PAX5 CONTROLS B LINEAGE SPECIFIC GENE EXPRESSION PROGRAM THROUGH ASSOCIATION WITH THE NUCLEAR MATRIX

by

Sang Yong Hong

ZHIXIN ZHANG, MENTOR
PETER D. BURROWS
CHRISTOPER A. KLUG
CHANDER RAMAN
DAVID M. BEDWELL

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

BIRMINGHAM, ALABAMA

2011
PAX5 CONTROLS B LINEAGE SPECIFIC GENE EXPRESSION PROGRAM THROUGH ASSOCIATION WITH THE NUCLEAR MATRIX

Sang Yong Hong

DEPARTMENT OF MICROBIOLOGY

ABSTRACT

Pax5 is an essential regulator for B lineage cell development and controls hundreds of positive and negative target genes. The nuclear matrix (NM) has long been proposed to provide a dynamic structural support for biological reactions inside of the nucleus, including DNA replication, RNA transcription, and RNA splicing. Although Pax5 binding sites have been identified in numerous Pax5 target genes, it is not clear how Pax5 controls so many target genes. In this dissertation, I demonstrate that Pax5 association with the NM is essential for regulation of B lineage specific gene expression program.

In part I, I found that the majority of the endogenous Pax5 protein in human and murine B lineage cells is associated with the NM, where it is distributed closely with the NM-bound RNA polymerase II and TATA box binding protein (TBP) at discrete foci. Lysine 67, 87, and 89 to Alanine mutation within Pax5 diminished its NM association and compromised Pax5-mediated global gene expression. Chromatin immunoprecipitation and NM precipitation assay results further showed that association with the NM is required for Pax5 to recruit the Cd19 promoter to the NM-bound RNA polymerase complex. Based on these results, we propose a model that Pax5 activates B lineage specific gene expression through recruiting its target genes to the NM-bound transcription centers.

In part II, I performed global proteomic analyses to compare NM-associated
proteins in wild type and $Pax5^{-/}$ pro-B cells. Interestingly, our results showed that wild type pro-B cells have established a B lineage specific NM infrastructure with the enrichment of important transcription factors for lymphocyte or B lineage development. By contrast, different groups of non-B lineage transcription factors are enriched in the NM of $Pax5^{-/}$ pro-B cells. Moreover, we characterized that Bcl11a-XL is a specific Pax5 target gene. Forced expression of Bcl11a-XL induces a subset of Pax5 target genes. These results indicate that Pax5 can reorganize the NM to fully control B lineage specific gene expression program.

Taken together, the results provide the first evidence that Pax5 controls B lineage specific gene expression through association with the NM and reorganization of the NM infrastructure.

Keywords: Pax5, Nuclear Matrix, Bcl11a-XL, B cell development
DEDICATION

To my wife, Hyang Mi Lee, my family, Sung Do Hong, Kyu In La, Hee Sun Hong, Hyun Sun Hong, and Jun Yong Hong, for their support and encouragements during all year of my education.
ACKNOWLEDGMENTS

This work was carried out in the Department of Microbiology, University of Alabama at Birmingham, during the years 2006-2011.

I owe my sincere gratitude to my mentor Dr. Zhixin (Jason) Zhang, for his great training in my research. His discussion and encouragement always inspire me to accomplish all my research.

I wish to thank Dr. Peter Burrows, who is my committee chair and gives me so much help during my PhD study.

I also owe my gratitude to my committee members Drs. Christopher Klug, David Bedwell, and Chander Raman for invaluable suggestion and support.

It is a great pleasure to work with my fellows in the lab, Wanqin Xie, Lin Huang, Erin, Jirong, Dallas Johns, Miles Lange, and Jing Liu. They are all good friends inside and outside the lab.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT .................................................................................................................. iii</td>
</tr>
<tr>
<td>DEDICATION ................................................................................................................ v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS ..................................................................................................... vi</td>
</tr>
<tr>
<td>LIST OF FIGURES ......................................................................................................... x</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS ............................................................................................... xi</td>
</tr>
</tbody>
</table>

## INTRODUCTION

1. B Cell Development in the Bone Marrow ..................................................................... 1

   *From hematopoietic stem cells to immature B cells in the bone marrow* ............... 1

   *Hematopoietic stem cells to multipotent progenitors* ........................................... 1

   *Multipotent progenitors to lymphoid-primed multipotent progenitors and Common lymphoid progenitors* ................................................................. 2

   *Common lymphoid progenitors (CLPs) to B cell precursors* .............................. 3

   *Signal molecules for early B cell development* ...................................................... 4

2. Transcription Factor Networks in Specification and Commitment ......................... 5

   *PU.1 (Sfpi1)* .............................................................................................................. 6

   *Ikarose* ..................................................................................................................... 7

   *c/EBPα and c/EBPβ* ............................................................................................... 8
RNA and heterogeneous nuclear ribonucleoproteins (hnRNP)s ..................34

The nuclear architecture ..........................................................................................35

The functional domains of the nucleus matrix .........................................................36

DNA replication foci in the nuclear matrix ...............................................................36

RNA transcription foci in the nuclear matrix ...........................................................38

Specked domains of RNA splicing factors in the nuclear matrix ......................39

A subgroup of nuclear matrix associated proteins ...............................................40

ASSOCIATION WITH THE NUCLEAR MATRIX IS ESSENTIAL FOR PAX5 TO
CONTROL B LINEAGE SPECIFIC GENE EXPRESSION PROGRAM ..................43

PAX5 REORGANIZES THE NUCLEAR MATRIX TO FACILITATE B LINEAGE
SPECIFIC GENE EXPRESSION PROGRAM ..........................................................80

CONCLUSION ........................................................................................................108

GENERAL REFERENCES .........................................................................................114
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ASSOCIATION WITH THE NUCLEAR MATRIX IS ESSENTIAL FOR PAX5 TO CONTROL B LINEAGE SPECIFIC GENE EXPRESSION PROGRAM</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Pax5 is associated with the NM in B lineage cells ........................................67</td>
</tr>
<tr>
<td>2</td>
<td>The N terminus of Pax5 is required for the association with the NM ..............69</td>
</tr>
<tr>
<td>3</td>
<td>The K 67 and K87/K89 within Pax5 PRD domain are required for the association with the NM ..................................................................................................................................................71</td>
</tr>
<tr>
<td>4</td>
<td>Lost NM association compromises Pax5-mediated gene expression .................73</td>
</tr>
<tr>
<td>5</td>
<td>Pax5 is essential to recruits CD19 locus to the NM-bound RNA pol II complex .76</td>
</tr>
<tr>
<td>6</td>
<td>Pax5 recruits target gene promoter to NM-bound transcription centers ..........78</td>
</tr>
</tbody>
</table>

| **PAX5 REORGANIZES THE NUCLEAR MATRIX TO FACILITATE B LINEAGE SPECIFIC GENE EXPRESSION PROGRAM** | |
| 1 | Global proteomic analyses of the NM-associated protein in wild type (2A) and $Pax5^{+/}$ (G5) pro-B cells.........................................................................................................................98 |
| 2 | Distribution of lineage specific NM-associated proteins in 2A and G5 cells ......100 |
| 3 | Pax5 induces Bcl11a-XL expression ......................................................................102 |
| 4 | Identification and characterization of Pax5 binding sites within the $Bcl11a$ gene promoter .................................................................................................................................104 |
| 5 | Pax5 and Bcl11a-XL control a subset of shared target gene expression ..........106 |
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AID</td>
<td>Activation Induced Deaminase</td>
</tr>
<tr>
<td>ALP</td>
<td>All Lymphoid Progenitor</td>
</tr>
<tr>
<td>BCR</td>
<td>B cell Receptor</td>
</tr>
<tr>
<td>BCL2</td>
<td>B cell Lymphoma 2</td>
</tr>
<tr>
<td>BCL11a</td>
<td>B cell Lymphoma11a</td>
</tr>
<tr>
<td>BLNK</td>
<td>B cell Linker</td>
</tr>
<tr>
<td>BLP</td>
<td>B cell-biased Progenitor</td>
</tr>
<tr>
<td>BM</td>
<td>Bone Marrow</td>
</tr>
<tr>
<td>Blimp-1</td>
<td>B Lymphocyte-induced Maturation Protein-1</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB binding Protein</td>
</tr>
<tr>
<td>CCL3</td>
<td>Chemokine (C-C motif) Ligand 3</td>
</tr>
<tr>
<td>CCL9</td>
<td>Chemokine (C-C motif) Ligand 9</td>
</tr>
<tr>
<td>CCR2</td>
<td>Chemokine (C-C motif) Receptor 2</td>
</tr>
<tr>
<td>CCR5</td>
<td>Chemokine (C-C motif) Receptor 5</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CLP</td>
<td>Common Lymphoid Progenitor</td>
</tr>
<tr>
<td>CMP</td>
<td>Common Myeloid Progenitor</td>
</tr>
<tr>
<td>CTCF</td>
<td>CCCTC-binding factor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>CTIP1</td>
<td>COUP-interacting Protein 1</td>
</tr>
<tr>
<td>CTIP2</td>
<td>COUP-interacting Protein 2</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic Cell</td>
</tr>
<tr>
<td>EBF</td>
<td>Early B cell Factor</td>
</tr>
<tr>
<td>ELP</td>
<td>Earliest Lymphocyte Progenitor</td>
</tr>
<tr>
<td>ETP</td>
<td>Early T-lineage Progenitor</td>
</tr>
<tr>
<td>ETS</td>
<td>E26 avian leukemia oncogene 1,5’ domain</td>
</tr>
<tr>
<td>EVI9</td>
<td>Ecotropic Viral Integration site 9</td>
</tr>
<tr>
<td>FLT3</td>
<td>Fms-related tyrosine kinase 3</td>
</tr>
<tr>
<td>FO</td>
<td>Follicular</td>
</tr>
<tr>
<td>FOXO1</td>
<td>Forkhead box protein O1</td>
</tr>
<tr>
<td>Gfi1</td>
<td>Growth factor independent 1</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescence Protein</td>
</tr>
<tr>
<td>HAS</td>
<td>Heat Stable Antigen</td>
</tr>
<tr>
<td>hnRNP U</td>
<td>Heterogeneous nuclear Ribonucleoprotein U</td>
</tr>
<tr>
<td>HSC</td>
<td>Hematopoietic Stem Cell</td>
</tr>
<tr>
<td>HTH</td>
<td>Helix-turn-Helix</td>
</tr>
<tr>
<td>Igα</td>
<td>Immunoglobulin α</td>
</tr>
<tr>
<td>Igβ</td>
<td>Immunoglobulin β</td>
</tr>
<tr>
<td>IgH</td>
<td>Immunoglobulin heavy chain</td>
</tr>
<tr>
<td>IgL</td>
<td>Immunoglobulin light chain</td>
</tr>
<tr>
<td>IL-7</td>
<td>Interleukin 7</td>
</tr>
</tbody>
</table>
IL-7Rα  Interleukin 7 receptor α chain
LFA-1  Lymphocyte Function-associated Antigen-1
LMPP  Lymphoid-primed Multipotent Progenitor
MAR  Matrix Attachment Region
M-CSFR  Macrophage Colony-stimulating Factor Receptor
MHC  Major Histocompatibility Complex
MPP  Multipotent Progenitor
MZ  Marginal Zone
NM  Nuclear Matrix
NM-IF  Nuclear Matrix Intermediate Filament
NuMA  Nuclear Mitotic Apparatus protein
NuRD  Nucleosome Remodeling and Deacetylase
PCAF  p300/CBP-associated factor
RAG  Recombination Activating Gene
RNP  Ribonucleoprotein
RSS  Recognition Signal Sequence
SAR  Scaffold Attachment Region
INTRODUCTION

B cell Development in the Bone Marrow

From hematopoietic stem cells to immature B cells in the bone marrow.

Hematopoietic stem cells (HSCs) to multipotent progenitors (MPPs)

B cells, like all specialized mature cell types of the blood, are generated from the multipotent and self-renewing hematopoietic stem cells (HSCs) through successive rounds of lineage fate restriction. HSCs reside in the fetal liver before birth or in the bone marrow after birth. It has been reported that the activation of the tyrosine kinase receptor c-Kit by its ligand stem cell factor (SCF) promotes the survival of long-term reconstituting HSCs in vitro (Domen and Weissman, 2000).

As an initial step for lymphocyte development, HSCs differentiate into multipotent progenitors (MPPs) that lose self renewal capacity and acquire Flt3 tyrosine kinase expression (Adolfsson et al., 2001; Adolfsson et al., 2005). Flt3 has been shown to be required for efficient hematopoietic cell differentiation because Flt3 ligand (FL) deficiency leads to a severe reduction of CLP, pro-B, and pre-B cells in the bone marrow (McKenna et al., 2000; Sitnicka et al., 2002). Even though MPPs that express Flt3 sustain the lymphoid restricted reconstitution ability but lose the myeloid potential, the MPP population still maintains multi-lineage differentiation potential (Adolfsson et al., 2001;
Adolfsson et al., 2005; Mansson et al., 2007).

