medicine, Ulster University, Altnagelvin Area Hospital, Derry/Londonderry
Examining committee	**Honar Cherif,** Associate Professor, MD Department of Medical Sciences, Section of Hematology, Uppsala University, Uppsala
	Berivan Baskin, Associate Professor Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala
	Johanna Ungerstedt, Associate Professor**,** MD Center for Hematology and Regenerative Medicine, Department of Medicine Huddinge, Division of Hematology, Karolinska Institutet, Stockholm

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals. Reprints were made with permission from the respective publishers.

I **Bhoi S**, Baliakas P, Cortese D, Mattsson M, Engvall M, Smedby KE, Juliusson G, Sutton LA, Mansouri L. *UGT2B17* expression: a novel prognostic marker within IGHV-mutated chronic lymphocytic leukemia? *Haematologica* 2016; 101(2):e63-5.

II **Bhoi S**, Ljungström V, Baliakas P, Mattsson M, Smedby KE, Juliusson G, Rosenquist R, Mansouri L. Prognostic impact of epigenetic classification in chronic lymphocytic leukemia: The case of subset #2. *Epigenetics* 2016; 2;11(6):449-55.

III Kopparapu PK*, **Bhoi S***, Mansouri L, Arabanian LS, Plevova K, Pospisilova S, Wasik AM, Croci GA, Sander B, Paulli M, Rosenquist R, Kanduri M. Epigenetic silencing of miR-26A1 in chronic lymphocytic leukemia and mantle cell lymphoma: Impact on EZH2 expression. *Epigenetics* 2016; 11(5):335-43.

IV Mansouri L*, **Bhoi S***, Castellano G, Sutton LA, Papakonstantinou N, Queirós A, Baliakas P, Ek S, Emruli VK, Plevova K, Ntoufa S, Davis Z, Young E, Göransson-Kultima H, Isaksson A, Smedby KE, Gaidano G, Langerak AW, Davi F, Rossi D, Oscier D, Pospisilova S, Ghia P, Campo E, Stamatopoulos K, Martín-Subero JI **, Rosenquist R**. Genome-wide DNA methylation profiling in chronic lymphocytic leukemia patients carrying stereotyped B-cell receptors. *Manuscript*.

*Equal first authors, **Equal senior authors

List of related papers published during the PhD period

1. Rosenquist R, Cortese D, **Bhoi S**, Mansouri L, Gunnarsson R. Prognostic markers and their clinical applicability in chronic lymphocytic leukemia: where do we stand? *Leuk Lymphoma* 2013; 54(11):2351-2364.

2. Mansouri L, Sutton L-A, Ljungström V, Bondza S, Arngården L, **Bhoi S**, Larsson J, Cortese D, Kalushkova A, Plevova K, Young E, Gunnarsson R, Falk-Sörqvist E, Lönn P, Muggen AF, Yan X-J, Sander B, Enblad G, Smedby KE, Juliusson G, Belessi C, Rung J, Chiorazzi N, Strefford JC, Langerak A W, Pospisilova S, Davi F, Hellström M, Jernberg-Wiklund H, Ghia P, Söderberg O, Stamatopoulos K, Nilsson M, Rosenquist R. Functional loss of IκBε leads to NF-κB deregulation in aggressive chronic lymphocytic leukemia. *J Exp Med* 2015; 212(6):833-843

Contents

Abbreviations

+12	Trisomy 12
AID	Activation induced cytidine deaminase
BcR	B-cell receptor
BL	Burkitt lymphoma
BTK	Bruton's tyrosine kinase
C	Constant
CCND1	Cyclin D1
CDR	Complementarity determining region
CLL	Chronic lymphocytic leukemia
CLLU1	CLL up-regulated gene 1
cs MBC	Class switched memory B-cell
CSR	Class-switch recombination
D	Diversity
DAC	5'-Aza-2'-deoxycytidine
DD	Differential display
del(11q)	Deletion of long arm of chromosome 11
del(13q)	Deletion of long arm of chromosome 13
del(17p)	Deletion of short arm of chromosome 17
DLBCL	Diffuse large B-cell lymphoma
FCR	Fludarabine, cyclophosphamide and rituximab
FISH	Fluorescence in situ hybridization
FL	Follicular lymphoma
GC	Germinal center
GEP	Gene expression profiling
GO	Gene ontology
HR	Hazard ratio
HSC	Hematopoietic stem cell
i-CLL	Intermediate CLL
IG	Immunoglobulin
IGH	Immunoglobulin heavy
IGHV	Immunoglobulin heavy variable
IL-7	Interleukin-7
IPI	International prognostic index
J	Joining
LDH	Lactate dehydrogenase
LPL	Lipoprotein lipase

MBC	Memory B-cell
MBL	Monoclonal B-cell lymphocytosis
MCL	Mantle cell lymphoma
MCL1	Myeloid cell leukemia 1
M-CLL	IGHV-mutated CLL
m-CLL	Memory B-cell-like CLL
MDR	Minimal deleted region
MIPI	MCL-International Prognostic Index
miR	Micro-RNA
MZ	Marginal zone
NBC	Naïve B cell
n-CLL	Naive B-cell-like CLL
ncs MBC	Non-class switched memory B-cell
NGS	Next-generation sequencing
NHL	Non-Hodgkin lymphomas
OS	Overall survival
PB	Peripheral blood
PCR	Polymerase chain reaction
PFS	Progression-free survival
PI3K	Phosphatidyl inositol 3 kinase
Pre-B-cell	Precursor B-cell
Pro-B-cell	Progenitor B-cell
RAG	Recombination activating gene
ROC	Receiver operating characteristics
RQ-PCR	Real-time quantitative PCR
RS	Richter's syndrome
RSS	Recombination signal sequences
SHM	Somatic hypermutation
SVM	Support vector machine
TCL1	T-cell leukemia protein 1a
TdT	Terminal deoxynucleotidyl transferase
TLR	Toll-like receptor
TTFT	Time-to-first-treatment
U-CLL	IGHV-unmutated CLL
VH	Variable heavy
WGBS	Whole-genome bisulfite sequencing
WHO	World Health Organization
ZAP70	Zeta chain associated protein kinase 70 kDa

Introduction

Lymphoid malignancies occur due to the malignant transformation of normal lymphocytes at various stages of differentiation and are considered as the sixth most common group of malignancies worldwide.[1] They are remarkably heterogeneous from a clinicobiological perspective and vary greatly with respect to their morphology, immunophenotype, molecular characteristics, and maturation stage at oncogenic transformation. According to the World Health Organization (WHO), malignant lymphomas can be broadly classified into three subtypes; mature B-cell lymphoma, T-cell lymphoma and Hodgkin lymphoma.[2,3] B-cell malignancies comprises approximately 90% of all lymphoid neoplasms throughout the world and includes lymphomas, such as diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), mantle cell lymphoma (MCL), marginal zone lymphoma and Burkitt lymphoma (BL), as well as some leukemias, such as chronic lymphocytic leukemia (CLL) and hairy cell leukemia.[4,5]

This thesis will focus on prognostic markers and methylation profiling in CLL and MCL. However, in order to appreciate the relevance of the findings from this thesis, a basic introduction to normal B-cell development and maturation is first provided alongside descriptions of the disease entities.

Normal B-cell development and maturation

B-cells are the principal cellular component of the adaptive immune system raising humoral immunity through the secretion of antibodies. Deriving from pluripotent hematopoietic stem cells (HSCs), the B-cells differentiation begins in primary lymphoid organs (fetal liver and bone marrow) to the secondary lymphoid organs (lymph nodes and spleen) as represented in Figure 1.[6,7] The developing B-cells final destiny is either as a memory B-cell expressing surface immunoglobulin (IG) or as an antibody producing plasma cell.[7]

The first stage of B-cell development involves the differentiation of HSCs into the earliest committed cell of the B-cell lineage, known as progenitor B-cell (pro B-cell), which is followed by differentiation into precursor B-cells (pre B-cells). Both stages are antigen-independent, however the differentiation of pro B-cells to pre B-cells depends on the support from

bone marrow stromal cells mediated through cytokines such as interleukin-7 (IL-7).[8]

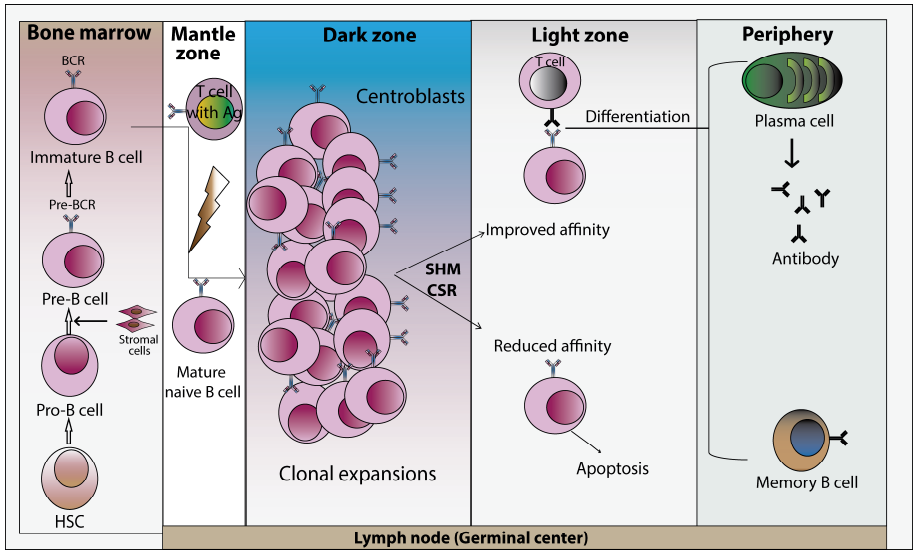

Figure 1. Schematic representation of normal B-cell development and maturation in the bone marrow and lymph nodes. Antigen activated B-cells migrate to the lymphoid organ where they undergo clonal expansion, followed by somatic hypermutation (SHM) and class-switch recombination (CSR) inside the dark zone of the germinal center (GC). B-cells with improved affinity undergo further differentiation into antibody producing plasma cells or memory B-cells. Modified from Edwards JC, Nat Rev Immunol, 2006 and Rickert RC, Nat Rev Immunol, 2013.

Ordered rearrangement of the IG heavy (IGH) locus initially takes place during the differentiation of pro B-cells to pre B-cells. The pre B-cells express a pre-B-cell receptor (pre-BcR) that consists of a heavy chain and a surrogate light chain.[9,10] Once functional heavy and light chain rearrangements are expressed, the pre B-cells differentiate into membrane IgM expressing immature B-cells.[11] Upon exiting the bone marrow and entering the peripheral blood the immature B-cells mature to näive B-cells, expressing both IgM and IgD. Activation of these B-cells occurs through antigen-antibody interactions facilitated by antigen presenting cells such as follicular dendritic cells.[12,13]

Antigen-activated B-cells then migrate into the primary follicle of secondary lymphoid organs, where they undergo clonal expansion in specialized compartments known as germinal centers (GC).[13,14] The activated B-cells at this stage are referred to as centroblasts. The GC architecture can be divided into four distinct zones known as the dark zone, the light zone, the mantle zone and the marginal zone (Figure 1). Within the dark zone, rapid

proliferation of centroblasts takes place coupled with genetic modification i.e. somatic hypermutation (SHM) and class-switch recombination (CSR) as discussed below.[14,15] These genetic modification prime the B-cell for its eventual faith with high affinity B-cells undergoing further differentiation to plasma cells or memory B-cells and reduced affinity B-cells undergoing apoptosis.[16]

Immunoglobulin gene rearrangements

The IG molecule is made up of two identical heavy and light chains, joined by disulfide bridges and non-covalent interactions (Figure 2).[17] Both heavy and light chains contain constant (C) regions and variable (V) regions. The C region is highly conserved, while the V region displays extensive variability and represents the antigen binding sites. The V region consists of three complementarity-determining regions (CDRs), CDR1, CDR2 and CDR3, of which the variable heavy (VH) CDR3 is the most hypervariable segment. IG rearrangements occurs through the joining of V, diversity (D) and joining (J) genes.[17] VDJ recombination plays a significant role in creating IG diversity, which is an integral aspect to the body's immune response.[17,18]

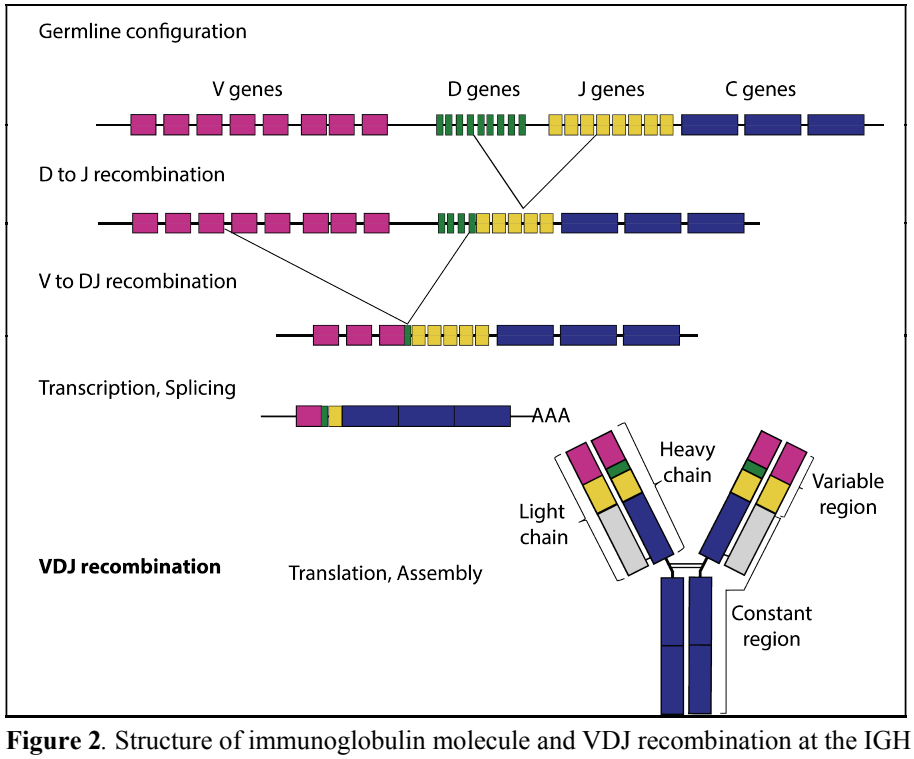

Figure 2. Structure of immunoglobulin molecule and VDJ recombination at the IGH locus.

Genes encoding the heavy chain are located on chromosome 14, i.e. the IGH locus, while two IG light chain loci exist, i.e. IG kappa (IGK) on chromosome 2 and IG lambda (IGL) on chromosome 22. The variable (IGHV), diversity (IGHD) and joining (IGHJ) genes are located at the IGH locus, whereas the light chain only contains variable (IGKV/IGLV) and joining (IGKJ/IGLJ) genes. The recombination process is initiated by two specific recombination activating gene (RAG) enzymes, known as RAG1 and RAG2 that introduce double stranded breaks at specific recombination signal sequences (RSS) flanking the coding V, D and J genes.[19-21] Several other enzymes are involved in the process of cleavage and ligation, such as DNA repair enzymes that remove unpaired nucleotides through exonuclease activity, and the DNA polymerase terminal deoxynucleotidyl transferase (TdT) that introduces nucleotides randomly at the free ends of DNA leading to additional IG diversity[22-24]; and, finally, ligase enzymes join the DNA strands to each other thus accomplishing the recombination process.[24]

The initial events involve the juxtaposition of an IGHD gene to an IGHJ gene leading to an IGHD-IGHJ rearrangement, followed by the joining of an IGHV gene to the IGHD-IGHJ complex, while the IGHC gene remains separated from the IGHV-D-J complex by an intronic region which is removed later through RNA splicing. This rearranged IGHV-D-J complex encodes the VH CDR3 of the IG molecule. The light chain rearrangement involves only one rearrangement step that joins IGKV/IGLV to IGKJ/IGLJ genes, as the IGK/IGL loci lack D genes (Figure 2).[21,25]

Somatic hypermutation and class-switch recombination

Affinity maturation of IG molecules is mediated through a highly specialized process of SHM, which randomly introduces single nucleotide mutations into the IG genes and takes place in the dark zone of GC (Figure 1).[26] This process is triggered by activation induced cytidine deaminase (AID), an enzyme that is highly expressed in activated B-cells in the GCs.[26,27] Acquired mutations are more commonly targeted at certain hotspots regions in CDR sequences consisting of specific amino acid motifs.[28-31] The process of SHM increases IG diversity and specificity.[27] Cells with disadvantageous mutations lead to reduced affinity resulting in apotosis.[32]

CSR is a mechanism through which switch in isotypic expression of IG molecules occurs, such as isotype IgM to isotype IgG, IgA or IgE, and is mostly restricted to the activated GC B-cells.[33] The C region of the IG molecule is targeted by CSR while the V region remains unchanged, hence CSR does not influence antigen specificity.[34] It occurs through a deletional recombination event that involves two different conserved nucleotide motifs

known as switch (S) regions, associated with the heavy chain C region and is mediated by a number of enzymes including AID.[33,35] During CSR the expressed heavy chain C region of the IgD (δ) or IgM (μ) molecule is replaced with the expression of the heavy chain C region of IgG (γ), IgA (α) or IgE (ε).[33,34]

Chronic lymphocytic leukemia

Background

CLL is a B-cell malignancy characterized by clonal expansion and accumulation of mature neoplastic B-cells in the bone marrow, peripheral blood and lymph nodes. The typical immunophenotypic profile includes the expression of surface markers CD5, CD19, CD20 and CD23.[36] CLL is the most common form of leukemia in adults in the Western world.[37,38] The median age at diagnosis is 71 years and more male than female are affected (ratio 2:1). The clinical diagnosis of CLL is determined by increased absolute lymphocyte count in peripheral blood (>5.0x10^9 cells/L) and the aforementioned clonal immunophenotype.[36] Therapeutic options in CLL have improved greatly in recent years including chemotherapy, monoclonal antibodies and small molecule inhibitors, however, the disease remains incurable.[39]

CLL is an extremely heterogeneous disease, both clinically and biologically, where some patients (15%) require treatment immediately after diagnosis, in contrast to other patients that have a considerably more indolent disease course, monitored by a "wait and watch" approach even for decades. The median survival time today is approximately 10 years.[40] The Rai and Binet staging systems are routinely employed in CLL and are mostly based on physical examination and standard laboratory blood tests. Both systems stratify patients into distinct stages with varying clinical outcome (i.e. Rai stage 0-IV and Binet stage A-C).[41,42] CLL is always preceded by a pre-leukemic condition known as monoclonal B-cell lymphocytosis (MBL).[43] Additionally, a minor proportion of CLL patients (~5%) will transform into a high-grade lymphoma, known as Richter's syndrome (RS), often associated with a very poor clinical outcome.[44]

Currently, fludarabine in combination with cyclophosphamide and rituximab (FCR) is considered the "gold standard" treatment regimen for medically fit patients with a more than 90% response rate.[39,45] However, patients with *TP53* aberrations are resistant to this treatment and hence should be considered for treatment with BcR inhibitors, such as ibrutinib which targets Bruton's tyrosine kinase (BTK) and idelalisib which targets phosphatidyl inositol 3 kinase (PI3K) or BCL2 inhibitors such as venetoclax.[46-50]

Prognostic markers in CLL

The clinical heterogeneity observed in CLL is considered to reflect the underlying genetic and epigenetic heterogeneity that contributes to the multifaceted pathobiology of the disease.[51] In the last decades, significant progress has been made in identifying novel biomarkers that can aid prediction of disease progression and improve risk stratification.[52-54] A considerable number of prognostic markers have been proposed in CLL (Figure 3), including both DNA and RNA-based markers, however only a few of these are currently being applied in routine clinical practice. For instance, the IGHV mutational status and certain recurrent genomic aberrations (more detailed below) have been shown to be particularly relevant and are commonly applied in the clinical setting.[55-57]

Thanks to advancement in next-generation sequencing (NGS) and array-based technologies, our understanding of the genomic, transcriptomic and epigenomic complexity of the disease is expanding and novel biomarkers have been uncovered. Using high-throughput sequencing technologies, mutations within *NOTCH1, SF3B1* and *BIRC3* were identified, improving risk stratification and refining prognosis in CLL.[58-63] Additionally, epigenetic alterations, such as aberrant DNA methylation, have in recent years been suggested to improve CLL prognostication, where the methylation statuses of a number of genes (e.g. *ZAP70, KIBRA, TWIST*) show significant influence on outcome prediction.[64-68] In the following sections, a more detailed description of relevant prognostic markers, at the genomic, transcriptomic and epigenomic level, will be outlined and discussed.

