THE ROLE OF SOXC TRANSCRIPTION FACTORS IN B-CELL DEVELOPMENT AND LYMPHOID MALIGNANCIES

Martin Lord

Stockholm 2017
To my family
ABSTRACT

Mantle cell lymphoma (MCL) accounts for 5-10% out of all Non-Hodgkin lymphomas (NHLs) and is one of the most aggressive forms of lymphomas with a median survival of less than 5 years. Currently, MCL is considered to be an incurable disease.

MCL is characterized by the t(11;14)(q13;q32) CCND1/IGH translocation that results in high expression of cyclin D1. This translocation takes place at the pre-B cell stage and is generally recognized as the hallmark and primary oncogenic event in the evolution of MCL. Recently, the neural transcription factor SRY (sex-determining region Y) box 11 (SOX11) gene was found to be expressed in over 90% of all MCLs. The SOX11 protein is not detected in the vast majority of other lymphomas or mature B-cells and its expression is independent of cyclin D1 status. Moreover, SOX11 has been proposed to have a functional role in the pathogenesis of MCL and may not only serve as a diagnostic biomarker.

In this thesis, the functional role of the SOXC genes (SOX4, SOX11 and SOX12) have been studied in several different ways, both in MCL primary samples/cell lines and in non-MCL related cells with focus on the SOX11 gene.

The SOXC transcription factors are known to compete for the same target genes. For the first time in MCL, the SOXC genes were quantified by qPCR in a set of MCL patients and MCL cell lines. As previously reported, SOX11 expression was high in MCL, but also SOX12 mRNA levels were found to be higher compared to non-malignant B-cells, whereas the expression levels of SOX4 varied. Further, expression of the SOXC genes correlated in SOX11 positive MCL (determined by immunohistochemistry). How SOX11 gene expression in MCL is regulated was also addressed by studying its promotor region. The promotor region of SOX11 was found to be hypomethylated in MCL patients and cell lines, but also in non-malignant B-cells indicating regulation by other epigenetic mechanisms than promotor methylation.

Fast and accurate differentiation between similar entities of lymphoma is important since MCL has a more aggressive clinical course. Although having certain distinctive phenotypical markers, MCL and B-cell chronic lymphocytic leukemia/small lymphocytic lymphoma (B-CLL/SLL) are both CD19+, CD20+ and usually CD5+, which could complicate diagnosis by flow cytometry. We developed a method to accurately implement SOX11 in the diagnostic flow panel that consistently detected SOX11 protein in ex vivo isolated MCL cells, but not in CLL/SLL. When conjugated SOX11-antibodies are available, this method could be implemented in the clinic for CLL/SLL with aberrant immune phenotypes or rare cyclin D1- MCLs.

The expression levels of SOX11 were further studied in a relatively large group of MCL patients (n=102) by qPCR to determine a cut-off for SOX11-negative MCL and to investigate how quantitative expression related to positivity/negativity by IHC. A cut-off was defined, which resulted in misclassification of only 2/102 by qPCR and IHC. However, for the IHC SOX11+ cases, the qPCR analysis was not able to find a natural cut-off that would identify cases with low expression. When grouping the samples based on expression (10% lowest expression versus the remaining cases), nodal disease was less frequent (p=0.01) and lymphocytosis more frequent (p=0.005) in the qPCR SOX11low-cases. Leukemic non-nodal MCL often expresses low levels of SOX11. The quartile of patients with the lowest SOX11 expression had significantly shorter overall survival in the group of patients who did not receive autologous stem cell transplantation.
Studies were conducted in primary murine B-cells and a murine pro-B cell line to study Sox11 oncogenic potential and role in differentiation in early B-cells. In the studied cell types, Sox11 did not per se act as an oncogene. Instead the rate of proliferation was reduced in the pro-B cell line and these cells changed morphology upon expressing the Sox11 gene. Gene expression analysis revealed upregulation of early cell cycle and cellular adhesion genes upon introduction of the Sox11 gene in the pro-B cells. Despite high similarity to Sox4 (important for B-cell survival and development), no obvious effect on selected B-cell differentiation stage associated genes were detected, which suggest that the effects of Sox11 are context dependent and might differ in murine pro-B cells compared to MCL and during embryogenesis.
LIST OF SCIENTIFIC PAPERS


CONTENTS

1 Background...................................................................................................................... 1
  1.1 B-cell biology ................................................................................................... 1
2 Mantle cell lymphoma .................................................................................................... 4
  2.1 Background....................................................................................................... 4
  2.2 Morphological subtypes.................................................................................... 4
  2.3 Phenotype ......................................................................................................... 4
  2.4 Cellular origin................................................................................................... 5
    2.4.1 Immunoglobulin status ......................................................................... 5
    2.4.2 Epigenetics ............................................................................................ 5
  2.5 Disease presentation ......................................................................................... 5
    2.5.1 Cyclin D1 in MCL ................................................................................ 7
    2.5.2 Cyclin D1-negative MCL ..................................................................... 7
    2.5.3 Secondary genetic aberrations .............................................................. 8
  2.6 Current standards of care .................................................................................. 9
3 Sry-related high-mobility-group box (SOX) transcription factors .............................. 11
  3.1 Background ..................................................................................................... 11
  3.2 The sox transcription factors .......................................................................... 11
  3.3 Molecular features .......................................................................................... 13
    3.3.1 DNA binding....................................................................................... 13
    3.3.2 TAD and NLS domain ........................................................................ 14
  3.4 The SOXC transcription factor group............................................................. 14
    3.4.1 Expression ........................................................................................... 14
    3.4.2 Features of the SOXC group.................................................................. 14
    3.4.3 Overlapping function in vivo ................................................................ 14
    3.4.4 SOXC expression in B-cells and cancer ............................................. 15
  3.5 The SOX11 transcription factor...................................................................... 15
4 The role of SOX11 in cancer ........................................................................................ 17
  4.1 Expression pattern .......................................................................................... 17
  4.2 Functional role in non-MCL derived tumors.................................................. 17
5 SOX11 in MCL ............................................................................................................. 18
  5.1 Background..................................................................................................... 18
  5.2 SOX11 expression in MCL, a double-edged sword? ..................................... 18
    5.2.1 Diagnosing MCL ................................................................................ 19
  5.3 Regulation....................................................................................................... 19
  5.4 SOX11 transcriptional targets in MCL........................................................... 20
  5.5 Functional role of SOX11 in MCL ................................................................. 21
    5.5.1 Discrepancies between oncogenic effects of SOX11 in a MCL model .................................................................................................................. 22
6 Comments on the methodology .................................................................................... 24
  6.1 Cell lines and primary cells ............................................................................ 24
  6.2 Inducible expression system ........................................................................... 24
  6.3 Gene expression .............................................................................................. 25
    6.3.1 Quantitative PCR ................................................................................ 25
    6.3.2 DNA microarray ................................................................................. 26
  6.4 Epigenetics....................................................................................................... 26
    6.4.1 McrBC-based methylation assay ........................................................ 27
    6.4.2 Bisulfite treatment and pyrosequencing analysis.................................. 27
  6.5 Protein Detection ............................................................................................ 27
    6.5.1 Western Blot ....................................................................................... 27
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>4C</td>
<td>Circularized Chromosome Conformation Capture</td>
</tr>
<tr>
<td>5-AZA</td>
<td>5-azacytidine</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>ASCT</td>
<td>Autologous stem cell transplantation</td>
</tr>
<tr>
<td>BCR</td>
<td>B-cell receptor</td>
</tr>
<tr>
<td>BL</td>
<td>Burkitt lymphoma</td>
</tr>
<tr>
<td>CCL</td>
<td>Centrocytic lymphoma</td>
</tr>
<tr>
<td>CD</td>
<td>Cluster of differentiation</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CLP</td>
<td>Common lymphoid progenitors</td>
</tr>
<tr>
<td>CMP</td>
<td>Common myeloid progenitors</td>
</tr>
<tr>
<td>cRNA</td>
<td>Complementary RNA</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic cell</td>
</tr>
<tr>
<td>DLBCL</td>
<td>Diffuse large B-cell lymphoma</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>FDC</td>
<td>Follicular dendritic cell</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>FL</td>
<td>Follicular lymphoma</td>
</tr>
<tr>
<td>FSC/SSC</td>
<td>Forward scatter/side scatter</td>
</tr>
<tr>
<td>GC</td>
<td>Germinal center</td>
</tr>
<tr>
<td>GEP</td>
<td>Gene expression profiling</td>
</tr>
<tr>
<td>HCL</td>
<td>Hairy cell leukemia</td>
</tr>
<tr>
<td>HMG</td>
<td>High-Mobility Group</td>
</tr>
<tr>
<td>ICC</td>
<td>Immunocytochemistry</td>
</tr>
<tr>
<td>IDL</td>
<td>Lymphocytic lymphoma of intermediate differentiation</td>
</tr>
<tr>
<td>IGH</td>
<td>Immunoglobulin heavy chain</td>
</tr>
<tr>
<td>IGL</td>
<td>Immunoglobulin light chain</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>ILL</td>
<td>Intermediate lymphocytic lymphoma</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosomal entry site</td>
</tr>
<tr>
<td>ISMCN</td>
<td>In situ mantle cell neoplasia</td>
</tr>
<tr>
<td>LPA</td>
<td>Lysoosphatidic acid</td>
</tr>
<tr>
<td>MCL</td>
<td>Mantle cell lymphoma</td>
</tr>
<tr>
<td>MFI</td>
<td>Mean fluorescence intensity</td>
</tr>
<tr>
<td>MPP</td>
<td>Multipotential progenitor</td>
</tr>
<tr>
<td>MZL</td>
<td>Marginal zone lymphoma</td>
</tr>
<tr>
<td>NGS</td>
<td>Next-generation sequencing</td>
</tr>
<tr>
<td>NHL</td>
<td>Non-Hodgkin lymphoma</td>
</tr>
<tr>
<td>NK-cell</td>
<td>Natural killer cell</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization sequence</td>
</tr>
<tr>
<td>nnMCL</td>
<td>Non-nodal MCL</td>
</tr>
<tr>
<td>OS</td>
<td>Overall survival</td>
</tr>
<tr>
<td>PHSC</td>
<td>Pluripotent Hematopoietic stem cell</td>
</tr>
<tr>
<td>PMBL</td>
<td>Primary mediastinal B-cell lymphoma</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RFI</td>
<td>Relative fold increase</td>
</tr>
<tr>
<td>RMA</td>
<td>Robust Multichip Analysis</td>
</tr>
<tr>
<td>SAHA</td>
<td>Suberoylanilide hydroxamic acid</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SHM</td>
<td>Somatic hypermutation</td>
</tr>
<tr>
<td>shRNA</td>
<td>Short-hairpin RNA</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>SOX</td>
<td>SRY (sex-determining region Y) box</td>
</tr>
<tr>
<td>TAD</td>
<td>Transactivation domain</td>
</tr>
<tr>
<td>TFH-cell</td>
<td>T follicular helper cell</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>WB</td>
<td>Western blot</td>
</tr>
<tr>
<td>XTT</td>
<td>2,3-Bis-(2-Methoxy-4-Nitro-5-Sulfophenyl)-2H-Tetrazolium-5-Carboxanilide</td>
</tr>
</tbody>
</table>
Comments about terminology
Gene names are referred to as XXX (human) and Xxx (mouse). Protein names are referred to as XXX (human) and XXX (mouse), but the text specifies the type of species. When describing the functional role of a transcription factor, it is always XXX.

“XXX expression” is an ambiguous term to define cases where both protein and mRNA levels are detected or discussed. For specific mRNA expression, the term “XXVXxx mRNA expression” is used instead.
1 BACKGROUND

1.1 B-CELL BIOLOGY

Gnathostomes (vertebrates with jaws), the metazoan lineage to which humans belong, originated 500 million years ago together with our adaptive immune system [1]. The adaptive immune system includes B (bursa/bone marrow) and T (thymus) cells, whose purpose is to recognize and respond to different antigenic epitopes or toxins. The delineation of the B-cell and T-cell lineages began in 1965 with experiments by Cooper et al. conducted in chicken (hence the name bursa, a lymphoid organ in birds) [2].

In humans, B-cells are derived from pluripotent hematopoietic stem cells (PHSCs). The hematopoietic system constantly generates a large number of specialized cell types from PHSCs with self-renewal potential. The PHSCs can give rise to all hematopoietic cells via sequential stages of differentiation described below (based on the following reviews [3, 4]). At each new stage, the progenitor cells encounter different binary choices from which they cannot return, and thus restrict their capacity to differentiate as they continue to progress through cellular development.

Differentiated PHSCs become multipotential progenitors (MPPs). These cells are able to differentiate into either early lymphoid progenitors (ELPs) or common myeloid progenitors (CMPs). The transcription factor PU.1 (encoded by \textit{ Spi1}) has been shown to be an important regulator at this stage as \textit{ Spi1-}knockout mice die absent of B/T-cells, monocytes and granulocytes prior to birth [5]. The CMPs are limited to the myeloid or erythroid lineages and cannot differentiate into any lymphocyte lineage. The ELPs further differentiate into common lymphoid progenitors (CLPs) with the capacity of differentiating into B-cells, T-cells, dendritic cells (DCs) or natural killer (NK) cells. In the specific context of B-cell development, successful differentiation from CLPs into immature B-cells and more differentiated B-cells requires expression of several transcription factors to activate genes important for further differentiation, especially SOX4 [6], E2A [7], PAX5 [8] and EBF [9]. Absence of these genes result in differentiation blockades at the pro/pre-B cell stage [10].