Three distinct subsets of the MPP were characterized based on the expression of Flt3 and vascular cell adhesion molecule (VCAM). These subsets showed different ability for the erythrocyte, macrophage, and granulocyte differentiation but had the same potential to give rise to lymphocyte in vivo (Lai and Kondo, 2006). The MPP was also divided into two subsets by utilizing PU.1 and GATA-1 transcription factor reporters. The GATA-1⁺ MPP had the potential to differentiate to myeloerythroid progenitors without lymphocytes differentiation whereas the PU.1⁺ MPP showed the granulocyte/monocyte/lymphoid potential without megakaryocyte/erythroid differentiation (Arinobu et al., 2007).

**Multipotent progenitors (MPPs) to lymphoid-primed multipotent progenitors (LMPPs) and common lymphoid progenitors (CLPs)**

Common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs) are known as two lineage-restricted populations that are derived from a common pool of multipotent progenitors (MPPs) and are important for hematopoietic cell differentiation (Akashi et al., 2000; Adolfsson et al., 2001; Kondo et al., 1997).

More detailed analyses of each stage of hematopoietic cell differentiation revealed that MPPs produce lymphoid-primed multipotent progenitors (LMPPs) (Adolfsson et al., 2005; Arinobu et al., 2007; Lai and Kondo, 2006; Yoshida et al., 2006; Pronk et al., 2007). LMPPs are subsequently able to differentiate to the earliest lymphocyte progenitors (ELPs) where RAG1 and RAG2 are expressed and D_{H}-J_{H} rearrangements occur at the immunoglobulin heavy-chain (IgH) locus (Igarashi et al., 2002). Recent studies suggested
that subset of ELP cells that express CCR9 or CD62L differentiates into the newly identified early T-lineage progenitor (ETP) in the thymus although the precise origin of ETPs is still controversial (Allman et al., 2003). Therefore, T cell can be committed at a stage before the CLP stage (Lai and Kondo, 2008; Rumfelt et al., 2006). In the bone marrow, ELPs generate CLPs that develop into four different cell types, i.e. B lymphocytes, T lymphocytes, natural killer (NK) cells, and dendritic cells (DC). However, CLPs may prefer to be progenitors of B lymphocytes and NK cells in vivo (Allman et al., 2003; Harman et al., 2006). CLPs lack all myeloid potential when tested in vivo but not in vitro (Rumfelt et al., 2006; Balciunaite et al., 2005; Kondo et al., 1997). Recently, more refined characterization of CLPs identified the transmembrane protein Ly6D as a marker of specified B cell progenitors. CLPs can be subdivided into Ly6D− all lymphoid progenitors (ALPs) or Ly6D+ B cell-biased progenitors (BLPs) because Ly6D is expressed with a bimodal pattern (Inlay et al., 2009).

**Common lymphoid progenitors (CLPs) to B cell precursors**

During B cell development in the bone marrow, CLPs differentiate into the first clearly identifiable B cell-specific progenitor, differently named CLP-2 cell, pre-pro B cell, or fraction A that expresses the B cell-associated marker B220 and activates many B cell-lineage-associated genes (Gounari et al., 2002; Martin et al., 2003). Induction of CD19 expression and $D_H$-$J_H$ rearrangement lead the B cell pathway into the early Pro-B cell stage (Li et al., 1996; Tudor et al., 2000). B cell lineage commitment can be identified by the expression of CD19. Productive $V_{H}$-$DJ_{H}$ recombination occurs at the late pro-B cell stage and IgM protein is expressed on the cell surface as part of the pre-B cell
receptor (pre-BCR). Signaling via the pre-BCR plays essential roles for allelic exclusion at the *IgH* locus, proliferative cell expansion, and transition from the pro-B to the pre-B cell stage, indicating that the pre-BCR acts as a checkpoint for the next stage of B cell development. The pre-B cells undergo immunoglobulin light-chain recombination (Meffre et al., 2000). A successfully rearranged light-chain gene is expressed on the cell surface together with the functional heavy chain and the signal transducing molecules, Igα and Igβ to form the B cell receptor (BCR). These immature IgM+ B cells emigrate from the bone marrow to the peripheral lymphoid organs for further development into mature B cells (Meffre et al., 2000).

*Signal molecules for early B cell development*

In mice, IL-7 signaling has been demonstrated to be required for guiding B cell development in the bone marrow. For example, *κ−/−* mice that lack a functional receptor of IL-7 generate normal numbers of CLP but have severe defects even in the differentiation into the earliest pro-B cell stage in the adult bone marrow. *IL-7Rα−/−* and *IL-7−/−* mice that lack a functional IL-7 receptor and IL-7 receptor ligand, respectively have similar defects in B cell development in the bone marrow (Miller et al., 2002; Carvalho et al., 2001). In double-mutant *flt3−/−IL-7Rα−/−* mice, there are no B-lymphocytes in the bone marrow (Vosshenrich et al., 2003). The block of early B cell development in these mice confirmed that the lymphoid cytokine IL-7 plays an important role for both pro-B cell survival and differentiation from the CLPs to the B cells (Miller et al., 2002). A recent study using conditional inactivation of Stat5 revealed its role in pro-B cells. Stat5 has been known to be a regulator of *IL-7Rα* signaling pathway. Pro-B cell development in *Stat5−/−* mice can
be rescued by transgenic expression of the prosurvival protein Bcl-2. The expression of EBF1 and Pax5 as well as V_H gene rearrangement are normal whereas Ig_K recombination is more abundant in these Stat5^- pro-B cells. Hence, Stat5 signaling controls cell survival and suppresses the premature Ig_K rearrangement in pro-B cells (Malin et al., 2010). These results reveal that the c-Kit, Flt3, IL-7R, and Stat5 signaling systems together account for the generation of all B-lymphocytes in the bone marrow of adult mice. However, the equivalent growth factor and signaling pathway for human early B lineage cells have not yet been identified.

Transcription Factor Networks in Specification and Commitment

B lineage specification and commitment require the coordinated processes of signaling cascades and transcriptional networks to initiate B cell-specific gene expression and repress other lineage determinants (Busslinger, 2004). Genetic ablation studies demonstrated that several transcription factors, including PU.1, Ikaros, Bcl11a, E2A, EBF1, and Pax5 are required for the direct progression of lymphoid progenitors into the B lineage and proper B cell development (Nutt and Kee, 2007). Previously, these factors were considered to act as a part of a linear transcriptional hierarchy. However, it has now become clear that the transcription factors function in a complex network, including cross-regulation, auto regulation, and positive and negative feedback loops (Nutt and Kee, 2007).
The *Sfpi1* gene encodes PU.1, a hematopoietic-specific member of the Ets family of transcription factors. *Sfpi1*−/− mice that die during late embryogenesis or shortly after birth have severe defects in myeloid and lymphoid differentiation (McKercher et al., 1996; Scott et al., 1994). Previous studies implicated that PU.1 regulated the choice between myeloid and B lymphocyte lineages (DeKoter and Singh, 2000). A low concentration of PU.1 protein stimulates B cell fate, whereas a higher concentration promotes myeloid differentiation and suppresses B cell development. Consistent with this result, PU.1 is differently expressed in B cells and macrophages (Nutt et al., 2005).

However, this model has been challenged as the PU.1 expression level is similar in HSCs, CLPs, and CMPs in the bone marrow of knockin mice carrying a GFP- *Sfpi1* reporter gene under translational IRES control. Difference in PU.1 expression seems to occurred after specification to the macrophage or B cell lineage (Back et al., 2005; Dakic et al., 2007). Conditional inactivation of PU.1 in adult bone marrow results in severely impaired hematopoiesis and the absence of an identifiable CLP population (Dakic et al., 2005). Interestingly, conditional deletion of PU.1 in committed B cells by *CD19*-cre shows a relatively normal B cell differentiation and function, suggesting that PU.1 is required for specification of lymphoid progenitors but is not necessary for further B lymphocyte development (Polli et al., 2005; Ye et al., 2005).

A recent work has shown that Growth factor independent 1 (Gfi1) protein displaces PU.1 from positive autoregulatory elements of *Sfpi1* and repress its expression (Spooner et al., 2009). The *Gfi1*−/− MPPs exhibit elevated levels of PU.1 and defects in B cell development. However, this reduced B cell potential can be reversed by ectopic
expression of Sfpi1-specific shRNAs. In addition, Ikaros upregulates Gfi1 and suppresses PU.1 expression in MPPs.

Ikaros

The Ikzf1 gene encodes the Ikaros family of transcription factors that are generated by alternative splicing (Georgopoulos et al., 1992; Hahm et al., 1994). The Ikaros protein that contains various Krüppel-like zinc fingers can function as transcriptional activator or repressors (Ng et al., 2007). Ikaros has been shown to be associated with two chromatin remodeling complexes, the nucleosome remodeling and deacetylase (NuRD) and the SWI-SNF complex (Kim et al., 1999; O'Neill et al., 2000). Ikaros is required for priming lymphoid lineage-specific genes and repressing self-renewal in HSCs (Ng et al., 2009). For example, Ikaros regulates Igll1 (λ5) gene expression by competing with EBF1 for DNA binding in a stage-specific manner. Expression of the pre-BCR component λ5 was upregulated by EBF1 and suppressed by Ikaros at the pre B cell stage (Thompson et al., 2007). Loss of Ikaros leads to the arrest of lymphoid cell differentiation as early as at the LMPP stage and its development is exclusively diverted into erythroid and myeloid lineages (Georgopoulos et al., 1994; Wang et al., 1996). Mutation of Ikaros also results in a defect in CLP differentiation, showing preferential commitment to the NK lineage in the expense of the B and T lineages (Papathanasiou et al., 2009). The Ikaros-deficient hematopoietic progenitors can generate CD19+ pro-B cells after reconstitution with EBF1. However, these pro-B cells are not committed to the B cell fate despite the normal expression of EBF1 and Pax5, suggesting that Ikaros promotes B cell identity (Reynaud et al., 2008). On the other hand,
a hypomorphic allele of Ikaros showed the defects in the transition from the pro-B cell to the pre-B cell and insufficiency in formation of IL-7-dependent pro-B cell colonies in vitro (Kirstetter et al., 2002).

\textit{c/EBP}α and \textit{c/EBP}β

The transcription factor C/EBPα is highly expressed in myeloid progenitors and their differentiated progeny (Akashi et al., 2000). The C/EBPα and C/EBPβ transcription factors are capable of reprogramming differentiated B cells into macrophages (Xie et al., 2004). Enforced expression of C/EBPα and C/EBPβ upregulates myeloid markers by inducing the transcription of myeloid-associated genes such as PU.1 and simultaneously downregulates B cell markers by inhibiting the expression of B cell specific genes such as E2A, EBF1, and Pax5,(Bussmann et al., 2009; Xie et al., 2004). However, the cellular context and culture conditions seem to affect this reprogramming experiment, as enforced expression of C/EBPα is unable to transdifferentiate committed pro-B cells into macrophages under culture conditions for lymphoid cell development (Heavey et al., 2003).

\textit{Bcl11a}

\textit{Identification of Bcl11a}

\textit{Bcl11a} was initially identified as an ecotropic viral integration site 9 (Evi9), a common site of retroviral integration in BXH murine myeloid leukemia (Nakamura et al., 2000). At the same time, \textit{Bcl11a} was also characterized as a chicken ovalbumin upstream promoter transcription factor (COUP-TF)-interacting protein 1, \textit{CTIP1}(Avram et al.,
2000). Separately, Bcl11a was identified as a gene involved in 4 cases of B-cell malignancy with t(2;14)(p13;q32.3) (Satterwhite et al., 2001). The same region (2p13) has been reported to be commonly overexpressed in B cell lymphoma and Hodgkin disease, suggesting that Bcl11a seems to be a proto-oncogene in human lymphocytes (Satterwhite et al., 2001).

The structure and expression of Bcl11a

The Bcl11a gene encodes a member of the Krüppel-like family of zinc finger proteins. Alternative splicing of the human Bcl11a transcript results in at least four isoforms translated into four protein isoforms, i.e. Bcl11a-XL (5.9 kb/125KD), Bcl11a-L (3.8 kb/100KD), Bcl11a-S (2.4 kb/35KD), and Bcl11a-XS (1.5 kb/25 kD). The Bcl11a-XL is the longest isoform that contains one unusual C2HC zinc finger, six Krüppel C2H2 zinc fingers, a proline rich domain, and an acid domain. The C2H2 zinc fingers are homologue of each other. The alternative isoforms contain the same regions encoded by exon 1 and 2 but are different at the carboxy-terminus with variable numbers of C2H2 zinc fingers.