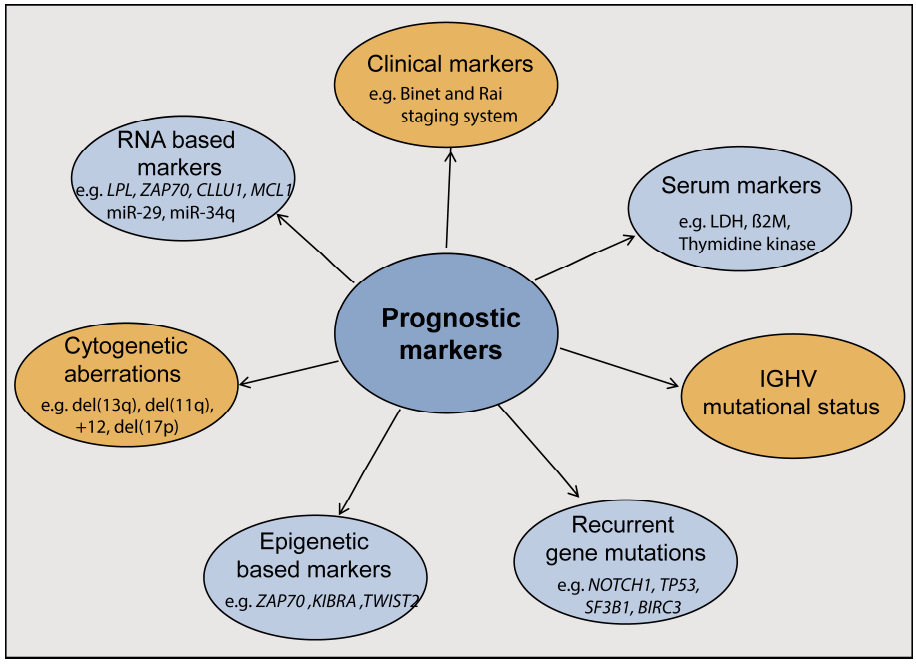

Figure 3. Various types of prognostic markers in CLL.

Considering the remarkable clinical heterogeneity observed in CLL, there is a crucial need for prognostic and predictive markers to stratify patients at an early stage of the disease and to help patient management and treatment decisions. In addition, prognostic and predictive markers may be useful in medical research as their biological function could provide further information regarding the pathogenesis of CLL and possible advancements in treatment approaches. The prognostic strength of new markers can be evaluated using different endpoints such as OS (the time span from date of diagnosis until death or last follow-up), TTFT (the time span from date of diagnosis until the date of initial treatment) or progression free survival (PFS; the time span from date of treatment until date of any sign of disease progression). Furthermore, multivariate analysis can be applied to test the prognostic capacity of the biomarker in relation to established markers.

In order to be applicable in the clinical setting, prognostic markers must fulfill certain criteria including (i) validation in independent and prospective studies; (ii) easily and reliably implemented in hospital laboratory settings; and (iii) stability over time.

IGHV gene mutational status

An important milestone in elucidating the pathobiology of CLL was the finding in 1999 that the SHM status of the IGHV genes clearly defined two

molecular subtypes of patients with divergent clinical outcome.[55,56] The subgroup of patients carrying unmutated IGHV genes (U-CLL, i.e. ≥98% identity to germline) followed an aggressive clinical course associated with shorter OS and higher risk for progressive disease, while the other subgroup of patients carrying mutated IGHV genes (M-CLL, i.e. <98% identity to germline) had a much more indolent disease course with long OS and lower risk of disease progression.[55,56] Following these two papers, many subsequent studies have confirmed the importance of IGHV gene sequence analysis in CLL with IGHV mutation status being the strongest prognostic marker in CLL.[69-71] Furthermore, U-CLL is associated with poor prognostic factors such as unfavorable genomic aberrations, resistance to therapy and recurrent mutations in certain genes (e.g. *NOTCH1*, *TP53*).[63,72,73] In contrast to genomic aberrations, the IGHV mutation status remains stable over time.[69,74,75] Finally, IGHV mutational status is included in the recently proposed international prognostic index for CLL (CLL-IPI), described below.[76]

Cytogenetic aberrations

Almost 80% of CLL patients harbor recurrent cytogenetic aberrations in certain chromosomal regions, as detected by fluorescence *in situ* hybridization (FISH), and the most common aberrations include del(13q), trisomy 12 (+12), del(11q) and del(17p), all of which are included in the diagnostic CLL-FISH panel used now a days.[57,77]

del(13q) is the most frequent genomic aberration detected, observed in roughly 55% of CLL patients.[57] As a sole aberration, it is found in 75-80% of patients with monoallelic deletion and is usually associated with a favorable prognosis.[57,78,79] In recent years, investigations into candidate genes present within 13q14 revealed two microRNAs, miR-15a/miR16-1, located in the minimal deleted region (MDR), which appear to negatively regulate expression of the anti-apoptotic gene BCL2.[80-82] However, larger 13q deletions were recently reported to correlate with a poorer outcome.[83,84]

+12, detected in 10-20% of CLL patients, is associated with an intermediate survival.[57,85] However, it was recently reported that the *NOTCH1* mutation frequency is higher in cases with +12 and such cases exhibited a worse clinical outcome.[57,59,86]

Patients with del(11q) more often have a progressive disease with inferior clinical outcome (found in 18-20% of patients).[57,85] The MDR includes the *ATM* gene with 8-30% of cases carrying biallelic aberrations (i.e. del(11q) and *ATM* mutations).[87] This implies potential involvement of other genes within the 11q-deleted region contributing to the pathobiology of 11q-

deleted CLL. One such potential player is *BIRC3,* which is frequently mutated in fludarabine-refractory CLL patients and is located near the *ATM* gene (further described below).[87-89]

del(17p) is considered to be the prognostic factor associated with the highest risk; 17p-deleted CLL patients generally show very short TTFT, poor response to current treatment regimens and dismal OS.[57,90] The deleted region on 17p encompasses the important cell cycle regulating gene, *TP53*, and a large proportion of 17p-deleted cases (~70%) also carry a *TP53* mutation on the second allele.[91] However, patients carrying only a *TP53* mutation (without del(17p)) display an equally poor clinical outcome as patients with del(17p).[92] Although a low frequency of *TP53* aberrations is seen at diagnosis, it increases considerably in relapsing patients or in patients with chemorefractory disease (>40%).[92-94] Today, screening for del(17p) and *TP53* mutations is recommended before starting any treatment regimen for CLL patients and at relapse.[95]

Using the hierarchical prognostic model proposed by Döhner *et al.* based on the aforementioned genomic aberrations, CLL patients can be classified into different risk groups in terms of TTFT and OS. More specifically, patients harboring del(17p) display the worst outcome, followed by 11q deleted patients, then cases with +12 and normal karyotype, whereas patients with isolated del(13q) patients have the longest TTFT and OS.[57]

Novel gene mutations

The advent of NGS technologies has provided a great opportunity to dissect the CLL genome at an unprecedented level and helped us to identify disease-associated genes and to study their role in the pathogenesis of CLL.[60,96-98] These techniques have also enabled ultra-deep sequencing to detect mutations at the sub-clonal level.[60,61,98,99] A number of recurrently mutated genes have been identified through recent NGS-based studies, including *NOTCH1, SF3B1, BIRC3* and *MYD88*, all of which target important signaling pathways with potential impact on CLL pathobiology.[60,61,89,96,97,100,101]

NOTCH1 is a ligand-activated transcription factor that encodes a transmembrane protein implicated in apoptosis, cell differentiation and proliferation.[63,102,103] The frequency of *NOTCH1* mutations ranges from 5-10 % in CLL and is reported to be higher in U-CLL.[54,86,104-107] The hotspot mutation consists of a 2 base-pair frameshift deletion (7544_7545delCT) in the C-terminal PEST domain, which accounts for up to 90% of mutations reported in CLL.[62] *NOTCH1*-mutated patients generally experience shorter

OS and PFS compared to wildtype patients. Furthermore, *NOTCH1* mutations are associated with advanced clinical stage, +12 and RS.[52,54,63,106]

SF3B1 is a core component of the spliceosome complex and acts as an important player in the splicing machinery.[108] *SF3B1* mutations have been reported in up to 20% of CLL patients with most mutations clustering in a hotspot region in the HEAT domain.[109] A number of independent studies reported *SF3B1* mutations association with U-CLL, advanced clinical stage, del(11q) and poor clinical outcome. [61,100,110,111] However, there is yet no clear evidence how *SF3B1* mutations contribute to the disease pathogenesis.

BIRC3 encodes a protein known as Baculoviral IAP repeat-containing protein 3 that functions as an inhibitor of the non-canonical NF-κB pathway.[89] The frequency of *BIRC3* mutations are quite low at diagnosis (<3%), but increases to more than 20% in fludarabine-refractory patients, alluding to an important role in development of chemo-refractory disease.[63,89]

MYD88 is an adaptor protein for the interleukin-1 receptor/toll-like receptor (TLR) signaling pathways. The frequency of *MYD88* mutations is relatively low in CLL (<3%).[60,63,101] The most common mutation identified involves a substitution within exon 5 (L265P), leading to constitutive activation of the NF-κB pathway and providing survival and proliferation signals to the mutant cells.[60] Although the prognostic relevance of *MYD88* is unclear, most patients carry mutated IGHV genes and have a more favorable disease profile.[101]

RNA-based prognostic markers

Microarray-based technology has enabled us to perform gene expression profiling (GEP) of various prognostic subgroups of CLL, whereby a number of genes with differential expression in different subgroups of patients were identified.[112-114] These differentially expressed genes, typically derived from comparisons of M-CLL and U-CLL, are usually been investigated both at the protein and mRNA level, in order to find suitable surrogate markers to the IGHV mutation status, previously consider a difficult marker to assess.

The most commonly analyzed RNA-based markers in CLL are lipoprotein lipase (*LPL*), zeta chain associated protein kinase 70kDa (*ZAP70*), CLL up-regulated gene 1 (*CLLU1*), myeloid cell leukemia 1 (*MCL1*) and T-cell leukemia protein 1a (*TCL1*), which will be described in the following sections. A novel RNA-based biomarker, *UGT2B17*, the focus of **paper I**, will also be discussed at the end of this section.

Lipoprotein lipase (*LPL*)

In two independent GEP studies, *LPL* was found as one of the most differentially expressed genes between U-CLL and M-CLL.[112,113] LPL plays a significant role in lipid metabolism and is normally expressed in adipose tissue, skeletal and cardiac muscles, and lactating mammary glands.[115] Since *LPL* is generally not expressed or is expressed at low levels in normal B-cells, it was an attractive marker to apply in CLL. High *LPL* expression was found to be associated with U-CLL, high-risk features (i.e. high CD38 and *ZAP70* expression and poor-risk genomic aberrations) and a poor clinical outcome.[116-120] Furthermore, our group has reported *LPL* as the strongest prognostic factor in a comparative analysis of various RNA-based markers.[121]

Zeta chain associated protein kinase 70kDa (*ZAP70*)

ZAP70 expression was one of the first markers identified through GEP of M-CLL and U-CLL, and several studies have reported that high mRNA or protein expression levels were associated with poor clinical oucome and other poor-prognostic markers in CLL such as U-CLL.[113,122-124] However, since T-cells also express ZAP70 this might influence the mRNA/protein levels in non-sorted CLL samples. The aberrant expression of *ZAP70* in U-CLL is thought to reflect a higher level of BcR signaling.[125,126] Of note, aberrant methylation of *ZAP70* is reported to be associated with poor outcome in CLL.[64,65]

CLL up-regulated gene 1 (*CLLU1*)

CLLU1 was identified through differential display (DD) technique applied to U-CLL and M-CLL that allowed identification of this novel transcript.[127] Interestingly, *CLLU1* was only expressed in CLL samples and was not detected in other hematological malignancies. High *CLLU1* expression was associated with U-CLL, advanced stage disease (Binet B/C), unfavorable genomic aberrations, and high ZAP70/CD38 expression.[128] However, the actual function of this gene is still unknown.

Myeloid cell leukemia 1 (*MCL1*)

MCL1 is a protein encoded by the *MCL1* gene belonging to the BCL2 family.[129] At the protein level, MCL1 expression was associated with other prognostic markers in CLL, such as IGHV status and CD38 expression, and was able to predict PFS and poor response to chemoimmunotherapy.[130,131] At the RNA level, high expression reportedly predicts shorter OS and TTFT in CLL and was observed in patients with partial or no response to chemotherapy.[132] However, we did not find any significant association of *MCL1* expression and clinical outcome in our aforementioned comparative study of RNA-based markers.[121]

UGT2B17

UGT2B17 is a phase-II metabolizing enzyme and a member of the UGT2B super-family which conjugates various endogenous compounds, in particular steroid hormones as well as several pharmaceutical drugs.[133,134] Through catalyzing the transfer of glucuronic acid from uridine diphosphoglucuronic acid to a variety of substrates, it detoxifies endogenous and exogenous steroid hormones and xenobiotics.[135] The *UGT2B17* gene is located on chromosome 4q13 and has mainly been investigated in solid tumors. In particular, it was found highly expressed in endometrial cancer tissue compared to normal tissue, while in prostate cancer, the UGT2B17 protein was expressed five times higher in lymph node metastasis as compared to benign control.[136,137] Recently, high mRNA expression of *UGT2B17* was correlated with high-risk CLL and associated with other established poor-prognostic markers (e.g. Binet stage, CD38 expression).[138] Moreover, the expression level increased after fludarabine treatment in the poor-responding group.[138] As this is the only report of *UGT2B17* in CLL, we investigated its prognostic relevance in a well-characterized CLL cohort in **paper I**.

Prognostic models

Though a plethora of biological and genetic markers have been proposed in CLL prognostication, their clinical applicability is still limited. However, efforts have been made during the recent years to construct prognostic models or prognostic indices that combine several clinically relevant markers, in order to stratify patients with distinct clinical outcomes.[139]

Integrated mutational and cytogenetic model

To this end, Rossi *et al.* has recently proposed a prognostic algorithm based on OS that integrates cytogenetic aberrations with recurrent gene mutations identifying four hierarchical risk groups in CLL i.e (i) high-risk group, patients harboring *TP53* and/or *BIRC3* aberrations; (ii) intermediate-risk group, patients harboring *NOTCH1* and/or *SF3B1* mutations and/or del(11q); (iii) low-risk group, patients with +12 or no cytogenetic aberrations; and, (iv) very low-risk group, patients with isolated del(13q). Furthermore, the model retained prognostic significance over time regardless of clonal evolution, making it particularly interesting for patient stratification.[59] However, in a recent study, our group did not find any difference in TTFT between the high vs. intermediate-risk group, intermediate vs. low-risk group or low vs. very low-risk group in early stage cases.[63]

CLL-international prognostic index (CLL-IPI)

More recently, a large co-operative effort has been made to develop an internationally applicable prognostic index for CLL patients (CLL-IPI) that

incorporates 5 independent factors, i.e. age, clinical stage, del(17p) and/or *TP53* mutation, IGHV mutation status and β2-microglobulin (B2M) level. This index stratified patients into four risk groups, i.e (i) low risk (score 0-1), intermediate risk (score 2-3), high risk (score 4-6) and very high risk (score 7-10) based on the weighted scoring of the independent factors.[76]

Although both these aforementioned prognostic models seem appealing, they still warrant large-scale validation in order to reach a consensus which prognostic model to apply for which setting.

Immunogenetics in CLL

BcR diversity and IG gene repertoire

As discussed in the first section of this thesis, IG gene rearrangements along with SHM and CSR enables the production of IGs with vast variability, so the probability of two healthy individuals expressing identical BcRs is almost negligible (1 in 1 trillion).[140,141] However, in recent years, immunogenetics studies revealed evidence of a restricted IGHV gene repertoire with overrepresentation of certain IGHV genes (i.e. IGHV1-69, IGHV4-34 and IGHV3-21) in CLL.[142,143] Furthermore, the SHM levels are not uniform among CLL patients utilizing different IGHV genes, eg. IGHV1-69 bear few, if any mutations, IGHV4-34 exhibit a heavy mutational load, while IGHV3-21 carry mixed mutational load.[144]

Stereotyped subset classification

Soon after the revelation of the restricted IGHV gene repertoire in CLL, it was further discovered that apparently unrelated CLL patients could carry highly similar or quasi-identical VH CDR3 sequences with shared amino acid motifs in their BcRs; referred to as "stereotyped" BcRs, providing strong evidence of antigen selection in CLL pathogenesis.[145-147] Thereafter, great effort were made in order to understand and characterize these stereotyped BcRs, which resulted in large multicenter studies with thousands of CLL patients reporting the finding that approximately 30% of all CLL patients can be assigned to a stereotyped subset based on their expression of specific stereotyped BcRs. Though more than a hundred stereotyped subsets have been identified in CLL, 19 major subsets have been characterized accounting for up to 40% of all stereotyped cases.[146] Stereotypy is observed among both M-CLL and U-CLL, however is more frequently seen in U-CLL cases. Mounting evidence suggests that patients assigned to a particular subset also share clinico-biological profiles, such as genomic aberrations,

gene expression, DNA methylation and micro-RNA (miRNA) profiles.[148-153] For instance, subset #2 (IGHV3-21/IGLV3-21), the largest subset accounting 3% of all CLL, includes both M-CLL and U-CLL patients. This subset is generally associated with an aggressive disease course with poor clinical outcome independent of IGHV mutational status.[146] Subset #1 (IGHV1/5/7/IGKV1-39) is the largest subset within U-CLL (2% of all CLL) and is associated with a particularly poor clinical outcome. Similarly, subset #8 (IGHV4-39/IGKV1(D)-39) is also associated with poor prognosis and reported to have the high risk for RS transformation.[154] In contrast, subset #4 (IGHV4-34/IGKV2-30) is the most frequent within M-CLL (1%) and associated with an especially indolent disease course that rarely is in need of treatment.[146] Notably, certain mutations are more frequent in certain subsets. For instance, subset #1 more frequently carry *TP53* and *NOTCH1* mutations, while subset #2 cases show high frequencies (45%) of *SF3B1* mutations.[151,152,155]

Antigens in CLL development

The findings of restricted IG repertoire and stereotyped BcRs above provided strong evidence for antigen involvement in disease initiation of CLL, however the exact way antigen selection causes CLL has remained largely unknown.[156] Attempts have been made to shed light on this issue, mostly based on recombinant antibodies, generated from CLL BcR IG sequences, where it is reported that U-CLL carry polyreactive BcR IGs that bind with a low affinity to molecular motifs present on apoptotic cells and bacteria, while M-CLL exhibits more restrictive antigen-binding properties.[157,158] These molecular motifs include cytoskeletal proteins vimentin, filamin B, cofilin-1 but also *Streptococcus pneumoniae* polysaccharides and oxidized low-density lipoprotein.[159]

Furthermore, antigen reactivity studies in stereotyped subsets have revealed that patients from the same subset demonstrated similar antigen reactivity profiles. Bergh *et al.* reported that subset #1 BcR IGs specifically bound to oxidized low-density lipoprotein.[160] Similarly, Chu *et al.* have shown that subset #6 BcRs bound specifically to non-muscle myosin heavy chain IIA (MYHIIA), which is an intracellular antigen associated with apoptosis.[161] Potential association between viral antigens (e.g. Epstein–Barr virus or Cytomegalovirus) and subset #4 patients has also been reported in CLL.[162] Based on these observations, it is postulated that CLL may arise from B-cells with dual function that not only maintain their ability to bind conserved bacterial epitopes but also functions as scavengers of apoptotic residues.[156]

Apart from extrinsic antigen(s), a recent study suggested that CLL BcRs can instigate cell autonomous signaling in an antigen independent manner

whereby the heavy-chain CDR3 recognizes distinct BcR epitopes.[163] More recently, another study reported that homotypic interactions between the BcR epitopes, as the basis for cell-autonomous signaling. Moreover, the internal epitopes as well as the interaction avidity were different for different subgroups of patients, which may explain the biological characteristics and the clinical features of these patient subgroups. Clinically aggressive subset #2 displayed low affinity interactions and fast-dissociating self-recognition of their BcRs, while indolent subset #4 experienced high affinity interaction and tighter self-recognition.[164]

Cell of origin

Over the last decades, several hypotheses have been put forth in order to define the candidate cell(s) of origin of CLL, however this still remains a matter of debate. Earlier studies suggested naïve $CD5^+$ B-cells as the cell of origin of CLL. Seen to the SHM differences between U-CLL and M-CLL this hypothesis was replaced with the "2-cell model", postulating the derivation of U-CLL from naïve B-cells while M-CLL, conversely, arising from antigen experienced memory B-cells.[55,56,165]

From GEP studies, it became evident that U-CLL and M-CLL had a common gene expression signature with few differences resembling the profile of antigen experienced B-cells, hence questioning that U-CLL derive from naïve B-cells.[112] In a more recent study, based on transcriptome analysis of normal B-cell subpopulations and CLL, a novel subset of $CD5^+$ post-GC B-cells that co-expressed CD27 and carried mutated IGHV genes was identified. This study suggested that the $CD5^+/CD27^+$ IGHV-M post-GC B-cell population gives rise to M-CLL, while the CD5+/CD27- would instead give rise to U-CLL.[166]

The characterization of the epigenomic landscape in CLL and normal B-cell subpopulations added further complexity to this line of research. In a recent large scale study Kulis *et al.* reported that the methylation profile of U-CLL was similar to that of naïve B-cells while the M-CLL profiles were akin to memory B-cells.[167] In another similar study, the authors postulated that the disease heterogeneity, based on methylation profiling, points to the origin from a continuum of maturation states manifested by the normal B-cell developmental stages.[168]

Although all aforementioned studies tried to solve the complex issue of the cell of origin of CLL, they have not yet been able to provide an exact identity of the cell(s) of origin.