Rearrangement of the immunoglobulin heavy chain (\textit{ IGH}) and the immunoglobulin light chain (\textit{ IGL}) gene segments by RAG1/2 occur at the pro-B cell stage after the late CLP stage. First is the recombination of the \textit{ IGHD} genes (diversity) to \textit{ IGHJ} genes (joining), which is followed by \textit{ IGHV} genes (variable) to \textit{ IGHDJ}, encoding the Ig \textit{\mu} heavy chain protein [11]. To validate that the Ig \textit{\mu} heavy chains are functional, a pre-B cell receptor (pre-BCR) is formed, constituted of the surrogate light chains \textit{\lambda}5 and VpreB and the Ig \textit{\mu} heavy chain [12]. Successful pre-BCR complex formation leads to proliferation of the parent cell and the rearrangement of \textit{ IGLV} to \textit{ IGLJ}, encoding the \kappa or \textit{\lambda} chains on the surface of the pre-B cell. The outcome of the \textit{ IGHVDJ} and \textit{ IGLVJ} rearrangement is an immature B-cell with a complete IgM displayed on the cell surface comprised of the Ig \textit{\mu} heavy chain and the \kappa or \textit{\lambda} chain proteins. Subsequently, the newly formed immature B-cells are selected not to react with self-antigens before permitted to access the blood circulation [13]. The cells gain access to primary lymphoid follicles where they start to mature and express IgD. At this stage, the transcriptional repressor BCL6 has an important function for germinal center (GC) formation, but also in preventing pre-mature activation and differentiation of B-cell in the GC [14].
Formation of the GC starts in the lymphoid follicle when the mature naïve B-cells become activated by an exogenous antigen. Following interaction with T-cells in the T-cell zone of the follicle, the activated B-cells can either differentiate into short-lived plasma cells in the lymph node (usually those with high-affinity BCR) or long-lived memory B-cells. The GC is divided into two compartments, the dark zone (affinity maturation) and the light zone (selection). In order to generate high-affinity antibodies, the B-cells undergo somatic hypermutation (SHM) of the \textit{IGH} and \textit{IGL} and immunoglobulin isotype-switching by the DNA-editing deaminase enzyme AID. Upon proliferation and SHM in the dark zone, the cells enter the light zone where selection is based on BCR-affinity to the immunizing antigen.

The B-cell can either encounter non-bound antigens or antigens presented on follicular dendritic cells (FDCs). The BCR binds and sequester free antigens into peptides before presenting them to T follicular helper cells (TFH-cells) on the surface in the context of its MHC-II complex. Hence, survival signals to the B-cells are provided by MHC-II and TFH-cell interaction and interaction with the antigen presenting FDCs and the BCR. High affinity of the BCR to the antigen results in outcompeting those B-cells with lower affinity in binding to FDCs and higher concentration of peptides presented on the surface to TFHs, generating more survival signals. Thus, the B-cells with the highest BCR affinity survive and are selected. These can either re-enter the dark zone for further SHM or differentiate into memory B-cells or plasma cells [15, 16]. Exit from the GC occurs from the light zone and is mainly promoted by reduction of BCL6. The plasma cell differentiation program is induced by the transcription factors BLIMP1, XBP1 and IRF4, resulting in the subsequent repression of \textit{BCL6} and \textit{PAX5} [14].

There are several lymphomas that arise from B-cells. A cancer derived from a B-lymphocyte is referred to as a B-cell lymphoma. The B-cell lymphomas often resemble specific stages of B-cell differentiation with respect to morphology, phenotype and gene expression profile, as illustrated for a selection of lymphomas in Figure 1.
2 MANTLE CELL LYMPHOMA

2.1 BACKGROUND
Mantle cell lymphoma (MCL) was first described as centrocytic lymphoma (CCL) in the early 1970s according to the Kiel classification (updated classification [17]).

MCL was originally considered to be exclusively stemming from naïve B-cells expressing CD5. However, newer findings now suggest that subsets of MCLs are derived from antigen-experienced B-cells [18-21]. Based on morphology, it was difficult to differentiate CCL from several closely related Non-Hodgkin lymphoma’s (NHLs), especially intermediate lymphocytic lymphoma (ILL)/lymphocytic lymphoma of intermediate differentiation (IDL) [22]. CCL and ILL/IDL were defined as a B-cell neoplasms consisting of small to medium sized lymphoid cells similar to centrocytes or cleaved follicular center cells [23] making the two entities very similar.

A breakthrough in recognizing MCL came after the discovery of the t(11;14)(q13;q32) \( CCND1/IGH \) translocation that results in overexpression of cyclin D1. Based on morphological, immunological and molecular data, Banks et al. [24] proposed the collective term mantle cell lymphoma (MCL) for CCL and several of its highly similar entities since they were all found to overexpress cyclin D1. MCL is also the term used by The World Health Organization (WHO) lymphoma classification [23].

Cyclin D1 was fundamental in facilitating differential diagnoses of MCL from morphologically similar lymphomas, but also exposing the vast morphological spectrum of the disease. This breakthrough in diagnosis of MCL was much later followed by discovery that the transcription factor \( SRY \) (sex-determining region Y) box 11 (SOX11) is overexpressed in more than 90% of all MCLs. Importantly, SOX11 expression is not detected in highly similar lymphomas and its expression is independent of cyclin D1 [25, 26].

2.2 MORPHOLOGICAL SUBTYPES
MCL have three different growth patterns: nodular, diffuse and mantle zone growth pattern [23]. Currently, based on cytology, there are four different variants of MCL recognized: classical, small cell, blastic/blastoid and pleomorphic (mantle zone; diffuse variant) [27] (reviewed in [28]). Classical MCL is frequently defined by small to medium sized lymphoid cells with irregular nucleus and dispersed chromatin structure; the small cell subtype is defined by small round lymphocytes, clumped chromatin and high resemblances to chronic lymphocytic leukemia (CLL); the aggressive blastoid variant is defined by high mitotic rate and their resemblance to lymphoblasts, and the pleomorphic subtype is characterized by often apparent nucleoli and cleaved nuclei [23, 28].

2.3 PHENOTYPE
In addition to cyclin D1+ and SOX11+ and intense expression of IgM and/or IgD, MCL cells express the following B-cell markers: CD19, CD20, CD22, CD79a, CD5, CD24, CD43, FMC7 and BCL2, but are generally negative for CD10, CD11c, CD23, BCL6 and CD200 [23, 28].
2.4 CELLULAR ORIGIN

2.4.1 Immunoglobulin status
Based on the large number of cases with unmutated IGHV, intense expression of IgM and/or IgD, CD5+ expression and mantle zone growth, the corresponding counterpart for MCL was hypothesized to stem from pre-germinal center naïve B-cells [29]. However, this concept has recently been challenged [18, 19, 21, 30, 31]. MCL has also been proposed to stem from a mature B-cell population discovered in human tonsils [32]. The subpopulation could be an intermediate of naïve and GC cells (IgD+CD38–CD23–FSehiCD71+). This particular subpopulation is mostly CD5-, whereas classical MCL is mostly CD5+.

Not all MCLs are completely antigen-inexperienced with unmutated IGHV. Based on established cut-off values (>2% difference from germ line identity), roughly 75% of all cases were unmutated [18]. However, when applying a more stringent criterion (not having a single somatic hypermutation) for defining unmutated MCL on the very same cases, merely 29.5% (238/807) were truly unmutated (TU) [18]. This was confirmed in later studies, resulting in 24% (43/177) and 31.5% (40/127) to be TU MCL cases [33, 34]. Navarro et al. [33] further reported enrichment in naïve B-cell signatures for TU compared with enrichment in memory B-cell signatures for highly IGHV mutated cases.

2.4.2 Epigenetics
The methylation profile of MCLs can also provide information to better understand different subpopulations. Although the two subgroups, one GC-inexperienced (ranging from naïve B-cells to pro-GC B-cells with low frequency of SHM) and one GC-experienced population, revealed by genome-wide DNA methylation were very heterogeneous it clearly reflected two distinct entities of MCL. [21]. Similar to previous studies with TU definition of IGHV for MCL [18, 33, 34], the majority of MCLs demonstrated a DNA methylation pattern more similar to that of antigen-experienced cells.

Another study [35] implemented epigenetics to delineate the cellular origin of MCL by comparing it to normal naïve and GC-experienced B-cells. Briefly, the study demonstrated that the epigenetically defined regions of open chromatin (H3K36me3, H3ac, and H3K4me1) for naïve B-cells and GC-experienced B-cells had a highly significant overlap with the corresponding gene expression profiles that distinguish MCL and Burkitt lymphoma (BL, a GC-experienced lymphoma). The genes with higher/lower chromatin marker expression in naïve B-cells corresponded to higher/lower expression for the same genes in MCL, suggesting that the cellular counterpart of the MCL likely is naïve B-cells.

2.5 DISEASE PRESENTATION
MCL only accounts for 5-10% out of all NHLs, but it is one of the most aggressive forms of lymphomas, mainly in older men [19, 36-38]. It is characterized by rapid relapse after treatment and resistance to therapy. An overview of the disease development is outlined in Figure 2. At diagnosis, the neoplastic cells are often already disseminated into lymph nodes, peripheral blood, bone marrow, spleen and the gastrointestinal tract (GI) [39, 40]. Patients with primary disease presentation in the lymph nodes (classical MCL) have significantly lower survival compared to primary appearance in the extra nodal sites (non-nodal MCL), such as the GI, head or neck and hematologic/reticuloendothelial systems (bone marrow and/or spleen) was emphasized in a recently published study [41]. Upon disease onset, the median survival is generally less than 5 years [42] with the most aggressive blastoid variant of MCL comprising 20–30% of all diagnosed cases [43].
Classical MCL can evolve into blastoid MCL and blastoid MCL can relapse as classical MCL; however, *de novo* blastoid MCLs still constitute the majority of cases [44].

Currently, MCL is considered to be an incurable disease; however, there are subsets of MCLs that exhibit a more indolent disease with longer survival and no need for immediate treatment [45-48]. This subgroup can constitute up to 30% of all MCLs [49]. According to the latest WHO classification of lymphoid neoplasms [50], there are now two separate indolent subtypes; *in situ* mantle cell neoplasia (ISMCN) and the leukemic non-nodal MCL (nnMCL). ISMCN is characterized by cyclin D1+ cells situated at the inner mantle zone of follicles of otherwise healthy lymph nodes or tissue, which very infrequently can evolve into classical MCL [51]. The classical aggressive MCL involves the lymph node, often express SOX11, and is associated with an aggressive disease progression (developing to blastoid or pleomorphic). In contrast, nnMCL often involves peripheral blood, bone marrow and occasionally the spleen and is considered to be more indolent. It is characterized by longer survival, hypermutated *IGHV*, low genetic complexity, downregulation of cell adhesion genes and genes of DNA damage repair pathways. This subtype is also often found to have low expression of SOX11 (reviewed by Jares et al. [52]). However, nnMCL can progress to an aggressive form after genetic alterations of *TP53* [53], and of importance, not all indolent MCLs are associated with low expression of SOX11 [54].

2.5.1 Cyclin D1 in MCL
Initially termed B-cell leukemia/lymphoma 1 (BCL1) based on the assumption of an oncogene existing at that locus, the t(11;14) BCL1/IGH breakpoint was first cloned from a, retrospectively, misdiagnosed CLL in 1984 [55]. It was later discovered that the active transcript of 11q13 cloned from parathyroid adenoma encoded a new G1 cyclin, denoted cyclin D1 (CCND1) [56], and subsequently it was established that CCND1 corresponds to the BCL1 gene [57]. Cyclin D1 promotes the transition from G1/S phase into S-phase and when gained at the pre-B-cell stage, the t(11;14)(q13;q32) translocation is generally regarded as the hallmark and primary oncogenic event in the evolution of MCL [58]. By this translocation the CCND1 gene becomes juxtaposed to the immunoglobulin heavy chain Eμ enhancer, resulting in aberrant cyclin D1 expression. Infrequently, the CCND1/IGH fusion gene is also amplified in MCL which has been associated with higher proliferation and worse clinical outcome [59, 60].

The intricacy of cyclin D1 in MCL increased after reports suggested that cyclin D1 may be involved in chromatin remodeling and chromosome stability, as well as having a role in supporting DNA repair. Reports of deletion of the E2F inhibitor RB1 further indicate potential for a wider function of cyclin D1 than merely overcoming the G1/S phase [61].

2.5.1.1 Cyclin D1 isoforms
Cyclin D1 exists as two forms, isoform a and isoform b [62].

The canonical cyclin D1 (isoform a) is a 36 kDa protein encoded by a mRNA transcript existing in two different lengths, 4.5 kb and 1.5 kb long, respectively. The 4.5 kb transcript encompasses a destabilizing 3'-UTR region allowing miRNAs to anneal, which leads to a reduced half-life compared to its truncated counterpart; 30 min compared with 3 hours [63].

Next-generation sequencing (NGS) data has provided a more detailed mutational status in MCL and uncovered that specific hotspots in the 3'-UTR region of exon 1 on CCND1 are mutated in 14-34% of all investigated cases [64, 65], which could prevent binding of miRNA. The longer half-life of the shorter or mutated transcript consequently resulted in increased cyclin D1 protein expression in the investigated MCLs [66]. Data from MCLs expressing the shorter CCND1 transcript also correlated to a more aggressive disease progression [63, 66-68]. Furthermore, this isoform has been associated with ibrutinib (inhibits the function of Bruton's tyrosine kinase) resistance in MCL patients [69].

The second isoform, cyclin D1b, is less frequent and deficient of the C-terminal domain that partly regulates nuclear export. Despite a potentially higher protein accumulation in the nucleus, this isoform has never been considered of great importance in the pathogenesis of MCL due its very moderate expression levels [70, 71]. Further reasons for cyclin D1b not outcompeting cyclin D1a when expressed in MCL are their highly similar half-life and distribution in the cell [72]. To be noted, retention of cyclin D1 in the nucleus is oncogenic in vivo as shown by a cyclin D1 mutant mimicking the truncated cyclin D1b isoform [73]. However, cells carrying the t(11;14) translocation can be found in the blood of healthy individuals not developing MCL, indicate that additional genetic aberrations are required to develop MCL [74, 75].

2.5.2 Cyclin D1-negative MCL
Albeit aberrant cyclin D1 expression is considered to be one of the principal features of MCL, there are MCLs that are negative for cyclin D1. These types of cases can now better be identified by SOX11 expression.
The importance of implementing gene expression profiling (GEP) in MCL was emphasized when investigating a rare group of difficult to classify lymphoma cases. Although negative for cyclin D1, they otherwise shared similar morphological, molecular and clinical characteristics as conventional cyclin D1+ MCL. The GEP-studies performed on these rare cyclin D1- lymphoma cases demonstrated that they shared secondary genetic alterations similar to those found in cyclin D1+ MCL (in addition to the observed clinical and molecular manifestations) [76, 77]. The notion that this rare subgroup of cyclin D1- cases actually could constitute a separate entity within MCL was further corroborated in a large study of 40 cyclin D1- cases diagnosed as MCL. Among the cyclin D1- MCLs, the \textit{CCND2} translocation was the most frequent chromosomal rearrangement resulting in high expression of cyclin D2 [77-85] (only one case has been reported for \textit{CCND3} translocation [85]).