Bcl11a is highly expressed in normal and malignant lymphoid tissues, including germinal center B cells, B-CLL, follicular lymphoma, and DLBC as well as fetal brain and plasmacytoid dendritic cells. However, the majority of multiple myeloma (MM), Hodgkin lymphoma (HL), myeloid and T-cell leukemia cells are negative for Bcl11a. Bcl11a-XL is localized exclusively within the paraspeckle subnuclear bodies. Bcl11a-L shows fewer nuclear dots and a more diffusion compared to Bcl11a-XL. Bcl11a-S is predominantly localized in the cytoplasm. However, it is translocated into the paraspeckle subnuclear bodies by coexpression with the XL or L form, suggesting that these isoforms
can interact with each other. The global gene expression profiles of different developmental stage of B cells, including hematopoietic stem cells, early B (E-B), pro-B, pre-B, and immature B cells in human adult bone marrow shows that Bcl11a transcript is turned on at the early B (E-B) cell stage before the expression of CD19 (Hystad et al., 2007).

The role of Bcl11a in B cell development

Bcl11a has been shown to be required in lymphoid development from the phenotypic analysis of Bcl11a−/− mice (Liu et al., 2003). Previous results showed that retroviral activation of Bcl11a led to the induction of myeloid leukemia (Nakamura et al., 2000). However, Bcl11a−/− mice had no defect in myeloid development. Mice in which Bcl11a has been knocked out had defects in early lymphoid development and died shortly after birth from unknown causes. There was no difference in the number of Max1+Gr1+ or Ter119+ cells between wild type and Bcl11a mutant embryos, suggesting that Bcl11a is not necessary for the development of the macrophage-granulocyte or erythroid lineages. However, there were few B220+IgM+, B220+IgM−, B220−CD19+, or B220+CD43+ in the Bcl11a−/− fetal liver cells, indicating that Bcl11a is required for B cell formation. Moreover, Bcl11a−/− fetal liver cells did not express essential genes for B cell development such as Ebf1, Pax5, Il7r, Cd19, Rag1, and Vpreb2. This mutant mice also did not undergo productive diversity (D)-joining (J) recombination at the Igh locus. The defects in B cells were intrinsic to the Bcl11a−/− hematopoietic progenitors as in wild type mice transplanted with Bcl11a−/− fetal liver cells, B cells were barely detected in the bone marrow, spleen, and peripheral blood. In the Bcl11a mutant thymuses, the number of TCRγδ T cells was
increased three to four fold whereas that of TCRαβ T cells was correspondingly reduced, indicating that Bcl11a is also required for fetal thymocyte development. However, the defects in thymic development were not as severe as B cell development in Bcl11a deficient mice. Notably, Bcl11a has been suggested to function as a non-autonomous T cell tumor suppressor gene from the experimental observation that Bcl11a mutant fetal liver cells induced thymic lymphomas of host origin. The highly increased transcription of Notch1 in the thymic lymphomas from mice transplanted with Bcl11a mutant fetal liver cells suggested that Bcl11a can regulate Notch signaling. Notch1 transcript was greatly increased in leukemic CD4⁺CD8⁺ T cells. Consequently, Bcl11a is required for lymphoid development in mice, especially early B cell differentiation and normal T cell development. However, it remains to be determined that Bcl11a is essential for the development of multipotent progenitors which are specifically involved in the B and T cell commitment and differentiation or it is required for the B cell or T cell maintenance (Liu et al., 2003).

**E2A**

E2A is included in the E-protein family of basic helix-loop-helix (bHLH) proteins. The E2A (Tcfe2a) gene encodes E12 and E47, two basic spliced variants that are required for B cell development beyond the pre-pro B cell stage. The inactivation of E2A blocks B cell differentiation at the pro-B cell stage and impairs the expression of the crucial factors for B cell development such as EBF1 and Pax5 (Bain et al., 1994; Zhuang et al., 1994; Borghesi et al., 2005). Reconstitution of EBF1 and Pax5 in E2A⁻/⁻ progenitors is capable of rescuing B cell differentiation, suggesting that the essential role of E2A in initiation of
B cell specific program is through activation of EBF1 and Pax5 (Bain et al., 1994; Seet et al., 2004; Kwon et al., 2008). Further examination of mice deficient for either E12 or E47 revealed that the E47 is essential for B cell lineage specification although both E12 and E47 are expressed in LMPPs and CLPs (Beck et al., 2009). Moreover, two highly conserved E2A transcription activation domains are dispensable for the rescue of B lymphopoiesis in $E^{2A-}$ MPPs but instead N terminus of E2A is required for this process (Bhalla et al., 2008). A conditional E2A mutagenesis showed its additional roles for B cell development. E2A is required for the development of the bone marrow pro-B, pre-B, and immature B cells in addition to germinal center B cells whereas it is dispensable for production of mature B cells and plasma cells in peripheral lymphoid organs (Kwon et al., 2008).

**EBF1**

EBF1 (encoded by *Ebf1*) is an essential transcription factor for the specification and progression of the B cell program (Lukin et al., 2008; Hagman et al., 1995; Sigvardsson et al., 2002). It contains an N-terminal novel zinc-coordination motif for DNA binding, helix-loop-helix (HLH) domain for dimerization, and a C-terminal activation domain. Loss of EBF1 leads to the arrest of B cell differentiation at an early stage before the onset of $IgH(D_{H-J_H})$ recombination and these cells fail to express many genes involved in B cell development, including *Cd79a* (*mb-1, Igα*), *Cd79b* (*Igβ, B29*), *Igll1* (*λ5*), *VpreB1*, and *Pax5* (Lin and Grosschedl, 1995). Retroviral transduction of EBF1 into HSCs induces differentiation toward the B cell lineage, suggesting that EBF1 is sufficient to activate the B cell lineage-specific gene expression program (Zhang et al., 2008).
Moreover, it has been reported that enforced expression of EBF1 is sufficient to rescue B cell potential from the progenitor cells that lack E2A, IL7, IL7-Rα, Ikaros, or PU.1 protein (Seet et al., 2004; Dias et al., 2005; Kikuchi et al., 2005; Medina et al., 2004; Reynaud et al., 2008). In these cases, ectopic expression of Pax5 failed to restore B lymphopoiesis, implying that EBF1 performs more functions than Pax5 activation for B cell differentiation.

*Ebf1* transcription is mediated by two differentially regulated promoters in B cells, which results in the expression of protein isoforms that differ by 14 amino acids in the N-terminus (Roessler et al., 2007). The distal *Ebf1* α promoter is partly controlled by IL7 signaling, E2A, and EBF1 (Roessler et al., 2007; Smith et al., 2002; Kikuchi et al., 2008). The presence of EBF1 binding sites in the distal *Ebf1* α promoter region implies an autoregulatory function for EBF1. Stat5, a downstream mediator of IL7 signaling can activate this endogenous *Ebf1* promoter in transfection assays. However, *Ebf1* transcription is also normal in the *IL7Rα*−/− and *Stat5α*−/− pro-B cells, suggesting that IL7 signaling may be required but is not essential for EBF1 expression (Roessler et al., 2007; Malin et al., 2010). The proximal *Ebf1* β promoter is upregulated by Pax5, Ets1, and PU.1. The promotion of *Ebf1* expression by Pax5 was suggested by the observation that enforced expression of Pax5 induced EBF1 in T cell progenitors (Fuxa et al., 2004). EBF1 upregulates *Pax5* expression and upregulated Pax5 induces EBF1 expression. Consequently, the activity of both *Ebf1* promoters is regulated by multiple feedback loops (Roessler et al., 2007; Smith et al., 2002; Kikuchi et al., 2008).

EBF1 expression is initiated at the CLP stage and primes B cell specific genes to establish the transcription network for B cell commitment (Zandi et al., 2008). In the
absence of EBF1, CLPs retain considerable myeloid potential. Conversely, ectopic expression of EBF1 in MPPs or Pax5−/− hematopoietic progenitors is sufficient to block myeloid or T cell potential. Thus, the developmental plasticity previously characterized in Pax5−/− cells may be partly due to the reduced level of EBF1 (Pongubala et al., 2008). The enforced expression of EBF1 in IL-7R−/− mice downregulates the E2A-inhibitory myeloid-promoting factors Id2 and Id3 whereas overexpression of either Id2 or Id3 leads to developmental arrest at pre-pro B cell stage, suggesting that EBF1 specifies B cell lineage through the inhibition of Id factors (Thal et al., 2009).

The genome wide ChIP sequencing analyses of EBF1 combined with gain- or loss-of-function transcriptome analyses revealed that 565 genes were occupied and transcriptionally regulated by EBF1 in early stage B cells. This large sets of the identified genes are involved in (pre)-B cell receptor signaling, Akt signaling, cell adhesion, and migration (Treiber et al., 2010). Of these targets, the transcription factor FoxO1 directly activates transcription of Rag1 and Rag2 through early B cell development (Amin and Schlissel, 2008). The conditional inactivation of FoxO1 demonstrates that Foxo1-dependent gene expression plays essential roles for different stages of B cell development by inducing IL7Rα, Rag1, Rag2, L-selectin, and AID (Dengler et al., 2008). Furthermore, ChIP-seq analyses of E2A identified cis-regulatory sequence within 100 bp upstream and downstream of E2A occupancy at pre-pro or pro B cell stage. This result showed that E2A, EBF1, and FoxO1 coordinated the DNA binding affinity to a subset of regulatory elements, suggesting that the global network of these factors plays an important role for B cell fate (Lin et al., 2010).
Pax5

**The Pax gene family and characterization of Pax5**

The Pax gene family is defined by the presence of a highly conserved 128-amino acid N-terminal paired domain (PD) that functions as a bi-partite DNA-binding region (Czerny et al., 1993; Mansouri et al., 1996; Chi and Epstein, 2002). In mammals, the identified nine members of the Pax family are divided into distinct classes or subgroup by the presence or absence of a paired-type homeodomain (HD), an octapeptide domain (OP), transactivation, and transrepression domain. For example, the subgroup of Pax2, 5, and 8 genes is characterized by including an octapeptide domain, partial homeodomain, and a C terminal domain that has both transactivation and inhibitory properties identified by deletion analyses (Dorfler and Busslinger, 1996).

Pax genes show dynamic expression patterns during development in different tissues (Mansouri et al., 1996). For example, Pax1 is expressed in the sclerotome and mutation of this gene impairs the murine skeletal development (Balling et al., 1988). Pax2 plays an important role for eye and kidney formation (Dressler et al., 1990; Sanyanusin et al., 1995). Expression of Pax3 is mainly detected in the neural tube, neural crest, and the limb muscle (Epstein et al., 1991; Tassabehji et al., 1993). Pax5 is expressed in B lymphocytes, the developing CNS, and adult testis (Adams et al., 1992). Pax6 is involved in the eye development (Hill et al., 1991; Jordan et al., 1992). Of the nine mammalian Pax transcription factors, only Pax5 is expressed within the hematopoietic system.

Pax5 was first identified as a nuclear DNA binding protein that recognized the promoter of an Igκ germline transcript, termed kappa locus protein (KLP) or recognized
the α switch region of the *Igh* locus, termed Sα DNA-binding protein (Sα-BP) (Weaver and Baltimore, 1987; Waters et al., 1989). Independently, Pax5 was identified as a B-cell-specific activator protein (BSAP) that interacted with the same DNA sequence recognized by the sea urchin transcription factor TSAP (Barberis et al., 1990). Protein purification and cDNA cloning showed that KLP, Sα-BP, and BSAP are identical and encoded by the *Pax5* gene.

*B lineage commitment and maintenance by Pax5*

The phenotypic analysis of *Pax5* deficient mice revealed that Pax5 plays an essential role in B cell differentiation. In this mice, B cell development was completely arrested at an early precursor stage (Urbanek et al., 1994). More detailed analysis of B cell phenotype of *Pax5*−/− mice showed that Pax5 was differently required for the fetal and adult B lymphopoiesis (Nutt et al., 1997). In the bone marrow of *Pax5*−/− mice, B cell development was arrested at the early pro-B, c-Kit+B220+ progenitor cell stage whereas in the fetal liver of *Pax5*−/− embryos, B lymphopoiesis was blocked before the appearance of B220+ progenitors. *Pax5* deficient pro-B cells can be cultured in the presence of interleukin-7 (IL)-7 cytokine and stromal ST2 cells (Nutt et al., 1997). Upon withdrawal of lymphoid cytokine IL-7 and stimulation of other-lineage appropriate cytokines, *Pax5* deficient pro-B cells are able to differentiate *in vitro* into functional macrophages, dendritic cells, granulocytes, osteoclasts, and natural killer cells, indicating that these cells are not committed to the B lineage and still retain the lymphomyeloid developmental potential. However, reconstitution of Pax5 protein in *Pax5*−/− pro-B cell by retroviral transduction rescues differentiation to the immature B cell and suppresses the
multi lineage potential. Thus, Pax5 plays an critical role in B lineage commitment that restrict alternative lineage choices (Nutt et al., 1999). After transplantation into recipient mice, $Pax5^{-/-}$ pro-B cells home to the bone marrow where they retain self-renewal (Rolink et al., 1999) and develop into several functional hematopoietic cell types \textit{in vivo} (Schaniel et al., 2002a). Although $Pax5^{-/-}$ pro-B cells possess self-renewal, long-term reconstitution, and hematopoietic multipotency like HSCs (Schaniel et al., 2002b), they are different from HSCs in that they more efficiently differentiate into lymphoid rather than myeloid (Rolink et al., 1999; Schaniel et al., 2002a). Interestingly, a previous study showed that E2A deficient lymphoid cell lines also had multilineage potential after long term culture \textit{in vitro} and such a finding correlated with the fact that these cells lack high expression of B cell specific markers as well as Pax5 expression (Ikawa et al., 2004).