Role of DNA methylation in CLL

Epigenetic changes, such as DNA methylation and histone modification, play a significant role in regulating gene expression. Epigenetic changes are not only implicated in regulating normal function and development together with genomic imprinting, but have also critical roles in many diseases.[169-171] It is now well-established that aberrant methylation is a common feature of many human diseases including cancer. More specifically, DNA methylation, which is a heritable epigenetic mark involving the addition of methyl group (-CH3) to the fifth carbon of a cytosine ring within CpG dinucleotides has been extensively studied in many human cancers including leukemia and lymphoma.

In normal B-cell development widespread hypomethylation takes place during differentiation from naïve B-cells to non-class switched memory B-cells (ncs MBC) or to class switched memory B-cells (cs MBC).[167] In CLL, the genome is globally hypomethylated compared to normal B-cells and recent epigenomic studies of normal B-cell subpopulations and CLL cells suggest massive hypomethylation mainly targeting gene body and enhancer regions.[167,172] Another study reported a similar finding of hypomethylation targeting enhancers and transcription factor binding motifs, while hypermethylation mainly involved transcribed genomic regions.[168]

A number of studies have reported on the role of altered epigenetic changes such as promoter hypermethylation in tumor initiation and disease progression and correlated findings with clinical outcome in CLL.[173] For example, aberrant methylation of tumor suppressor genes (e.g. *SERP1, IDH4*), genes implicated in apoptosis (e.g. *DAPK*), prognostic genes (e.g. *ZAP70, CD38, LPL, CLLU1*) and other disease-associated genes (e.g. *KIBRA, HOXA4, BTG4* and *TWIST2*), have been correlated with disease outcome in CLL.[64-66,174-179] Well-known prognostic genes, such as *ZAP70* and *LPL*, are reportedly regulated through methylation. More specifically, reports suggest a hypomethylated state of the *LPL* gene is linked to the high expression levels seen in U-CLL patients.[174,179] Similarly, hypomethylation of the highly conserved intronic region of *ZAP70* in U-CLL was reported to be responsible for the high expression level.[180] Furthermore, methylation of *HOXA4* and *KIBRA* was associated with poor-prognostic markers such as unmutated IGHV genes and high CD38 expression.[66-68] On the other hand, a well-known transcription factor, *TWIST2*, a silencer of p53, was reportedly hypermethylated in M-CLL patients and associated with good prognosis.[178]

With the advent of high-throughput sequencing and array-based technology, such as whole-genome bisulfite sequencing (WGBS) and Infinium HumanMethylation450 BeadChip arrays (450K), distinct methylation

profiles were unraveled in prognostic subgroups of CLL and underlying epigenetic events contributing to CLL development and pathogenesis were identified.[167,168,181] Using 450K arrays, our group identified the distinct methylation profiles within U-CLL and M-CLL and further reported a relatively stable methylation pattern overtime and with similar profiles in the resting and proliferative compartments.[181] Furthermore, our study confirmed the global hypomethylation of the CLL genome as compared to normal controls derived from peripheral blood mononuclear cells and sorted B-cells.[181] Moreover, we have investigated the global DNA methylation profiles of three prototypic CLL subsets, i.e. subsets #1, #2 and #4, using 27K arrays that interrogate 27 578 CpG sites covering 14 495 genes, and reported subset-biased DNA methylation profiles. The differentially methylated genes identified from the subset comparison were enriched in immune response and Toll-like receptor signaling.[149]

As mentioned, efforts have recently been made to identify the disease specific epigenetic changes in CLL by investigating the CLL methylome in the context of normal B-cell differentiation. Two such studies reported that the epigenetic changes identified during the normal B-cell differentiation process were highly overlapping with those found in CLL, highlighting that changes observed in CLL are for the most part not disease-specific but instead reflect the B-cell maturation stage at transformation.[168,182] In addition, the existence of intra-tumor methylation heterogeneity in CLL also correlated with genetic subclonal complexity and further supported the co-evolution of novel methylation patterns and genetic alterations in CLL.[183]

Epigenetic classification and risk-stratification in CLL

According to a recent report by Kulis *et al.* DNA methylation profiling was able to distinguish three clinico-biological subgroups of CLL.[167] Using WGBS and 450K methylation arrays, the authors compared the methylation signature of normal B-cell subpopulations (NBCs, cs MBCs and ncs MBCs) with CLL cells and proposed the existence of three subgroups which they termed as memory-like CLL (m-CLL), naïve-like CLL (n-CLL) and a third intermediate CLL subgroup (i-CLL). Importantly, the three epigenetic subgroups showed distinct clinicobiological profiles. More specifically, m-CLL exhibits a favorable disease outcome compared to the other two subgroups. In multivariate analysis, the methylation signature remained as an independent prognostic marker together with CD38 expression and lactate dehydrogenase (LDH) level.[167] The existence of the three epigenetic subgroups with markedly different clinical outcome was confirmed by another study by Oakes *et al*, where they instead named the subgroups as low programmed CLL (LP-CLL), intermediate programmed CLL (ip-CLL) and high programmed CLL (HP-CLL).[168]

Taking into considerations these observations, DNA methylation signatures of CLL subtypes might correlate with the epigenetic imprint of their respective cell of origin and hence affect the clinico-biological profile of the disease. Therefore, implementation of methylation signatures into patient stratification and subgrouping of patients seems to be plausible. These reports emphasized the presence of a novel third group, however it was not clear why this group had an intermediate methylation profile.

Five CpG signature

Following this study, Queiros *et al*. proposed a prognostic epigenetic signature consisting of 5 CpG sites that more readily identifies m-CLL, n-CLL and i-CLL.[184] They applied a support vector machine approach to build a prediction model that identified these 5 methylation marks, which classified CLL patients into the three epigenetic subgroups with high accuracy (98.7%). These 5 CpGs corresponds to the promoter region of *SCARF1* (cg00869668), the gene body of *B3GNTL1* (cg11472422), *CTBP2* (cg17014214) and *TNF* (cg09637172), and an intergenic region on chromosome 14 (cg03462096). Genetic and functional characteristics of the 5 CpGs are presented in Table 1. The methylation status of the 5 CpGs was assessed using bisulfite pyrosequencing (discussed later in methods section). They investigated the efficacy of the epigenetic signature in 211 CLL patients with subsequent validation in an independent cohort of 97 patients. Furthermore, the stability of the epigenetic signature was evaluated by analyzing sequential samples from 27 patients revealing stabile DNA methylation profiles overtime.[184] This signature was the focus of **paper II.**

Table 1. Genetic and functional features of the five CpGs.

Characteristics	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5
Probe ID	Cg03462096	Cg09637172	Cg11472422	Cg17014214	Cg00869668
Gene name		*TNF*	*B3GNTL1*	*CTBP2*	*SCARF1*
Gene related region	Intergenic region	Gene body	Gene body	Gene body	Promoter region
Chromosome	14	6	17	10	17
Function/implication in cancer		cytokine secreted by macrophages regulation of cell proliferation, differentiation and apoptosis[185] plasma level of TNF is higher in CLL compared to healthy control[186] higher level associated with advanced Rai and Binet stage[186] implicated in tumor suppression[187]	described as putative glycosyl-transferase involved in transferring glycosyl groups[188]	acts as transcriptional corepressors in mammals[189] *CTBP2* has been reported to promote cancer cell migration in mammals[190] overexpression has been correlated to poor prognosis in breast and prostate cancer[191,192] knocking down of the gene leads to reduction in cell proliferation and induced apoptosis[192]	regulates the uptake of chemically modified low density lipoproteins expressed highly in some cancers including lymphoma

Micro-RNA deregulation in CLL

miRNAs are a group of small non-coding RNAs, which are usually ~21 nucleotide long, that play a significant role in a number of biological processes including regulation of gene expression by targeting the mRNA or by inhibiting their translation. However, their deregulation can lead to aberrant expression of many genes, which in turn promotes tumor progression. Several studies have shown miRNA deregulation in CLL as an important event in the pathobiology of the disease.[193-198] Different miRNA are reported to distinguish normal B-cells from CLL, segregate aggressive CLL from indolent forms, predict disease progression, as well as distinguish

refractory from responding patients.[199-201] For instance, high miR-21 expression in poor-prognostic CLL patients is associated with dismal outcome compared to patients with low expression.[202] Similarly, high miR-155 levels were reported in aggressive CLL cases.[203]

Though the exact mechanisms behind miRNA deregulation are not yet properly understood, epigenetic mechanisms such as aberrant promoter methylation are suggested to play a role in CLL.[204] Recently, our group reported aberrant hypermethylation of miR34b/c, located in the commonly deleted region on chromosome 11q, in ~48% of CLL cases where the expression levels were inversely correlated with the methylation levels.[205] Promoter hypermethylation of miR-9 family members were also reported in CLL, where overexpression of miR-9 resulted in reduced cell proliferation and enhanced apoptosis together with downregulation of the NF-κB pathway, adding credence to its tumor suppressor role in CLL.[206] Furthermore, aberrant methylation of the miR-708 promoter seen in CLL which in turn leads to enhanced activation of the NF-κB pathway.[207]

More recently, high expression of miR-26A was reported to associate with advanced Binet stage, inferior TTFT and *TP53* aberrations in CLL, while its inhibition resulted in increased apoptosis in primary CLL cells, suggesting a tumor suppressor role in CLL.[208] Moreover, restoration of miR26A expression using the BET bromodomain inhibitor, JQ1, or the EZH2 inhibitor, DZNep, led to suppressed growth in aggressive lymphoma cells. Combined treatment of both aforementioned drugs rendered disruption of MYC activation that resulted in a greater restoration of miR26A.[209] This study further reported that *MYC* recruited *EZH2* to the miR-26A promoter that resulted in repression of miR-26a expression in aggressive lymphoma cell lines and primary lymphoma cells.[209]

Mantle Cell lymphoma

Background

MCL is a mature B-cell malignancy characterized by the B-cell markers CD19, CD20, CD22, CD79, with co-expression of the T-cell antigens CD5 and CD43.[3,210] It comprises ~ 6% of all non-Hodgkin lymphomas (NHLs), is considered as a disease of the elderly, with a median age of diagnosis of 68 years, and has a male predominance (3:1).[211,212] The clinical course is usually aggressive with short OS (median survival 3-5 years) and frequent disease relapses. The patients often present with advanced stage and a disseminated disease.[213] However, MCL patients with a more indolent disease and longer survival times (>7-10 years) have also been identified. This indolent subgroup of patients usually carries hypermutated IGHV genes, a non-complex karyotype and displays low SOX11 expression.[214,215]

The translocation t(11;14)(q13,q32) is the major genetic hallmark of MCL, found in most but not all cases and leads to constitutive overexpression of cyclin D1 (*CCND1*).[216,217] Interestingly, a small subset (<5%) of cases lack this translocations and are negative for *CCND1,* instead they express *CCND2* or *CCND3*.[213,218] However, it has been reported that this aberration cannot alone induce MCL in mice models.[213,219] Thus far, a high number of secondary genetic alterations have been reported in MCL and implicated in various pathways and processes, such as cell proliferation, DNA repair and apoptosis.[213] *ATM* mutations are reported in 42-55% of cases, usually associated with del(11q), and are the most common secondary genetic alterations in MCL.[220-222] *TP53* is commonly mutated in MCL patients and found in 19 to 28% of cases.[221,223,224] Recurrent genomic aberrations involving gains of 3q, 8q and 12q and losses of 9p, 9q, and 17p are also frequent in MCL.[213] NGS has expanded the knowledge of genes and pathways involved in the pathogenesis of MCL. For instance, novel mutations involving i) activating mutations in *NOTCH1*, detected in ~10% of cases[220,225] ii) mutations in chromatin modifiers, such as *WHSC1* in 10% cases and *MEF2B* in 3% cases[220,221] and iii) mutations affecting the NF-κB pathway, detected in 10%-15% of cases, including *BIRC3* (mutated in 6-10% cases).[220,221,226]

The vast majority of patients with MCL require treatment upon diagnosis. There is a great diversity regarding the first line therapeutic choices for MCL patients. In general young and fit patients are candidates for treatment with either Hyper-CVAD (or similar regimens) alone or combined chemoimmunotherapy followed by high-dose chemotherapy and autologous stem cell transplantation.[227,228] For older or more fragile patients combined chemoimmunotherapy is the treatment of choice.[227,228] Unfortunately, a great number of MCL patients will eventually relapse. Recently, a number of novel agents such as proteasome, mTOR, histone deacetylase and more importantly BcR signaling inhibitors have shown promising results.[229-232]

Prognostic parameters in MCL

The percentage of Ki-67-positive tumor cells, i.e. the Ki-67 index, measuring the tumor cell proliferation rate, is considered as one of the most powerful single prognostic parameter for OS in MCL. Cases with high proliferative rate are associated with significantly poorer clinical outcome compared to the cases with low proliferative rate.[233,234] The proliferation gene expression signature model proposed by Rosenwald *et al*. involves 20 genes related to cell proliferation (i.e. mitosis, cell cycle and DNA replication) and is able to predict OS in a reproducible manner.[233]

Recently, a MCL-specific prognostic index, the MCL-International Prognostic Index (MIPI), designed based on four prognostic factors, i.e. age, performance status, LDH levels, and white blood cell (WBC) counts; resulted in improved patient stratification.[235,236] In addition, Ki-67 was reported to improve risk stratification when incorporated with the MIPI.[235,237]

TP53 and *NOTCH1* mutations have been associated with poor clinical outcome in MCL.[220,221,223-225] Most of the other recurrent genomic aberrations did not hold in multivariate analysis and may hence merely reflect tumor proliferation and clinical aggressiveness.

Immunogenetics in MCL

The IGHV mutational rate in MCL is generally lower than CLL and the mutational status does not correlate to clinical outcome, as in CLL.[238] A recent immunogenetic study involving 807 MCL patients reported that a large proportion (~70%) of patients showed minimally mutated and borderline mutated IGHV genes, while up to 30% of patients carried unmutated IGHV gene.[239] Furthermore, biased usage of certain IGHV genes such as IGHV3-21, IGHV3-23 and IGHV4-34 was observed,[239-242] and 10%

of cases displayed MCL-specific stereotyped VH CDR3 sequences, which strongly suggests antigen involvement in disease pathogenesis also in MCL.[239] Hence, the stereotyped BcRs found in MCL are different from those reported in CLL, therefore indicating distinct antigenic selection in MCL.[146]

microRNA deregulation in MCL

In recent years, deregulations of miRNA have been implicated in MCL pathogenesis.[243-245] GEP and miRNA profiling in MCL has revealed a role of miRNAs in regulating key pathways such as the CD40, mitogen-activated protein kinase and NF-κB pathways.[243] miRNA profiling in *CCND1* positive and negative MCL cases, compared with other aggressive lymphoma and normal B-cells, revealed a 19 miRNA classifier that could discriminate MCL cases from other aggressive lymphomas.[246] Another study reported that downregulation of miR-29 family members (miR-29A, miR-29B, and miR-29C) was associated with shorter OS in MCL.[245]

Limited studies exist regarding the epigenetic silencing of tumor suppressor miRNAs in MCL. Recently, hypermethylation of tumor suppressor miRNA-155-3p was observed in MCL and other NHL subtypes which lead to its downregulation and subsequent upregulation of LT-β, an upstream activator of non-canonical NF-κB pathway.[247] As discussed earlier, restoration of miR-26A expression results in suppressed lymphoma growth in aggressive lymphoma cells and primary lymphoma cells including MCL, which further supports the tumor suppressor role of miR26A1.[209] We have investigated the functional role and the mechanism behind miR26A1 deregulation in CLL and MCL in **paper III** in this thesis.

Techniques to assess DNA methylation

Over the past two decades, great advancement has been achieved in technologies for analyzing DNA methylation patterns both at global and local (single CpG) levels. As such, a plentitude of methylation analysis techniques are readily available today, revolutionizing our perspective of the methylome of both normal B-cell subpopulations and various lymphoid malignancies.

Bisulfite conversion is the basis for most of the DNA methylation analysis techniques and it involves the treatment of genomic DNA with sodium bisulfite, converting all unmethylated cytosines into uracil while all methylated cytosines remain unchanged.[248] Following conversion, the methylation status of the DNA can be determined by techniques such as methylation specific PCR, microarray, WGBS and pyrosequencing. Microarrays and pyrosequencing were the techniques predominately used in this thesis.

Microarray-based methods provide a more global view of the epigenome, and interrogate the methylation status at predefined CpG sites hybridized to the arrays. The 450K BeadChip array designed by Illumina allows the interrogation of >485,000 CpG sites across the whole genome and includes CpG sites present in the CpG islands/shores/shelves/open sea, non-coding RNA, CpGs surrounding the transcription start sites (-200 bp to -1,500 bp and 5'-UTRs) of coding genes, gene bodies and 3'-UTRs as well as CpGs in intergenic regions.[249] More recently, Illumina launched the 850K array that allows interrogation of over 850,000 CpG sites, also including 333,265 CpG sites located in enhancer regions.[250] The aforementioned arrays are based on bisulfite conversion of DNA and the main principle behind the analysis is single base primer extension and ligation using linker primers (locus specific oligonucleotide primers) attached to two different bead types corresponding to methylated and unmethylated CpG locus, differentiating unconverted methylated DNA from converted unmethylated DNA. The fluorescent light intensity emitted from the single base primer extension phase is measured and a DNA methylation value called “beta value” is assigned for each CpG locus. The beta value ranges from 0 to 1, where 0 corresponds to completely unmethylated while 1 denotes completely methylated CpGs.

Though microarrays render genome wide information regarding the methylation status of a large number of CpG sites, they are limited to those that are hybridized to the arrays. NGS technologies such as WGBS can overcome this problem by assessing the DNA methylation status with a single base pair resolution and with genome wide coverage (~10^8 CpG sites per sample genome). It combines bisulfite conversion of DNA with high-throughput sequencing to provide a complete overview of the methylome

Nevertheless, findings from these aforementioned techniques often need to be verified by more quantitative methods such as pyrosequencing. Pyrosequencing is a sequencing technique based on the principle "sequencing by synthesis" that provides a quantitative methylation analysis of the region of interest containing single or multiple CpGs.[251-253] The pyrosequencing reaction is a multistep reaction based on the transformation of pyrophosphate (PPi) into detectable light that is proportional to the number of nucleotides incorporated.[254] Like most other DNA methylation analysis methods, the pyrosequencing method is also based on bisulfite converted DNA. Briefly, after the bisulfite conversion of DNA, the specific region of interest is PCR amplified, denatured to single stranded DNA and captured by streptavidin beads, followed by annealing with the biotinylated sequencing primer. Sequencing continues with the sequential addition of nucleotides by DNA-polymerase and PPi is released which is converted into ATP by the enzyme ATP sulfurylase. This reaction is followed by the emission of a light signal catalyzed by the enzyme luciferase and the emitted light signal can be recorded by a charge-coupled device camera (CCD) and can be seen as a peak in the form of a pyrogram. Furthermore, the height of each peak in the pyrogram corresponds to the percentage of methylation at each CpG site and it is determined from the T and C ratio at that particular CpG.[253-255] A schematic representation of the pyrosequencing technique is presented in Figure 4.