2.5.3 Secondary genetic aberrations
Classical MCL acquires a high degree of genomic aberrations and is characterized by genomic instability. The (11;14)(q13;q32) translocation is essential for the initiation of the disease, however, cyclin D1 expression is not \textit{per se} oncogenic as it has been detected in up to 1-2% of healthy individuals [86]. Further, knockdown of cyclin D1 has only minor effects on proliferation and survival in MCL cell lines [87]. Moreover, transgenic mice models overexpressing the most frequent isoform of cyclin D1 required additional oncogenic hits, such as \textit{MYC}, to develop a B-cell lymphoma [88, 89]. Highlighted in several reviews ([19, 36, 52, 58, 90, 91]), genetic deletions, mutations or amplifications frequently include genes of the \textit{CDKN2A/CDKN2B/CDK4/RB1} and the \textit{CDKN2B/MDM2/TP53} signaling pathways by targeting important pro-survival or pro-apoptotic genes, such as \textit{BCL2} (overexpressed), \textit{BCL2L11} (deleted) and \textit{FBX025} (deleted). Increased cell proliferation resulting from downregulation of genes (\textit{MOBKL2B}, \textit{MOBKL2A} and \textit{LATS1}) involved in the Hippo-pathway has also been reported in up to 38% of MCL [92].

High proliferation is one of the best predictors for inferior survival in MCL [93]. This high proliferation rate is largely based on several genetic aberrations that involve impairment of cell cycle functions, DNA damage responses and regulated cell death pathways [36]. There is also a signature cluster of proliferation-associated genes constructed from expression profiling data collected from a large number of MCLs that accurately predicted shorter survival for patients expressing that set of genes [68].

As summarized in Table 1 from whole exome/targeted sequencing and copy-number studies, the pathogenesis of MCL involves several genetic alterations influencing genes associated with cell cycle, apoptosis, DNA repair, NOTCH signaling, BCR and NF-κB signaling, epigenetic modifiers, RNA and ribosomes [68, 76, 91, 94]. Since cyclin D1 overexpression is not sufficient for oncogenic transformation of naïve B-cells into MCL, it was postulated from data in [90] that the t(11;14) translocation is followed by hitherto unknown alterations that after various latency can evoke a state referred to as in “\textit{in situ}” MCL, now denoted ISMNC [50]. ISMNC is phenotypically and morphologically similar to MCL and is assumed to have fewer secondary genomic alterations; however, no further studies have yet been performed to confirm this. Additional acquired genetic modifications in a fraction of these cells would result in more malignant clones with increased genomic instability.
Impaired DNA damage response is another major factor underlying the aggressiveness of MCL. The ATM gene in particular, critical for promoting cell cycle arrest in response to double-strand breaks, is very important and has been reported deleted or mutated in 11-61% of MCLs (11–57% deletions, 41–61% mutations) [91]. Another important gene is TP53, which is also frequently mutated or deleted in MCL in 14-22% of cases [35, 64, 95] with 17p deletions found in 32% of cases [95].

Recent NGS studies were able to detect NOTCH1 and NOTCH2 mutations in a relatively low number of MCLs (5% and 5-14%, respectively) [64, 65], as well as genes in the NF-κB pathway, such as CARD11 (6%) [96]. The NF-κB pathway is also altered by mutations and epigenetically silencing of TNFAIP3 [97] together with BIRC3 [35, 96].

Tumor cell proliferation (due to secondary genetic alterations) and specifically TP53 mutations and TP53 overexpression are associated with disease aggressiveness in MCL [98]. TP53 mutational status was recently reported to identify younger MCL patients who were not benefited from heavy chemoimmunotherapy [99].

2.6 CURRENT STANDARDS OF CARE
The current treatment regime for MCL is treatment with anti-CD20 antibodies and combination chemotherapy. Younger patients can receive high dose chemotherapy and autologous stem cell transplantation (ASCT). Patients who are ineligible for such intense treatment can be treated with targeted therapy, for instance ibrutinib [100, 101].

<table>
<thead>
<tr>
<th>Pathway/process</th>
<th>Gene</th>
<th>Alterations in MCL (% range) (n = 624 from seven WES/targeted studies and n = 219 from six copy-number studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell cycle</td>
<td>CCND1 *</td>
<td>95% rearr, 14–34% mut 3% rearr 25–55% del 19–37% del 10–36% del 3–17% gain 6–12% gain 0</td>
</tr>
<tr>
<td></td>
<td>CCND2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RB1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LATS1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CDKN2A, CDKN2B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BCL2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BMI1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BRAF</td>
<td></td>
</tr>
<tr>
<td>Apoptosis</td>
<td>FBXO25</td>
<td>17–34% del 6–32% gain, 1% rearr</td>
</tr>
<tr>
<td></td>
<td>MYC</td>
<td></td>
</tr>
<tr>
<td>DNA repair and integrity</td>
<td>TP53 *</td>
<td>21–45% del, 14–31% mut 11–57% del, 41–61% mut 1% mut &lt;1% rearr</td>
</tr>
<tr>
<td></td>
<td>ATM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>POT1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TERT</td>
<td></td>
</tr>
<tr>
<td>NOTCH signaling</td>
<td>NOTCH1</td>
<td>5–14% mut 5% mut</td>
</tr>
<tr>
<td></td>
<td>NOTCH2</td>
<td></td>
</tr>
<tr>
<td>BCR/NF-κB signaling</td>
<td>BIRC3</td>
<td>11–57% del, 6–10% mut 3–15% mut 7% mut 7% mut 5% mut 2–3% mut 0 0</td>
</tr>
<tr>
<td></td>
<td>CARD11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TLR2 a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TRAF2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NFKBIE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAP3K14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MYD88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EGR2</td>
<td></td>
</tr>
<tr>
<td>Epigenetic modifiers</td>
<td>KMT2D</td>
<td>12–23% mut 5–16% mut 10–13% mut 8% mut 0 0</td>
</tr>
<tr>
<td></td>
<td>KMT2C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NSD2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SMARCA4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ARID1A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SETD2</td>
<td></td>
</tr>
<tr>
<td>RNA and ribosomes</td>
<td>SF3B1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>XPO1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>RPS15</td>
<td>0</td>
</tr>
<tr>
<td>Other pathways</td>
<td>UBR5</td>
<td>7–18% mut 3–15% mut 3–7% mut 0</td>
</tr>
<tr>
<td></td>
<td>S1PR1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MEF2B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IKZF3</td>
<td></td>
</tr>
</tbody>
</table>
3 SRY-RELATED HIGH-MOBILITY-GROUP BOX (SOX) TRANSCRIPTION FACTORS

3.1 BACKGROUND
Genes that have similar functions and sharing similar physiological regulation often share the same short regulatory sequence, which differentiate them from other genes. Transcription factors are regulatory proteins that are able to specifically interact with those DNA sequences and induce or repress transcription of the targeted genes [102].

3.2 THE SOX TRANSCRIPTION FACTORS
The sex-determining region on chromosome Y, the SRY transcription factor, was discovered in the 1990s [103, 104]. In addition to identifying the gene accountable for male differentiation, it also paved the way for the discovery of a whole family of highly specific transcription factors denoted as SOX, an acronym for “SRY-related HMG box” [105]. SOX genes were defined as genes sharing a consensus sequence coding for the conserved RPMNAFMVV motif found within the High-Mobility Group (HMG) domain of all known SOX transcription factors [105]. Ten years after the discovery of the SRY gene, the definite number of SOX transcription factors amounted to 20 distinct genes, which excluded all previously counted orthologues [106]. Grouping based on sequence homology further divided them into eight distinct subgroups (A-H) [107]. The main role of the SOX genes is to regulate cell fate during development, tissue homeostasis and regeneration of stem and progenitor cells [108].

The expression of SOX genes is generally tissue and developmental stage-specific, however the low number of SOX genes indicates that the genes are pleiotropic. Thus, the gene may act differently based on cell context and developmental stage. The Sox10 gene is involved in neural crest formation, however for early neural crest development Sox10 can be substituted with Sox8 (both members of the SoxE group). This stage is not affected in mice with a mutated Sox10 either. However, in the later differentiation stage (enteric nervous system), this mutation severely affects melanocyte development [109] demonstrating that this stage specifically requires Sox10.

The HMG box domain of SOX binds to the DNA and induces a conformational change, forming an L-shaped structure with the DNA, which allows other previously non-adjacent DNA binding sites and their transcription factors to come within a close proximity to the SOX protein. This stabilizes and forms a transcriptional complex as outlined in Figure 3.

SOX proteins regulate transcription of their target genes in three different ways [107, 110, 111], namely via transactivation (TA) (Figure 3.1), transrepression (TR) (Figure 3.2) or structural stabilization (Figure 3.3). TA is performed by the C-terminal region that interacts with a transcriptional co-activator, which in turn binds to the TATA-binding protein (TBP) and TBP-associated factors (TAFs). Transrepression is also performed by the C-terminus, but with a transrepression domain (TR). The TR interacts with a transcriptional co-repressor that blocks gene transcription. There are only a limited number of SOX proteins that solely act as structural proteins for constructing the enhanceosomes.
Figure 3: Transcriptional regulation of SOX protein by binding to the DNA and induce a conformational change via 1) transactivation 2) transrepression or 3) transcriptional complex stabilization (modified from Lefèbvre, V. et al. Int J Biochem Cell Biol, 2007. 39(12): p. 2195-214).
3.3 MOLECULAR FEATURES

3.3.1 DNA binding
The SOX proteins bind a short linear DNA consensus motif that can be found extensively throughout the genome and their binding affinity to DNA is generally lower than for other transcription factors ($K_d$ of $10^{-7}$-$10^{-9}$ compared with $K_d$ of $10^{-9}$-$10^{-11}$) [107]. Together, this complicates target specificity. However, the SOX proteins can acquire selectivity by the structural changes induced when the HMG domain binds to the DNA [112, 113] (discussed in Kondoh et al. [114] and below).

3.3.1.1 HMG domain
All SOX proteins include the roughly 80 amino acid residues long and highly conserved HMG domain of three $\alpha$-helices that form a L-shaped structure [115-117] that interacts with the minor groove of the DNA helix via their shared consensus motif, (A/T)(A/T)CAA(A/T)G [118-121]. Although belonging to the same family, the HMG domain of the SOX proteins does not share more than 50% identity for the different groups [107]. However, under certain conditions they are interchangeable, for example substituting the HMG domain of SRY with that of SOX3 or SOX9 did not result in any loss of function [122]. Despite their high similarity, the different SOX proteins do not always bind to the same targets. By compressing the major groove while widening the minor groove the interaction has a major structural impact on the conformation of the DNA helix (Figure 4). The conformational change facilitates recruitment and association of cell-specific proteins and transcription factors required for the transcription otherwise positioned out of reach on the DNA (Figure 3) [123]. Another reason for gene specificity is the flanking region of the binding sites that can influence the binding affinity of different SOX proteins (SOX4, SOX9 and SOX10) [120, 124, 125]. Further, the context dependent interaction could also account for how these similar proteins may have a diverse range of functions. During eye development, the lens-specific enhancer element DC5 is a target for SOX2 and PAX6 that form a stable and specific tertiary enhancer complex. Alone each factor was able to bind to the enhancer, but not able to activate transcription of target genes. Instead, PAX5 and SOX2 together have to interact with the sequence to form a tertiary complex that positions the proteins in a particular spatial organization for activation [126].

![Figure 4: Binding of two HMG domains (displayed in opposite direction) to DNA in the minor groove that induce a conformational change (modified from Lefebvre, V. et al. Int J Biochem Cell Biol, 2007. 39(12): p. 2195-214).](image)
3.3.2 TAD and NLS domain
In addition to the HMG domain, other molecular features of several of the SOX proteins are a C-terminal transactivation or repression domain [127-129], a nuclear localization sequence (NLS) situated in the HMG [130] and dimerization domains [130]. Of note, the transactivation domain (TAD) can also be positioned centrally. When interacting with different transcription factors, the TAD is a key factor for initiating and maintaining the protein heterodimerization state by interacting with the DNA-binding domain of the partner protein [112, 113]. SOX proteins are able to compete for the same target genes, as shown for the SOXB group where the SOXB2 proteins [131] gradually replace SOXB1 proteins. Since the C-terminal domain of SOXB1 is a transcriptional repressor, whereas SOXB2 is a transcriptional activator [132], the latter culminates transcriptional activity from a repressed to an activated state. Hence, the steady-state levels of highly similar SOX proteins could have significant biological impact on gene regulation.

3.4 THE SOXC TRANSCRIPTION FACTOR GROUP

3.4.1 Expression
In higher order species, homology grouping identifies SOX11 (474 amino acids, human), together with SOX4 (441 amino acids, human) and SOX12 (315 amino acids, human), as members of the SOXC group of single exon genes [105, 128].

The SOXC genes are highly conserved and concomitantly expressed at elevated levels during embryogenesis and in neuronal and mesenchymal progenitor cells [128, 133-140]. Postnatally, as development progress, expression of the SOXC family members is reduced or restricted to specific cell lineages. Expression of SOX12 has a more unconstrained pattern with low mRNA levels detected in most adult tissues [141]. In adult mice, Sox11 mRNA expression is detected in neuronal tissue [142] and pancreatic islet cells [143] whereas Sox4 mRNA expression is restricted to B/T-cells and gonad cells [127].

3.4.2 Features of the SOXC group
The HMG domains of the SOXC group share 84% identity [128] and are known as transcriptional activators with a C-terminally positioned transactivation domain which enables all members to activate or interact with proteins in different transcriptional complexes. In the SOXC group, the TAD is comprised of the last 33 residues in the protein, but share less identity (67%). Both SOX4 and SOX11 demonstrate a very high sequence similarity in the TAD compared to SOX12. Nevertheless, all three members function as transactivators; SOX11 the most potent, SOX4 intermediate and SOX12 the weakest [128, 144, 145]. The SOX4 and SOX12 proteins bind DNA with higher affinity on electrophoretic mobility shift assay (EMSA) than SOX11 [128, 146].