It has been investigated whether Pax5 can directly force early uncommitted progenitors such as HSCs and MPPs to the B lineage cells. Ectopic expression of Pax5 in HSCs by retroviral transduction or pan-hematopoietic expression of Pax5 under the control of the \textit{Ikaros} locus is unable to divert HSCs and erythro-myeloid progenitors into the B cell pathway. By contrast, the premature expression of Pax5 is able to promote B cell development at the expense of T lymphopoiesis (Souabni et al., 2002; Cotta et al., 2003). Unlike the case of HSCs, the ectopic expression of Pax5 in $E2A^{-/-}$ progenitors can promote pro-B cell development in the fetal liver and the bone marrow by activating the B lineage specific gene expression program (Kwon et al., 2008). These studies strongly suggest that the function of Pax5 for B lineage commitment depends on the developmental stage within the hematopoietic system and the collaboration with other lymphoid factors.
Conditional Pax5 inactivation reveals that Pax5 is required for maintaining B cell identity and function throughout B lymphopoiesis (Horcher et al., 2001; Mikkola et al., 2002). For example, the floxed Pax5 allele is conditionally inactivated by B cell specific CD19 cre or interferon inducible Mx cre expression, which leads to severe defects in B cell development, including loss of mature B cells, downregulation of surface IgD expression, altered cell surface proteins, loss of B cell specific gene expression, inefficient lymphoblast formation, and reduced IgG secretion (Horcher et al., 2001). Moreover, cre-mediated Pax5 inactivation in committed Pro-B cells leads to the regaining of multilineage potential to differentiate into macrophages in vitro and reconstitute T cell development in vivo, demonstrating that Pax5 is required for both the initiation of B lineage specific gene expression program and the maintenance of early B cell development (Mikkola et al., 2002). Notably, even mature B cells dedifferentiate to the uncommitted progenitors and then give rise to functional T cells that contain rearranged Ig heavy and light chain genes after conditional inactivation of Pax5 (Cobaleda et al., 2007). Thus, these results confirm that continuous expression of Pax5 is essential for maintenance of B lineage commitment.

Pax5 expression is reduced during terminal plasma cell differentiation (Barberis et al., 1990). Many genes repressed by Pax5 are reactivated in plasma cells, including Cd28 and Ccr2 that are required for normal plasma cell functions (Delogu et al., 2006). These results suggested that loss of Pax5 during terminal differentiation turns on the transcription program for the plasma cell differentiation. Blimp1 has been known to be important for plasma cell differentiation (Shaffer et al., 2002). Blimp1 binds to the Pax5 promoter and represses its promoter activity (Lin et al., 2002). On the other hand, Blimp1
Prdm1 is expressed after conditional Pax5 inactivation. Hence, Prdm1 is transcribed and stabilizes the transcription program for plasma cell differentiation after its release from Pax5-mediated repression.

**The regulation of Pax5 target genes for B cell commitment and maintenance**

During B lymphocyte lineage commitment, Pax5 performs a dual role by activating B cell specific genes and repressing other lineage-associated genes (Nutt et al., 1998; Nutt et al., 1999). This Pax5-mediated gene repression was first noted from B cell-specific downregulation of M-CSFR (Nutt et al., 1999) and Notch1 (Souabni et al., 2002) that are responsive to myeloid cytokine M-CSF and T-cell-inducing Notch1 ligands, respectively. Their expression in Pax5−/− pro-B cells provides a molecular explanation for the role of Pax5 in the lineage plasticity. Recently, 110 Pax5-repressed genes were identified by cDNA microarray analysis, which were normally expressed in non-B cell lineages and were involved in receptor signaling, cell adhesion, migration, nuclear processes, and cellular metabolism (Delogu et al., 2006). Among these genes, the cell surface proteins downregulated by Pax5 include Flt3, Ly6a (Sca1), subunit of Ramp1, and Gp49b. Enforced expression of Flt3 in wild type hematopoietic cells significantly impairs B cell development by diverting progenitors into the dendritic cell (DC) pathway (Holmes et al., 2006). Moreover, Flt3 is characterized as a direct Pax5 target gene as Pax5 binds to its promoter regions and mediate its repression (Holmes et al., 2006; Tagoh et al., 2006). Intracellular signaling molecules such as Grap2 and NTAL are repressed by Pax5. Thus, these results suggested that Pax5 prevented the committed B lymphocytes from responding to the signal transduction pathway of early progenitors or other
hematopoietic lineage cells (Delogu et al., 2006).

The B lymphocytes adhere and move to specific niches within the bone marrow during development (Tokoyoda et al., 2004), indicating that lymphoid progenitors alter the expression of proteins involved in adhesion and migration on B cell commitment. Pax5 represses the expression of the CCL3, CCL9, CCR2, CCR5, LFA-1, CD47, and Tm4sf2. Ectopic CCL3 expression from the CD19 locus during B cell development leads to increased osteoclast formation and bone loss, suggesting that Pax5-mediated gene repression is essential for normal B cell development (Delogu et al., 2006).

Pax5-activated target genes encode essential components of the (pre) BCR signaling pathway, including the receptor signaling chain Igα (CD79, mb-1) (Fitzsimmons et al., 1996), the costimulatory receptor CD19 and CD21 (Kozmik et al., 1992; Horcher et al., 2001), the inhibitory coreceptor CD72 (Ying et al., 1998) and the central adaptor protein BLNK (SLP65) (Schebesta et al., 2002). CD19 and BLNK are considered critical Pax5 target genes as B cell development is completely arrested at the pro-B to pre-B cell transition in CD19−/−BLNK−/− mice (Hayashi et al., 2003).

Pax5 activates the expression of a number of transcription factors important for B cell differentiation, including SpiB, Aiolos, Id3, IRF4, and RIF8 and then initiates a cascade of transcriptional events. Furthermore, Pax5 induces genes involved in the regulatory network of B cells such as the transcription factor Lef1 (Nutt et al., 1998), Ebf1 (Roessler et al., 2007), and coactivator CIITA (Mikkola et al., 2002). Interestingly, a recent study showed that ectopic expression of EBF1 in Pax5−/− pro-B cells was able to block differentiation to other lineages such as T cells or myeloid cells (Pongubala et al., 2008). EBF1 is able to activate the Pax5 gene and conversely, Pax5 enhances Ebf1.
expression. Therefore, these results imply that the developmental plasticity in \(Pax5^{+}\) pro-B cells is partially due to the reduced level of EBF1 expression (Pongubala et al., 2008). Moreover, global EBF1 target gene profiles show that a third of previously characterized Pax5 target genes are also EBF1 target genes (Treiber et al., 2010), suggesting that the transcriptional network for B cell commitment and maintenance is not a simple linear cascade but involves in multiple feedback loops (Mandel and Grosschedl, 2010).

Cell-surface molecules such as Bst1, Cd44, Cd55, Cd97, Sdc4, and Tnfrsf19 are upregulated by Pax5. Intracellular molecules activated by Pax5 are Bcar3, Capn2, Eps8, Fhod3, Gsn, Myh10, Mylip, Nedd9, and Pard3. Both grouped genes have been reported to be involved in cell migration and adhesion in various different cell types. Consistent with this observation, an adhesion and migration assay verifies that Pax5 controls adhesion and migration properties of B lymphocytes (Schebesta et al., 2007).

Recently, streptavidin-mediated Chip-chip analyses using the pro-B cells that express biotinylated Pax5 \textit{in vivo} identified the direct Pax5 target genes (McManus et al., 2011). Pax5 directly controls the expression of its target genes by binding to the promoter or enhancer region and recruiting chromatin modifying factors (McManus et al., 2011).

\textit{Regulation of Pax5 transcription}

The potent activities of Pax5 for B cell identity and function suggest that its expression must be tightly regulated throughout lymphopoiesis. It has been reported that Pax5 is biallyleically transcribed at all stages of B cell development using \(Pax5^{\text{GFP/\text{hCd2}}}\) knockin mice that carrying a noninvasive insertion of a GFP or human \(Cd2\) reporter gene under the translational control of IRES elements in the 3’ untranslated region of Pax5.
(Fuxa and Busslinger, 2007). The expression of Pax5 is fully turned on only during the transition from uncommitted pre-pro-B cells to committed pro-B cells in adult B lymphopoiesis. Subsequently, Pax5 expression is remarkably stable throughout the B cell lineage from the pro-B cell stage until its downregulation in plasma cells (Fuxa and Busslinger, 2007). However, little is known about how Pax5 gene transcription is regulated. A recent study revealed that Pax5 expression in B lymphoid was regulated by its tissue specific enhancer in intron 5 in combination with its promoter (Decker et al., 2009). This potent enhancer is silenced in embryonic stem cells but is activated in multipotent hematopoietic and lymphoid progenitors. Its activity is regulated by multiple transcription factors, including PU.1, IRF4, IRF8, and NF-κB (Decker et al., 2009). By contrast, Pax5 promoter region is activated only at the onset of pro-B cell development. EBF1 is required for the formation of hypersensitivity sites at the Pax5 promoter region and directly binds to these sites, indicating that EBF1 plays an important role for inducing active chromatin configuration at the Pax5 promoter. It was also known that Pax5 transcription is mediated by two distinct promoters followed by the production of two distinct mRNAs through alternative splicing of 5’ exons. These data suggest that expression of Pax5 occurs in a stepwise fashion during early lymphopoiesis by remodeling the enhancer region in the MPP stage and subsequently activating the promoter at the pro B cell stage.

**Transcriptional activities of Pax5 and interacting proteins**

The transcriptional activity of Pax5 is differently regulated by its interaction with distinct partner proteins. For example, N-terminus of Pax5 recruits Ets to a low affinity
Ets-binding site within the mb-1 promoter region by directly interacting with the Ets domain (Fitzsimmons et al., 1996). This cooperative interaction alters the DNA sequence recognition of Ets (Garvie et al., 2001) and results in increased mb-1 transcription (Nutt et al., 1998). The partial homeodomain of Pax5 interacts with TATA-binding protein (TBP) and unphosphorylated retinoblastoma (RB), suggesting that Pax5 can directly bind basal transcriptional machinery and its activity is affected by multifunctional RB protein (Eberhard and Busslinger, 1999). Daxx protein also binds the partial homeodomain of Pax5 and can modulate its transcriptional activity as a coactivator or copressor. Daxx-mediated its coactivation involves in the recruitment of CREB binding protein (CBP), one of histone acetyltransferase (HAT) family (Emelyanov et al., 2002). In B cell lines, Pax5 can be associated with Ada2β known as transcriptional adaptor protein, together with either Gcn5, one of HAT family or Brg1, one catalytic component of chromosomal remodeling Swi/Snf complex. Coexpression of Gcn5 and Ada2β elevated the endogenous CD19 transcription (Barlev et al., 2003). By contrast, Grg4b, one member of Groucho copressor family binds to the octapeptide motif of Pax5 and efficiently represses Pax5-mediated transcriptional activation (Eberhard et al., 2000). Recently, Pax5 was shown to recruit multiple complexes of chromatin remodeling, histone modifying, and basal transcription factors, including Brg1, BAF170, BAF57, TAF6, TBP, PTIP, CBP, and NcoR1 to its target genes (McManus et al., 2011). In addition, p300, one of HAT family interacts with C-terminal region of Pax5 and subsequently acetylates Pax5, which plays an important role for the enhancement of its transcriptional activity (He et al., 2011).
The role of Pax5 for IgH V<sub>H</sub>-DJ<sub>H</sub> recombination

Functional Ig<sub>H</sub> gene is rearranged from discontinuous variable (V), diversity (D) and Joining (J) gene segments (Bassing et al., 2002). V(D)J recombination at the Ig<sub>H</sub> locus occurs sequentially with D<sub>H</sub>-J<sub>H</sub> rearrangement followed by V<sub>H</sub>-DJ<sub>H</sub> recombination, which is regulated at several steps by numerous transcription factors (Bassing et al., 2002).