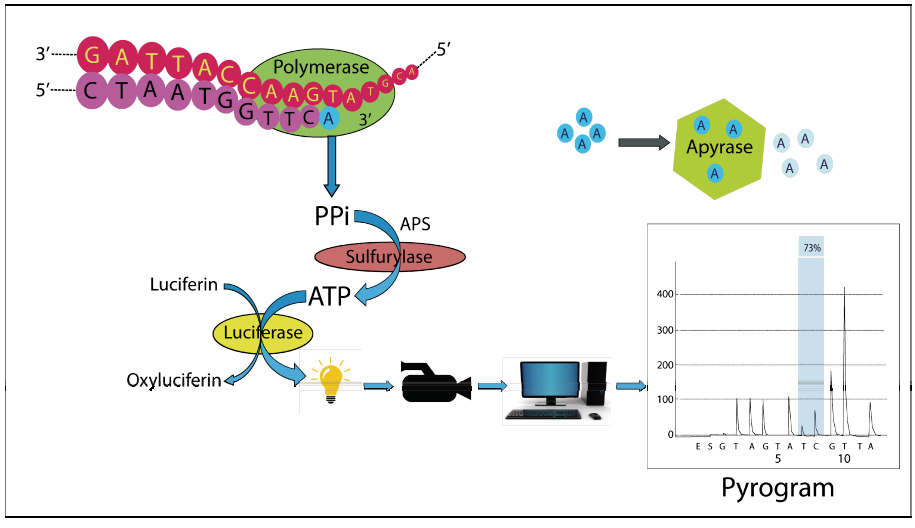

Figure 4. Schematic representation of the pyrosequencing technique. Pyrosequencing is initiated by the incorporation of dNTP to the complementary DNA strand, one at a time and the reaction is catalyzed by the DNA polymerase enzyme. The released PPi is then converted to ATP by the ATP sulfurylase enzyme in presence of the substrate adenosine 5′ phosphosulfate (APS) and visible light is generated due to the conversion of luciferin to oxyluciferin catalyzed by the enzyme luciferase. The light is then detected by a CCD camera and the resulting peaks can be seen in the form of a pyrogram that corresponds to the number of nucleotides incorporated. Unincorporated nucleotides and excess ATP are degraded by the enzyme apyrase. The percentage of methylation is presented in a box above the pyrogram and it ranges from "0-100%", where 0% = completely unmethylated and 100% = completely methylated.

Present investigations

Thesis aims

The main aim of this thesis was to identify and validate novel RNA and epigenetic based prognostic markers in mature B-cell malignancies, particularly in CLL and MCL, and to explore the global DNA methylation landscape in major stereotyped CLL subsets.

More specifically, the aims were as follows:

I- To validate the prognostic relevance of a recently proposed RNA based marker, *UGT2B17,* in a large population-based CLL cohort. We assessed the expression level of *UGT2B17* in 253 patients using real-time quantitative PCR (RQ-PCR) analysis.

II- To validate and further investigate the prognostic significance of the recently proposed epigenetic 5 CpG signature in a well-characterized CLL cohort including the clinically aggressive stereotyped subset #2. We assessed the methylation status of these 5 CpG sites in 135 patients using pyrosequencing.

III- To gain insight into the functional role and the mechanism behind deregulation of miR26A1 in CLL and MCL. We investigated the methylation status and expression level of miR26A1 in well-characterized CLL (n=70) and MCL (n=65) patient cohorts using RQ-PCR and pyrosequencing.

IV- To investigate the DNA methylation profiles of major stereotyped subsets of CLL in the context of non-subset CLL and sorted normal B-cell populations at different stages of B-cell differentiation. We applied 450K methylation arrays to 176 CLL cases assigned to one of 8 major stereotyped subsets (#1-8), non-subset CLL (n=325) and normal B-cell subpopulations (n=74).

Patients and Methods

Patients

All CLL patients were diagnosed according to the iwCLL guidelines and MCL patients were diagnosed according to the WHO classification.[3,256] Informed consent was collected from all patients according to the Declaration of Helsinki with local ethics review committees for all collaborating institutions granting ethical approval. For CLL, peripheral blood (PB) was the main source of material and for MCL; samples were obtained from lymph nodes. The tumor cell percentage was >90% in paper **I, II** and **IV**, and >70% in paper **III**. In **paper I**, 253 CLL patients from the Scandinavian population-based case-control study known as SCALE (Scandinavian Lymphoma Etiology) were investigated. **In paper II**, 135 CLL patients from the aforementioned cohort were included. In **paper III**, 70 CLL cases obtained from the Biobank at Uppsala University Hospital in Sweden and Brno University Hospital in the Czech Republic were included. Two different MCL cohorts were investigated: i) sorted lymph node samples from 24 MCL cases were obtained from the Biobank at Karolinska University Hospital in Sweden and analyzed using 450K arrays, and ii) a validation cohort that included 38 lymph node samples collected from the Department of Molecular Medicine at University of Pavia in Italy was investigated. In **paper IV**, 176 CLL patients assigned to one of eight major stereotyped subsets (#1-8) collected from collaborating institutions in the Czech Republic, Greece, France, Italy, the Netherlands, Sweden, Spain and the United Kingdom were included. A second non-subset cohort consisting of 325 Spanish CLL patients was investigated as reference. All samples were sorted by either fluorescence-activated cell sorting (FACS) or magnetic bead sorting. Furthermore, 10 different sorted normal B-cell subpopulations isolated from peripheral blood, bone marrow and tonsils of healthy individuals were analyzed for comparison.

Methods

RQ-PCR analysis

In **papers I** and **III**, expression levels of the investigated genes were quantified using RQ-PCR. Total RNA was extracted using RNA easy Mini kit (Qiagen, Hilden, Germany) or miRNeasy Mini Kit (Qiagen, Hilden, Germany). RNA quality was assessed using the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and concentration was measured using Nanodrop 2000 (Thermo Fisher Scientific, Wilmington,USA) or Qubit fluorometric technology (Life Technologies, Waltham, MA,USA). Reverse transcription was performed using the M-MLV reverse transcriptase Kit (Invitrogen, Carlsbad, CA) and random hexamers (Thermo Scientific, Pittsburgh, PA) or TaqMan micro RNA transcription kit (Applied Biosystems, Foster City, USA) according to the manufacturer's protocol. RQ-PCR analysis was performed using SYBR Green master mix (Thermo Scientific) or TaqMan gene expression assay for miR-26A1 (Applied Biosystems, Foster City, USA) and reactions were run on a Stratagene MX 3005p instrument (Agilent Technologies, Santa Clara, CA) or 7900HT Fast Real-time PCR System instrument (Applied Biosystems, Warrington, UK). In **paper I**, beta-actin was used as endogenous control while in **paper III**, RNU6B was used.

Pyrosequencing

In **papers II** and **III**, pyrosequencing was applied to assess the methylation level of the 5 CpGs included in the epigenetic signature and the miR26A1 target region (127 bp) containing 4 CpG sites including the CpG site (ILMNID: cg26054057) that was found to be differentially methylated in primary CLL and MCL samples using 450K arrays.

First, bisulfite conversion of DNA was performed using the EZ DNA Methylation-GoldTM Kit (Zymo Research Corporation, CA, USA) according to manufacturer's protocol and converted DNA was subjected to PCR amplification using a forward and 5´ biotinylated reverse primer pair. Primers were designed using the PyroMark^{TM} software. Amplified PCR products were immobilized to streptavidin sepharose beads, denatured and annealed to a sequencing primer followed by methylation analysis performed on the PyroMarkTM Q24 instrument (QIAGEN, Sweden).

Epigenetic classification

In **paper II**, methylation values of the 5 CpGs were subjected to epigenetic classification using the support vector machine (SVM) model and the R software environment as previously described.[184] DNA methylation values of each CLL cases to be classified was uploaded into the R script and then the SVM function was run in order to classify the cases.

In-vitro functional characterization of miR26A

All methods described below refer to **paper III**, where we investigated the role of miR26A1 silencing using both CLL and MCL cell lines.

Cell culture- Two CLL-derived (HG3 and MEC1) and two MCL-derived (Z138 and GRANTA 519) cell lines were used.

Transient transfections- miR26A1 oligo (has-miR-26a1) and miR mimic negative (mirVana miRNA mimic negative control) were purchased from Life Technologies. Transient transfections of CLL and MCL cell lines were carried out on an Amaxa Nucleofection Device II (Nucleofector 2b device, Lonza group AG) according to the manufacturer's instructions.

DNA methyl inhibitor treatment- CLL and MCL cell lines were treated with different concentration of DNA methyl inhibitor 5'-za-2'-deoxycytidine (DAC). In order to analyze miR26A1 methylation and expression level, cells were harvested and DNA and RNA were extracted.

Western blot analysis- Nuclear extracts of transfected CLL and MCL cell lines were extracted using the NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Scientific) according to manufacturer's protocol and lysates were run on sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel and then transferred to nitrocellulose membranes. Membranes were incubated with primary and secondary antibodies, followed by washing with phosphate-buffered saline according to standard protocol. Immunoreactive proteins were visualized with SuperSignal West Dura Extended Duration Substrate (Thermo Scientific) on the ChemiDoc XRSC (Bio-Rad) instrument.

FACS analysis- The FITC Annexin V Apoptosis Detection Kit I (BD Biosciences) was used to measure apoptosis of transfected cells according to the manufacturer's protocol followed by flow cytometry analysis using a BD FACSAria cell sorter and analyzed using the FACSDiva version 6 software (BD Biosciences).

Methylation array analysis

In **papers III** and **IV**, the HumanMethylation450 Beadchip array (Illumina, San Diego, USA) was applied to study the global DNA methylation profiles in MCL, stereotyped subset and non-subset cases as well as a series of normal B-cell subpopulations spanning the different stages of B-cell development. All DNA samples were subjected to bi-sulfite conversion according to standard EZ DNA Methylation-GoldTM Kit (Zymo Research Corporation, CA, USA) and converted DNA was hybridized to the HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) according to manufacturers' protocols. Bioinformatic analysis was conducted using the "R" based Illumina Methylation Analyzer (IMA) or minfi packages as previously described.[181,182,257] A "beta value" was assigned to each detected CpG which is defined as the ratio of fluorescent signals from the methylated alleles to the sum of the signals from the methylated and unmethylated alleles and ranges from 0 (completely unmethylated) to 1 (completely methylated).

Raw methylation array data was quantile normalized, arcsin transformed and an empirical Bayes moderated t-test was applied in order to identify the differentially methylated sites between CLL subsets. The p-values were adjusted using the Benjamini and Hochberg method and a level of $P < 0.05$ was taken as cut-off. In order to detect CpG sites with a greater methylation difference among subsets, an additional filter using an average geometric difference of 0.30 was applied. Unsupervised principal component analysis was carried out to assess global methylation pattern in stereotyped subsets and non-subset CLL in the background of normal B-cell populations. Supervised hierarchical clustering analysis was performed to assess the methylation differences among the investigated subsets. Gene ontology (GO) enrichment analysis and functional annotation clustering of differentially methylated genes was done using Database for Annotation, Visualization and Integrated Discovery (DAVID) tools version 6.7.

Statistical analysis

In **papers I** and **III** receiver operating characteristics (ROC) curve analysis was applied to determine the expression and methylation cutoff value and patients were classified as low or high expressing or low or high methylated cases. Throughout **papers I-IV**, bivariate comparisons were made using t-test while one-way ANOVA and $\chi 2$ tests were used to assess differences among multiple data sets. In **papers I-III**, Kaplan-Meier survival analysis was carried out to construct survival curves for OS and TTFT. OS was calculated from date of diagnosis until date of death or last follow-up, while TTFT was calculated from date of diagnosis until the date of initial

treatment. Statistical differences in OS and TTFT between studied subgroups were assessed using the Log-rank test. Multivariate Cox-regression analysis was employed in order to test the ability of a given prognostic factor to independently predict clinical outcome, taking into account the effect of the other remaining factors. Correlations between expression levels were assessed using Spearman's rank correlation co-efficient. In **paper II**, the Genesis software was used to perform hierarchal clustering analysis and to generate heatmaps. Statistical significance level was defined as a p-value <0.05. All statistical analyses were carried out using Statistica Software 12.0 (Stat Soft, Tulsa, OK).

Result and Discussion

Paper I: *UGT2B17* expression: a novel prognostic marker within IGHV-mutated CLL?

Based on a recent publication suggesting *UGT2B17* as a potential prognostic marker in CLL, we decided to investigate the prognostic significance of *UGT2B17* mRNA expression in a large well-characterized cohort (n=253) using RQ-PCR analysis.[138,258] We also investigated the stability of *UGT2B17* by studying its expression level overtime in 91 cases.

As defined by ROC curve analysis and the median survival, a total of 117 patients (46%) showed high *UGT2B17* expression and the remaining 136 had low expression or no detectable expression (54%). We observed a significant association between high *UGT2B17* expression and other established poor-prognostic markers, such as U-CLL and advanced clinical stage (Binet B/C) as compared to cases with low expression ($p<0.001$ each). Moreover, in our study we found higher *UGT2B17* expression to be significantly correlated with del(11q) and del(17p) ($p=0.005$ and $p=0.029$, respectively), whereas a negative correlation of high *UGT2B17* expression and del(17p) and no association with del(11q) was reported in the previous study.[138]

Furthermore, high expression of *UGT2B17* correlated with shorter TTFT and OS ($p<0.001$ each) in the entire cohort and also in specific subgroups of patients, such as Binet stage A and M-CLL patients. In multivariate analysis, together with other established markers such as IGHV mutational status and unfavorable genetics [*TP53* abnormalities, i.e. del(17p) and/or *TP53* mutations, *NOTCH1* and *SF3B1* mutations], *UGT2B17* expression remained significant as an independent marker both for OS and TTFT.

As a prime finding from our study, we observed that 30% of M-CLL cases, which were negative for CD38 expression and devoid of any poor-prognostic genetic aberration, and thus expected to have a favorable prognosis, instead exhibited high *UGT2B17* expression and displayed a poor clinical outcome. Hence, *UGT2B17* mRNA expression analysis identified a subgroup of patients within M-CLL with progressive disease for which no established prognostic marker has been found until now (Figure 5A).

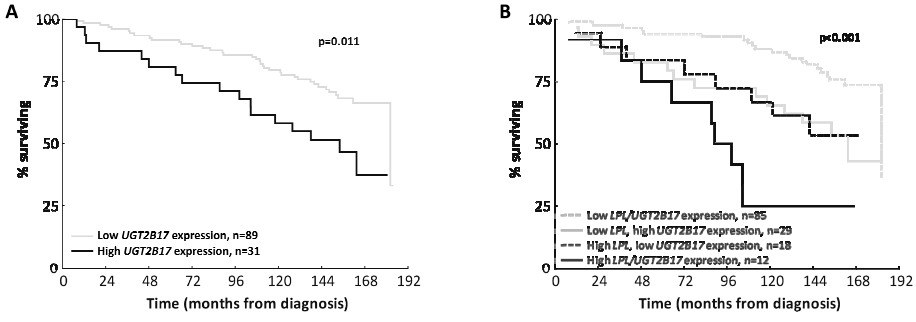

Figure 5. Prognostic impact of high *UGT2B17* expression on (A) OS for M-CLL cases negative for poor-prognostic genomic aberrations [del(17p), del(11q) and trisomy 12], novel mutations (*SF3B1, NOTCH1* and *TP53*) and CD38 expression; (B) Impact of *UGT2B17* and LPL co-expression on OS in M-CLL cases.

Furthermore, we investigated another RNA marker, *LPL*, in an effort to integrate both markers to identify cases with high risk of disease progression. From this analysis, we observed that patients with high expression of both genes exhibited the lowest median survival (7.4 years) compared to cases with low expression of both genes (15.2 years), whereas patients with isolated expression of either gene displayed an intermediate survival (Figure 5B). Therefore, incorporation of *LPL* with *UGT2B17* expression appears to further stratify patients with poor outcome in an otherwise favorable group of patients such as M-CLL.

In order to adopt RNA-based markers such as *UGT2B17* into everyday clinical practice our findings first need to be validated in larger patient series, in particular in prospective studies. In addition, investigation of the biological background behind increased *UGT2B17* expression level needs to be resolved. Nevertheless, in our overtime analysis, we observed that *UGT2B17* expression remained stable over several years (range, 5.0-8.1 years), which highlights the potential utility of *UGT2B17* as a suitable prognostic marker in CLL.

Paper II: Prognostic impact of epigenetic classification in CLL

Based on the similarity of the methylation signatures to normal B-cell populations, an epigenetic classification of CLL was recently proposed that subdivided CLL patients into three subgroups, named as m-CLL, n-CLL and i-CLL[167]. In a subsequent study, the same group proposed a 5 CpG signature that was able to classify the patients into the aforementioned epigenetic subgroups with great accuracy.[184] In this study, we investigated the prognostic significance of the proposed 5 CpG model in a well-characterized cohort (n=135), including both U-CLL and M-CLL, but also clinically aggressive stereotyped subset #2 patients (mixed IGHV mutation status), and studied its association with other established prognostic markers. The percentage of DNA methylation of 5 CpGs were assessed by pyrosequencing and samples were subjected to epigenetic classification using the SVM model as previously described.[184]

Fifty-seven samples (42.2%) were classified as m-CLL, 29 samples (21.5%) as i-CLL, while 49 samples (36.3%) were classified as n-CLL. Interestingly, as a novel finding, we observed that 11/12 subset #2 cases were classified as i-CLL (the remaining case was classified as m-CLL). Similar to the previous report, we observed an enrichment of del(11q) in n-CLL and i-CLL cases (29.8% and 29.6%, respectively). i-CLL cases tended to have a higher frequency of del(11q) (29.6% vs. 9.1%, p=0.10) and a lower frequency of *SF3B1* mutations (6.8% vs. 21.7%, p=0.12) compared to n-CLL in our material, however these discrepancies can be due to relatively low numbers of i-CLL patients analyzed.[184] As expected, the methylation patterns of the 5 CpGs within the three subgroups differed considerably; while n-CLL showed methylation of 4/5 CpG sites and the opposite was seen for m-CLL, i-CLL had a more variable methylation profile for these sites. Based on hierarchical clustering, we observed that within the i-CLL group, subset #2 cases did not cluster as one distinct group (Figure 6).

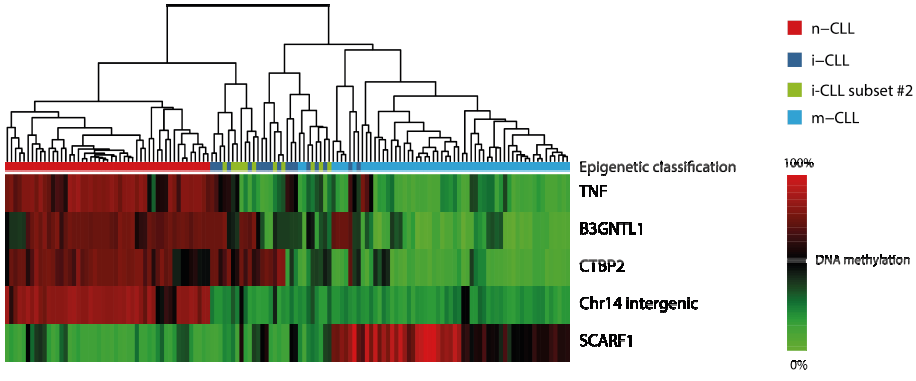

Figure 6. Hierarchical clustering of CLL samples into three epigenetic subgroups based on the methylation pattern of 5 CpG sites.

In accordance with previous findings, survival analysis of the three epigenetic subgroups revealed that n-CLL patients displayed a significantly worse outcome, contrasting m-CLL patients that had a more favorable prognosis, whereas i-CLL patients demonstrated an intermediate outcome.[184] Interestingly, when i-CLL subset #2 cases were compared to other i-CLL cases, the former tended to have a shorter OS (p=0.05), while the latter group had a longer OS compared to n-CLL cases (Figure 7A). Similarly for TTFT, only non-subset #2 i-CLL was found to be significantly different from n-CLL patients and no difference was observed between subset #2 i-CLL and n-CLL (Figure 7B). Moreover, the epigenetic classification remained as an independent marker in multivariate analysis both for OS and TTFT, also in presence of IGHV status.