3.4.3 Overlapping function in vivo
Knockout experiments in mice have demonstrated that Sox4 and Sox11 have non-reciprocal functions during development, as Sox4−/− mice die at E14 and Sox11−/− succumbs hours post birth [6, 147]). In contrary, Sox12−/− mice survive embryogenesis and are born alive, almost indistinguishable from their wild-type littermates [145].

The SoxC genes are functionally redundant up until E8.5 of embryogenesis when investigated in SoxC−/− (Sox4−/− Sox11−/− Sox12−/−) mice [148]. Past that stage, the study demonstrated how expressions of the SoxC genes were of different significance during the
developmental stages; the mice displayed different severity of malformation dependent on how the \textit{SoxC} genes were deleted. From combinatorial deletion of the genes, the study concluded that \textit{Sox4} has a more essential role during early development by observations that \textit{Sox4\textsuperscript{−/−}11\textsuperscript{−/−}} mice exhibited less malformations than \textit{Sox4\textsuperscript{−/−}11\textsuperscript{+/−}} mice. The impact of \textit{Sox12} was less pronounced. Although, the \textit{SoxC\textsuperscript{−/−}} mice had more marked organ malformations in addition to increased cell death compared with the \textit{Sox4\textsuperscript{−/−}Sox11\textsuperscript{−/−}} mice [149]. The observed effects are indicative for some compensatory effect by \textit{Sox12}, but not very strong.

3.4.4 \textbf{SOXC expression in B-cells and cancer}

During hematopoiesis, SOX4 is necessary for pro-B cell survival [150-152]. While SOX4 is present in lymphocytes and vital for B-cell differentiation by regulating the transition from CLPs to early B-cells [150], the role of SOX11 in B-cells is still not established. SOX11 is expressed in MCL, but has yet no defined role in hematopoiesis and is only rarely expressed in normal B-cells [153]. SOX12 is yet to be investigated.

One fundamental function of the SOXC proteins is supporting cell survival, which is closely related to cancer and the relevance of the \textit{SOXC} genes has gained more interest as more information from whole genome/exome data suggests a role in tumorigenesis. For example, among 40 published cancer microarray data sets, \textit{SOX4} was one of 67 genes that were reported as a cancer signature gene when comparing cancer tissue with healthy tissue [154]. Aberrant expression of \textit{SOX4} and \textit{SOX11} (the \textit{SOXC}) genes is reported in several different tumors.

In addition to maintaining survival of pro-B cells [150-152], SOX4 has been ascribed an oncogenic role in acute lymphoblastic leukemia (ALL) [155, 156]. However, in non B-cells, SOX4 has also been shown to interact and stabilize TP53 by preventing Mdm2-mediated TP53-ubiquitination, and thereby preventing tumorigenesis. This demonstrates a context dependent function of the \textit{SOX4} gene. Depending on cell type, the \textit{SOXC} genes could have either a pro-apoptotic or anti-apoptotic effect. A summary of different SOX4 expressing tumors uncover how differently the very same SOXC transcription factor acts in different biological contexts [157]. However, for certain cancers, such as medulloblastoma, high expression of SOX4 in the tumor correlated to both a favorable [158] and an unfavorable [159] disease progression. This can likely be due to several other factors than solely SOX4.

3.5 \textbf{THE SOX11 TRANSCRIPTION FACTOR}

Human \textit{SOX11} was discovered, cloned and mapped to chromosome 2 at position p25 (2p25) using fluorescence \textit{in situ} hybridization nearly 20 years ago [160].

The structural features of \textit{SOX11} are very similar to the other \textit{SOXC} genes, but there are certain differences, especially between the domains. The HMG domain is followed by an acid and proline-glutamine rich region located centrally [105] with the TAD forming an uninterrupted $\alpha$-helical structure [146] comprised of a serine rich and a highly conserved C-terminal region [105]. When elucidating the transcriptional regulation of \textit{SOX11}, removal of different parts of its C-terminal region demonstrated an auto-inhibiting function as its ability to interact with DNA increased in EMSA [128]. This is the reason for its lower DNA affinity compared with SOX4 and SOX12. Posttranslational modifications could also influence binding capacity, but it has only been shown \textit{in vitro} for SOX9 and SOXE proteins [107].
The NLS of SOX11 is responsible for nuclear import, but also reduces caspase-6 activity, as demonstrated after removal of the domain. Among different SOX proteins, SOX11 could also reduce caspase-6 activity with the highest efficacy, indicating its ability to abrogate cell death under certain conditions [161].

During embryogenesis, Sox11 transcripts are detected in various tissues and organs, including central and peripheral nervous system, branchial arches, genital tubercle, limbs, eyes, ears, mammary buds, nasal invagination and somites [162]. In order to investigate the effect of Sox11 during embryogenesis, different knockout models have been established (Sox11−/− [147] and SoxC−/− [148]). Postnatally, the Sox11−/− mice succumbed from cyanosis as a consequence of severe cardiac defects or pulmonary insufficiency due to significant hypoplasia of the lungs. The newly born mice had 23% lower birth weight. Other recurrent defects involved absence of spleen, underdeveloped lung, stomach and pancreas in addition to skeletal and craniofacial malformations. The nervous system did not exhibit as profound malformations as other tissues. The neural cells did not exhibit a significant reduction in proliferation following Sox11 inactivation. A possible theory could be that the highly homologous Sox4 is found co-expressed in the nervous system and could partly compensate for the loss of Sox11 [148].

These experiments demonstrate the importance of Sox11 in regulating cell survival during organ growth and neurogenesis in vivo [139, 142, 145, 163]. After establishing its role as a master regulator during development in mice, de novo mutation in the human orthologue have now been found to be that is responsible for certain congenital/developmental diseases, such as Coffin-Siris syndrome [164]. SOX11 is also directly involved in a unique case with 2p25 duplication and CHARGE syndrome [165].

While SOX11 is necessary for embryogenesis, postnatal expression is restricted in most tissues [133, 148, 166]. In non-malignant tissue, Tubb3 [128], important for axon guidance [167] and Tead2 [148], important for the Hippo signaling pathway [168] are so far the only known and validated genes that are directly targeted by SOX11. The formation of transcriptional complex to induce gene activity is dependent on binding partner. However, there are only few known proteins identified and validated as binding partners to SOX11: BRN1 (yeast-two-hybrid screening) [169] and NGN1 and BRN2 (co-immunoprecipitation) [170].

The cellular background and context are critical for the function of the SOX transcription factors. SOX11 can, in addition to SOX2 and SOX10, under correct binding conditions synergistically interact and cooperate with different POU domain proteins to enhance transcriptional activity: SOX2 with Oct-3/4 [171], SOX10 with Tst-1/Oct6/SCIP [172] and SOX11 with BRN1 [172]. Unsuitable binding partners to SOX11 were unable to activate transcription, or even lead to a reduced response as they interacted [172]. Transcriptional activity of SOX11 could theoretically also be influenced by the concentration and availability of the other SOXC members. Due to its higher DNA affinity over SOX11, but lower TAD activation capacity [128], the most substantial effect would be with SOX12. Hence, having a higher DNA binding affinity, SOX12 could potentially interfere with the genes SOX11 normally target.

Neural cells are one of the most studied cell types with regard to Sox11 and overexpression of Sox11 has been reported to have an inductive effect on pan-neural markers and the neural cytoskeleton [133, 142]. However, not all neural cells are the same. Overexpressing of Sox11 in two different types of retinal ganglion cells (RGCs), α-RGCs and non-α-RGCs, resulted in two contrasting outcomes; cell death and axon regeneration, respectively [134,
Moreover, adult neural precursor cells do not differentiate in the absence of Sox11; in contrast, gain of Sox11 initiates adult neurogenesis in immature neurons [142, 174] as well as regeneration of damaged peripheral nerves [175] and ensuring sensory neuron outgrow [176].

4 THE ROLE OF SOX11 IN CANCER

4.1 EXPRESSION PATTERN

Nuclear staining of SOX11 can be detected in MCL [25, 26], but it is also expressed in subsets of medulloblastoma [158], malignant glioma [177], ovarian cancer [178] and breast cancer [179]. Regarding lymphoproliferative diseases other than MCL, certain BL, subsets of precursor B/T-cell lymphoblastic neoplasia, especially B-ALL with the TEL-AML1 fusion or E2A rearrangement, and hairy cell leukemia (HCL) also express SOX11 (protein or mRNA) [180, 181]. However, transcript levels of SOX11 are otherwise very low in other lymphoproliferative diseases, including most ALL and BL, CLL, follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL) and primary mediastinal B-cell lymphoma (PMBL) [181]. SOX11 expression was also shown to be low in nnMCL, a subset suggested to be correlated with indolent disease [33, 45, 181, 182].

In conclusion, high expression of SOX11 has been reported in different B-cell lymphoid malignancies and solid tumors with the expression associated with differential clinical outcome [54, 153, 183-186], again illustrating how context dependent the SOX11 gene is.

4.2 FUNCTIONAL ROLE IN NON-MCL DERIVED TUMORS

The role of SOX11 in cancer varies, but the capacity to proliferate and survive appears to be influenced in several cancers expressing SOX11.

As an example, SOX11 is upregulated in certain gliomas (however, only very few cases in the study) [177]. Xenotransplanted mice with Sox11 expressing glioma cell lines had longer survival (60 days compared to 45 days for the non-Sox11 expressing cell lines) [183]. Thus, in this context Sox11 has a tumor suppressive effect by differentiating the cells, compared to the proposed oncogenic effect of SOX11 in the human gliomas. Moreover, in certain ovarian cancer, high SOX11 expression has shown to correlate with a longer recurrence-free survival [178]. Therefore, its functional role is not clear in most cancers. For MCL, its functional role is even more ambiguous with several well-conducted studies reporting conflicting data [187, 188].
5 SOX11 IN MCL

5.1 BACKGROUND
MCL is predominantly diagnosed by morphology and (11;14)(q13;q32), or cyclin D1 overexpression. However, after recognizing SOX11 to be expressed in over 90% of all MCLs [25, 26, 45, 180, 189, 190], the transcription factor has gained considerable interest as a potential biomarker and potential master regulator in MCL. The SOX11 protein is not expressed in the vast majority of lymphomas or mature B-cells and its expression is independent of cyclin D1 status. Given the aggressive disease progression and recent advancements in treatment regimens tailored for MCL, it is important to correctly diagnose the disease. Taken together, SOX11 can facilitate differential diagnosis and identifying MCL, which is especially useful for certain morphologically and phenotypical similar lymphomas, such as CLL and MZL [25, 26, 190].

5.2 SOX11 EXPRESSION IN MCL, A DOUBLE-EDGED SWORD?
Within the field of MCL, the role of SOX11 as a diagnostic marker is undisputed; however, its role as a prognostic marker is not. The reliability of SOX11 as a diagnostic marker has been verified in numerous cohorts (to mention some of the studies) [25, 33, 45, 54, 153, 180, 185, 190, 191]. However, not all of these studies are coherent with regard to clinical outcome, definition of indolent disease, percentage of cases with nodal presentation and SOX11 expression. In some studies, SOX11- MCLs had longer survival [33, 45, 53, 182], whereas in other studies SOX11- MCLs had shorter survival [26, 54, 153, 185, 191, 192]. The reasons for the inconsistency are still not fully understood. Nevertheless, there are certain parameters to consider when interpreting these conflicting results:

1) Detection specificity (antibody/primers used)
2) Stratification (from where are the patients collected)
   a. Clinical behaviour
   b. Treatment
   c. Geographical area

The sensitivity and specificity of the different SOX11 antibodies may differ [153, 193, 194]. Clinical behavior is also important as there have been suggested to exist several different entities of MCL [50], particularly the classical MCL situated in the lymph node (low grade of SHM, high degree of genetic alterations and aggressive behavior) and leukemic non-nodal MCL (high grade of SHM, low degree of genetic alterations and less aggressive behavior).

In the studies of Fernandez et al. [45] and Navarro et al. [33], the SOX11- cases used were characterized as nnMCL, and as previously described, this entity is much more indolent than the classical type of MCL. In Lord et al. [153], samples were largely population based and not a coherently treated collection. Both the study by Nygren et al. [54] and Nordström et al. [185] included several SOX11- cases with strong TP53-positivity, a factor previously shown to be associated with shorter overall survival (OS). The SOX11- cases could be derived from indolent nnMCLs that later acquire 17p/TP53 mutations, however Nordström et al. [185] emphasizes that no indolent MCLs were included in their study. Meggendorfer et al. [192] reported that in MCLs (t(11;14)-positive and t(11;14)-negative) SOX11+ patients showed a more indolent course compared with SOX11- MCLs. Further, SOX11+/t(11;14)- cases had a more adverse prognosis than SOX11+/t(11;14)+ MCL patients.
5.2.1 Diagnosing MCL

Although MCL can be diagnosed by morphology, phenotype and t(11;14)(q13;32) or cyclin D overexpression, SOX11 protein is now routinely used in the diagnostic work-up. Detection of SOX11 is primarily performed by IHC/ICC, but there is now a validated protocol for detecting expression using flow cytometry [195]. This is practical for cytology and blood samples. Detection of SOX11 mRNA expression by quantitative polymerase chain reaction (qPCR) can also be used in MCL and has high concordance to SOX11 detection by immunohistochemistry (IHC) [153]. Additionally, SOX11 mRNA expression correlates with the t(11;14)(q13;32) translocation in patient samples [192, 196].

5.3 REGULATION

As previously mentioned, MCL is characterized by a high degree of chromosomal instability and genetic aberrations. However, several studies with large numbers of MCLs investigated have not detected any genetic alterations in or near 2p25 where SOX11 gene is located [59, 92].

Since no apparent genetic reason is behind the observed overexpression of SOX11 in MCL, it is expected to be epigenetically regulated. Interestingly, compared to the majority of BL, CLL, FL, DLBCL, breast cancer, ovarian cancer, lung cancer and brain cancer cases that are all negative for SOX11, the SOX11 promotor is not heavily methylated in MCL (cell lines and primary samples) [181, 197-200]. The vast majority of investigated B-cells have hypomethylated promotor region, but are still epigenetically silenced compared to MCL.