Pax5 plays an important role in regulating the Ig<sub>H</sub> V<sub>H</sub>-DJ<sub>H</sub> recombination (Nutt et al., 1997). In Pax5<sup>−/−</sup> pro-B cells, D<sub>H</sub>-J<sub>H</sub> recombination occurs normally, but V<sub>H</sub>-DJ<sub>H</sub> rearrangements involving the V<sub>H</sub>J558 genes are reduced 100-fold compared to wild-type pro-B cells. The analyses of entire 2.5-mega base mouse Ig<sub>H</sub> V (V<sub>H</sub>) region sequences show that V<sub>H</sub>J558 genes are the largest family of V<sub>H</sub> gene families, span more than half of the V<sub>H</sub> locus, and are located at relatively distal end from domains consisting of the D<sub>H</sub>, J<sub>H</sub> and C<sub>H</sub> gene segments (Johnston et al., 2006). The V<sub>H</sub>-DJ<sub>H</sub> recombination occurs more efficiently in the V<sub>H</sub> genes that are more proximal to the DJC<sub>H</sub> in Pax5<sup>−/−</sup> pro-B cells (Hesslein et al., 2003). Analyses of histone acetylation and germ-line transcription in Pax5<sup>−/−</sup> pro-B showed that the distal V<sub>H</sub>J558 genes are as accessible as the proximal V<sub>H</sub>7183 genes, suggesting that distal V<sub>H</sub>-DJ<sub>H</sub> recombination requires Pax5-dependent regulatory mechanism in addition to chromatin accessibility (Hesslein et al., 2003).

Indeed, Pax5 has been shown to control the spatial organization of the Ig<sub>H</sub> locus within nucleus. Pax5 induces the contraction of the Ig<sub>H</sub> locus in wild-type pro-B cells (Fuxa et al., 2004; Kosak et al., 2002) by looping of individual Ig<sub>H</sub> subdomain (Roldan et al., 2005; Sayegh et al., 2005), which facilitates V<sub>H</sub>-DJ<sub>H</sub> recombination between distal V<sub>H</sub> and proximal D<sub>H</sub> gene segments. However, in Pax5<sup>−/−</sup> pro-B cells, Ig<sub>H</sub> locus is under the
extended state without contraction, which prevents distal $V_H$-$DJ_H$ recombination due to the physical separation between the distal $V_H$ genes and the $DJ_H$ domain (Fuxa et al., 2004). Pax5 has also been involved in the regulation of $IgH$ chromatin by removing repressive H3K9me2 modification that is inversely correlated with V(D)J recombination of $IgH$ gene segments (Johnson et al., 2004).

Moreover, Pax5 is able to recruit the RAG1-RAG2 protein complex to the coding regions of mouse and human $V_H$ genes that contain multiple Pax5 binding sites, which enhances RAG-mediated cleavage of VH recognition signal sequences *in vitro* (Zhang et al., 2006).

A recent study revealed that Pax5-activated intergenic repeat (PAIR) elements were interspersed in the distal $V_H$ gene clusters and contained conserved DNA binding sites for Pax5, E2A, CTCF, and Rad21. The PAIR elements were mapped as active chromatin regions and were reversely transcribed in a Pax5-dependent manner at the pro-B cell stage. However, PAIR elements did not bind Pax5 and were not reversely transcribed after pro-B cell stage, suggesting that PAIR elements were involved in the regulation of distal $V_H$-$DJ_H$ recombination at this stage (Ebert et al., 2011). Targeted disruption of intergenic control region 1 (IGCR1) further indicated that the CTCF binding region serves as an insulator to separate the $IgH$ locus into different regions (Guo et al., 2011).
The Nuclear Matrix

*History and definition of the nuclear matrix*

In the past, the cell nucleus was shown to contain only chromatin and nucleoli due to the limitations of analytic methods. The nuclear space was considered to be occupied by a translucent gel or fluid, termed nuclear sap, nucleoplasm or karyolymph. However, the application of electron microscopy to the cell nucleus provided a new view for the cell nucleus. For the first time, these electron micrographs showed that the cell nucleus contained heterogeneous structures in the inter-chromatin space between condensed chromatin.

*Fibrogranular ribonucleoprotein network of the intact nucleus*

The early electron micrographs showed the basic morphology of the nucleus. Dense chromatin (heterochromatin) was located along the nuclear periphery and at other interior areas. The interchromatin space was dispersed among dense chromatin and was considered to contain the diffused chromatin (euchromatin) and fibrogranules of 30-50 Å (Bernhard and GRANBOULAN, 1963; Smetana et al., 1971). However, it was generally difficult to distinguish this fibrogranules from the diffused chromatin fibers by early method of electron microscope. The development of the ethylenediaminetetraacetic acid (EDTA)-regressive staining technique was an important breakthrough to overcome this problem (Bernhard, 1969; Monneron and Bernhard, 1969). This EDTA-regressive staining selectively bleached the DNA-containing chromatin whereas maintained significant contrast of the nonchromatin (especially RNA-containing) structures. This
staining clearly revealed that the fibrogranules were distinct from chromatin and instead contained RNA. This ribonucleoprotein (RNP)-containing network that was made up of both granules and irregular fibers was observed throughout the nuclear interior except the regions of dense heterochromatin. Later, this RNP network was shown with high resolution by the electron spectroscopic image in unstained sections of fixed and unextracted cells (Hendzel et al., 1999). This fibrogranular ribonuclearprotein network was defined as the nuclear matrix. Since Don Fawcett defined the nuclear matrix as the non-chromatin structures of the nucleus readily observed by the electron microscope in unextracted cells in 1966 (FAWCETT, 1966), many other names or definitions have been used to designate this network structure.

_Ultrastructure of the nuclear matrix_

The nuclear matrix seems to be composed of at least two structurally distinct regions: the nuclear lamina, a protein shell built primarily by the lamin protein A, B, and C (Franke, 1987), and the internal nuclear matrix network. Both regions are linked together by highly structured fibers. The nuclear lamina on the shell of the nuclear matrix is connected to the extensive network of intermediate filaments that are components of structural proteins in cytoplasm. This nuclear matrix and intermediate filament are integrated into structural complex to support the intact cell, termed the nuclear matrix-intermediate filament scaffold (NM-IF). A multicellular NM-IF is considered to underlie tissue architecture and function as a linker between cell organization and tissue morphology.

Conventional electron micrographs showed that the cell nucleus in intact cells
formed a double membrane at the outer nuclear boundary (Nigg, 1989). The treatment of detergent such as Triton X-100 extracts the lipids in these membranes and a single proteinaceous layer, nuclear laminar remains forming the outer boundary of the nuclear matrix. The nuclear lamina also attaches to nuclear pores, forming the complex structure often referred to the pore-lamina complex (Gerace et al., 1984). The intermediate filament appears to be connected to the lamina at the nuclear pores by filamentous crossbridges whose components have not been identified (Carmo-Fonseca et al., 1988).

As shown by the EDTA-regressive staining, the nuclear matrix contains a large amount of RNA. When the nuclear matrix is isolated with RNase inhibitors, the nuclear matrix retains more than 70% of total nuclear RNA (Fey et al., 1986a; Fey and Penman, 1988; He et al., 1990). It is not clear whether this RNA is a component of the nuclear matrix structure. However, it has been reported that this RNA plays an important role for chromatin organization as either RNase A digestion or transcription inhibition dramatically impairs the chromatic structure (Nickerson et al., 1989).

*Isolation method of the nuclear matrix*

The nuclear matrix has been characterized to be the structural network of fibrogranular ribonucleoprotein of the intact nucleus under the electron microscopy. Nuclear matrix structures consider to be ubiquitous in the eukaryote and have been identified in Saccharomyces (Amati and Gasser, 1988), plants (Hall, Jr. et al., 1991), Drosophilar (Fisher et al., 1982), and mammals (Berezney and Coffey, 1974). After the nuclear matrix was observed in the eukaryotic nucleus, a lot of efforts have been taken to identify the common components consisting of the nuclear matrix.
As an initial step to characterize the components of nuclear matrix, various biochemical protocols for the nuclear matrix isolation have been developed by different groups. Although the molecular basis of the nuclear matrix is still not known, many candidates have been proposed as the components of the nuclear matrix. Components of nuclear matrix are expected to have certain properties: (1) the continuous but dynamic structures, (2) interaction with other nuclear components, (3) evolutionarily conserved components and (4) abundant in the nuclear matrix preparation.

Insoluble subnuclear components resistant to high salt extraction was isolated as early as in 1942 (Mayer and Gulick, 1942). The nuclear matrix was initially isolated from this high salt extraction procedure. Berezney and Coffey (Berezney and Coffey, 1974; Berezney and Coffey, 1975; Berezney and Coffey, 1977) developed this nuclear matrix-isolation procedure by combining the nuclease digestion and the high salt extraction. Briefly, chromatin was digested by deoxyribonuclease (DNase) I and then was removed with 2M NaCl extraction. The remnant stable proteinaceous skeleton was the nuclear matrix. Electron microscopy revealed that this structure was composed of fibers connected to a nuclear lamina. This nuclear matrix isolation protocol included an RNase digestion step and could not isolate a complete RNP network.

Since this pioneer study from Berezney and Coffey, many nuclear matrix isolation protocols have been developed and applied. Berenzney-Coffey method was modified by using different salts, ionic strengths, or enzymes (Kaufmann et al., 1981). For example, digested chromatin was removed from the nucleus by using detergents such as lithium 3,5-diiodosalicylate (LIS) without hypertonic salt concentration (Mirkovitch et al., 1984). The agarose-embedded nuclei were digested with nuclease under the physiological buffer
condition and electrophoretic fields were used to remove this digested chromatin (Jackson and Cook, 1985; Jackson and Cook, 1986; Jackson and Cook, 1988).

In the nuclear matrix isolated without RNase digestion, RNA can be released in the form of RNP particles by subsequent RNase treatment or sheering force (Fey et al., 1986b). Therefore, several early nuclear matrix isolation approaches were used to isolate nuclear RNP structures. The RNP particles released from the nuclear matrix were similar to those released from nuclei destroyed by mechanical force such as sonication because they were connected to remnant fibrils (Berezney, 1980).

Nuclear matrix preparation can be evaluated by how well the ultrastructure of the nuclear RNP network is preserved. Penman and co-workers developed relatively gentle isolation procedures that effectively extracted chromatin and soluble material from nuclei and left an apparently well-preserved nuclear structure (Fey et al., 1986a; He et al., 1990; Capco et al., 1982). In this method, the whole cells were extracted first with Triton X-100 in a physiological buffer. Then, Tween 40 and Deoxycholate in a low ionic strength buffer removed the cytoskeleton proteins except for the intermediate filaments. The remaining membrane-free nuclei that connected to the intermediate filaments at the surface lamina were digested with DNase I and then were extracted with 0.25 M ammonium sulfate to remove digested chromatin. More than 97% of nuclear DNA and almost all histones were removed through these steps. The resulting structure was an RNA-containing nuclear matrix as it still contained about 75% of nuclear RNA.

Components of the nuclear matrix

Matrix-Associated Region (MAR)-binding proteins.
It has been studied that DNA in interphase (Vogelstein et al., 1980) and mitotic chromosomes (Paulson and Laemmli, 1977) is organized in loops of 5-200kb by attachment to the nuclear matrix (interphase) and chromosome scaffold (metaphase), respectively. These DNA regions are known as matrix attachment regions (MARs) (Cockerill and Garrard, 1986) or scaffold attachment regions (SARs) (Mirkovitch et al., 1984). The nuclear structural proteins, A- and B-type lamins have been reported to bind MARs in vitro (Luderus et al., 1992; Luderus et al., 1994).

Scaffold attachment factor A (SAF-A) known as a p120 (von Kries et al., 1994) or a sp120 (Tsutsui et al., 1993) is identical to heterogeneous nuclear ribonucleoprotein U (hnRNPU) (Fackelmayer et al., 1994). This protein can bind to multiple MARs and mediate DNA loop formation in vitro, suggesting SAF-A plays important roles for both chromosomal organization and hnRNA metabolism (Romig et al., 1992).

Two cell type-specific MAR-binding proteins, special AT-rich binding protein 1 (SATB1) and 2 (STAB2), have been identified. SATB1 protein is expressed predominantly in thymocytes and is essential for T cell development (Dickinson et al., 1997). SATB1 knockout mice die soon after birth and have multiple defects in T cell development (Alvarez et al., 2000). SATB1-binding MARs are localized at the base of chromatin loop and are anchored to the nuclear matrix (de, I et al., 1998). On the other hand, SATB2 protein can bind MARs on the immunoglobulin μ locus and enhance gene expression in pre-B cell lines (Dobreva et al., 2003). Disruption of SATB2 leads to defects in bone development (Dobreva et al., 2006).
Actin

Actin that plays important roles in cytoarchitecture is also detected in the nucleus and the nuclear matrix. Some of the nuclear actin were likely to be in filamentous form (Amankwah and De, 1994) and depolymerization of this actin filaments led to rapid release of labeled RNA from the nuclear matrix. Actin was observed to be closely associated with small nuclear ribonucleoproteins (snRNPs) and pre-mRNAs (Nakayasu and Ueda, 1985). Moreover, it was shown to colocalize with shRNP and SC-35, a splicing factor in PC12 cell nuclei, depending on the state of cell differentiation (Sahlas et al., 1993).