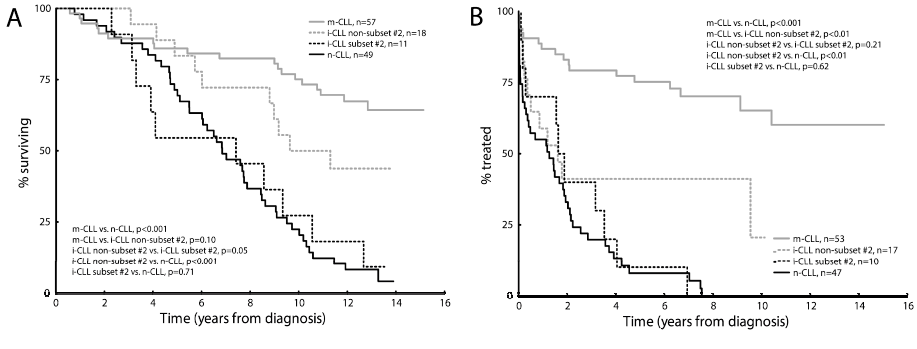

Figure 7. Prognostic impact of epigenetic classification in CLL on OS (A) and TTFT (B) when sub-classifying subset #2 and non-subset #2 i-CLL.

In summary, as a novel finding, we report for the first time that poor-prognostic subset #2 patients are predominantly classified as i-CLL, although they display poor clinical outcome similar to clinically aggressive n-CLL patients, highlighting that the proposed model might not be suitable for this group of patients. The limitation of the study is the number of patients being investigated and our findings need to be further validated in larger, and preferably prospective, studies in order to implement this classification in the clinical setting. Furthermore, the methodology to assess the epigenetic signature needs to be standardized and harmonized.

Paper III: Epigenetic silencing of miR-26A1 in CLL and MCL

Recently, the regulatory effect of miR26A1 on EZH2 expression has been reported in lung cancer and nasopharyngeal carcinoma, where overexpression of miR26A1 resulted in *EZH2* downregulation, inhibition of cell proliferation and increased apoptosis.[259,260] Furthermore, miR26A1 expression was shown to be lower in CLL and MCL samples as compared to normal B-cells,[245,261] however, the mechanism underlying its deregulation and its potential functional role in disease pathogenesis still remains unexplored. In order to address this issue, we investigated the expression level and methylation status of miR26A1 in two well-characterized CLL (n=70) and MCL (n=65) patient cohorts using RQ-PCR analysis and pyrosequencing, respectively.

In accordance to our recent 450K methylation array data in CLL, we observed a differential miR26A1 methylation between U-CLL and M-CLL, where U-CLL displayed significantly higher methylation levels compared to M-CLL ($p < 0.001$). Interestingly, high miR26A1 methylation was also associated with shorter OS in CLL patients (Figure 8B). Accordingly, we observed that 89% of M-CLL cases with >10-year survival displayed lower methylation levels, while 83% U-CLL cases with <10-year survival demonstrated higher methylation levels. In contrast, in MCL, all cases showed a more uniform miR26A1 hypermethylation, irrespective of being high (HP) or low proliferative (LP), defined by Ki-67 expression above or below 30%, indicating a more uniform silencing in MCL (Figure 8A).

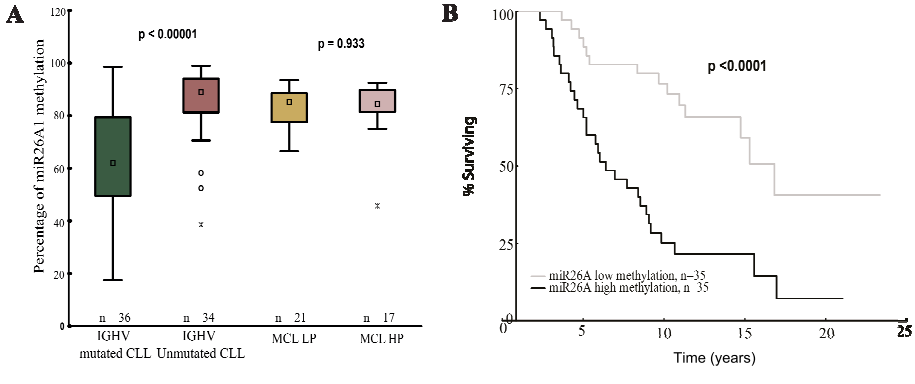

Figure 8. miR26A1 methylation in CLL and MCL. (A) Box plot showing percentage of miR26A1 meth-ylation level in CLL and MCL patient samples, as assessed by pyrosequencing. (B) OS curve for CLL patients with high and low miR26A1 methylation.

From the expression analysis, we observed very low miR26A1 expression levels in both U-CLL/M-CLL and HP/LP MCL compared to normal B-cells. Accordingly, no differential miR26A1 expression was observed between U-CLL and M-CLL cases or between HP and LP MCL cases. Similarly, no differences in OS were observed in relation to miR26A1 expression.

Furthermore, we noted a significant negative correlation ($p = 0.04$, $r = -0.45$) between miR26A1 expression and *EZH2* mRNA expression in CLL samples, which supported a potential link between them. When miR26A1 was overexpressed in CLL and MCL cell lines, we observed a significant downregulation of EZH2 protein levels as assessed by Western blot and flow cytometry analysis, which suggests that miR26A1 targets directly EZH2 in both CLL and MCL (Figure 9). Moreover, overexpression of miR26A1 resulted in increased apoptosis in leukemic cell lines, indicating its tumor suppressor role in B-cell malignancies. Finally, we observed increased miR26A1 expression and decreased *EZH2* expression after methyl inhibitor DAC treatment to CLL/MCL cell lines, suggesting a direct impact of miR26A methylation on its expression and an inverse correlation between miR26A1 and *EZH2*.

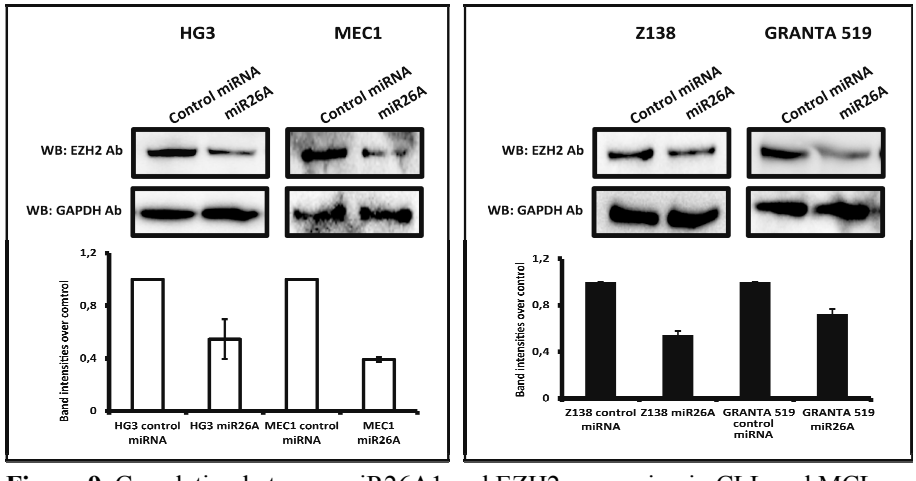

Figure 9. Correlation between miR26A1 and EZH2 expression in CLL and MCL cell lines. Western blot analysis of EZH2 expression in miR26A1 and control mimic miRNA transfected CLL (left panel) and MCL (Right panel) cell lines. GAPDH was used as internal loading control. Histogram below Western blot show band intensities based on two independent transfections.

Taken together, our study highlights the important role of epigenetic silencing of miR26A1 in CLL and MCL pathobiology and further supports its direct regulatory role on EZH2 expression. Moreover, results from the overexpression analysis further support the tumor suppressor role of miR26A1 in B-cell malignancies. Similar to **paper II**, the prognostic impact of miR26A1

methylation in CLL need to be further verified in a larger series of patients, in particular in the context of other prognostic markers.

Paper IV: DNA methylation profiling in stereotyped subsets of CLL

As mentioned, high-resolution microarrays and WGBS studies have unraveled the DNA methylation landscape in prognostic subgroups of CLL.[167,168,181,182,262] However, the characterization of the DNA methylation profile of stereotyped subsets of CLL is largely unresolved, since only few studies have so far been reported.[149,262] Therefore, using 450K methylation arrays, we set out to investigate the DNA methylome in major stereotyped subsets (#1-8, n=176) in relation to non-subset CLL (n=325) as well as various normal B-cell populations (n=74) grouped as precursor, GC inexperienced and GC experienced B-cells.

Overall, unsupervised principal component analysis of subset and non-subset CLL, the latter grouped based on its epigenetic classification, revealed that all U-CLL and M-CLL subsets clustered with n-CLL and m-CLL respectively, while subset #2 clustered with i-CLL, implying common cellular origin (Figure 10A). In line with this, supervised analysis revealed for the most part only a limited number of differentially methylated CpG sites when comparing each U-CLL or M-CLL subset with non-subset cases, with subset #5 as a rare exception, and warranting further investigation.

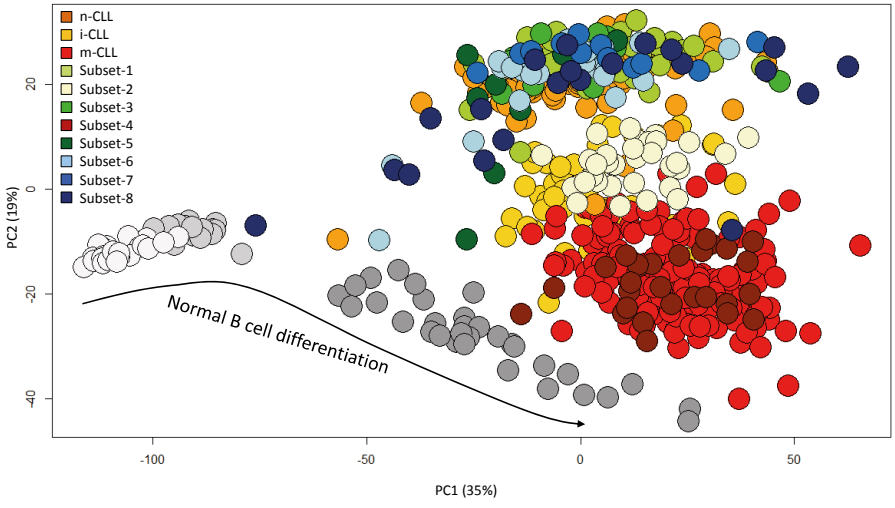

Figure 10. PCA of subset and non-subset CLL cases in the background of normal B-cell populations grouped as precursor, NBC and GC/post-GC. The non-subset CLL are grouped based on their epigenetic classification, i.e. i-CLL, m-CLL and n-CLL.

We previously reported that subset #2 cases were predominantly classified as i-CLL (**Paper II**), though they did not cluster separately in hierarchical clustering analysis, which might be due to low number of cases investigated and the fact that only 5 CpG sites were investigated. In contrast and based on array data, we here observed that almost all subset #2 cases clustered separately from i-CLL in supervised analysis (Figure 11). This highlights that subset #2 may comprise a distinct subgroup within i-CLL.

The number of epigenetic changes that a tumor acquires compared to its cellular origin, "epigenetic burden", has recently been suggested as a powerful predictor of clinical aggressiveness.[263] Bearing this in mind, we took a step further and calculated the epigenetic burden for specific subsets vs. their non-subset cases matched by epigenetic subgroup. We observed a significant difference in the epigenetic burden amongst the various subset comparisons; more specifically, in U-CLL subsets #1 vs. n-CLL (72K vs. 67K, $p<0.05$) and #5 vs. n-CLL (57K vs. 67K, $p<0.01$) and subset #2 vs. i-CLL (76K vs. 68K, $p=0.001$). Furthermore, subset #2 cases frequently carry del(11q) and harbor *SF3B1* mutations, however, neither the IGHV mutation status nor the presence of del(11q) or *SF3B1* mutations had any impact on the epigenetic burden within subset #2. No difference was observed between the M-CLL subset #4 vs. m-CLL (83K vs. 82K, p=not significant).

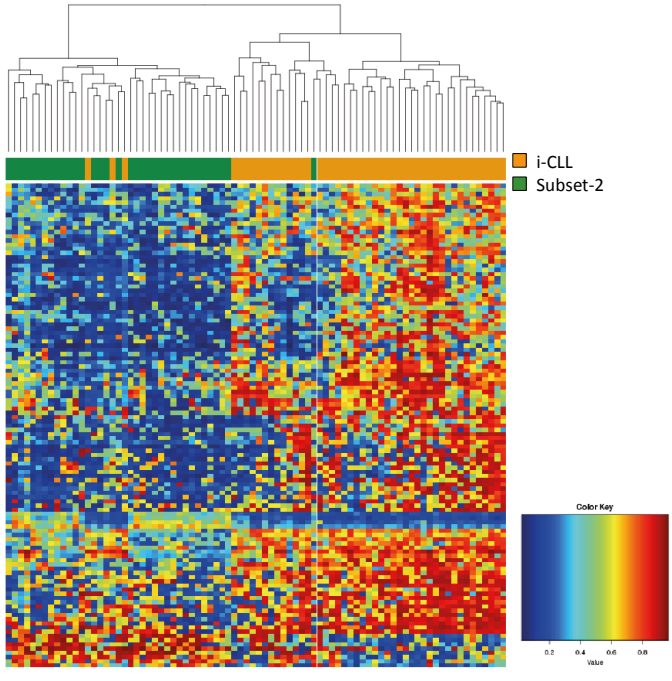

Figure 11. Supervised hierarchical clustering analysis of subset #2 and non-subset i-CLL.

In summary, based on their epigenetic profiles, U-CLL and M-CLL subsets generally clustered with n-CLL and m-CLL subgroups, respectively, implying common cellular origins. In contrast, subset #2 emerged as the first defined member of the i-CLL group, which in turn alludes to a distinct cellular origin for subset #2 and i-CLL patients. However, the differentially methylated gene signatures between specific subsets and their respective epigenetic subgroup, especially between #2 and i-CLL warrants further investigation.

Concluding Remarks

Over the past decades, significant advancements have been achieved in the field of biological and clinical research of lymphoid malignancies, which have provided us deeper insight into the pathobiology of these diseases assisting in further refining the lymphoma classification. Moreover, great efforts have been made to identify novel prognostic and predictive markers that could improve patient stratification and guide the clinicians in selecting therapy and predicting treatment responses.

Considering the remarkable clinical heterogeneity observed in CLL, there is an unmet need with regards to new biomarkers and as a result, a plethora of molecular prognostic markers has been proposed during the last decades. However, the prognostic relevance for the majority of them is restricted to U-CLL patients and there is a deficit of markers for M-CLL patients that usually lack poor-prognostic factors such as high-risk genetic aberrations and recurrent gene mutations. In our validation study of a recently proposed RNA-based marker *UGT2B17,*[138] we not only confirmed its prognostic relevance in a population-based cohort, but also for the first time reported its prognostic significance among M-CLL patients. More specifically, 30% of M-CLL patients expressed high levels of *UGT2B17*, despite lacking any other established poor-prognostic marker, and had a worse clinical outcome. Furthermore, we observed that the risk stratification of M-CLL patients could be further improved by the incorporation of another RNA-based marker i.e. *LPL*. Hence, this study provides a rationale for the applicability of RNA-based markers in patient stratification, in particular for M-CLL patients although our findings need to be confirmed in larger patient series.

Recently, studies of the epigenomic landscape have revealed that CLL can be classified into three epigenetic subgroups, i.e. m-CLL, n-CLL and i-CLL, based on their DNA methylation similarities with normal B-cell populations.[167] Furthermore, these epigenetic subgroups exhibit markedly different clinico-biological profiles. Subsequently, the 5 CpG signature that correspond to the promoter region of *SCARF1*, the gene body of *B3GNTL1*, *CTBP2* and *TNF* and an intergenic region on chromosome 14 can accurately classify the aforementioned subgroups.[184] Although the proposed epigenetic classification largely mirrors the IGHV mutational status, it has still been an open question if the i-CLL group is merely a mix of outliers from the M-

CLL and U-CLL subgroups or if it comprises separate subgroups. We could confirm the proposed epigenetic 5 CpG signature as an independent marker in multivariate analysis for both TTFT and OS. Despite the relatively low number of subset #2 cases investigated, we identified that this subgroup belongs to the i-CLL subgroup, hence supporting the existence of a "real" i-CLL subgroup. However, subset #2 patients had a similarly poor outcome as n-CLL patients, indicating the epigenetic classification might not be applicable for this subgroup of patients from a prognostic point of view.

Downregulation of miR26A1 has been reported in various B-cell malignancies including CLL and MCL.[245,261] However, the molecular mechanisms behind miR26A1 deregulation and its associated role in these malignancies is not yet fully elucidated. We showed that DNA hypermethylation resulted in silencing of miR26A1 in both CLL and MCL and higher methylation levels were associated with a worse prognosis in CLL. Moreover, *in vitro* data revealed that overexpression of miR26A1 resulted in downregulation of *EZH2* expression that subsequently rendered increased apoptosis, further supporting its tumor suppressor role in the disease pathogenesis of CLL and MCL.

In recent years, several methylation studies have reported the differential methylation profiles in prognostic subgroups of CLL, such as U-CLL and M-CLL, however the DNA methylation landscape in stereotyped subsets remains largely unexplored. By applying methylation arrays to stereotyped subsets and non-subset cases clustering of cases generally followed their IGHV mutational status; more specifically, all U-CLL subsets clustered with n-CLL, subset #4 clustered with m-CLL, while subset #2 clustered separately with i-CLL. Furthermore, when epigenetic burden was compared between specific subsets vs. non-subset cases, significant differences was observed, in particular for subsets #1 and #2, in line with their very poor prognosis. Interestingly, further analysis of i-CLL and subset #2 cases revealed that the latter group emerged as the first defined member of the former group; indicating a different cellular origin for these subgroups.

Altogether, this thesis identified a new poor-prognostic group of M-CLL patients expressing *UGT2B17,* defined the first member of the i-CLL epigenetic subgroup, highlighted the possible mechanism behind miR26A1 deregulation in CLL and MCL and unraveled the DNA methylation landscape in stereotyped subset of CLL. This will eventually lead to further refinement of the disease classification and risk stratification for patients diagnosed with these two B-cell malignancies.

Acknowledgments

This work was carried out at the Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Sweden. It would not have been possible to conduct this research work without the financial support of the **Swedish Research Council** and the **Swedish Cancer Society**.

First of all, I would like to express my sincere gratitude to my supervisor, Professor **Richard Rosenquist Brandell** for giving me the opportunity to work in his group as a Ph.D student. Thank you for introducing me to the world of leukemia. You have always encouraged me to think independently, to express myself and to build up my confidence level. Whenever I came to you with any kind of problem, you have solved it immediately. You are not only the best supervisor I have ever seen, but also a wonderful human being. I always wonder how you manage so many responsibilities! You will always be in my memory as the best mentor, kind and generous person. Thank you for all your support, guidance and encouragement during the last 5-years.

I would like to thank my co-supervisor **Larry Mansouri** for his continuous guidance to make all my projects work. You are a very sincere and calm person. Thanks for your patience with me for teaching all the statistical analysis and introducing me to the Statistica software. You have always encouraged me to think independently and draw conclusions. Thank s for your healthy criticism and discussion which helped me do better over the years.

To the present and former members of the **Rosenquist** group; **Lesley**, thank you for your guidance and I really admire your dedication to work and of course your linguistic skill. **Panos**, thank you for all the help and guidance that you have given to me during the years in all aspects in life. You are one of the best colleagues that I have ever met. **Diego**, thank you for your unconditional help with everything that starts from lab to my home. Thank you for your moral support during my thesis writing. **Emma**, thank you for your care, concern and help with my thesis. I have learned many things from you, especially the organizing skill and I appreciate your straight forward and practical thinking. **Viktor**, thank you for being so kind and helpful throughout the last 5 years. I missed your presence in the office during my writing. **Aron**, thank you for being so nice and supportive. **Mattias**, thank

you for your valuable input into my publication. **Karin**, it was nice to know you and thank you for being part of the group. **Nikki**, I always hoped you would stay in our group when I joined, so that I could learn from you. Thank you for your help and guidance in my initial phase in the lab, especially when I was struggling with pyrosequencing! **Millaray**, thank you for being so kind and I appreciate your positive thinking.

To Professor **Lena Claesson-Welsh**, thank you very much for giving me the golden opportunity to resume my research career in your group. You have encouraged, motivated and guided me in best possible way. Thank you for believing in me. I am grateful for your huge and continuous support to our family.

To **Meena Kanduri**, who is not only our collaborator but also a good friend, thank you very much for all your guidance, support and encouragement that you have given me throughout these years. Thank you for being patience with me to teach the pyrosequencing technique (over phone…). I always admire your multitasking ability and technical skills. **Pradeep**, it was nice meeting you and working with you.