Further verification that the promotor region of SOX11 is not methylated was performed by 5-azacytidine (5-AZA), an inhibitor of DNA methyltransferase. No increased expression was detected as would have been the case for a heavily methylated promotor region. Hence, SOX11 expression is not governed by promotor methylation in MCL. Expression is instead regulated by histone modification via activating histone mark [181, 200], as demonstrated for two SOX11- cell lines (one hypomethylated MCL cell line and one hypermethylated BL cell line). Inhibiting histone deacetylases significantly elevated SOX11 expression for both cell lines, showing how the treatment re-activated the SOX11 gene.

Lastly, circularized chromosome conformation capture (4C)-Seq data recently identified a distant enhancer element binding to the SOX11 locus in 3D. This region, positioned 650 kb downstream of SOX11 differed in methylation status for SOX11+ and SOX11- MCL (Figure 5). Bisulfite sequencing showed it to be de novo demethylated in SOX11+ compared with SOX11- MCLs and cell lines [21].

In conclusion, SOX11 can potentially be regulated by promotor methylation in other lymphoma and certain solid cancer types as have been shown in the above studies [199]. However, in the particular context of MCL, the SOX11 is regulated by histone modification [181] and potentially also a distant enhancer element [21]. Interestingly, in the hypermethylated BL cell line Raji, SOX11 can be re-activated by both the de-methylation agent 5-AZA and the histone deacetylase inhibitor SAHA.
5.4 SOX11 TRANSCRIPTIONAL TARGETS IN MCL

So far, three different studies have conducted chromatin immunoprecipitation (ChIP) experiments to discover transcriptional targets of SOX11. The initial study was based on genes identified after SOX11-specific small interfering RNA (siRNA) knockdown followed by gene expression profiling after 20 hours in the MCL cell line Granta519. The 26 differentially expressed genes were then validated in data sets from primary MCLs, out of which DBN1, SETMAR and HIG2 significantly correlated to SOX11. They were next confirmed to bind by ChIP-qPCR [201]. A similar approach was performed for Granta519, Z138 and JeKo-1, but with longer time of exposure to SOX11-specific siRNA (48 hours) and RNA-seq. The three cell lines encompassed 2799 differentially expressed genes. ChIP-Seq in Granta519 identified 1912 unique SOX11-bound genes out of which SMAD3, TGFBR1, WNT4, NLK and PRKACA were validated by ChIP-qPCR [191]. In a third study, 2790 differentially expressed genes were found when gene expression profiling was performed on Z138 cells transduced with SOX11-specific short hairpin RNA (shRNA) for stable knockdown. ChIP-Chip in Z138 identified 1133 (1132) genes bound by SOX11. When overlaying GEP and ChIP-Chip data, 147 genes were bound by SOX11. From them PAX5, MSI2, HSPD1, SUV39H2, and SEPT2 were validated by ChIP-qPCR [188].

When the authors in the studies cited above [188, 191] performed GO-term and KEGG Pathway analysis on the genes identified in each separate SOX11-binding study, there were two main findings. Kuo et al. [191] found pathways related to cancer (WNT and TGF-beta pathway genes ranked highest) [191], whereas Vegliante et al. [188] found hematopoiesis and hematological system development to be the key pathways highly enriched. Overlapping pathways for these two studies were proliferation and tissue invasion/hematological system development and function. However, the actual intersect between the two studies only consists of 96 genes (Figure 6) after I compared the two data sets. A GO-term analysis of the overlapping genes only shows enrichment of a few processes (negative regulation of transcription and brain and developmental processes).
5.5 FUNCTIONAL ROLE OF SOX11 IN MCL

In spite of recent breakthroughs in MCL, the molecular mechanism of SOX11 in MCL and its potential binding partners in the transcriptional complex is still not fully understood.

Transient siRNA-knockdown of SOX11 in Granta519 altered a group of genes involved in cell shape and motility, but there were no detectible changes in cell morphology [201] or increased proliferation [197]. Genes involved in WNT/TGF-β pathways are also direct targets of SOX11 (ChIP-Seq) and differentially expressed after transient siRNA-knockdown of SOX11 in Granta519 and other MCL cell lines [191]. Dysregulated WNT signaling can lead to malignant transformation of cells and WNT-signaling is often found to be upregulated in MCL [202]. However, when Kuo et al. [191] overexpressed SOX11 in the same cell lines, members of the WNT pathway were functionally repressed. Moreover, cells showed a significantly reduced proliferation rate with an increased proportion of cells in the G2/M-phase. This is in agreement with a previous report where overexpression reduced proliferation rate in Granta519 and Z138, and upregulation of TGF-β, BRCA1 and SMAD2/3 in Granta519 [197].

In another study, the functional roles of SOX11 were interrogated after stable knockdown of SOX11 by shRNA in Z138 cells. Their gene signature abruptly shifted to that of one more similar to the plasma cell differentiation program when compared to the control cells, which represented that of the mature B-cell program. Moreover, downregulation of SOX11 additionally reduced cell proliferation and the size of engrafted tumors [188], contrary to results from when overexpressing SOX11 in the same cell line [191]. Downregulation of SOX11 protein expression leads to decreased PAX5 activation and upregulated BLIMP1. PAX5 is needed to establish the B-cell identity [203], but it also binds and repress the promotor of the BLIMP1 gene, leading to block of plasma cell differentiation [204]. Hence, low SOX11 could promote differentiation from a mature B-cell into a more plasmacytic differentiated subtype [188]. SOX11 has also been speculated to prevent naïve B-cells carrying the t(11;14)(q13;q32) from leaving the mantle zone since it is able to bind to the promotor of BCL6 with a repressive effect as shown by luciferase assays [205].
Therefore, overexpression would repress transcription of BCL6, which stops the cells from entering the GC and preventing them from going through SHM and further differentiation steps. Investigations in the same SOX11 shRNA-knockdown model system further showed that SOX11 can induce angiogenesis by transcriptionally activating PDGFA [206]. Reported in Conrotto et al. [187], upon SOX11 downregulation, ATX is one of the most highly upregulated genes. The gene is involved in producing a lysophosphatidic acid (LPA) responsible for stimulating angiogenesis and often found in tumors [207], so angiogenesis could potentially also be negatively affected by SOX11 in MCL. However, there are no in vitro or in vivo studies supporting this. Further, cell counting and 3H-Thymidine incorporation demonstrated an increased proliferation rate upon downregulation of SOX11 protein, as well as more cells in the S-phase. Engraftment of these SOX11 knockdown cells into mice also led to shorter survival compared to mice engrafted with corresponding control cells.

In summary, SOX11 directly binds and influence genes critical for cell shape and motility, proliferation, cell cycle, apoptosis, B-cell differentiation, angiogenesis and tumor microenvironment. The proposed mechanistic model on MCL pathogenesis could theoretically partly be explained by how SOX11 regulates PAX5 and BLIMP1, which are important for GC-formation and B-cell differentiation.

5.5.1 Discrepancies between oncogenic effects of SOX11 in a MCL model
Based on data from Vegliante et al. [188], SOX11 has been proposed to function as an oncogene in MCL [208]. However, studies regarding the influence of SOX11 on cell proliferation in MCL cell lines have yielded conflicting results [187, 188, 191, 197]. Overexpression of SOX11 in Z138 and Granta519 resulted in decreased proliferation [191, 197], whereas transient knockdown increased proliferation [197]. Notably, two well-conducted and experimentally very similar studies reported contrasting results. One study reported that stable knockdown of SOX11 (using shRNA in Z138) resulted in decreased proliferation [188], whereas the other study reported increased proliferation [187]. Both studies also investigated tumor growth in xenografts with the same conflicting results. However, different mice strains were used (NOD-SCID and CB17-SCID) and different injection methods were used (intravenously in the tail vein and subcutaneously into the lower dorsum). This could influence the tumor engraftment, but requires further validation of the in vitro data.

Regarding the in vitro data, despite that both studies verified downregulation on transcriptional and protein level, only 10 genes overlapped when I investigated the intersect of the reported genes for the two studies using the genes that were reported as significantly differentiated (Figure 7). The small overlap indicates a significant difference in SOX11 gene regulation for the two cell lines that should not be due to the different expression platforms used. Normalization was performed with the Robust Multichip Analysis (RMA) algorithm and adjusted for false discovery rate (FDR). Expression is to some extent analyzed differently, the old HG-U133 Plus 2 (Vegliante et al. [188]) implements a Perfect Match (PM) and Mismatch Probe (MM) to remove noise, whereas the Human Gene 1.0 ST Array only implements PM (Conrotto et al. [187]). The PM and MM ratio is used to determine non-specific binding by inserting a mismatch in the central position of the probe [209].

The relevance of cellular context and expression levels have previously been discussed for several SOX proteins. Here, the cellular context is highly similar (equal cell line, Z138) if
no genetic drift has occurred that would explain their dissimilar response to reduced levels of SOX11 protein in the cell. Z138 has high levels of SOX11 protein, but the cell line was derived from the more rare cases within the MCL entity (a MCL in leukemic phase that also has extremely prominent splenomegaly) [210]. Reduced levels of SOX11 protein were clearly observed in both studies, but how pronounced the reduction was for the studies might not be identical. Highly speculative, potentially the SOX11 expression levels have a larger influence than initially considered.

Hence, additional studies are needed in order to clarify the molecular function of SOX11 in MCL. The clinical relevance of SOX11 levels was addressed in Paper III by investigating how SOX11 mRNA expression correlated to different clinical parameters.

Figure 7: The intersect of differentially expressed genes (FDR<0.05) reported for the above mentioned studies. Only 10 genes are found to be overlapping for the two studies [187 and 188]. Data plotted in R.
6 COMMENTS ON THE METHODOLOGY

6.1 CELL LINES AND PRIMARY CELLS

Cell lines predominantly used in this thesis include the MCL cell lines (TP53-status indicated with wt or m [145, 211, 212]): Granta519(wt), JVM2(wt) and JeKo1(m) were obtained from DSMZ; Z138(wt) and Mino(m) were obtained from ATCC; Rec1(wt) was a kind gift from Dr. Christian Bastard, Ronan, France. Additionally, one murine pro-B cell line, Ba/F3 [213] was used.

Primary lymphoma cells were isolated from tumor biopsies, mainly lymph nodes and tonsils, followed by filtration to remove cell debris and clumps and washed in PBS. Lymphocytes derived from bone marrow and blood were isolated by Ficoll separation. Cells were viability frozen in 10% DMSO, 40% FBS and 50% RPMI-GlutaMax medium at -150°C.

For all experiments, cell lines were cultured in RPMI 1640 GlutaMAX™ (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS), 50 µg/ml gentamicin (Invitrogen) at 37°C and 5% CO2. For Ba/F3, 5% of IL-3 supernatant from transfected X63 cells previously described by Karasuyama et al [214] was also added. For ecotropic pseudotyped virus production, Phoenix cells (293T) were used. Cells were maintained in complete DMEM medium and in OptiMEM medium during transfections.

The different MCL cell lines are all positive for cyclin D1, but they differ in morphology and proliferation rate. The level of SOX11 overexpression also varies between the cell lines. Importantly, SOX11 expression is not detected in JVM2, which provides a good negative control when validating experiments based on different detection methods (e.g. western blot or flow cytometry).

6.2 INDUCIBLE EXPRESSION SYSTEM

The complete coding sequence encompassing specific cloning sites with a Kozak sequence were generated from purified plasmids containing the mouse Sox11 gene (Q7M6Y2 plasmid kindly provided by Dr. Veronique Lefebvre, Dept. of Cell Biology, Cleveland clinic, Cleveland, OH, USA). Also used, but not shown in this thesis, were human full-length CDS or TAD truncated Sox11 and SOX11, in addition to Myc, Bmi1 and Bcl2l1.

To investigate the influence of the Sox11 gene in early B-cells without strong oncogenic drivers, the non-malignant Sox11-negative Ba/F3 cell line was used. Previously, Ba/F3 cells have been used when attempting to establish a blastoid MCL murine model with cyclin D1 overexpression [215]. However, as a MCL model, it does not recapitulate the phenotype and all the secondary genetic alterations. Instead the cell line was used to investigate the impact of Sox11 on early B-cells with regards to global gene expression, viability, differentiation and tumorigenic capacity. In order to investigate the effect of SOX11 protein expression a doxycycline-inducible murine retroviral vector with an internal ribosomal entry site (IRES) was employed to transduce primary murine B-cells and the Ba/F3 cell line. The IRES enables co-expression of Sox11 and enhanced GFP [216] in one vector.
This allows expression to be indirectly monitored by flow cytometry. Using an inducible expression system allows for cells from the same culture to be studied in the experiments (they come from the same culture before the experiment and not cultured separately as the case for constitutive overexpression/downregulation models).

6.3 GENE EXPRESSION

In this thesis, two different methods were employed to investigate gene expression (as measured by mRNA transcript levels), quantitative PCR and DNA microarray.

6.3.1 Quantitative PCR

qPCR was conducted using SYBR green, a fluorescent dye which intercalates with each new copy of double-stranded DNA (dsDNA) [217] produced from the PCR reaction. As SYBR green binds all dsDNA, the increase in fluorescence is proportional to the amount of synthesized DNA. The main advantage using a fluorescent dye that can intercalate with any dsDNA is that no specific probe binding to the target is required, instead well-designed PCR primers are used, which often is faster and less costly. However, emphasis is on well-designed primers and characterized PCR product as any dsDNA will be detected. We used SYBR green to analyze gene expression in Paper I-IV by the ΔΔCt method by having β-actin as reference and tonsil B-cells as control to calculate the ΔΔCt value and to determine the relative fold increase (RFI). Drawbacks with qPCR in general are sample preparation, template amount, primer specificity and non-proportional fluorescence to amount of produced PCR product since longer fragments risk binding more SYBR green. Therefore, it is important normalize expression using genes, known as housekeeping genes, with stable expression over different samples and treatment groups. We used ACTB, which has shown to be a stable housekeeping gene [218] when investigating all primary MCLs (Paper I-III) and for our McrBC-based methylation assay (Paper I).