Nuclear actin has been reported to involve in RNA metabolism. Microinjection of actin-binding proteins or actin antibodies into amphibian oocyte nucleus showed that the transcription by RNA polymerase II was inhibited at the lampbrush chromosomes (Scheer et al., 1984). A recent work showed that actin-hnRNP U (SAF-A) complex were associated with the phosphorylated C- terminal domain (CTD) of RNA polymerase II, which played an important role for RNA pol II-mediated transcription (Kukalev et al., 2005). Actin was also involved in the transcription initiation by RNA polymerase II (Hofmann et al., 2004).

Actin-related nuclear proteins such as myosin and structural protein 4.1 can bind to NuMA, suggesting that the nuclear structures related to actin were linked to those related to NuMA (Krauss et al., 1997). In addition, actin and actin related proteins have been found in chromatin-remodeling complexes (Rando et al., 2000).
**Intermediate filament proteins**

The intermediate filament is tightly associated with the nuclear matrix, forming a huge network termed the nuclear matrix-intermediate filament (NM-IF). Extending this observation, it has been proposed that the nuclear matrix, the cytoskeleton, and the extracellular matrix composed one large, continuous structure that may contribute to cell to cell signaling (Ingber, 1993).

Previous studies showed that the fibrils of the nuclear matrix strongly resembled to the intermediate filaments (He et al., 1990). However, antibodies against all known intermediate filaments did not recognize the nuclear matrix proteins prepared from HeLa cells except lamin proteins, suggesting that the intermediate filament-resembling structure may represent new intermediate family (Belgrader et al., 1991).

It has been suggested that intermediate filament-like proteins were the main component of the nuclear matrix in plants (Frederick et al., 1992). In addition, a cytokeratin-like protein were found in nuclear matrix prepared from regenerating rat liver (Aligue et al., 1990).

**Nuclear-Mitotic Apparatus Protein and other coiled-coil proteins**

The nuclear-mitotic apparatus (NuMA) protein was first identified by Pettijohn group (Lydersen and Pettijohn, 1980). NuMA was inside the nucleus in interphase cells but moved to the spindle poles in mitotic cells. During mitosis, NuMA appeared to be required for the spindle microtube formation (Kallajoki et al., 1991; Tousson et al., 1991) because microinjection of antibodies against NuMA blocked mitosis and led to the generation of daughter cells with micronuclei (Kallajoki et al., 1993).
NuMA has been shown to be a component of the nuclear matrix (Lydersen and Pettijohn, 1980; Kallajoki et al., 1991). It colocalized with splicing factors (Sm proteins) and snRNP complexes in the nuclear matrix prepared from HeLa cells (Zeng et al., 1994). The splicing complexes should be reconstituted with pre-mRNA to bind to NuMA in vitro. In the other hands, NuMA can assemble into multi-arm oligomers in vitro or in the overexpressed cells (Gueth-Hallonet et al., 1998; Harborth et al., 1999). These results suggested that NuMA functioned as a mediator between RNA processing and nucleoskeleton (Zeng et al., 1994).

The nuclear filament-related protein (NUF1), a related coiled-coil protein (110kDa) was characterized in yeast (Mirzayan et al., 1992) and was assumed to be a component of the nuclear matrix. Also, in mammalian cells, a homologue protein of NUF1 was identified. This protein was dispersed in the cytoplasm during mitosis differing from NuMA. In addition, a nuclear matrix protein (W511) was reported to have homology to the α-helical regions of intermediate filaments.

**RNA and heterogeneous nuclear ribonucleoproteins (hnRNP)**

The nuclear matrix contained about 70% of the nuclear RNA when it was prepared without RNase (He et al., 1990). Furthermore, it has been reported that the nuclear matrix was sensitive to RNase digestion and underwent structural rearrangement after transcriptional inhibitor or activator treatment (Fey et al., 1986b). These observations suggested that RNA had a potential structural role with certain RNA binding proteins. One example of a specific nuclear matrix-associated and non-protein-coding hnRNA is XIST, a polyadenylated RNA that is involved in higher-order chromatin
structure and epigenetic regulation (Clemson et al., 1996; Panning and Jaenisch, 1998).

Many hnRNP proteins were found in the nuclear matrix (Dreyfuss et al., 1984; Verheijen et al., 1986b; Verheijen et al., 1988) and constituted the filaments of the internal nuclear matrix with pre-mRNA (He et al., 1991). For example, a subset of nuclear matrix proteins was released by treatment with 2 M NaCl, RNase A, and DTT. The analyses by 2D electrophoresis identified these proteins as hnRNP proteins and the nucleolar protein B23 (Mattern et al., 1996).

The nuclear architecture

The mammalian cell nucleus is a three-dimensional structure composed of condensed chromatin (heterochromatin), interchromatic regions (euchromatin and nonchromatin nuclear matrix structures), nucleolar compartments, and a surrounding double-membraned nuclear envelope that contains nuclear pore complexes. Furthermore, high-resolution and computer-aided microscopic approaches led to the detection of the discrete domains of genomic organization and function. For example, in situ hybridization (FISH) approaches have been developed to detect specific DNA (or RNA) sequences in single cells. This technique can distinguish each chromosome, combined with different colored fluorochromes. Amazingly, individual interphase chromosomes occupy their own territory without overlapping and are restricted to a limited domain within the mammalian cell nucleus, termed chromosome territory (CT). In addition to the highly organized chromosome, protein components also show patterned arrangements, forming subnuclear compartment in the nucleus (Dundr and Misteli, 2001). Apparently, interphase chromosomes remain compact and are highly organized in discrete and non-
overlapping three dimensional territories inside the nucleus (Haaf and Schmid, 1991; Cremer et al., 1996). The distribution of chromosomal territories was preserved intact on the nuclear matrix isolated in the absence of nuclease digestion (Ma et al., 1999).

*The functional domain of the nuclear matrix*

Nucleoli were the nuclear domains that were well known as the sites of ribosomal RNA synthesis, processing, and assembly into ribosomal subunits, although other functions were still investigated (Fischer et al., 1991). In addition, Several types of nonnucleolar domains inside of nucleus were characterized by the size, the number per nucleus, and the presence of specific nuclear components. These spatially distinct, functional nuclear domains included RNA transcription sites, DAN replication sites, speckled domains enriched in RNA splicing factors, coiled bodies and chromosomal territories (Nickerson et al., 1995; Spector, 1996).

Accumulating studies showed that the nuclear matrix played important roles in the spatial and functional organization of the interphase nucleus (van der, 2000). Notably, the spatial distribution of nascent DNA(Nakayasu and Berezney, 1989), nascent RNA, and splicing factors still remained detectable and unaltered after removing most of chromatin by DNase I digestion and high salt extraction from sodium tetrathionate-stabilizing cells. These results suggested that the machineries for DNA replication, RNA transcription, and RNA processing were associated with the nuclear matrix.

*DNA replication foci in the nuclear matrix*

The replication foci were one of nonnucleolar domains seen in the nuclear matrix.
A few hundred discrete replication foci were observed in Eukaryote nucleus (Nakamura et al., 1986; Nakayasu and Berezney, 1989) and DNA-replication sites were showed in stable cell line expressing low level of PCNA fused to GFP (Leonhardt et al., 2000). The size and distribution of these foci were able to be altered depending on the cell cycle phase. For example, these replication foci became ordered during S phase of the cell cycle (D'Andrea et al., 1983). Moreover, these foci contained components required for DNA modification and cell cycle control such as DNA polymerase α (Bensch et al., 1982), replication protein A (a 70-kDa subunit) (Cardoso et al., 1993), DNA ligase (Lasko et al., 1990), proliferating cell nuclear antigen (PCNA) (Bravo and donald-Bravo, 1987), DNA methyltransferase (Leonhardt et al., 1992), as well as cyclin A, and cyclin-dependent kinase 2 (cdk2) (Cardoso et al., 1993).

To get higher resolution image for DNA replication foci, nucleotide analogs have been used to label nascent DNA. Briefly, 5’-bromodeoxyuridine triphosphate (BrdUTP) can be incorporated into nascent DNA during DNA synthesis and be recognized by anti-BrdU antibodies. DNA replication foci were extensively labeled, showing that DNA synthesis was not randomly diffused (Nakamura et al., 1986; Nakayasu and Berezney, 1989; O'Keefe et al., 1992). Notably, after extraction of most of chromatin, remnant DNA and some replication factors were unaltered in spatial arrangement, indicating that they were associated with the nuclear matrix (Nakayasu and Berezney, 1989). These “replication factories” within the nuclear matrix-filament network has also been visualized using resinless section microscopy (Hozak et al., 1993).

A recent study showed that DNA replication factories were stably anchored in the living cell nucleus (Leonhardt et al., 2000). Time-lapse examination during S phase
reveals that replication foci did not show any directional movement, consistent with the previous results that replication complexes were stably anchored on a non-chromatin structure. Alteration of the distribution of replication sites occurred through disassembly and new assembly of replication factors, not by movement, merge, or division (Leonhardt et al., 2000).

**RNA transcription foci in the nuclear matrix**

At the early electron microscopy experiments, EDTA-regressive staining was combined with autoradiography to detect the nascent (nonnucleolar) RNA in the nuclear matrix. Actively synthesized heterogeneous nuclear RNA (hnRNA) was mainly detected in the interchromatin granules and perichromatin fibrils (Fakan and Hughes, 1989). Both structures contained polyadenylated RNA, hnRNP, small nuclear RNPs (snRNPs), and non-snRNP splicing factors (Spector et al., 1991).

Labeling the nascent RNA with Brd-UTP showed that newly synthesized RNA was distributed into over 2000 well-defined foci throughout the nucleoplasm except for the nucleoli in the cultured mammalian cells (Wansink et al., 1993; Jackson et al., 1998; Wei et al., 1999). These transcription foci contained newly synthesized RNA, RNA polymerase I and II, and various transcription factors (Mancini et al., 1999; Stenoien et al., 2000; van Wijnen et al., 1993). RNA transcription foci retained their three dimensional distribution in the nuclear matrix after removing most of chromatin by DNase digestion (Jackson and Cook, 1985; Wei et al., 1999; Dickinson et al., 1990; Wansink et al., 1996; Xing and Lawrence, 1991). These sites were not generally colocalized with the speckled domains where splicing factors were accumulated.
(Wansink et al., 1993).

Recently, transcription foci have been shown to be remained even in the absence of transcription initiation and elongation, indicating that RNA polymerase II complexes were not simply accumulated in the actively transcribing genes but existed as independent nuclear subcompartments (Mitchell and Fraser, 2008).

**Specked domains of RNA splicing factors in nuclear matrix**

Speckled domains were known to be highly enriched in molecules involved in RNA splicing. Splicing factor speckles were first detected by staining with autoimmune patient sera that reacted with shRNPs (Spector et al., 1983). Speckle domains were also highly concentrated with non-snRNP splicing factors that belonged to the SR family (Spector et al., 1991). Later, more proteins that constituted these speckled domains were identified, including the U1 snRNP-associated 70-kDa protein (Verheijen et al., 1986a) and a 64-kDa autoantigen with sequence similarity to the 65-kDa subunit of U2AF (Imai et al., 1993).

20-50 RNA splicing domains were found in mammalian nuclei after staining with specific probes for pre-mRNA splicing factors. The splicing factors were present in relatively high concentration in the interchromatin granule clusters previously observed by electron microscopy in unfractionated cells (Monneron and Bernhard, 1969; Spector, 1996). It was also reported that nuclear matrix-associated splicesome can undergo RNA splicing (Zeitlin et al., 1987; Zeitlin et al., 1989). The nuclear matrix proteins SRm160 and SRm300 were identified to be required for *in vitro* RNA-splicing reactions and were present in interchromatin granule clusters (Blencowe et al., 1998; Blencowe et al., 2000).
A subgroup of the nuclear matrix associated proteins

Many factors required for chromatin remodeling and gene activation were associated with the nuclear matrix. For example, histone acetyltransferases (HATs) and histone deacetylase (HDACs) known for coactivators and coprepressors, respectively have been reported to be associated with the internal nuclear matrix to regulate their target gene expression (Hendzel et al., 1994; Sun et al., 1999; Hendzel et al., 1991). Human BRM and BRGH1, components of SWI/SNP complex that involved in the chromatin remodeling during gene activation were found to be associated with the nuclear matrix and enriched in active chromatin (Reyes et al., 1997).