I would like to thank our collaborators **Kostas** (also part of our extended MolHem group) and **Inaki**, it was a great pleasure to work with you and to have your expert input to my projects. Thank you **Nikos**, it was great working with you.

I am very much thankful to **Christina Magnusson** and **Helene Norlin** for guiding me in the right direction throughout my Ph.D. Christina, you are like an angel for every Ph.D student at IGP. Thank you very much for your positive and kind words. I would also like to thank the head of the department Professor **Karin Forsberg Nilsson** for your support to all Ph.D students.

I would like to thank all the members of the Molecular Pathology group, especially **Magnus**, **Monica**, **Johanna**, **Linnea** and **Jennifer** for your help and allowing me to work in your lab. To the members of Professor Helena Jernberg Wiklund's group; **Antonia**, you have always been a very good friend and I am heartily thankful to you for all the help and guidance that you have given me from the very beginning and until now. **Alba**, **Mohammad**, **Prasoon, Charlotte** and **Lotta**, thanks for being so supportive throughout the last 5 years. Thanks to all the members at the SNP&SEQ and BioVis facilities.

Mimmi, **Charlotte**, **Pernilla** and **Lotta**, you all are amazingly helpful people at Rudbeck Laboratory. Thank you all for your help and support

during these years. Thanks to all my friends at IGP for making my stay enjoyable; **Oriol**, **Ross**, **Eric**, **Veronica**, **Matko**, **Argyris**, **Kiki**, **Svea**, **Anne**, **My**, **Linda**, **Lucy**, **Tanja**, **Khayrun**, **Colin**, **Bong**, **Vivek** , **Akhtar**, **Yuan** and anyone I missed.

To **Sonia**, thank you for your great guidance while my stay in Lena´s lab. You are one of the most dedicated and enthusiastic persons I have met. I have enjoyed the work under your supervision. **Frank**, thank you for your guidance towards the end part of the HRG project.

I would like to thank Professor **Sudha Bhattacharya** and **Alok Bhattacharya** for giving me the opportunity to start my research career at Jawaharlal Nehru University, New Delhi, India. Thank you for allowing me to work in your lab where I have learned my basic molecular biology techniques.

To Professor **Renu Tuteja**, ICGEB, New Delhi, thank you for accepting me as a senior research fellow in your Malaria Research group where I have learned techniques related to protein purification.

A big thank you to Dr. **Jitendra Badhai**, you are one of the biggest inspirations not only for me but also for many other people who had made their research journey to Sweden. Thank you very much for all your invaluable suggestions and support for my research career. Thank you dear **Jyoti** and little **Reyansh** for all the love and care and the fun time we spent together.

I would like to thank all the IGP Indians present and past members for nice chats and discussion during lunch and coffee breaks. Thanks to **Kundu da** (**Snehangshu**), **Soumi**, **Prathyusha**, **Rajesh**, **Sathish**, **Priya**, **Prasoon**, **Gaurav**, **Sanhita**, and **Tanmoy**.

My ‘‘**little India**’’ in **Uppsala**. Thanks to all my friends for making my Uppsala chapter amazingly beautiful. Thanks all for all the kids birthday parties, Diwali, Christmas and New Year celebrations. Thank you, **PK Singh** uncle ji for all the love and care. **Marja**, though you are Finnish, I always feel like, you are an Indian. Thank you for all the support and love you have given to us. The Marathi community in Uppsala; **Varsha, Sanjay,** ***Swaroop*** and ***Shweta***, **Prajkta**, **Prasad**, ***Vedu*** and ***Rudransh***, **Snehal**, **Ganesh**, and ***Nil***. It was pleasure to know all of you. Thank you all for the discussions and delicious food at the parties. **Lakshmi, Nikhil**, and ***Khushi***, **Rashmi**, **Sarosh**, ***Maanvi,*** and ***Chirag*** thank you for being so kind and supportive all the time. **Deepesh** and **Priya** thank you for all your help. **Shalini** and **Surya,** thank you for all the fun we made in our girls day

out. **Sarbasis**, **Gargi** and ***Swastika***, thank you for the time we spent in Norway.

Sulena, Kali, and cutest ***Aura*** you are our extended family in Sweden. I will not forget the way you welcome us in Uppsala and shared all your experiences and teaching me every tiny thing about the city. Thank you for everything. **Smitha**, **Nimesh**, ***Adi***, and ***Drithi,*** I do not have words to describe your care and love. **Smitha**, thank you for being so caring and supportive in these years. I badly missed your invaluable moral support during this stressful period. **Navya, Sandeep**, and ***Aarav***, perhaps you are the longest companion for us in Uppsala so far. Thank you so much for all your love, affection, care and support in every respect, delicious food, all the chats, discussion and Friday fun parties together. I would cherish those moments.

To my family in Mursund: I would like to thank my father-in-law, **Jagannath Padhan** and mother-in-law, **Padmabati Padhan**, for your unconditional care and support for us. **Amulya dada** and **Anjali nani**, you both have always supported me even in the tough periods of my life. I think, words are not sufficient to express my gratitude towards you. **Surendra** and **Alaka**, thank you for your support, love and care. **Surendra**, I really appreciate your unconditional support and help for all our official work back home. You are the legal advisor for us in every aspect ☺. **Madhu dada** and **Manjari nani**, you both are the best example of tolerance and social relationship. I really wonder how you could manage to entertain so many relatives!! Our adorable "bacha party"; ***Sunita***, ***Subrata***, ***Roli*** and ***Ronu***, thank you for making our holidays in India most memorable.

To my beautiful sisters, **Sabita** and **Liza**, thank you for your unconditional support. I think, we all have made our parents proud and raised our family name to a respectable level in our society. **Sagara da** and **Prashanta**, thank you for being so supportive and caring. We are lucky to have you in our family. Special thanks to "bacha party", ***Pranshu***, ***Ashu*** and ***Prayash***, for making my holidays most enjoyable in India. To my uncle **Surendra Bhoi,** thank you for your support and encouragement.

To my father, **Makunda Bhoi**, I think it is not fair just to describe your incredible support and encouragement in a few lines. You are the biggest inspiration for me to pursue my Ph.D. You always dreamed and even proceeded towards your Ph.D, but due to your responsibilities towards the family, you had to quit. I hope you will be proud of your daughter, who has fulfilled your dream. The surname that I inherited from you will always be in my publications. You have always provided me the best education, starting

from class I and encouraged me to do my best. Thank you from the bottom of my heart for everything!

To my mother, **Sarojini Bhoi**, I do not have words to express my gratitude to you. You are the one who understand me best! Whatever may be the situation you have always supported me. All the sacrifices that you made for us will never pay off. Thank you for all your unconditional care, love, support, blessing and tremendous patience!

Finally, to my beloved family, two apples of my eyes, **Lora** (**Anubha**) and **Adriyaan**. You are the two most precious gifts of my life. You are my stress releasing factor and my energy booster. Sorry from the bottom of my heart, not being able to spent time with you during this period. Last but not the least, my dearest hubby **Narendra**, without you, I could not have imagined to do a Ph.D in my life. You are the best friend and also best critic of mine. Your unconditional support and encouragement has brought me here. Thank you for being the **best husband** and **best papa** in the world.

All four of us deserve this Ph.D and I think we do have very valid reason for that. **Love you all** ☺

Sujata Bhoi

References

1 Morton, L. M. *et al.* Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood.* **110** (2), 695-708, doi:10.1182/blood-2006-11-051672, (2007).

2 Swerdlow, S. H. *et al.* The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood.* **127** (20), 2375-2390, doi:10.1182/blood-2016-01-643569, (2016).

3 Swerdlow, S. H. *et al.* WHO Classification of Tumors of Haematopoietic and Lymphoid Tissues. *IARC: Lyon.* (2008).

4 Armitage, J. O. & Weisenburger, D. D. New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. *J Clin Oncol.* **16** (8), 2780-2795, doi:10.1200/JCO.1998.16.8.2780, (1998).

5 Tonon, G. & Anderson, K. C. in *The Molecular Basis of Cancer (Third Edition)* eds Peter M. Howley, Mark A. Israel, Joe W. Gray, & Craig B. Thompson) 351-359 (W.B. Saunders, 2008).

6 Ikuta, K., Uchida, N., Friedman, J. & Weissman, I. L. Lymphocyte development from stem cells. *Annu Rev Immunol.* **10** 759-783, doi:10.1146/annurev.iy.10.040192.003551, (1992).

7 LeBien, T. W. & Tedder, T. F. B lymphocytes: how they develop and function. *Blood.* **112** (5), 1570-1580, doi:10.1182/blood-2008-02-078071, (2008).

8 Corfe, S. A. & Paige, C. J. The many roles of IL-7 in B cell development; mediator of survival, proliferation and differentiation. *Semin Immunol.* **24** (3), 198-208, doi:10.1016/j.smim.2012.02.001, (2012).

9 Samitas, K., Lotvall, J. & Bossios, A. B cells: from early development to regulating allergic diseases. *Arch Immunol Ther Exp (Warsz).* **58** (3), 209-225, doi:10.1007/s00005-010-0073-2, (2010).

10 Melchers, F. *et al.* The surrogate light chain in B-cell development. *Immunol Today.* **14** (2), 60-68, doi:10.1016/0167-5699(93)90060-X, (1993).

11 DeFranco, A. L. Structure and function of the B cell antigen receptor. *Annu Rev Cell Biol.* **9** 377-410, doi:10.1146/annurev.cb.09.110193.002113, (1993).

12 Meffre, E., Casellas, R. & Nussenzweig, M. C. Antibody regulation of B cell development. *Nat Immunol.* **1** (5), 379-385, doi:10.1038/80816, (2000).

13 Manser, T. Textbook germinal centers? *J Immunol.* **172** (6), 3369-3375 (2004).

14 MacLennan, I. C. Somatic mutation. From the dark zone to the light. *Curr Biol.* **4** (1), 70-72 (1994).

15 Kuppers, R., Klein, U., Hansmann, M. L. & Rajewsky, K. Cellular origin of human B-cell lymphomas. *N Engl J Med.* **341** (20), 1520-1529, doi:10.1056/NEJM199911113412007, (1999).

16 Guzman-Rojas, L., Sims-Mourtada, J. C., Rangel, R. & Martinez-Valdez, H. Life and death within germinal centres: a double-edged sword. *Immunology.* **107** (2), 167-175 (2002).
17 Weigert, M. *et al.* The joining of V and J gene segments creates antibody diversity. *Nature.* **283** (5746), 497-499 (1980).
18 Tonegawa, S. Somatic generation of antibody diversity. *Nature.* **302** (5909), 575-581 (1983).
19 McBlane, J. F. *et al.* Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell.* **83** (3), 387-395 (1995).
20 Market, E. & Papavasiliou, F. N. V(D)J recombination and the evolution of the adaptive immune system. *PLoS Biol.* **1** (1), E16, doi:10.1371/journal.pbio.0000016, (2003).
21 Alt, F. W. *et al.* VDJ recombination. *Immunol Today.* **13** (8), 306-314, doi:10.1016/0167-5699(92)90043-7, (1992).
22 Komori, T., Okada, A., Stewart, V. & Alt, F. W. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science.* **261** (5125), 1171-1175 (1993).
23 Benedict, C. L., Gilfillan, S., Thai, T. H. & Kearney, J. F. Terminal deoxynucleotidyl transferase and repertoire development. *Immunol Rev.* **175** 150-157 (2000).
24 Bassing, C. H., Swat, W. & Alt, F. W. The mechanism and regulation of chromosomal V(D)J recombination. *Cell.* **109 Suppl** S45-55 (2002).
25 Schatz, D. G., Oettinger, M. A. & Schlissel, M. S. V(D)J recombination: molecular biology and regulation. *Annu Rev Immunol.* **10** 359-383, doi:10.1146/annurev.iy.10.040192.002043, (1992).
26 Peled, J. U. *et al.* The biochemistry of somatic hypermutation. *Annu Rev Immunol.* **26** 481-511, doi:10.1146/annurev.immunol.26.021607.090236, (2008).
27 Di Noia, J. M. & Neuberger, M. S. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem.* **76** 1-22, doi:10.1146/annurev.biochem.76.061705.090740, (2007).
28 Milstein, C., Neuberger, M. S. & Staden, R. Both DNA strands of antibody genes are hypermutation targets. *Proc Natl Acad Sci U S A.* **95** (15), 8791-8794 (1998).
29 Rogozin, I. B. & Diaz, M. Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. *J Immunol.* **172** (6), 3382-3384 (2004).
30 Rogozin, I. B. & Pavlov, Y. I. Theoretical analysis of mutation hotspots and their DNA sequence context specificity. *Mutat Res.* **544** (1), 65-85 (2003).
31 Shapiro, G. S. & Wysocki, L. J. DNA target motifs of somatic mutagenesis in antibody genes. *Crit Rev Immunol.* **22** (3), 183-200 (2002).
32 Hartley, S. B. *et al.* Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. *Cell.* **72** (3), 325-335 (1993).
33 Stavnezer, J. & Schrader, C. E. IgH chain class switch recombination: mechanism and regulation. *J Immunol.* **193** (11), 5370-5378, doi:10.4049/jimmunol.1401849, (2014).
34 Xu, Z., Zan, H., Pone, E. J., Mai, T. & Casali, P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol.* **12** (7), 517-531, doi:10.1038/nri3216, (2012).

35 Durandy, A. Activation-induced cytidine deaminase: a dual role in class-switch recombination and somatic hypermutation. *Eur J Immunol.* **33** (8), 2069-2073, doi:10.1002/eji.200324133, (2003).
36 Hallek, M. *et al.* Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood.* **111** (12), 5446-5456, doi:10.1182/blood-2007-06-093906, (2008).
37 Zenz, T., Mertens, D., Kuppers, R., Dohner, H. & Stilgenbauer, S. From pathogenesis to treatment of chronic lymphocytic leukaemia. *Nat Rev Cancer.* **10** (1), 37-50, doi:10.1038/nrc2764nrc2764 [pii], (2010).
38 Rozman, C. & Montserrat, E. Chronic lymphocytic leukemia. *N Engl J Med.* **333** (16), 1052-1057, doi:10.1056/NEJM199510193331606, (1995).
39 Woyach, J. A. & Johnson, A. J. Targeted therapies in CLL: mechanisms of resistance and strategies for management. *Blood.* **126** (4), 471-477, doi:10.1182/blood-2015-03-585075, (2015).
40 Hallek, M. Chronic lymphocytic leukemia: 2015 Update on diagnosis, risk stratification, and treatment. *Am J Hematol.* **90** (5), 446-460, doi:10.1002/ajh.23979, (2015).
41 Binet, J. L. *et al.* A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer.* **48** (1), 198-206 (1981).
42 Rai, K. R. *et al.* Clinical staging of chronic lymphocytic leukemia. *Blood.* **46** (2), 219-234 (1975).
43 Strati, P. & Shanafelt, T. D. Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification. *Blood.* **126** (4), 454-462, doi:10.1182/blood-2015-02-585059, (2015).
44 Tsimberidou, A. M., Keating, M. J. & Wierda, W. G. Richter's transformation in chronic lymphocytic leukemia. *Curr Hematol Malig Rep.* **2** (4), 265-271, doi:10.1007/s11899-007-0036-9, (2007).
45 Byrd, J. C. *et al.* Randomized phase 2 study of fludarabine with concurrent versus sequential treatment with rituximab in symptomatic, untreated patients with B-cell chronic lymphocytic leukemia: results from Cancer and Leukemia Group B 9712 (CALGB 9712). *Blood.* **101** (1), 6-14, doi:10.1182/blood-2002-04-1258, (2003).
46 Brown, J. R. *et al.* Idelalisib, an inhibitor of phosphatidylinositol 3-kinase p110delta, for relapsed/refractory chronic lymphocytic leukemia. *Blood.* **123** (22), 3390-3397, doi:10.1182/blood-2013-11-535047, (2014).
47 Byrd, J. C., O'Brien, S. & James, D. F. Ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med.* **369** (13), 1278-1279, doi:10.1056/NEJMc1309710, (2013).
48 Byrd, J. C. *et al.* Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med.* **369** (1), 32-42, doi:10.1056/NEJMoa1215637, (2013).
49 Furman, R. R. *et al.* Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *N Engl J Med.* **370** (11), 997-1007, doi:10.1056/NEJMoa1315226, (2014).
50 Roberts, A. W. *et al.* Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. *N Engl J Med.* **374** (4), 311-322, doi:10.1056/NEJMoa1513257, (2016).

51 Guieze, R. & Wu, C. J. Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia. *Blood.* **126** (4), 445-453, doi:10.1182/blood-2015-02-585042, (2015).
52 Schuh, A. *et al.* Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood.* **120** (20), 4191-4196, doi:10.1182/blood-2012-05-433540, (2012).
53 Rosenquist, R., Cortese, D., Bhoi, S., Mansouri, L. & Gunnarsson, R. Prognostic markers and their clinical applicability in chronic lymphocytic leukemia: where do we stand? *Leuk Lymphoma.* **54** (11), 2351-2364, doi:10.3109/10428194.2013.783913, (2013).
54 Mansouri, L. *et al.* NOTCH1 and SF3B1 mutations can be added to the hierarchical prognostic classification in chronic lymphocytic leukemia. *Leukemia.* **27** (2), 512-514, doi:10.1038/leu.2012.307 leu2012307 [pii], (2013).
55 Damle, R. N. *et al.* Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood.* **94** (6), 1840-1847 (1999).
56 Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood.* **94** (6), 1848-1854 (1999).
57 Dohner, H. *et al.* Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med.* **343** (26), 1910-1916 (2000).
58 Cortese, D. *et al.* On the way towards a 'CLL prognostic index': focus on TP53, BIRC3, SF3B1, NOTCH1 and MYD88 in a population-based cohort. *Leukemia.* **28** (3), 710-713, doi:10.1038/leu.2013.333 leu2013333 [pii], (2014).
59 Rossi, D. *et al.* Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood.* **121** (8), 1403-1412, doi:10.1182/blood-2012-09-458265 blood-2012-09-458265 [pii], (2013).
60 Puente, X. S. *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature.* **475** (7354), 101-105, doi:10.1038/nature10113 nature10113 [pii], (2011).
61 Quesada, V. *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet.* **44** (1), 47-52, doi:10.1038/ng.1032ng.1032 [pii], (2012).
62 Quesada, V. *et al.* The genomic landscape of chronic lymphocytic leukemia: clinical implications. *BMC Med.* **11** 124, doi:10.1186/1741-7015-11-1241741-7015-11-124 [pii], (2013).
63 Baliakas, P. *et al.* Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia.* **29** (2), 329-336, doi:10.1038/leu.2014.196leu2014196 [pii], (2015).
64 Claus, R. *et al.* Validation of ZAP-70 methylation and its relative significance in predicting outcome in chronic lymphocytic leukemia. *Blood.* **124** (1), 42-48, doi:10.1182/blood-2014-02-555722, (2014).
65 Claus, R. *et al.* Quantitative DNA methylation analysis identifies a single CpG dinucleotide important for ZAP-70 expression and predictive of prognosis in chronic lymphocytic leukemia. *J Clin Oncol.* **30** (20), 2483-2491, doi:10.1200/JCO.2011.39.3090, (2012).
66 Irving, L. *et al.* Methylation markers identify high risk patients in IGHV mutated chronic lymphocytic leukemia. *Epigenetics.* **6** (3), 300-306 (2011).