All SOX11/Sox11 gene expression experiments were conducted in biological and technical triplicates for Paper I-IV and evaluated using the SOX11- cell lines JVM2 and Ba/F3. Primary SOX11- MCLs by IHC were also used for this.

When analyzing over 100 primary samples for quantitative levels of SOX11 expression, qPCR is a relatively fast and robust method. It gives a better resolution of quantification than relying on IHC stained tissue, which is more of a binary method (even though the methods of detection has improved with computer based quantification). The qPCR method allows quantitative analysis of a specific marker, which can be correlated to clinical parameters (Paper III). The major drawback of performing this type of study using qPCR is the inability to explore expression of a large set of genes at the same time for the same sample. The alternatives are DNA microarray and DNA/RNA-sequencing (the latter more expensive, but also more informative). They provide expression levels for all genes in addition to the gene of interest, which is a major advantage when knowing how complex MCL is. However, the specific purpose was to investigate how SOX11 mRNA levels correlated to clinical parameters, making qPCR a reasonable method for this type of study.
6.3.2 DNA microarray

For **Paper IV**, the Affymetrix GeneChip™ Mouse Gene 1.0 ST Array for complete expression profile of mRNA and long intergenic non-coding RNA transcripts (lincRNA) was used. Here, the array was used to assess changes in global gene expression to *Sox11* expression in a non-malignant B-cell that significantly changed the morphology of the cells. The Affymetrix concept is based on *in situ*-synthesized high-density oligonucleotide probe DNA microarrays where short oligonucleotides at specific coordinates are synthesized on the array. Multiple probes per gene are used as well as perfect match probes (100% sequence identity to the gene) with two sets of negative control probes: “antigenomic background probes” (sequences with no match in genome) and “genomic background probes” (sequences unlikely to be transcribed). RNA extracted from *Sox11* and *SOX11* transduced Ba/F3 cells, normal Ba/F3 cells, all cultured with and without doxycycline, were converted to biotin labelled complementary RNA (cRNA) and hybridized to the array.

Microarrays provide a straightforward comparison and normalization between samples compared to RNA-sequencing, which could have certain algorithmic and logistical challenges in relation to data analysis and data storage. However, RNA-sequencing also provides a way to find novel transcripts, allele-specific expression and splice junctions since it does not rely on any genome annotation for probe design [219]. Thus, it could be feasible as an initial comparison to first investigate differences by microarray and then perform a more comprehensive study by sequencing. We were interested in the particular differences in gene expression levels between ON and OFF conditions to find genes responsible for the observed morphological effect. What changed when turning on *Sox11* to that extent that the cells started to form aggregates and stopped dividing? The downside is difficulties to detect low abundant gene transcripts. The relationship between probe intensity and transcript levels in relation to its affinity during hybridization is also a concern. Moreover, the expression is only relative and not absolute. Although certain drawbacks, implementing the microarray approach in this context is feasible as a first step since we were not primarily searching for novel transcripts. Instead we implemented it to recognize significantly differentially expressed genes. In the future, specific patient samples with those genes could then potentially be *de novo* sequenced for a deeper understanding. In turn, the collected result from those experiments could be validated by microarrays to reduce costs. One approach to reduce cost would be to use them not interchangeable, but for well-defined purposes. In particular, microarrays could be used to find interesting cases among a large sample collection. The interesting cases could then be investigated by RNA/DNA-Seq and finally validated by microarray in large cohorts.

6.4 EPIGENETICS

In **Paper II**, three different approaches were used to study whether or not *SOX11* expression is regulated via promotor methylation. As suggested when choosing a suitable method for DNA methylation analysis for a known gene [220], the following methods were used. The use of two experimentally different concepts also supports the results. The shortcoming of these methods are that neither restriction or bisulfite conversion are able to discriminate between 5-methylcytosine and 5-hydroxymethylcytosine DNA [221] (info below).
6.4.1 McrBC-based methylation assay
DNA was treated with McrBC endonuclease, which digest DNA at methylation sites (5-methylcytosine, 5-hydroxymethylcytosine or N4-methylcytosine) [222]. Cleaved DNA indicates higher degree of methylation, detected by difference in ΔCt between McrBC-treated and control. Hence, methylated CpGs had higher ΔCt value for their respective PCR fragment.

6.4.2 Bisulfite treatment and pyrosequencing analysis
Sodium bisulfite treatment of DNA deaminates unmodified cytosines to uracil while leaving 5-methylcytosine or 5-hydroxymethylcytosine DNA. In the subsequent PCR amplification reaction of the selected promotor fragment, the uracils are amplified as thymines whereas modified cytosines are resistant and get amplified as cytosines [223]. Next is detection by pyrosequencing. Based on the principle of sequencing by synthesis, the nucleotide sequence of the fragments are detected (for an in-depth description [224]). The ratio based on signal intensity for incorporated dATP and dGTP determines to what extent the fragment is methylated. Given that pyrosequencing is very sensitive and can detect subtle differences in methylation, tumor samples are highly suitable as they often are not very homogenous [220].

6.5 PROTEIN DETECTION
Several different methods are available to detect and identify protein. In this thesis western blot (WB), immunohistochemistry (IHC), immunocytochemistry (ICC) and flow cytometry were used.

6.5.1 Western Blot
WB was used in Paper I, II and IV. The technique is based on separating proteins by size that is proportional to their charge and detecting those using antibodies in a later stage. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), where denatured and negatively charged proteins were separated according to size as smaller proteins migrate faster through the gel than larger proteins (towards a positively charged electrode), was used to separate proteins. The proteins were transferred onto a polyvinylidene fluoride (PVDF) membrane, blocked and stained with a primary antibody, followed by a secondary antibody. Enhanced chemiluminescence (ECL) solution was used for detection of the specific protein. Two different SOX11 antibodies, one polyclonal (pAb) (Paper I, II and IV) and one monoclonal (mAb) (Paper II and III) were used in this thesis. Recognizing only one epitope, greater specificity is achieved with a monoclonal antibody compared with a polyclonal antibody that also can suffer from batch-to-batch variance. Also, working with the highly similar SOXC proteins, cross-reactive could be an issue (however, recognizing multiple epitopes could be beneficial for co-immunoprecipitation where the epitope could be masked). The same batch of pAbs were used for experiments and specificity validated by knockdown of SOX11 by siRNA and the use of JVM2 as negative control. The downside of using chemiluminescent detection for WB is the risk of signal saturation and it is not the best quantitative method, although very sensitive. Alternative, fluorescently labelled secondary antibodies can be used for detection. It allows for multiplexing and a better quantitative detection, since the fluorescent signal is directly proportional to the amount of protein. However, WB is still best viewed as a semi-quantitative method to detected target protein. Lastly, WB further requires a reference protein that is well selected and do not change upon treatment/manipulation (similar to qPCR).

6.5.2 Immunohistochemistry (IHC) and Immunocytochemistry (ICC)
IHC and ICC are important to validate SOX11 expression and its nuclear localization in MCL cells in patient tissue or primary cell samples. IHC staining for SOX11 was...
performed on whole sections of paraffin-embedded patient biopsies by a Bond Max robot (detailed description in [54]). We evaluated the pAb and mAb SOX11 antibodies in Paper III, 7 cases previously negative for SOX11 with the pAb became positive with the mAb. This difference could be due to higher specificity of the mAb, which can detect lower amount of protein in MCL cases. As WB, IHC should also be considered as a semi-quantitative method to detected target protein. Staining for fluorescent microscopy was used to verify nuclear specify of the MRQ-58 SOX11 mAb. The staining procedures require optimization (fixation, permeabilization etc.) and good controls, especially negative controls (both for the primary antibody and the target tissue/cell) since non-specific staining can be a problem.

6.5.3 Flow cytometry
Cells are injected into the flow cytometer where they are aligned into single cells in a narrow flow-channel after which they are exposed to a laser and excited light is captured by detectors. When quantifying protein expression, intracellular flow cytometry provides certain advantages over WB and IHC. In addition to simplicity and reduced time, the main advantage is the possibility to employ more markers for phenotyping of cells expressing SOX11. In Paper II, we demonstrate that protein expression measured by flow cytometry correlates to that of WB, but it gives a higher resolution as fluorescent intensity was used compared to chemiluminescence. Drawbacks are permeabilization methods, which require optimization and might not always work for fragile cells. This is something we noticed in Paper II; Triton X-100 gave higher SOX11 intensities, but did not preserve the forward scatter/side scatter (FSC/SSC) distribution to the same extent as saponin (the other solution we investigated). The negative control cell line, JVM2 was also significantly affected by this permeabilization and could therefore not be used. Knockdown by siRNA or shRNA to validate SOX11 serves as a good substitute. Saponin interacts with membrane cholesterol and removes it, resulting in holes in the membrane while Triton X-100 is non-selective and thus potentially tougher on the cell membrane [225].

6.6 VIABILITY AND PROLIFERATION ASSAYS
Cell viability measurements have been performed in Paper I-IV. They all have their pros and cons, and this is the reason to use all of them in Paper IV when studying viability and proliferation.

6.6.1 Trypan blue exclusion
Trypan blue exclusion is the most common method for counting living cells and it is based on the concept of an intact cell membrane exclude the trypan dye. Hence, trypan blue measure membrane integrity. For the cell lines used in Paper I-IV, trypan blue worked reliably (without an automated cell counter, the observer subjectively determines if the cell is blue or not) to observe stained and unstained cells. However, primary cells rapidly turned blue despite being alive. It could be due to temporary weaknesses of the membrane integrity, either as consequence of damage during the isolation procedure or because they are more permeable by nature. So, for primary cells especially, multiple rounds of counting were performed for every experiment.
6.6.2 XTT
XTT (2,3-Bis-(2-Methoxy-4-Nitro-5-Sulfophenyl)-2H-Tetrazolium-5-Carboxanilide) is an easy and fast method for assessing cell viability based on mitochondrial activity of viable cells (used for Paper IV). It does not require any radioactive labelling. The water soluble yellow colored tetrazolium salt, XTT, is cleaved by mitochondrial dehydrogenases and forms an orange formazan dye. Only metabolically active cells are able to perform this conversion. The dye is then measured by spectrophotometry, which directly correlates to total metabolic activity and indirectly to the sum of viable cells. The disadvantage of XTT is that the assay measures metabolic activity and this is not always the same as viability or proliferation. Cells could be alive, but have a reduced metabolic activity as shown for several cell types. Contrarily, cells could be equipped with very active metabolic enzymes and co-factors. Therefore, using the correct control is of importance and understanding the limitations of what conclusions that can be drawn from the results.

6.6.3 ³H-Thymidine incorporation
³H-Thymidine incorporation was used for Paper IV when investigating proliferation. It is one of the more reliable methods to measure proliferation since radiolabeled ³H-Thymidine is incorporated into newly synthesized DNA strands, hence it directly corresponds to DNA synthesis. Incorporation is also an advantage compared to the XTT assay that relies on metabolic activity, which can be disturbed by different compounds. For detection, labelled DNA is recovered from the cells and radioactivity measured by a scintillation beta-counter.

6.7 BIOINFORMATICS
Bioinformatics and statistical analysis were performed in R and based on the Bioconductor platform and its packages.

source("https://bioconductor.org/biocLite.R")
biocLite("biomaRt")

For microarray analysis performed, CEL files were imported in to the transcriptome analysis console program TAC (v.1.4.1) and normalized by RMA. Pair-wise comparisons between expression values for different samples were interrogated for differential gene expression by two-tailed unpaired Student's t-test and the results were adjusted for multiple testing by false discovery rate (FDR). Significantly differentially expressed genes were genes with a FDR q-value ≤ 0.05.

For comparison between Ba/F3 data and other human expression studies orthology conversion of mouse gene symbols to official human gene symbols was performed via the biomaRt bioconductor package (biomaRt v2.30.0, Bioconductor v3.4, R v3.3.2) with the Ensembl annotation database for mouse (useMart("ensembl", dataset="mmusculus_gene_ensembl")) and human (useMart("ensembl", dataset="hsapiens_gene_ensembl")).

The cell cycle database Cyclobase (http://www.cyclebase.org) was used to retrieve and identify cell cycle related genes for Paper IV. It is an online database that combines data from mRNA expression profiling and quantitative proteomics, but also microscopy-based knockdown screens [226].

The method section in Paper IV describes more specific analysis used in this thesis.
7 OVERALL AIM
The overall aim of this thesis focuses on understanding the biological function and clinical relevance of the neural transcription factor SOX11 in mantle cell lymphoma and non-malignant B-cells.

7.1 SPECIFIC AIMS

**Paper 1:** To define the expression of the SOXC group of transcription factors in MCL and non-malignant B-cells, and how the SOX11 gene is epigenetically regulated.

**Paper 2:** To develop a method for accurate differential diagnosis of lymphoma entities similar to MCL by flow cytometry.

**Paper 3:** To investigate how SOX11 expression levels correlate to clinical parameters and if there is a natural cut-off that can be used to define negative cases.

**Paper 4:** To investigate how SOX11 regulates cell fate and differentiation in early non-malignant B-cells
8 RESULTS AND DISCUSSION

8.1 PAPER I
SOXC transcription factors in mantle cell lymphoma: the role of promoter methylation in SOX11 expression

Analysis of the relative expression of the SOXC genes, SOX11, SOX4 and SOX12, was performed on primary cyclin D1+ MCLs having the t(11;14)(q13;q32) translocation, and in non-malignant tissue from tonsil, lymph node and spleen. Out of 29 MCL samples examined by IHC, 27 samples could be assessed and 24/27 displayed SOX11 nuclear positivity. Non-malignant lymphoid tissue and 3 primary MCLs were negative for SOX11.

Expression levels for the SOXC genes for all 29 samples, non-malignant lymphoid tissue and the four MCL cell lines Granta519, Rec1, JeKo-1 and JVM2 were determined by qPCR. High SOX11 mRNA levels were detected in all primary MCLs compared to non-malignant B-cells, except the 3/27 immunohistochemistry in the negative cases, which only expressed minute levels. In contrast, expression of SOX4 was variable over the cohort: 13/29 samples expressed less mRNA than non-malignant B-cells, 6/29 samples similar levels of mRNA as non-malignant B-cells and 10/29 higher levels of mRNA than non-malignant B-cells. We found higher expression of SOX12 in 27/29 samples compared to non-malignant B-cells, which was independent on SOX11 status. The expression of SOXC varied in the 4 MCL cell lines. In SOX11+ MCLs, the mRNA levels of SOX11 and SOX4 mRNA, as well as SOXI1 and SOX12, significantly correlated (this was only moderate when including all MCLs and non-malignant tissues).