A number of transcription factors have been found to be associated with the nuclear matrix (van Wijnen et al., 1993), suggesting that actively transcribed genes may be tightly associated with the nuclear matrix, while inactive loci may not. For example, previous studies showed that Runx1 and 2 functionally were associated with the nuclear matrix (Harrington et al., 2002). Both proteins contained the nuclear matrix-targeting signal in their C terminal regions that were very well conserved in human, mouse, rat, and chicken and their nuclear matrix-targeting signal was sufficient to target heterologous proteins to the nuclear matrix (Zeng et al., 1997). Preserving their DNA binding activities, compromising their nuclear matrix-targeting capabilities dramatically affected their transcriptional activities, leading to profound biological consequences. The disruption of the nuclear matrix-targeting signal within Runx1 prevented myeloid differentiation, leading to the transformed and leukemia phenotype in mice. When the nuclear matrix-targeting signal within Runx2 was removed in mice, the bone formation was impaired due to the defects in maturation of osteoclast (Choi et al., 2001; Vradii et al., 2005).
These results provided the direct evidence that association with the nuclear matrix was critical for the target gene expression.

Studies on the nuclear matrix showed that the protein composition of the nuclear matrix depended on the cell differentiation state and varied with the cell type. This was first observed by analyses of the nuclear matrix proteins in human cell lines (Fey and Penman, 1988). Different cell types in prostate cancer cell lines expressed distinct nuclear matrix proteins (Getzenberg and Coffey, 1990; Partin et al., 1993). Using in vitro cell culture system, the nuclear matrix proteins were examined at different developmental stage of fetal rat osteoblasts. The result showed that the nuclear matrix proteins were changed during development (Dworetzky et al., 1990).

Some transcription factors can be differentially exchanged between the nuclear matrix and extractable soluble compartments of nucleus, depending on the various tissue types and growth conditions (van Wijnen et al., 1993). These findings implied that the association with the nuclear matrix can be flexible and there are specific mechanisms to regulate the exchange between the nuclear matrix and soluble compartments of nucleus. Post-translational modification was one of the mechanisms to regulate the association with nuclear matrix. For example, the tumor suppressor Rb protein was associated with the nuclear matrix in a relatively hypophosphorylated status but highly phosphorylated form was released from the nuclear matrix (Mancini et al., 1994). By contrast, p73, homologue of the p53 tumor-suppressor protein was phosphorylated by c-Abl and predominantly found in association with the nuclear matrix (Ben-Yehoyada et al., 2003). After acetylation by GCN5 and P/CAF, the CDK9 is specifically found in the insoluble nuclear matrix (Sabo et al., 2008). PDLIM2 can bind to p65 and promote p65
polyubiquitination as a nuclear ubiquitin E3 ligase. Interestingly, PDLIM2 targeted p65 to the nuclear matrix where polyubiquitinated p65 was degraded by proteasome. Translocation of p65 into nuclear matrix by PDLIM2 seemed to be mediated by nuclear matrix-interacting proteins such as actinin and actin (Tanaka et al., 2007). These examples suggested that association of different proteins with the nuclear matrix may control their many functions.
ASSOCIATION WITH THE NUCLEAR MATRIX IS ESSENTIAL FOR PAX5 TO
CONTROL B LINEAGE GENE EXPRESSION PROGRAM

BY

SANG YONG HONG, TI HE, LIN HUANG, WANQIN XIE, KAI ZHONG DUAN,
ZHIHONG YU, PETER D BURROWS, KAIHONG SU, AND ZHIXIN ZHANG

In preparation for Journal of Biological Chemistry

Format adapted for dissertation
ABSTRACT

The nuclear matrix (NM) has long been proposed to provide a dynamic structural support for various biological reactions inside of the nuclei, including DNA replication, RNA transcription, and RNA splicing. Pax5 is an essential regulator for B lineage cell development, which controls the B lineage gene expression program by inducing and repressing hundreds target genes. Here, we show that majority of the endogenous Pax5 proteins in human and murine B lineage cells are associated with the NM, where they are distributed closely with the NM bound RNA polymerase II and the TATA box binding protein (TBP). Detailed analyses showed that the N-terminal of Pax5 is essential for NM targeting. In particular, mutations of Lysine 67, 87, and 89 residues within Pax5 to Alanine diminished its NM association and compromised Pax5-mediated global regulation of B lineage gene expression. Chromatin immunoprecipitation (ChIP) results further show that association with the NM is required for Pax5 to recruit the Cd19 promoter to the NM bound RNA polymerase complex. Based on these results, we propose an experimental model that Pax5 activates B lineage specific target gene expression through binding to the promoter or regulatory regions and recruiting these target genes to the NM bound transcription centers.
**INTRODUCTION**

Pax5 is a B lineage specific transcription factor required for B cell commitment and maintenance (1). The important function of Pax5 in B lineage cell development was demonstrated by the studies in *Pax5*-/- mice. Loss of Pax5 results in severe defects in generating B220+ progenitors in fetal liver (2) and developing CD19+ cells in adult bone marrow (3). Pax5 activates many B lineage specific genes, including *Cd19, Blnk, Mb-1 (Igα), VpreB*, and λ5(4), which partially explains the defective B cell development in *Pax5*-/- mice. Detailed comparisons of the gene expression profiles between the wild type and *Pax5*-/- pro B cells reveal that Pax5 activates the expression of hundreds of target genes encoding proteins for B cell signaling, adhesion, migration, antigen presentation, and germinal center B cell formation (5). On the other hand, Pax5 represses many lineage or developmental stage inappropriate genes, including *Flt3* (6), *Sca-1, Gp49b, Grap2, IgJ chain, Blimp-1*, and *Cd28* (7), which is required at all stages of B cell development. *Pax5*-/- pro B cells can differentiate into other lineage cells such as NK cells, dendritic cells, macrophages, osteoclasts, and granulocytes under appropriate culture conditions *in vitro* (8). When *Pax5*-/- pro B cells were transferred into mouse, they could migrate to the bone marrow, undergo self-renewal, and redifferentiate into all major hematopoietic lineages (9). After conditional deletion of Pax5, mature B cells lose their identity and prematurely express plasma cell specific genes (10).
Pax5 belongs to the Pax family of transcriptional regulators that play important roles in different developmental systems. Pax5 has a conserved “paired box” (PRD) DNA binding domain (11), a relatively conserved octapeptide motif (O), a partial homeodomain (HD), transactivation domain (TA), and transrepression domain (TR) (12). Pax5 interacts with multiple cellular factors by its functional domains to regulate transcriptional activities. Its partial homeodomain (HD) interacts with TATA box binding protein (TBP), and retinoblastoma (RB) protein (13). The N-terminus of the paired domain (PRD) can interact with Ets-1 (14) and form ternary complexes to bind to the Mb-1 promoter region (15). It has been shown that Pax5 interacts with the histone acetyl-transferases, such as CBP (16), p300, or the SAGA complex (17). Pax5 is acetylated by p300 at several lysine residues, which is important for its activity (18). Through these interactions, Pax5 recruits chromatin modifying factors to facilitate gene expression in committed B cells (19). On the other hand, Pax5 interacts with a corepressor Grg4, a part of the histone deacetylase complex to represses Pax5-mediated transcriptional function (20).

Accumulating evidences indicate that the eukaryotic cell nucleus is spatially and functionally compartmentalized into distinct subnuclear territories (21). The nuclear matrix (NM) refers to the highly branched intermediate filament scaffold with different sized foci after removing soluble proteins and chromatin DNA, originally observed under the electron microscope (22;23). Even though the molecular identities of the basic structural components of the NM is still under debate (24), it is believed that the NM provides a dynamic structural support for many biological reactions inside the nuclei, including DNA replication, RNA transcription, and RNA splicing (25). It has been shown
that newly-synthesized DNA is enriched in the NM together with DNA replication
centers anchored on the NM (26). Using a similar strategy, newly-made RNA transcripts
were also found at discrete foci that co-localize with the NM bound RNA polymerase II
and the transcription machinery (27). A recent study further showed that the transcription
machinery exists as a preassembled complex anchored on the NM independent of
transcription initiation and elongation (28). The RNA splicing factors are also enriched in
the NM, as visualized under the electron microscope (29). On the other hand, it has been
known for a long time that chromatin DNA forms various sized loops anchored on the
NM through particular DNA regions known as Scaffold Attachment regions (SARs) or
Matrix Attachment Regions (MARs) (30;31). MARs were first found on both the 5’ and
3’ flanking regions of the immunoglobulin Eμ and Igκ gene intrinsic enhancers (32;33).
The identified MARs contain stretches of AT-rich regions but contain no clear motifs in
the sequence. MARs play important roles in organization of chromatin structure and
regulation of gene expression. A number of transcription factors have been found to be
associated with the NM (34). For example, previous studies showed that Runx1 and
Runx2 associate with the NM (35). Both proteins contain nuclear matrix targeting
signal(NMTS) in their C terminal regions (36). Preserving their DNA binding activities,
compromising their NM-targeting capabilities dramatically affects their transcriptional
activities, leading to profound biological consequences (37).

In this current study, we show that Pax5 is associated with the NM where they are
distributed into discrete foci with NM bound RNA polymerase II and TBP proteins. The
association with NM is essential for Pax5 to control the B lineage gene expression
program. Moreover, Pax5 recruits the Cdl9 locus to the NM bound RNA polymerase
complex. Based on these results, we proposed an experimental model that Pax5 controls target gene expression through association with the NM and recruitment of target gene loci to the NM bound transcription machinery.
RESULTS

Pax5 is associated with the NM in B lineage cells.

Accumulating studies indicate that the NM provides a dynamic structural support for many biological reactions inside the nuclei including DNA replication, RNA transcription, and RNA splicing. It has been shown that several transcription factors are associated with the NM to fulfill their functions. To understand how Pax5 controls so many biological events, we analyzed the distribution of Pax5 is associated with the NM in different cellular fractions prepared for human and mouse B lineage cells. Our initial Western blot results showed that more than 50% of the endogenous Pax5 proteins are distributed in the NM fraction in human B lineage EU12 cells as well as murine pre B 2A cells (Fig. 1A and 1B). Under the same experimental condition, only a small fraction of E12 or E47 proteins were detected in the NM faction. In agreement with previous observation, almost all the RNA polymerase II proteins are distributed in the NM fraction. As positive controls, NM-associated protein Satb2 and the nuclear membrane protein Lamin A/C are mainly presented in the NM fraction (Fig. 1A and B). Proteins known to be distributed in other cellular fractions are also analyzed to monitor the qualities of different cellular fraction. The association of Pax5 with the NM is further confirmed in a NM reassemble experiment. After dialysis to remove the urea in the NM fraction, Pax5 proteins are really reassembled into the mesh precipitates together with Pol
II, Satb2, and Lamin A/C (Fig. 1A and B). The distribution of Pax5 was further analyzed by immunofluorescence studies. In fixed and permeabilized EU12 cells, Pax5 proteins are mainly distributed in the nuclei. When the EU12 was prepared on slides, Pax5 proteins are remained on the NM, where they are distributed at discrete foci partially colocalizing with the remnant Lamin A/C (Fig. 1C). Moreover, Double staining of Pax5 and Pol II or TBP showed that Pax5 proteins are distributed closely to the RNA pol II and TBD foci on the NM slides (Fig. 1D and E). The close distribution of Pax5 with the TBP foci can be visualized with 3-dimensional images generated with pictures taken at different Z-stacks. Taken together, these results showed for the first time that Pax5 proteins are associated with the NM, where they closely distributed at foci containing Pol II and TBP in human B lineage cells.

*The N-terminal of Pax5 is necessary for association with the NM*

Pax5 has several well defined functional domains, including a DNA binding domain (PRD), an octapeptide motif (O), a partial homeodomain (HD), a transactivation domain (TA), and a transrepression domain (TD). To identify the functional domains within Pax5 that are responsible for NM targeting, different Pax5 truncation constructs were generated and analyzed in HEK293 cells. Using these Pax5 truncation constructs, we found that ΔC1, ΔC3, ΔC, PRD, PRDΔHD, and ΔHD mutant proteins are still capable of associating with the NM, whereas ΔB and ΔB2 proteins lose their abilities to associate with the NM, indicating that the N-terminal of Pax5 is required for NM targeting (Fig. 2A). These results were confirmed by direct fluorescent analysis using corresponding GFP fusion Pax5 constructs. In fixed whole cells, all these mutant Pax5 proteins are
localized in the nuclei. After preparing the NM on slides, the GFP-fusing ΔB and ΔB2 constructs are no longer detectable on the NM slides (Fig. 2B). Taken together, these results suggest that the N-terminal region of Pax5 is required for NM targeting.