67 Shinawi, T. *et al.* KIBRA gene methylation is associated with unfavorable biological prognostic parameters in chronic lymphocytic leukemia. *Epigenetics.* **7** (3), 211-215, doi:10.4161/epi.7.3.19222, (2012).
68 Hill, V. K. *et al.* Frequent epigenetic inactivation of KIBRA, an upstream member of the Salvador/Warts/Hippo (SWH) tumor suppressor network, is associated with specific genetic event in B-cell acute lymphocytic leukemia. *Epigenetics.* **6** (3), 326-332 (2011).
69 Oscier, D. G. *et al.* Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. *Blood.* **100** (4), 1177-1184 (2002).
70 Montserrat, E. New prognostic markers in CLL. *Hematology Am Soc Hematol Educ Program.* 279-284, doi:10.1182/asheducation-2006.1.279, (2006).
71 Stilgenbauer, S., Bullinger, L., Lichter, P., Dohner, H. & German, C. L. L. S. G. C. l. l. Genetics of chronic lymphocytic leukemia: genomic aberrations and V(H) gene mutation status in pathogenesis and clinical course. *Leukemia.* **16** (6), 993-1007, doi:10.1038/sj.leu.2402537, (2002).
72 Stilgenbauer, S., Bullinger, L., Lichter, P. & Dohner, H. Genetics of chronic lymphocytic leukemia: genomic aberrations and V(H) gene mutation status in pathogenesis and clinical course. *Leukemia.* **16** (6), 993-1007 (2002).
73 Stilgenbauer, S. *et al.* Clonal evolution in chronic lymphocytic leukemia: acquisition of high-risk genomic aberrations associated with unmutated VH, resistance to therapy, and short survival. *Haematologica.* **92** (9), 1242-1245, doi:10.3324/921242 [pii], (2007).
74 Parker, T. L. & Strout, M. P. Chronic lymphocytic leukemia: prognostic factors and impact on treatment. *Discov Med.* **11** (57), 115-123 (2011).
75 Bomben, R. *et al.* Molecular and clinical features of chronic lymphocytic leukaemia with stereotyped B cell receptors: results from an Italian multicentre study. *Br J Haematol.* **144** (4), 492-506 (2009).
76 An international prognostic index for patients with chronic lymphocytic leukaemia (CLL-IPI): a meta-analysis of individual patient data. *The Lancet Oncology.* **17** (6), 779-790, doi:10.1016/S1470-2045(16)30029-8.
77 Juliusson, G. *et al.* Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. *N Engl J Med.* **323** (11), 720-724 (1990).
78 Shanafelt, T. D. *et al.* Prospective evaluation of clonal evolution during long-term follow-up of patients with untreated early-stage chronic lymphocytic leukemia. *J Clin Oncol.* **24** (28), 4634-4641, doi:10.1200/JCO.2006.06.9492, (2006).
79 Migliazza, A. *et al.* Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. *Blood.* **97** (7), 2098-2104 (2001).
80 Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc Natl Acad Sci U S A.* **102** (39), 13944-13949, doi:0506654102 [pii]10.1073/pnas.0506654102 [doi], (2005).
81 Mertens, D. *et al.* 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** (11), 4116-4121 (2002).
82 Mertens, D. *et al.* Allelic silencing at the tumor-suppressor locus 13q14.3 suggests an epigenetic tumor-suppressor mechanism. *Proc Natl Acad Sci U S A.* **103** (20), 7741-7746, doi:10.1073/pnas.0600494103, (2006).

83 Klein, U. *et al.* The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell.* **17** (1), 28-40, doi:10.1016/j.ccr.2009.11.019, (2010).
84 Hernandez, J. A. *et al.* A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. *Haematologica.* **94** (3), 364-371 (2009).
85 Puiggros, A., Blanco, G. & Espinet, B. Genetic abnormalities in chronic lymphocytic leukemia: where we are and where we go. *Biomed Res Int.* **2014** 435983, doi:10.1155/2014/435983, (2014).
86 Balatti, V. *et al.* NOTCH1 mutations in CLL associated with trisomy 12. *Blood.* **119** (2), 329-331, doi:10.1182/blood-2011-10-386144, (2012).
87 Rose-Zerilli, M. J. *et al.* ATM mutation rather than BIRC3 deletion and/or mutation predicts reduced survival in 11q-deleted chronic lymphocytic leukemia: data from the UK LRF CLL4 trial. *Haematologica.* **99** (4), 736-742, doi:10.3324/haematol.2013.098574, (2014).
88 Ouillette, P. *et al.* Incidence and clinical implications of ATM aberrations in chronic lymphocytic leukemia. *Genes Chromosomes Cancer.* **51** (12), 1125-1132, doi:10.1002/gcc.21997, (2012).
89 Rossi, D. *et al.* Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. *Blood.* **119** (12), 2854-2862, doi:10.1182/blood-2011-12-395673blood-2011-12-395673 [pii], (2012).
90 Dohner, H. *et al.* p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. *Blood.* **85** (6), 1580-1589 (1995).
91 Foa, R., Del Giudice, I., Guarini, A., Rossi, D. & Gaidano, G. Clinical implications of the molecular genetics of chronic lymphocytic leukemia. *Haematologica.* **98** (5), 675-685, doi:10.3324/haematol.2012.069369, (2013).
92 Zenz, T. *et al.* TP53 mutation and survival in chronic lymphocytic leukemia. *J Clin Oncol.* **28** (29), 4473-4479, doi:10.1200/JCO.2009.27.8762 JCO.2009.27.8762 [pii], (2010).
93 Zenz, T. *et al.* Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. *Blood.* **114** (13), 2589-2597 (2009).
94 Delgado, J. *et al.* Chronic lymphocytic leukaemia with 17p deletion: a retrospective analysis of prognostic factors and therapy results. *Br J Haematol.* **157** (1), 67-74, doi:10.1111/j.1365-2141.2011.09000.x, (2012).
95 Rosenquist, R. *et al.* Clinical impact of recurrently mutated genes on lymphoma diagnostics: state-of-the-art and beyond. *Haematologica.* **101** (9), 1002-1009, doi:10.3324/haematol.2015.134510, (2016).
96 Wang, L. *et al.* SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med.* **365** (26), 2497-2506, doi:10.1056/NEJMoa1109016, (2011).
97 Fabbri, G. *et al.* Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med.* **208** (7), 1389-1401, doi:10.1084/jem.20110921jem.20110921 [pii], (2011).
98 Puente, X. S. *et al.* Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature.* **526** (7574), 519-524, doi:10.1038/nature14666, (2015).
99 Rossi, D. *et al.* Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood.* **123** (14), 2139-2147, doi:10.1182/blood-2013-11-539726, (2014).

100 Wan, Y. & Wu, C. J. SF3B1 mutations in chronic lymphocytic leukemia. *Blood.* **121** (23), 4627-4634, doi:10.1182/blood-2013-02-427641, (2013).
101 Baliakas, P. *et al.* Prognostic relevance of MYD88 mutations in CLL: the jury is still out. *Blood.* **126** (8), 1043-1044, doi:10.1182/blood-2015-05-648634, (2015).
102 Koch, U. & Radtke, F. Notch in T-ALL: new players in a complex disease. *Trends Immunol.* **32** (9), 434-442, doi:10.1016/j.it.2011.06.005, (2011).
103 Aster, J. C., Blacklow, S. C. & Pear, W. S. Notch signalling in T-cell lymphoblastic leukaemia/lymphoma and other haematological malignancies. *J Pathol.* **223** (2), 262-273, doi:10.1002/path.2789, (2011).
104 Jeromin, S. *et al.* SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia.* **28** (1), 108-117, doi:10.1038/leu.2013.263, (2014).
105 Oscier, D. G. *et al.* The clinical significance of NOTCH1 and SF3B1 mutations in the UK LRF CLL4 trial. *Blood.* **121** (3), 468-475, doi:10.1182/blood-2012-05-429282, (2013).
106 Villamor, N. *et al.* NOTCH1 mutations identify a genetic subgroup of chronic lymphocytic leukemia patients with high risk of transformation and poor outcome. *Leukemia.* **27** (5), 1100-1106, doi:10.1038/leu.2012.357 leu2012357 [pii], (2013).
107 Del Giudice, I. *et al.* NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. *Haematologica.* **97** (3), 437-441, doi:10.3324/haematol.2011.060129, (2012).
108 Will, C. L. & Luhrmann, R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol.* **3** (7), doi:10.1101/cshperspect.a003707, (2011).
109 Jeromin, S. *et al.* SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia.* doi:10.1038/leu.2013.263 leu2013263 [pii], (2013).
110 Landau, D. A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell.* **152** (4), 714-726, doi:10.1016/j.cell.2013.01.019 S0092-8674(13)00071-8 [pii], (2013).
111 Rossi, D. *et al.* Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood.* **118** (26), 6904-6908, doi:10.1182/blood-2011-08-373159, (2011).
112 Klein, U. *et al.* Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J Exp Med.* **194** (11), 1625-1638 (2001).
113 Rosenwald, A. *et al.* Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med.* **194** (11), 1639-1647 (2001).
114 Sevov, M., Rosenquist, R. & Mansouri, L. RNA-based markers as prognostic factors in chronic lymphocytic leukemia. *Expert Rev Hematol.* **5** (1), 69-79, doi:10.1586/ehm.11.80, (2012).
115 Mead, J. R., Irvine, S. A. & Ramji, D. P. Lipoprotein lipase: structure, function, regulation, and role in disease. *J Mol Med.* **80** (12), 753-769 (2002).
116 Heintel, D. *et al.* High expression of lipoprotein lipase in poor risk B-cell chronic lymphocytic leukemia. *Leukemia.* **19** (7), 1216-1223 (2005).
117 Nuckel, H. *et al.* Lipoprotein lipase expression is a novel prognostic factor in B-cell chronic lymphocytic leukemia. *Leuk Lymphoma.* **47** (6), 1053-1061 (2006).

118 van't Veer, M. B. *et al.* The predictive value of lipoprotein lipase for survival in chronic lymphocytic leukemia. *Haematologica.* **91** (1), 56-63 (2006).
119 Van Bockstaele, F. *et al.* Lipoprotein lipase mRNA expression in whole blood is a prognostic marker in B cell chronic lymphocytic leukemia. *Clin Chem.* **53** (2), 204-212 (2007).
120 Mansouri, M. *et al.* Lipoprotein lipase is differentially expressed in prognostic subsets of chronic lymphocytic leukemia but displays invariably low catalytical activity. *Leuk Res.* **34** (3), 301-306, doi:S0145-2126(09)00382-8 [pii]10.1016/j.leukres.2009.07.032, (2010).
121 Kaderi, M. A. *et al.* LPL is the strongest prognostic factor in a comparative analysis of RNA-based markers in early chronic lymphocytic leukemia. *Haematologica.* **96** (8), 1153-1160, doi:10.3324/haematol.2010.039396 haematol.2010.039396 [pii], (2011).
122 Wiestner, A. *et al.* ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. *Blood.* **101** (12), 4944-4951 (2003).
123 Durig, J. *et al.* ZAP-70 expression is a prognostic factor in chronic lymphocytic leukemia. *Leukemia.* **17** (12), 2426-2434 (2003).
124 Orchard, J. A. *et al.* ZAP-70 expression and prognosis in chronic lymphocytic leukaemia. *Lancet.* **363** (9403), 105-111, doi:10.1016/S0140-6736(03)15260-9, (2004).
125 Chen, L. *et al.* Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. *Blood.* **100** (13), 4609-4614, doi:10.1182/blood-2002-06-1683, (2002).
126 Chen, L. *et al.* ZAP-70 directly enhances IgM signaling in chronic lymphocytic leukemia. *Blood.* **105** (5), 2036-2041, doi:10.1182/blood-2004-05-1715, (2005).
127 Buhl, A. M. *et al.* Identification of a gene on chromosome 12q22 uniquely overexpressed in chronic lymphocytic leukemia. *Blood.* **107** (7), 2904-2911 (2006).
128 Buhl, A. M. *et al.* CLLU1 expression levels predict time to initiation of therapy and overall survival in chronic lymphocytic leukemia. *Eur J Haematol.* **76** (6), 455-464, doi:10.1111/j.0902-4441.2005.t01-1-EJH2530.x, (2006).
129 Bae, J., Leo, C. P., Hsu, S. Y. & Hsueh, A. J. MCL-1S, a splicing variant of the antiapoptotic BCL-2 family member MCL-1, encodes a proapoptotic protein possessing only the BH3 domain. *J Biol Chem.* **275** (33), 25255-25261, doi:10.1074/jbc.M909826199, (2000).
130 Pepper, C. *et al.* Mcl-1 expression has in vitro and in vivo significance in chronic lymphocytic leukemia and is associated with other poor prognostic markers. *Blood.* **112** (9), 3807-3817 (2008).
131 Awan, F. T. *et al.* Mcl-1 expression predicts progression-free survival in chronic lymphocytic leukemia patients treated with pentostatin, cyclophosphamide, and rituximab. *Blood.* **113** (3), 535-537, doi:10.1182/blood-2008-08-173450, (2009).
132 Saxena, A. *et al.* Mcl-1 and Bcl-2/Bax ratio are associated with treatment response but not with Rai stage in B-cell chronic lymphocytic leukemia. *Am J Hematol.* **75** (1), 22-33, doi:10.1002/ajh.10453, (2004).
133 Beaulieu, M. *et al.* Chromosomal localization, structure, and regulation of the UGT2B17 gene, encoding a C19 steroid metabolizing enzyme. *DNA Cell Biol.* **16** (10), 1143-1154 (1997).

134 Beaulieu, M., Levesque, E., Hum, D. W. & Belanger, A. Isolation and characterization of a novel cDNA encoding a human UDP-glucuronosyltransferase active on C19 steroids. *J Biol Chem.* **271** (37), 22855-22862 (1996).
135 Turgeon, D., Carrier, J. S., Chouinard, S. & Belanger, A. Glucuronidation activity of the UGT2B17 enzyme toward xenobiotics. *Drug Metab Dispos.* **31** (5), 670-676 (2003).
136 Hirata, H. *et al.* Function of UDP-glucuronosyltransferase 2B17 (UGT2B17) is involved in endometrial cancer. *Carcinogenesis.* **31** (9), 1620-1626, doi:10.1093/carcin/bgq124 bgq124 [pii], (2010).
137 Paquet, S. *et al.* Differential expression of the androgen-conjugating UGT2B15 and UGT2B17 enzymes in prostate tumor cells during cancer progression. *J Clin Endocrinol Metab.* **97** (3), E428-432, doi:10.1210/jc.2011-2064, (2012).
138 Gruber, M. *et al.* Overexpression of uridine diphospho glucuronosyltransferase 2B17 in high-risk chronic lymphocytic leukemia. *Blood.* **121** (7), 1175-1183, doi:10.1182/blood-2012-08-447359 blood-2012-08-447359 [pii], (2013).
139 Baliakas, P., Mattsson, M., Stamatopoulos, K. & Rosenquist, R. Prognostic indices in chronic lymphocytic leukaemia: where do we stand how do we proceed? *J Intern Med.* **279** (4), 347-357, doi:10.1111/joim.12455, (2016).
140 Li, Z., Woo, C. J., Iglesias-Ussel, M. D., Ronai, D. & Scharff, M. D. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev.* **18** (1), 1-11, doi:10.1101/gad.1161904, (2004).
141 Tonegawa, S. Somatic generation of immune diversity. *Biosci Rep.* **8** (1), 3-26 (1988).
142 Fais, F. *et al.* Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest.* **102** (8), 1515-1525 (1998).
143 Schroeder, H. W., Jr. & Dighiero, G. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunol Today.* **15** (6), 288-294, doi:10.1016/0167-5699(94)90009-4, (1994).
144 Dono, M. *et al.* Evidence for progenitors of chronic lymphocytic leukemia B cells that undergo intraclonal differentiation and diversification. *Blood.* **87** (4), 1586-1594 (1996).
145 Murray, F. *et al.* Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood.* **111** (3), 1524-1533, doi:10.1182/blood-2007-07-099564, (2008).
146 Agathangelidis, A. *et al.* Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. *Blood.* **119** (19), 4467-4475, doi:10.1182/blood-2011-11-393694 blood-2011-11-393694 [pii], (2012).
147 Messmer, B. T. *et al.* Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J Exp Med.* **200** (4), 519-525 (2004).
148 Marincevic, M. *et al.* High-density screening reveals a different spectrum of genomic aberrations in chronic lymphocytic leukemia patients with 'stereotyped' IGHV3-21 and IGHV4-34 B-cell receptors. *Haematologica.* **95** (9), 1519-1525, doi:haematol.2009.021014 [pii] 10.3324/haematol.2009.021014, (2010).

149 Kanduri, M. *et al.* Distinct transcriptional control in major immunogenetic subsets of chronic lymphocytic leukemia exhibiting subset-biased global DNA methylation profiles. *Epigenetics.* **7** (12), 1435-1442, doi:10.4161/epi.22901, (2012).
150 Marincevic, M. *et al.* Distinct gene expression profiles in subsets of chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. *Haematologica.* **95** (12), 2072-2079, doi:10.3324/haematol.2010.028639, (2010).
151 Sutton, L. A. *et al.* Different spectra of recurrent gene mutations in subsets of chronic lymphocytic leukemia harboring stereotyped B-cell receptors. *Haematologica.* **101** (8), 959-967, doi:10.3324/haematol.2016.141812, (2016).
152 Strefford, J. C. *et al.* Distinct patterns of novel gene mutations in poor-prognostic stereotyped subsets of chronic lymphocytic leukemia: the case of SF3B1 and subset #2. *Leukemia.* **27** (11), 2196-2199, doi:10.1038/leu.2013.98 leu201398 [pii], (2013).
153 Papakonstantinou, N. *et al.* Differential microRNA profiles and their functional implications in different immunogenetic subsets of chronic lymphocytic leukemia. *Mol Med.* **19** 115-123, doi:10.2119/molmed.2013.00005, (2013).
154 Rossi, D. *et al.* Stereotyped B-Cell Receptor Is an Independent Risk Factor of Chronic Lymphocytic Leukemia Transformation to Richter Syndrome. *Clinical Cancer Research.* **15** (13), 4415-4422, doi:10.1158/1078-0432.ccr-08-3266, (2009).
155 Baliakas, P. *et al.* Clinical effect of stereotyped B-cell receptor immunoglobulins in chronic lymphocytic leukaemia: a retrospective multicentre study. *The Lancet Haematology.* **1** (2), e74-e84 (2014).
156 Sutton, L. A. *et al.* Antigen selection in B-cell lymphomas--tracing the evidence. *Semin Cancer Biol.* **23** (6), 399-409, doi:10.1016/j.semcancer.2013.07.006 S1044-579X(13)00076-X [pii], (2013).
157 Herve, M. *et al.* Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J Clin Invest.* **115** (6), 1636-1643 (2005).
158 Catera, R. *et al.* Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol Med.* **14** (11-12), 665-674, doi:10.2119/2008-00102.Catera, (2008).
159 Lanemo Myhrinder, A. *et al.* A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood.* **111** (7), 3838-3848, doi:10.1182/blood-2007-11-125450, (2008).
160 Bergh, A. C. *et al.* Silenced B-cell receptor response to autoantigen in a poor-prognostic subset of chronic lymphocytic leukemia. *Haematologica.* **99** (11), 1722-1730, doi:10.3324/haematol.2014.106054, (2014).
161 Chu, C. C. *et al.* Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood.* **112** (13), 5122-5129, doi:10.1182/blood-2008-06-162024, (2008).
162 Kostareli, E. *et al.* Molecular evidence for EBV and CMV persistence in a subset of patients with chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. *Leukemia.* **23** (5), 919-924 (2009).
163 Duhren-von Minden, M. *et al.* Chronic lymphocytic leukaemia is driven by antigen-independent cell-autonomous signalling. *Nature.* **489** (7415), 309-312, doi:10.1038/nature11309, (2012).