The degree of promoter methylation for SOX11 in MCL was investigated by McrBC endonuclease digesting of DNA at methylation sites and bisulfite promoter sequencing. Both methods demonstrated low levels of promoter methylation for primary MCLs, irrespectively of SOX11 status. Thus, within the entity of MCL, there was no correlation between SOX11 expression pattern and the methylation status of investigated CpG regions.

After confirming low promoter methylation status, we further established non-promotor regulated expression with a demethylating agent, 5-azacitidine (5-AZA). No increase in SOX11 expression was observed for Granta519, Rec1, JeKo-1 and JVM2 cells when treated over 144 h with 5-AZA. Instead, 5-AZA decreased SOX11 expression by promoter methylation independent mechanisms in Granta519 and Rec1 (SOX11+ MCL cell lines). The SOX11+ MCL cell lines Granta519, Rec1 and JeKo-1 did not show any change in SOX4 mRNA levels. However, SOX4 mRNA levels increased 23-fold for JVM2 (SOX11-). The SOX12 mRNA levels remained unchanged for all cell lines.

Discussion
In Paper I, we investigated the mRNA levels for each member of the group to obtain a better understanding of their potential role in MCL. Patients diagnosed as SOX11+ by IHC also expressed high levels of mRNA. However, expression of the other SOXC members differed. SOX4, expression varied between the different patients compared to SOX12, which was overexpressed in most cases. Hitherto SOX12 has not been reported in MCL, but since SOX11 and SOX12 compete for the same binding targets with SOX12 having higher affinity and lower TAD activity (in vitro) the implication of its expression should be investigated. It could be relevant in cases with very high SOX12 expression and potentially have an impact in MCL. Even though expression of the two other SOXC genes was significantly lower, there is a possibility of co-regulation of the genes in MCL, as their expression positively correlated when investigated in all SOX11+ MCLs. Two previous
reports have investigated how $SOX11$ is regulated in MCL (Gustavsson et al. [197] and Vegliante et al. [181]). Our results corroborate the results from both of these studies. In the context of MCL, the $SOX11$ promoter region is hypomethylated (independent of SOX11 status defined by IHC), but not regulated by promotor methylation. Expression is instead regulated by histone modification. However, a later study with 42 different cancer cell lines (10 MCLs, not methylated) [199] showed that the $SOX11$ promotor region is variably methylated in other types of lymphoma and in cancers. The promotor region was not heavily methylated from our McrBC-based and pyrosequencing assays. However, to further confirm that SOX11 expression is not regulated via promotor methylation, we treated the cells with 5-AZA. As predicted, the mRNA or protein levels of $SOX11$ did not increase, which indicates that the promotor is not methylated. However, mRNA and protein levels decreased for two SOX11+ MCL cell lines. Certain toxicity has been reported for the 5-AZA drug [227], which mechanism might stress the two affected cell lines. Granta519 and Rec1 both have wild-type $TP53$ compared to JeKo-1. In our data, $SOX11$ mRNA levels did not increase by 5-AZA in Raji (SOX11- BL cell line), whereas the histone deacetylase (HDAC) inhibitor SAHA induced $SOX11$ mRNA expression in JVM2 and Raji [181]. From our study, we conclude that SOX11 expression is not regulated via promotor methylation. Instead, other mechanisms are responsible, most probably by histone modification, but potentially also by microRNAs or distant enhancer regions.

**Conclusion**

In addition to $SOX11$, $SOX12$ is also overexpressed in MCL (independent of SOX11 status by IHC). This finding could be of interest as its transactivation activity is lower than that of SOX11 and in *vitro* data [128] suggests that the SOXC proteins, when co-expressed in the same cell, compete for the same target genes. We can further conclude that the $SOX11$ promoter region in MCL is hypomethylated in SOX11+ MCL and non-malignant lymphoid tissue. The few SOX11- MCLs were heterogeneously methylated, but at low levels. Based on these observations, $SOX11$ gene activation is not regulated by promotor methylation in MCL or non-malignant lymphoid tissue.
8.2 PAPER II

Flow cytometric analysis of SOX11: a new diagnostic method for distinguishing B-cell chronic lymphocytic leukemia/small lymphocytic lymphoma from mantle cell lymphoma

For Paper II, a flow cytometric protocol for intracellular staining of SOX11 was developed. The purpose was to be able to integrate this method in a flow cytometry panel for lymphoma diagnosis. This would allow for a fast discrimination between MCL and other B-cell lymphomas that share a certain morphological and phenotypical overlap, such as B-CLL, which do not express SOX11.

In this study, the mouse monoclonal anti-SOX11 MRQ-58 antibody was used in six MCL cell lines (Granta519, JeKo-1, JVM2, Rec1, Z138 and Mino) and primary lymphoma cells. SOX11 specificity was verified after 24 h transient knockdown by siRNA and subsequent detection by WB in two SOX11+ MCL cell lines (Granta519 and Z138). A single band between 50 and 75 kDa indicating SOX11 was observed for the non-treated and scrambled siRNA-treated samples, whereas the SOX11 siRNA-treated samples had significantly reduced levels of the corresponding band. The subcellular localization of SOX11 was confirmed to be in the nucleus (highly specific nuclear staining) for the SOX11+ MCL cell lines by confocal fluorescent microscopy.

Two different types of permeabilization methods were optimized for Rec1 (the cell line with lowest SOX11) expression: Triton X-100 at 0.1% in PBS or saponin 0.1%. The saponin permeabilization better preserved FSC/SSC distribution, but Triton X-100 offered higher SOX11 staining intensities. A higher fluorescent intensity compared to their isotype control staining was observed for SOX11+ cell lines when applying the staining protocol. The SOX11- cell line JVM2 was too sensitive for this type of permeabilization and could not be included. SOX11 distribution varied for Granta519 (detected by both flow cytometry and confocal microscopy). SOX11 expression levels could reliably be quantified using flow cytometry with the geometric mean fluorescence intensity (MFI) values for flow cytometry, which correlated with the density measurements of the single band detected by WB using the MRQ-58 antibody (Pearson's correlation coefficient 0.97, \( p < 0.007 \)). There was also a strong correlation between protein and mRNA levels, analyzed as MFI and \( 2^{-\Delta\Delta C_t} \), respectively (Spearman's correlation coefficient 0.75, \( p < 0.001 \)). When applying the protocol on B-CLL and MCL primary samples, the MFI was obtained on the monotypic \( \lambda \) or \( \kappa \) positive population, whereas SOX11 mRNA levels were based on all mononuclear cells isolated from the tissue without enrichment for tumor cells.

Permeabilization of tumor cells from MCL (12/12 cyclin D1+, 11/12 SOX11+ by IHC/ICC) and CLL (8 patients) was successful for all cases. The SOX11 MFI ranged from 12.6-87.3 in the samples defined as SOX11+ by IHC/ICC, whereas the MFI for the B-CLL samples varied between -2.3 and 1.4. The SOX11- MCL sample had a MFI of 2.2. This protocol successfully distinguished the MCLs from CLLs based on SOX11 as measured by flow cytometric intracellular staining.

Discussion

In Paper II, we investigated how well SOX11 could be implemented in a potential clinical setting to separate MCL from CLL. Normally, immune phenotyping by flow cytometry can discriminate between MCL and CLL (differential expression of surface immunoglobulins and the CD79b, CD20, FMC7, CD23 and CD200 markers). However, for certain cases there might be phenotypical overlaps, which pose a problem for diagnosis.
Since both prognosis and therapy differ between MCL and CLL, a precise diagnosis is of clinical importance. Other mature B-cell lymphomas, such as marginal zone lymphomas (MZL) may sometimes be CD5+ and thus be difficult to classify by flow cytometry. SOX11 is not expressed in CLL, or most other mature B-cell lymphomas, making it a good marker for MCL in a flow cytometry panel.

A SOX11 polyclonal antibody was used for detecting SOX11 in Paper I. Recently several new antibodies have become available. For this project we used the mouse monoclonal MRQ-58 antibody that previously has been shown to have the highest specificity of tested SOX11 antibodies (we also validated the antibody by siRNA, our SOX11- cell line JVM2 and by confocal microscopy for nuclear localization specificity). An advantage with this antibody is that it does not to cross-react with the highly similar SOX4 [194] that also could be expressed in MCL, as shown in Paper I. SOX11 analysis by flow cytometry correctly identified all MCL cell lines and patient samples from the CLL patient samples (also negative by IHC). However, the number of cases included was too small to determine a cut-off level for SOX11-negativity by MFI. SOX11- MCLs will be difficult to differentiate from highly similar CLL cases, but that is regardless of method as long as no other exclusive markers for the different entities are used. Protein quantification of SOX11 by MFI was accurate when correlating to WB intensity, but MFI also correlated to mRNA levels. In addition to protein quantification, assessment of SOX11 by flow cytometry also allows for a more rapid analysis than IHC. It also permits simultaneous analysis of multiple markers allowing further investigation of phenotypical differences within the tumor cell populations expressing SOX11.

**Conclusion**

This method provides a new diagnostic tool for MCL that allows accurate quantification of SOX11 levels. The method is still not used in routine diagnostics due to lack of a conjugated SOX11 monoclonal antibody, however, routine labs have asked about details regarding the method. Quantifying SOX11 can also be used for several different research applications that combine SOX11 expression with phenotypical markers.
The utility of mRNA analysis in defining SOX11 expression levels in mantle cell lymphoma and reactive lymph nodes

To examine how SOX11 mRNA levels correlate to certain clinical parameters, 102 MCL cases (lymph node biopsies n=80, tonsils n=6, bone marrow cells n=4, spleens n=3, peripheral blood cells n=7, pleural cells n=1 and gastrointestinal biopsy n=1) and 16 reactive lymph nodes were investigated by qPCR. We further investigated if there is a natural SOX11 expression cut-off for defining positive or negative MCL cases by comparing qPCR and IHC data from the same patients.

Additionally, the performance of the newly implemented mouse monoclonal anti-SOX11 MRQ-58 antibody was compared with that of the previously used polyclonal anti-SOX11 HPA000536 antibody. Several of the cases that were negative with the HPA000536 antibody turned out to be positive after re-analyzing them with the MRQ-58 antibody. The new SOX11 IHC-positivity was more in line with the corresponding mRNA levels for the same cases.

For the 102 MCL cases, the relative fold increase (RFI) of SOX11 ranged from 0.01 to 7419.41. There was no correlation between tumor cell content and SOX11 RFI values (p=0.73). To study any potential clinical significance of low versus high levels of SOX11 mRNA expression in MCL, samples were grouped based on expression and correlated to clinical parameters. Clinical and pathological characteristics for the 10% cases with the lowest SOX11 mRNA versus the remaining cases were only significantly different for two features, nodal disease (22% in low vs 68% in high, p=0.01) and lymphocytosis (60% in low and 16% in high, p=0.005). The quartile of patients with the lowest SOX11 mRNA levels had significantly shorter OS in the group of 73 patients who had not received ASCT.

A potential SOX11 mRNA cut-off level in reactive lymph nodes was examined for 16 samples, showing a RFI ranging from 1.26 to 58.35. Although significantly lower than for all MCLs (p<0.0001), there was no significant difference from the 10% of cases with lowest SOX11 mRNA (p=0.24). IHC of the reactive lymph nodes also detected a few small SOX11+ cells (6/16) in the mantle zone of CD20+, but cyclin D1- cells.

The RFI grand mean for MCL samples and reactive lymph nodes was 5.7. Applying this criteria, 94/102 MCL samples and 2/16 reactive lymph nodes were above. A RFI above 5.7 was detected in 1/8 of the SOX11- MCLs (IHC) and 93/94 (98.9%) of the SOX11+ MCLs (IHC). Using the above cut-off value resulted in misclassification of only 2/102 by qPCR and IHC. However, for the SOX11+ cases by IHC, the qPCR analysis was not able to find a natural cut-off that would recognize cases with low expression.

Discussion
In Paper III, we wanted to investigate how SOX11 expression levels correlated to clinical parameters and to define a cut-off value for SOX11- MCL. Knowing how the different SOX proteins are able to compete for the same targets, potentially the difference in survival reported for SOX11+ MCLs in large cohort studies depends on the actual SOX11 levels in the tumor. IHC/ICC studies on large cohorts have been conducted. A weakness of IHC is that it is a semi-quantitative technique for quantification and subjective due to interobserver variation with no defined cut-offs. Here, we used IHC/ICC as a binary tool (SOX11 +/−) and to directly compare our two previously used anti-SOX11 polyclonal HPA000536 (Paper I) and monoclonal MRQ-58 (Paper II) antibodies with each other and to qPCR. Despite the large span among the investigated MCLs (RFI ranged from 0.01 to 7419.41),
the RFI values for all SOX11+ MCLs (IHC) were normally distributed. The SOX11 RFI could thus not identify a natural cut-off for cases with low RFI. Moreover, there was no correlation between tumor cell content and SOX11 RFI. However, among patients who had not received ASCT, those with the lowest SOX11 RFI had a significantly shorter OS. Of note, the cases were selected based on available material and the patients were not uniformly treated. Nonetheless, it is an interesting finding that deviates from the common view of SOX11 being an oncogene in MCL. SOX11 also had a reducing effect on proliferation/viability in Paper IV, which has been reported for MCL cell lines in some studies [187, 197, 228]. The SOX11 RFI low group was not significantly different from that of reactive lymph nodes. Using a grand mean for a SOX11 RFI cut-off resulted in a discordant classification of only 2/102 MCLs by qPCR and IHC in our study, showing that both methods are reliable. An advantage of IHC is the possibility to detect where the SOX11+ cells are localized in the tissue, since some of the reactive lymph nodes contained a few CD20+ and SOX11+ cells situated in the mantle zones. It also emphasizes the importance of scoring SOX11+ only in the tumor area. Finally, we were able to show that low SOX11 mRNA levels were significantly associated with less nodal disease and higher lymphocytosis compared to high SOX11 mRNA levels. This is in accordance with IHC data stating that low SOX11 expression often is associated with non-nodal leukemic disease [194].