164 Minici, C. *et al.* Distinct homotypic B-cell receptor interactions shape the outcome of chronic lymphocytic leukaemia. *Nat Commun.* **8** 15746, doi:10.1038/ncomms15746, (2017).
165 Chiorazzi, N. & Ferrarini, M. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. *Blood.* **117** (6), 1781-1791, doi:10.1182/blood-2010-07-155663, (2011).
166 Seifert, M. *et al.* Cellular origin and pathophysiology of chronic lymphocytic leukemia. *J Exp Med.* **209** (12), 2183-2198, doi:10.1084/jem.20120833, (2012).
167 Kulis, M. *et al.* Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet.* **44** (11), 1236-1242, doi:10.1038/ng.2443, (2012).
168 Oakes, C. C. *et al.* DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat Genet.* **48** (3), 253-264, doi:10.1038/ng.3488, (2016).
169 Vecchio, L., Seke Etet, P. F., Kipanyula, M. J., Krampera, M. & Nwabo Kamdje, A. H. Importance of epigenetic changes in cancer etiology, pathogenesis, clinical profiling, and treatment: what can be learned from hematologic malignancies? *Biochim Biophys Acta.* **1836** (1), 90-104, doi:10.1016/j.bbcan.2013.04.001, (2013).
170 Fedoriw, A., Mugford, J. & Magnuson, T. Genomic imprinting and epigenetic control of development. *Cold Spring Harb Perspect Biol.* **4** (7), a008136, doi:10.1101/cshperspect.a008136, (2012).
171 Sharma, S., Kelly, T. K. & Jones, P. A. Epigenetics in cancer. *Carcinogenesis.* **31** (1), 27-36, doi:10.1093/carcin/bgp220, (2010).
172 Wahlfors, J. *et al.* Genomic hypomethylation in human chronic lymphocytic leukemia. *Blood.* **80** (8), 2074-2080 (1992).
173 Rush, L. J. *et al.* Epigenetic profiling in chronic lymphocytic leukemia reveals novel methylation targets. *Cancer Res.* **64** (7), 2424-2433 (2004).
174 Moreno, P. *et al.* Lipoprotein lipase expression in unmutated CLL patients is the consequence of a demethylation process induced by the microenvironment. *Leukemia.* **27** (3), 721-725, doi:10.1038/leu.2012.212, (2013).
175 Raval, A. *et al.* Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. *Cell.* **129** (5), 879-890, doi:10.1016/j.cell.2007.03.043, (2007).
176 Liu, T. H. *et al.* CpG island methylation and expression of the secreted frizzled-related protein gene family in chronic lymphocytic leukemia. *Cancer Res.* **66** (2), 653-658, doi:10.1158/0008-5472.CAN-05-3712, (2006).
177 Chen, S. S. *et al.* Silencing of the inhibitor of DNA binding protein 4 (ID4) contributes to the pathogenesis of mouse and human CLL. *Blood.* **117** (3), 862-871, doi:10.1182/blood-2010-05-284638, (2011).
178 Raval, A. *et al.* TWIST2 demonstrates differential methylation in immunoglobulin variable heavy chain mutated and unmutated chronic lymphocytic leukemia. *J Clin Oncol.* **23** (17), 3877-3885, doi:10.1200/JCO.2005.02.196, (2005).
179 Abreu, C. *et al.* Methylation status regulates lipoprotein lipase expression in chronic lymphocytic leukemia. *Leuk Lymphoma.* **54** (8), 1844-1848, doi:10.3109/10428194.2013.796057, (2013).
180 Corcoran, M. *et al.* ZAP-70 methylation status is associated with ZAP-70 expression status in chronic lymphocytic leukemia. *Haematologica.* **90** (8), 1078-1088 (2005).

181 Cahill, N. *et al.* 450K-array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. *Leukemia.* **27** (1), 150-158, doi:10.1038/leu.2012.245, (2013).
182 Kulis, M. *et al.* Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet.* **47** (7), 746-756, doi:10.1038/ng.3291, (2015).
183 Oakes, C. C. *et al.* Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov.* **4** (3), 348-361, doi:10.1158/2159-8290.CD-13-0349, (2014).
184 Queiros, A. C. *et al.* A B-cell epigenetic signature defines three biologic subgroups of chronic lymphocytic leukemia with clinical impact. *Leukemia.* **29** (3), 598-605, doi:10.1038/leu.2014.252, (2015).
185 Locksley, R. M., Killeen, N. & Lenardo, M. J. The TNF and TNF receptor superfamilies: integrating mammalian biology. *Cell.* **104** (4), 487-501 (2001).
186 Ferrajoli, A. *et al.* The clinical significance of tumor necrosis factor-alpha plasma level in patients having chronic lymphocytic leukemia. *Blood.* **100** (4), 1215-1219 (2002).
187 van Horssen, R., Ten Hagen, T. L. & Eggermont, A. M. TNF-alpha in cancer treatment: molecular insights, antitumor effects, and clinical utility. *Oncologist.* **11** (4), 397-408, doi:10.1634/theoncologist.11-4-397, (2006).
188 Zheng, H. *et al.* Characterization of a cDNA encoding a protein with limited similarity to beta1, 3-N-acetylglucosaminyltransferase. *Mol Biol Rep.* **31** (3), 171-175 (2004).
189 Furusawa, T., Moribe, H., Kondoh, H. & Higashi, Y. Identification of CtBP1 and CtBP2 as corepressors of zinc finger-homeodomain factor deltaEF1. *Mol Cell Biol.* **19** (12), 8581-8590 (1999).
190 Paliwal, S., Ho, N., Parker, D. & Grossman, S. R. CtBP2 Promotes Human Cancer Cell Migration by Transcriptional Activation of Tiam1. *Genes Cancer.* **3** (7-8), 481-490, doi:10.1177/1947601912463695, (2012).
191 Liu, X., Yao, N., Qian, J. & Huang, H. High expression and prognostic role of CAP1 and CtBP2 in breast carcinoma: associated with E-cadherin and cell proliferation. *Med Oncol.* **31** (3), 878, doi:10.1007/s12032-014-0878-7, (2014).
192 Zhang, C., Gao, C., Xu, Y. & Zhang, Z. CtBP2 could promote prostate cancer cell proliferation through c-Myc signaling. *Gene.* **546** (1), 73-79, doi:10.1016/j.gene.2014.05.032, (2014).
193 Calin, G. A. & Croce, C. M. Genomics of chronic lymphocytic leukemia microRNAs as new players with clinical significance. *Semin Oncol.* **33** (2), 167-173 (2006).
194 Calin, G. A. *et al.* A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med.* **353** (17), 1793-1801 (2005).
195 Calin, G. A. *et al.* Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A.* **101** (9), 2999-3004 (2004).
196 Stamatopoulos, B. *et al.* microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. *Blood.* **113** (21), 5237-5245, doi:10.1182/blood-2008-11-189407, (2009).

197 Mraz, M., Pospisilova, S., Malinova, K., Slapak, I. & Mayer, J. MicroRNAs in chronic lymphocytic leukemia pathogenesis and disease subtypes. *Leuk Lymphoma.* **50** (3), 506-509, doi:10.1080/10428190902763517, (2009).
198 Mraz, M. *et al.* miR-34a, miR-29c and miR-17-5p are downregulated in CLL patients with TP53 abnormalities. *Leukemia.* **23** (6), 1159-1163, doi:10.1038/leu.2008.377, (2009).
199 Ferracin, M. *et al.* MicroRNAs involvement in fludarabine refractory chronic lymphocytic leukemia. *Mol Cancer.* **9** 123, doi:10.1186/1476-4598-9-123, (2010).
200 Calin, G. A. *et al.* MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A.* **101** (32), 11755-11760 (2004).
201 Balatti, V., Pekarky, Y., Rizzotto, L. & Croce, C. M. miR deregulation in CLL. *Adv Exp Med Biol.* **792** 309-325, doi:10.1007/978-1-4614-8051-8_14, (2013).
202 Rossi, S. *et al.* microRNA fingerprinting of CLL patients with chromosome 17p deletion identify a miR-21 score that stratifies early survival. *Blood.* **116** (6), 945-952, doi:10.1182/blood-2010-01-263889, (2010).
203 Cui, B. *et al.* MicroRNA-155 influences B-cell receptor signaling and associates with aggressive disease in chronic lymphocytic leukemia. *Blood.* **124** (4), 546-554, doi:10.1182/blood-2014-03-559690, (2014).
204 Baer, C. *et al.* Extensive promoter DNA hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia. *Cancer Res.* **72** (15), 3775-3785, doi:10.1158/0008-5472.CAN-12-0803, (2012).
205 Deneberg, S. *et al.* microRNA-34b/c on chromosome 11q23 is aberrantly methylated in chronic lymphocytic leukemia. *Epigenetics.* **9** (6), 910-917, doi:10.4161/epi.28603, (2014).
206 Wang, L. Q. *et al.* Epigenetic inactivation of miR-9 family microRNAs in chronic lymphocytic leukemia--implications on constitutive activation of NFkappaB pathway. *Mol Cancer.* **12** 173, doi:10.1186/1476-4598-12-173, (2013).
207 Baer, C. *et al.* Epigenetic silencing of miR-708 enhances NF-kappaB signaling in chronic lymphocytic leukemia. *Int J Cancer.* **137** (6), 1352-1361, doi:10.1002/ijc.29491, (2015).
208 Zou, Z. J. *et al.* miR-26a and miR-214 down-regulate expression of the PTEN gene in chronic lymphocytic leukemia, but not PTEN mutation or promoter methylation. *Oncotarget.* **6** (2), 1276-1285, doi:10.18632/oncotarget.2626, (2015).
209 Zhao, X. *et al.* Disruption of the MYC-miRNA-EZH2 loop to suppress aggressive B-cell lymphoma survival and clonogenicity. *Leukemia.* **27** (12), 2341-2350, doi:10.1038/leu.2013.94, (2013).
210 Jares, P. & Campo, E. Advances in the understanding of mantle cell lymphoma. *Br J Haematol.* **142** (2), 149-165, doi:10.1111/j.1365-2141.2008.07124.x, (2008).
211 Cheah, C. Y., Seymour, J. F. & Wang, M. L. Mantle Cell Lymphoma. *J Clin Oncol.* **34** (11), 1256-1269, doi:10.1200/JCO.2015.63.5904, (2016).
212 Perez-Galan, P., Dreyling, M. & Wiestner, A. Mantle cell lymphoma: biology, pathogenesis, and the molecular basis of treatment in the genomic era. *Blood.* **117** (1), 26-38, doi:10.1182/blood-2010-04-189977, (2011).

213 Jares, P., Colomer, D. & Campo, E. Genetic and molecular pathogenesis of mantle cell lymphoma: perspectives for new targeted therapeutics. *Nat Rev Cancer.* **7** (10), 750-762, doi:10.1038/nrc2230, (2007).
214 Fernandez, V. *et al.* Genomic and gene expression profiling defines indolent forms of mantle cell lymphoma. *Cancer Res.* **70** (4), 1408-1418, doi:10.1158/0008-5472.CAN-09-3419, (2010).
215 Ondrejka, S. L., Lai, R., Smith, S. D. & Hsi, E. D. Indolent mantle cell leukemia: a clinicopathological variant characterized by isolated lymphocytosis, interstitial bone marrow involvement, kappa light chain restriction, and good prognosis. *Haematologica.* **96** (8), 1121-1127, doi:10.3324/haematol.2010.036277, (2011).
216 Jares, P., Colomer, D. & Campo, E. Molecular pathogenesis of mantle cell lymphoma. *J Clin Invest.* **122** (10), 3416-3423, doi:10.1172/JCI61272, (2012).
217 Raffeld, M. & Jaffe, E. S. bcl-1, t(11;14), and mantle cell-derived lymphomas. *Blood.* **78** (2), 259-263 (1991).
218 Fu, K. *et al.* Cyclin D1-negative mantle cell lymphoma: a clinicopathologic study based on gene expression profiling. *Blood.* **106** (13), 4315-4321, doi:10.1182/blood-2005-04-1753, (2005).
219 Gladden, A. B., Woolery, R., Aggarwal, P., Wasik, M. A. & Diehl, J. A. Expression of constitutively nuclear cyclin D1 in murine lymphocytes induces B-cell lymphoma. *Oncogene.* **25** (7), 998-1007, doi:10.1038/sj.onc.1209147, (2006).
220 Bea, S. *et al.* Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proc Natl Acad Sci U S A.* **110** (45), 18250-18255, doi:10.1073/pnas.1314608110, (2013).
221 Campo, E. & Rule, S. Mantle cell lymphoma: evolving management strategies. *Blood.* **125** (1), 48-55, doi:10.1182/blood-2014-05-521898, (2015).
222 Zhang, J. *et al.* The genomic landscape of mantle cell lymphoma is related to the epigenetically determined chromatin state of normal B cells. *Blood.* **123** (19), 2988-2996, doi:10.1182/blood-2013-07-517177, (2014).
223 Greiner, T. C. *et al.* Mutation and genomic deletion status of ataxia telangiectasia mutated (ATM) and p53 confer specific gene expression profiles in mantle cell lymphoma. *Proc Natl Acad Sci U S A.* **103** (7), 2352-2357, doi:10.1073/pnas.0510441103, (2006).
224 Halldorsdottir, A. M. *et al.* Impact of TP53 mutation and 17p deletion in mantle cell lymphoma. *Leukemia.* **25** (12), 1904-1908, doi:10.1038/leu.2011.162, (2011).
225 Kridel, R. *et al.* Whole transcriptome sequencing reveals recurrent NOTCH1 mutations in mantle cell lymphoma. *Blood.* **119** (9), 1963-1971, doi:10.1182/blood-2011-11-391474, (2012).
226 Rahal, R. *et al.* Pharmacological and genomic profiling identifies NF-kappaB-targeted treatment strategies for mantle cell lymphoma. *Nat Med.* **20** (1), 87-92, doi:10.1038/nm.3435, (2014).
227 Fakhri, B. & Kahl, B. Current and emerging treatment options for mantle cell lymphoma. *Ther Adv Hematol.* **8** (8), 223-234, doi:10.1177/2040620717719616, (2017).
228 Brody, J. & Advani, R. Treatment of mantle cell lymphoma: current approach and future directions. *Crit Rev Oncol Hematol.* **58** (3), 257-265, doi:10.1016/j.critrevonc.2005.10.001, (2006).

229 Park, I. W., Reddy, M. V., Reddy, E. P. & Groopman, J. E. Evaluation of novel cell cycle inhibitors in mantle cell lymphoma. *Oncogene.* **26** (38), 5635-5642, doi:10.1038/sj.onc.1210350, (2007).
230 O'Connor, O. A. *et al.* Phase II clinical experience with the novel proteasome inhibitor bortezomib in patients with indolent non-Hodgkin's lymphoma and mantle cell lymphoma. *J Clin Oncol.* **23** (4), 676-684, doi:10.1200/JCO.2005.02.050, (2005).
231 Heider, U. *et al.* Synergistic interaction of the histone deacetylase inhibitor SAHA with the proteasome inhibitor bortezomib in mantle cell lymphoma. *Eur J Haematol.* **80** (2), 133-142, doi:10.1111/j.1600-0609.2007.00995.x, (2008).
232 Bernard, S. *et al.* Inhibitors of BCR signalling interrupt the survival signal mediated by the micro-environment in mantle cell lymphoma. *Int J Cancer.* **136** (12), 2761-2774, doi:10.1002/ijc.29326, (2015).
233 Rosenwald, A. *et al.* The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. *Cancer Cell.* **3** (2), 185-197 (2003).
234 Tiemann, M. *et al.* Histopathology, cell proliferation indices and clinical outcome in 304 patients with mantle cell lymphoma (MCL): a clinicopathological study from the European MCL Network. *Br J Haematol.* **131** (1), 29-38, doi:10.1111/j.1365-2141.2005.05716.x, (2005).
235 Hoster, E. *et al.* A new prognostic index (MIPI) for patients with advanced-stage mantle cell lymphoma. *Blood.* **111** (2), 558-565, doi:10.1182/blood-2007-06-095331, (2008).
236 Geisler, C. H. *et al.* The Mantle Cell Lymphoma International Prognostic Index (MIPI) is superior to the International Prognostic Index (IPI) in predicting survival following intensive first-line immunochemotherapy and autologous stem cell transplantation (ASCT). *Blood.* **115** (8), 1530-1533, doi:10.1182/blood-2009-08-236570, (2010).
237 Hoster, E. *et al.* Prognostic Value of Ki-67 Index, Cytology, and Growth Pattern in Mantle-Cell Lymphoma: Results From Randomized Trials of the European Mantle Cell Lymphoma Network. *J Clin Oncol.* **34** (12), 1386-1394, doi:10.1200/JCO.2015.63.8387, (2016).
238 Kienle, D. *et al.* VH mutation status and VDJ rearrangement structure in mantle cell lymphoma: correlation with genomic aberrations, clinical characteristics, and outcome. *Blood.* **102** (8), 3003-3009, doi:10.1182/blood-2003-05-1383, (2003).
239 Hadzidimitriou, A. *et al.* Is there a role for antigen selection in mantle cell lymphoma? Immunogenetic support from a series of 807 cases. *Blood.* **118** (11), 3088-3095, doi:10.1182/blood-2011-03-343434, (2011).
240 Thorselius, M. *et al.* Somatic hypermutation and V(H) gene usage in mantle cell lymphoma. *Eur J Haematol.* **68** (4), 217-224 (2002).
241 Walsh, S. H. *et al.* Mutated VH genes and preferential VH3-21 use define new subsets of mantle cell lymphoma. *Blood.* **101** (10), 4047-4054, doi:10.1182/blood-2002-11-3479, (2003).
242 Camacho, F. I. *et al.* Molecular heterogeneity in MCL defined by the use of specific VH genes and the frequency of somatic mutations. *Blood.* **101** (10), 4042-4046, doi:10.1182/blood-2002-11-3456, (2003).
243 Di Lisio, L. *et al.* Mantle cell lymphoma: transcriptional regulation by microRNAs. *Leukemia.* **24** (7), 1335-1342, doi:10.1038/leu.2010.91, (2010).

244 Navarro, A. *et al.* MicroRNA expression, chromosomal alterations, and immunoglobulin variable heavy chain hypermutations in Mantle cell lymphomas. *Cancer Res.* **69** (17), 7071-7078, doi:10.1158/0008-5472.CAN-09-1095, (2009).
245 Zhao, J. J. *et al.* microRNA expression profile and identification of miR-29 as a prognostic marker and pathogenetic factor by targeting CDK6 in mantle cell lymphoma. *Blood.* **115** (13), 2630-2639, doi:10.1182/blood-2009-09-243147, (2010).
246 Iqbal, J. *et al.* Genome-wide miRNA profiling of mantle cell lymphoma reveals a distinct subgroup with poor prognosis. *Blood.* **119** (21), 4939-4948, doi:10.1182/blood-2011-07-370122, (2012).
247 Yim, R. L. *et al.* Methylation of miR-155-3p in mantle cell lymphoma and other non-Hodgkin's lymphomas. *Oncotarget.* **5** (20), 9770-9782, doi:10.18632/oncotarget.2390, (2014).
248 Clark, S. J., Statham, A., Stirzaker, C., Molloy, P. L. & Frommer, M. DNA methylation: bisulphite modification and analysis. *Nat Protoc.* **1** (5), 2353-2364, doi:10.1038/nprot.2006.324, (2006).
249 Sandoval, J. *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics.* **6** (6), 692-702 (2011).
250 https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html.
251 Dupont, J. M., Tost, J., Jammes, H. & Gut, I. G. De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal Biochem.* **333** (1), 119-127, doi:10.1016/j.ab.2004.05.007, (2004).
252 Mikeska, T., Felsberg, J., Hewitt, C. A. & Dobrovic, A. Analysing DNA methylation using bisulphite pyrosequencing. *Methods Mol Biol.* **791** 33-53, doi:10.1007/978-1-61779-316-5_4, (2011).
253 Tost, J. & Gut, I. G. Analysis of gene-specific DNA methylation patterns by pyrosequencing technology. *Methods Mol Biol.* **373** 89-102, doi:10.1385/1-59745-377-3:89, (2007).
254 Ronaghi, M., Uhlen, M. & Nyren, P. A sequencing method based on real-time pyrophosphate. *Science.* **281** (5375), 363, 365 (1998).
255 Ronaghi, M. Pyrosequencing sheds light on DNA sequencing. *Genome Res.* **11** (1), 3-11 (2001).
256 Hallek, M. *et al.* Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood.* **111** (12), 5446-5456, doi:10.1182/blood-2007-06-093906, (2008).
257 Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* **30** (10), 1363-1369, doi:10.1093/bioinformatics/btu049, (2014).
258 Smedby, K. E. *et al.* Ultraviolet radiation exposure and risk of malignant lymphomas. *J Natl Cancer Inst.* **97** (3), 199-209 (2005).
259 Sander, S. *et al.* MYC stimulates EZH2 expression by repression of its negative regulator miR-26a. *Blood.* **112** (10), 4202-4212, doi:10.1182/blood-2008-03-147645, (2008).
260 Lu, J. *et al.* MiR-26a inhibits cell growth and tumorigenesis of nasopharyngeal carcinoma through repression of EZH2. *Cancer Res.* **71** (1), 225-233, doi:10.1158/0008-5472.CAN-10-1850, (2011).

261 Zhang, X. *et al.* Coordinated silencing of MYC-mediated miR-29 by HDAC3 and EZH2 as a therapeutic target of histone modification in aggressive B-Cell lymphomas. *Cancer Cell.* **22** (4), 506-523, doi:10.1016/j.ccr.2012.09.003, (2012).
262 Kanduri, M. *et al.* Differential genome-wide array-based methylation profiles in prognostic subsets of chronic lymphocytic leukemia. *Blood.* **115** (2), 296-305, doi:10.1182/blood-2009-07-232868, (2010).
263 Queiros, A. C. *et al.* Decoding the DNA Methylome of Mantle Cell Lymphoma in the Light of the Entire B Cell Lineage. *Cancer Cell.* **30** (5), 806-821, doi:10.1016/j.ccell.2016.09.014, (2016).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1363

Editor: The Dean of the Faculty of Medicine

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

Distribution: publications.uu.se
urn:nbn:se:uu:diva-328616

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2017