**Conclusion**

An important finding regarding transcriptional levels of SOX11 is that there seem to be a gradient and cut-off when it comes to expression: completely negative, weak, intermediate and strong. However, low expression and high expression only correlated to two clinical parameters. Nevertheless, applying an agreed SOX11 mRNA cut-off or stratify groups based on expression levels could benefit future cohort studies to better compare patients. MCLs should only be defined as positive if SOX11 is detected within the tumor region by IHC, independently of weak or variable SOX11 expression.
8.4 PAPER IV

Sox11 promotes phenotypical changes and alters the global gene expression pattern in pro-B cells

The effect of expressing Sox11 was investigated in non-malignant B-cells. A TetON system was used to induce ectopic Sox11 mRNA expression in the IL-3 dependent pro-B cell line Ba/F3. SOX11 protein expression in this cellular context clearly influenced cell morphology. Expression of Sox11 was induced for 72 h by culture in doxycycline supplemented medium (Sox11-ON). During a culture period of 72 h, the cells formed large aggregates, which was not observed in their non-induced counterpart (Sox11-OFF) without doxycycline.

Viability and proliferation by trypan blue exclusion, XTT, ³H-Thymidine incorporation and PI-staining demonstrated significant differences between the Sox11-ON and Sox11-OFF cells.

Whole genome expression profiling by microarray after 72 h of Sox11 induction demonstrated that the altered phenotype observed in Sox11-ON cells also induced profound alterations in gene expression. When comparing Sox11-ON and Sox11-OFF cells, 7980 genes showed significantly altered transcript levels (FDR q-value ≤ 0.05 and 871 of those altered more than 1.5-fold (533 genes up and 338 genes down). The gene showing the highest fold change, Mmp8, is involved in proteolysis. Two other highly expressed genes were the protocadherin beta genes Pcdhb17 and Pcdhb16. Genes promoting cell proliferation and survival were among the more significantly downregulated genes together with certain cell adhesion genes.

The leading edge from the Gene Set Enrichment Analysis (GSEA) resulted in 199 enriched gene sets for Sox11-ON and 31 enriched gene sets for Sox11-OFF (FDR q-value ≤ 0.05). Functions characterizing earlier stages of the cell cycle were attributable to 113/199 gene sets with positive enrichment score in Sox11-ON cells, whereas 21/199 gene sets were associated with cell cycle stages later than that of the S-phase. The Sox11-OFF cells showed a significantly higher enrichment score for metabolic activity (defined by gene sets related to lysosome, ribosome, cellular and cation homeostasis). Immune system associated gene sets were also significantly downregulated in Sox11-ON cells.

When using genes from Cyclebase 3.0 [226], the Sox11-ON and Sox11-OFF cells displayed a significant difference in microarray mean intensity values for genes associated with early phases as compared to later phases of the cell cycle.

Sox11 did not evoke significant changes in expression of key genes associated with later stages of B-cell differentiation when expressed in Ba/F3 cells. However, the transcript levels for two pro-B cell restricted genes Id1 and Tal in Sox11-ON (FDR q-value ≤ 0.05, Fold Change: 1.2 and 1.3, respectively) significantly increased upon Sox11 induction.

Gene expression data from sorted murine B-cell populations [229] and Ba/F3 Sox11-ON/OFF revealed that the gene expression profile of Sox11-ON cells is different from both Sox11-OFF and non-transduced Ba/F3 cells as well as cells at other B-cell differentiation stages. In spite of large impacts on global gene expression by Sox11, no significant changes in expression were observed for key B-cell stage genes.
**Discussion**

In Paper IV, we investigated the potential role of SOX11 in hematopoiesis or lymphopoiesis. SOX4 is important for survival of pro-B cells, but nothing is known about SOX11. In MCL, SOX11 has also been suggested to function as an oncogene [188]. However, studies regarding the influence of SOX11 on cell proliferation in MCL cell lines and tumor growth in xenografts have yielded conflicting results [187, 188]. We found that ectopic expression of Sox11 in Ba/F3 pro-B cells is associated with profound phenotypic changes resulting in cell aggregation and reduced cell proliferation. When investigating the minimal requirement of transforming a healthy murine B-cell to a cancer cell using combinations of Myc, BclXL and Bmi1 (a concept developed by Alf Grandien), cells were not “immortalized” when replacing any of these genes with different Sox11 constructs (data not shown). Instead, we focused on investigating the role of Sox11 in early B-cells. Gene expression analysis showed an enrichment of differentially expressed genes associated with earlier phases of the cell cycle in Sox11-ON cells. We found that Sox11 expression induced homotypic aggregation in Ba/F3 cells, potentially by protocadherin family members. Of interest, Pcdhb17 is proposed to block cell migration and reduce cell proliferation [230, 231] and methylation of PCDHB17 has been correlated with increased risk of relapses and higher mortality in BCP ALL [232]. The two genes S100a8 and S100a9 also showed high transcript levels in Sox11-ON cells. Consistent with our study, increased adhesion, reduced migration and impaired tumor growth and reduced transcript levels of the immunomodulatory cytokine CCL3 have been observed when overexpressing S100A8/S100A9 in human cervix carcinoma-derived cells [233]. Sell was significantly downregulated in Sox11-ON cells; SELL has also been reported to be downregulated in MCL (that normally express SOX11) [234]. We showed by functional and gene expression data that SOX11 in Ba/F3 cells promote homotypic aggregation and expression of genes associated with cell to cell adhesion, rather than genes associated with cell migration. In MCL, SOX11 has been suggested to affect B-cell differentiation by decreasing PRDM1 (encoding for BLIMP1) and BCL6, and increase PAX5 expression, which affected surface levels of plasmacytic B-cell markers [188]. We could not detect any significant changes in mRNA levels for Pax5, Prdm1 or Bcl6 in response to Sox11 overexpression in Ba/F3 cells. This could potentially be because of their distinct cellular background and the function of SOX11 can be very context dependent. Reduced proliferation has also been shown after overexpressing of SOX11 in MCL cell lines [191, 197], similar to our findings.

**Conclusions**

Sox11 was not able to replace any of the minimal genes required to immortalize a healthy murine B-cell, indicating that it does not per se act as an oncogene. In early B-cells, the transcription factor impacted the morphology and proliferation rate of the cells, but not genes associated with B-cell differentiation previously reported. In the context of Ba/F3 cells, Sox11 may be involved in blockade of B-cell differentiation, but mainly Sox11 seems to be involved in pathways that are not directly related to B-cell differentiation.
9 CONCLUSIONS
SOX11 belongs to the highly homologous group of SOXC genes and can act redundantly in vitro and in vivo under certain conditions.

From the different experiments and analysis conducted in this thesis, a definite function of SOX11 and role in MCL or B-cells remain elusive. The two other SOXC members, SOX4 and SOX12 are also expressed in MCL, but significantly less. The reason for activating SOX11 in MCL is not established, but from this thesis it can be concluded that expression is not activated via methylation of its promoter. In both patients and cell lines, the degree of promotor methylation is low and demethylating of MCL cells cannot activate expression of the gene. Instead it is regulated by histone modification (potentially also microRNAs) in the context of MCL.

The use of SOX11 as a highly specific marker for MCL by different methods has been demonstrated here by IHC, flow cytometry and qPCR. As SOX11+ by IHC has been reported to be associated with both positive and negative clinical outcomes, we quantified the levels of SOX11 expression. High expression correlated to longer OS, but there was no correlation between the degree of gene expression and clinical parameters that could impact survival. We could further show that SOX11 mRNA detection by qPCR and SOX11 protein by the gold standard, IHC, corresponded very well.

The idea that SOX11 could have a role as an oncogenic driver was addressed experimentally. Sox11 was not found to be a strong oncogenic driver in primary murine B-cells or when transduced into the progenitor B-cell line Ba/F3. Ba/F3 cells, which are themselves not oncogenic, were chosen as a model to investigate the potential effect of SOX11 expression in B-cells at the stage of B-cell differentiation at which t(11;14)(q13;q32) is thought to occur. In this model, Sox11 did not activate any genes implicated in further B-cell differentiation or related to oncogenic potential. We rather observed pronounced morphological changes and reduced proliferation. Genes associated with adhesion were altered as well as several neural genes, some of which also have been reported in MCL. As illustrated in Figure 8, the lack of a coherent gene expression pattern when comparing SOX11 knockdown studies in MCL cell lines [187, 188, 201] and overexpression of Sox11 in Ba/F3 cells indicates a context dependent function of this transcription factor.
Figure 8: Intersect between the significantly differentiated genes (FDR<0.05) reported for the three above mentioned studies [187, 188 and 201]. Data plotted in R.
9.1 FUTURE PERSPECTIVES
To delineate the role of SOX11 in the MCL pathogenesis, a more coherent system on classifying positive and negative SOX11 MCL is clearly needed. Currently, data from different cohorts have used different inclusion criteria, e.g. including nmMCLs and not including nnMCLs. Furthermore, a clear definition of nmMCL is lacking. Without this information it will be difficult to assess the actual contribution of SOX11 in MCL pathogenesis and clinical behavior.

With the new flow cytometry staining procedure for SOX11, different subpopulations within MCLs can now be investigated.

SOX11 regulation is known to be associated with histone modification, and now potentially a distant enhancer element. Less is known about its actual physical binding partners. For SOX proteins in general their binding partners are very important for stabilizing the transcriptional complex and thus, their potential for activation and repression of genes. Finding out those binding partners in MCL would be of great importance for further understanding of the role of SOX11 in MCL.

From this thesis and published data, the function of the SOX11 gene seems to be pluralistic and contextual. Depending on expression level, highly similar SOX proteins with different functions (activation and repression) have been shown to be able to target the same genes, but induce different effect. Highly speculative, but worth to study would be to investigate the effect of overexpressing SOX12 in MCL to observe if its expression can interfere with SOX11.

To really set-apart the function of SOX11 in MCL cells would be a SOX11-concentration-dependent experiment that considers both the effects of high expression and low expression of the gene in the cells. One way would be the use of an inducible overexpression system and an inducible downregulation system for the same cell line where the levels of SOX11 are titrated up and down, respectively. Different type of cell lines similar to classical (e.g. JeKo-1 or Mino) and blastoid (e.g. Granta519 or Z138) MCL should be used. Cells experiencing a series of increasing/decreasing SOX11 protein levels would then be analyzed by RNA-seq and placed in the context of already known transcriptional targets in MCL.
10 ACKNOWLEDGEMENTS

Birgitta Sander, my very inspiring supervisor. I’m grateful to you for giving me the opportunity to join your group and work with lymphoma, a disease that I lost my grandfather to when I was young. I didn’t know what it was at the time, I only knew that I wanted to cure it, thank you for giving me the chance to make a small contribution to the cancer field. You’ve always been fair, encouraging and positive. I still remember the recruitment process with the assignment regarding the research article and how much I wanted to work for a boss who recruited employees in that way. Thank you for always taking the time to talk to me when I came with questions or discussions, I usually left our meetings in a much more positive mindset. Further, you have allowed and encouraged me to develop my own ideas for projects, which has been very stimulating. Thank you for being a great supervisor and person, and for making our small research group a joy to be a part of.

Birger Christensson, my knowledgeable and “underfundiga” co-supervisor. Thank you for all the advice and scientific input, you have answers or solutions to almost any situation. You always come with interesting ideas to investigate. Thanks to my other co-supervisor Alf Grandien for teaching me about B-cells, cloning, virus and FACS. Your help was deeply appreciated. My mentor Edvard Smith, I did not need you much, but it was good to know that you were there just in case, thank you.

To my two “SOXC” group members: Agata Wasik, my 10% super-dobra boss and 100% caring friend, and helpful colleague. Thanks for taking me under your wings when I first started and didn’t know much. You’ve always been very supportive and helped me to improve within the field of science, both in the lab and outside, I really appreciated it. /Your first PhD student to supervise Martin Magali Merrien, my always so very caring and positive KID-sister, I hope your ears are intact after listening to all my French and arguing about why you don’t get Swish. It has been the utmost pleasure to have “Miss Soleil” or “Liten, men Tuff” (kört barn har många namn) as my colleague and friend during the time here.

Sougat Misra, my (coffee) brother from another mother, for always being there for me when I’ve needed, both at work and outside. Your dedication to science and the patients benefiting from your strive is an inspiration. You’re the closest I’ve come to a big brother and you’ve meant a lot to me. /Rishi’s favorite “play uncle”

I would also like to thank Gustav Arvidsson, it has been really fun to collaborate and discuss science and non-science related topics. I really appreciate how you introduced me to R, no more Excel for me.

Thanks for making the time here since I started fun. I’ve enjoyed being hugged, a kickboxing-victim and getting my hair cut by office scissors: Anja ”Schmetterling I” Reithmeier, Lisa Arodin and Christina Patlaka. To Rim Jawad, thank you for all your kindness and concern for me (and all the laughs!).


42
Hortlund, Davit Bzhalava, Marita Wallenberg Lundgren, Filip Mundt, Adam Szulkin, Vicky Chatzakos and Dilruba Ahmed.

To Åsa-Lena Dackland and the FACS lab, thanks for being helpful when we receive primary samples.

Andrej Weintraub, thank you for the time I spent at the doctoral research council as a PhD representative under your supervision. The experience was very educational.

I also would like to thank the head of the division of Pathology Göran Andersson and all the members working here for making it an enjoyable workplace.

Finally, thanks to my family for all the support and help during the years. I wouldn’t have made it without you. To my mom, Karin, and dad, Göran, thank you for the constant support and encouragement throughout the years. What you have taught me carries through in everything I do. I am grateful for having you.

Thanks to my little brother (and neighbor), Erik, for all the good times we’ve shared and for being one of the two good guys in our building.

Slutligen, tack för all din omtanke om mig Farmor, den har betytt mycket för mig.
11 REFERENCES


86. Lecluse, Y., et al., t(11;14)-positive clones can persist over a long period of time in the peripheral blood of healthy individuals. Leukemia, 2009. 23(6): p. 1190-3.