*The Lysine residues 67, 87, and 89 are required for association of Pax5 with the NM*

The N-terminal of Pax5 contains the highly conserved PRD DNA binding domains. Deletion of the Pax5 N-terminal region in the ΔB and ΔB2 constructs destroys the PRD DNA binding domain of Pax5, which makes these constructs to be inappropriate for functional analysis to determine if the NM association is necessary for Pax5 function. Currently the only well characterized NM targeting signal (NMTS) is from Runx family members. Using the motif base search (MEME), we found that the loop region (amino acid 67-103) within Pax5 PRD domain contains a structure similar to the well defined NMTS in all the RUNX family members (Fig. 3A). The PRD domain contains two DNA binding motifs separated by a loop. The PRD region is highly conserved among Pax5 members in different species, not only for the two DNA binding motifs but also for the loop region (Fig. 3B and C). Based on these finding, we selectively targeted a serial of amino acid residues within the Pax5 loop region to determine if this region is important for NM targeting (Fig. 3D). Among these sites directed mutants, the K67A, P80R, and K87/89A constructs have defects in the association with the NM when analyzed in HEK293 cells (Fig. 3E). The K94A mutant marginally affected NM association However, S77R, G84R, K98A constructs are still able to associate with the NM. These results were confirmed by direct fluorescence analyses using corresponding GFP fusion Pax5 constructs. All the mutant Pax5 are localized in the nuclei of fixed whole cells, the K67A
and K87A are barely detected on the NM prepared on the sides (Fig. 3F). When the K67A and K87/89A mutant Pax5 constructs were expressed in Pax5<sup>−/−</sup> pro-B cells, they also failed to associate with the NM (Supplementary Fig. 2D). Taken together, these results identified a potential region within the Pax5 PRD domain which is important for Pax5 NM association.

*The association with the NM is required for Pax5 to control B lineage specific gene expression.*

Our previous studies showed that the K67, 87, and 89 residues are necessary for the induction of several Pax5 target genes when reconstituted in Pax5<sup>−/−</sup> pro B cells. We further performed microarray analysis and analyzed global gene expression to determine the effects of the K67A, K87/89A constructs. Compared to Pax5<sup>−/−</sup> pre B cells reconstituted with the wild type Pax5, the cells reconstituted with the K67A or the K87/89A mutant Pax5 showed dramatic reduction of the global target gene expression. Almost all well known Pax5 target genes were not induced or repressed by the K67A or K87/89A mutant constructs (Fig. 4A and B) although comparable proteins are expressed in transduced G5 cells. These results showed that association with the NM is essential for Pax5 to control global target gene expression.

*Pax5 recruits the CD19 gene promoter to the NM-bound RNA polymerase II complex.*

It is starting to be realized that the basal transcriptional machinery is organized as large complexes anchored on the NM. To understand how the association with the NM plays such an important role for Pax5 to control global target gene expression, we first
performed a series of traditional Chromatin immunoprecipitation (ChIP) assays by anti-Pax5 or anti-RNA polymerase II antibodies using 2A (Pax5+/+ murine pre-B cells), G5 (Pax5−/− murine pro B cells), G5-K67A (G5 cells reconstituted with K67A mutant Pax5), and G5-K87/89A (G5 cells reconstituted with K87A/89A mutant Pax5). The CD19 gene is a well characterized Pax5 target gene with two identical Pax5 binding sites at its promoter region. The ChIP samples were analyzed by realtime PCR using 8 pairs of primers covering 9 regions spanning mouse CD19 gene locus (Fig. 5A). As expected, signals corresponding to the site b primer set are enriched in ChIP samples prepared with anti-Pax5 and anti-RNA polymerase II antibodies in 2A cells but not in G5 cells, or G5 cells reconstituted with K67A, K87A/89A mutants. The tight correlation of the enrichment of signals from the CD19 promoter region in both anti-Pax5 and anti-RNA polymerase II ChIP samples indicate that Pax5 binding to the CD19 promoter region is responsible for Pol II occupation (Fig. 5B). Interestingly, in G5 cells reconstituted with the K67A or K87A/89A mutant Pax5, signals spanning the entire CD19 locus are slightly evaluated in the anti-Pax5 ChIP samples but not in the anti-RNA polymerase II ChIP samples, indicating that Pax5 association with the NM is required for Pol II occupation at the CD19 promoter. To further determine if the CD19 locus is associated with the NM and if such association is dependent on Pax5, we purified NM bound and protected DNA samples from 2A and G5 cells after extensive DNase I digestion and washing with 2M NaCl buffer. The results showed that the promoter and C-terminal region within CD19 locus are associated with the NM in 2A cells but not in G5 cells (Fig. 5C). Taken together, we concluded that Pax5 association with the NM is essential to recruit the CD19 locus to the NM bound transcription machinery.
DISCUSSION

Pax5 is an essential regulator for B lineage cell development. It activates more than 200 target genes involving in various functions during B lineage development and differentiation. Simultaneously, Pax5 represses up to 110 lineage or developmental stage inappropriate genes. Pax5 also plays an important role in regulating the B lineage specific $V_H$-$DJ_H$ recombination through contracting the $IgH$ locus spanning up to 2 Mb of DNA templates prior to V-DJ recombination and interacting with the RAG complexes. Although Pax5 binding sites have been identified in the promoter or regulatory regions of many Pax5 target genes, VH gene coding regions (38), and intergenic regions within the $IgH$ locus, it is not clear how Pax5 conducts multiple functions in the nucleus.

Accumulating studies indicate that the NM provides a dynamic structural support for numerous biological processes inside the nuclei, including DNA replication, RNA transcription, and RNA splicing. Recent studies further showed that the transcription machinery exists as preassembled complex anchored on the NM independently of transcription initiation and elongation (28). RNA polymerase complex associates with the NM independently of transcription initiation and elongation. Correspondingly, many transcription factors are found to be associated with the NM to execute their regulatory function.
Based on these information and reasons, we analyzed whether Pax5 associates with the NM. Our results showed that majority of endogenous Pax5 are distributed in the NM in human and mouse B lineage cells together with well known NM associated proteins SATB2 and LaminA/C. In contrasts, several transcription factors for B cell development, such as E47 and E12 are not distributed in the NM fraction, suggesting that association with the NM is a specific feature for Pax5. Immunofluorescence analyses of NM prepared on slides showed that Pax5 is localized closely to many transcription foci marked by RNA Polymerase II and TBP. It has been well documented that the transcription machinery are anchored on the NM. Currently, it is not clear why there are about half of the Pax5 proteins are associated with the NM and the other half are in the soluble fractions. Pax5 may have the potential to move from the soluble fraction to the NM. Such movement may recruit DNA template to the NM.

Analyses of a serial N-and C-terminal truncation Pax5 constructs indicated that the PRD domain is required for NM targeting although deletion of the C-terminal of Pax5 also affects NM association. Currently, the only well characterized NMTS is from Runx family members. Using the motif base search (MEME), we found that the Pax5 65AA-98AA region displays structural similarity with the Runx NMTS. This region connects two DNA binding domains within PRD. However, this region is highly conserved in all Pax5 proteins from almost all species and in all Pax members in human. We specifically focused on this region and tested a serial of site-directed mutant. Among them, the K67A, P80R, and K87A/89A showed reduced or diminished NM association. The P80R mutant has been identified previously from childhood acute lymphoblastic leukemia (ALL) patient. The P80R mutant partially lost association with the NM and shows reduced
activity on the induction of CD19 expression when reconstituted into Pax5-/− pro B cells. These results suggested that alteration of Pax5 association with the NM may cause abnormal B cell development and leukemia. Global gene expression profile showed that K67A or K87A/K89A mutants failed to activate or inhibit many target genes, indicating that the association with the NM is essential for Pax5 to control the B lineage specific gene expression.

Based on our current understanding, Pax5 binding to its target gene promoter region recruit the RNA pol II to this promoter to initiate transcription. Results from our conventional ChIP confirmed that RNA Pol II occupied CD19 promoter region in 2A cells but not in G5 cells. Now, it becomes clear that RNA Pol II and the basal transcription machinery are physically anchored on the NM scaffold, rather than moving around to find promoters. So, we propose that Pax5 binding to the CD19 promoter region recruits the CD19 promoter to the NM-bound RNA Pol II. Indeed, our NM precipitation studies provided the direct evidence that the CD19 promoter is associated with the NM in 2A cells, but in G5 cells.

In summary, results presented in this study demonstrate that Pax5 association with the NM is essential for regulation of the B lineage specific gene expression program. We also propose that Pax5 binding to its target gene promoters recruit them to the NM-bound transcription center for expression.
MATERIAL AND METHOD

Cell lines

Human EU12 cells are maintained as described previously (39). Abelson virus transformed murine 2A and G5 cells were cultured in RPMI 1640 medium (Invitrogen) supplemented with 10% heat inactivated FBS (Invitrogen), 100 units/ml penicillin/streptomycin, 2 mM L-glutamate and 50 μM β-mercaptoethanol. HEK 293 cells and NIH 3T3 cells were maintained in DMEM medium supplemented with 10% heat inactivated FBS, 100 unit/ml penicillin/streptomycin, 2 mM L-glutamate and 10 mM Hepes.

Nuclear matrix preparation

Different cellular fractions were prepared as described previously with minor modification (40;41). Briefly, cells were harvested, washed twice with ice-cold PBS, and then extracted in CSK buffer (10 mM Pipes pH 6.8, 100 mM NaCl, 300 mM sucrose, 3 mM MgCl2, 1 mM EGTA, 0.5% Triton X-100, 1 mM DTT, and 1.2 mM PMSF) on ice for 5 min. After a centrifugation at 600 g for 5 min, the supernatant was collected as the soluble protein fraction (SO). The remaining pellet was extracted in low salt extraction buffer (42.5 mM Tris-HCl pH 8.3, 8.5 mM NaCl, 2.6 mM MgCl2, 1.2 mM PMSF, 1% Tween 40, 0.5% Deoxycholic acid) on ice for 5 min. After centrifugation at 9000 rpm for
1 min, the supernatant was collected as the cytoskeleton fraction (CK). The remaining pellet was resuspended in digestion buffer (10 mM Pipes pH 6.8, 50 mM NaCl, 300 mM sucrose, 3 mM MgCl2, 1 mM EGTA, 100 unit/ml DNase I, 0.5% Triton X-100, 1 mM DTT, and 1.2 mM PMSF) for 30 min at room temperature. Ammonium sulfate was added to a final concentration of 0.25 M to extract soluble nuclear proteins for 5 min at room temperature. After centrifugation at 13000 rpm for 5 min, the supernatant was collected as the chromatin fraction (CH). The pellet was further washed with 2 M NaCl buffer (2 M NaCl, 10 mM Pipes pH6.8, 10 mM EDTA, 1.2 mM PMSF) and collected as the 2 M wash fraction (2 M) after centrifugation. The remaining pellet was solubulized in 8 M Urea buffer (8 M Urea, 10 mM Tris-HCl, pH7.5) and referred as the nuclear matrix fraction (NM).

**In vitro NM reassemble assay**

NM proteins dissolved in 8 M urea buffer can be reassembled into mesh precipitants after dialysis (42). The NM fraction were transferred into a dialysis cylinder (10 KDa) and dialyzed against PBS at 4 °C overnight with constant stirring. After dialysis, the samples were transferred to a fresh eppendorf tube and centrifuged at 13000 rpm for 5 min at 4 °C. The resulting pellet contains the reassembled NM, which can be redissolved in 8 M urea buffer for Western blot analysis.

**Construction of Pax5 truncation mutants**

A serial of Pax5 truncation constructs and site-directed mutation constructs were generated by PCR using specific primers. All the Pax5 cDNA fragments were subcloned
either in pcDNA3.1, retroviral vector, pEBF vector for expression of GST fusion protein in mamalian cells, or pCMV-GFP vector to express GFP fusion proteins.

Transfection

HEK 293 cells were transfected using Polyethylenimine (PEI) method. Briefly, HEK 293 cells were seeded one day prior to transfection in the 100 mm plates at about 50% confluency. 10µg DNA and 30µl PEI solution (1mg/ml) were added to each tube with 500 µl serum free DMEM media and then gently mixed together. After incubation at room temperature for 20 min, the DNA/PEI mixture mix was added into each plate. The transfection efficiency is routinely monitored by flow cytometry analysis after transfection of pCMV-GFP plasmid.

Retroviral transduction

Different Pax5 cDNA fragments were subcloned into the pMI retroviral vector in front of the IRES-GFP expression cassette. Recombinant retrovirus was produced by co-transfection of the retroviral expression vectors together with the GP and ECO booster vectors into semi confluent ΦNX packaging cells. Recombinant retrovirus was collected from the culture supernatant 48 hours after transfection and filtered through a 0.45 um filter. The resulting live virus was used to transduce Pax5−/− mouse pro B cells with the help of polybrene (4 µg/ml). The culture medium was changed to polybrene-free RPMI in the following day. The transduction efficiency was monitored by FACS analysis of the frequency of GFP positive cells. To monitor CD19 expression on the cell surface, the retrovirus transduced cells (10^6 cells) were pellet and washed with FACS buffer (PBS
supplemented with 2% fetal bovine serum) twice and stained with APC-conjugated anti-mouse CD19 antibodies on ice for 15 minutes. After washing twice with FACS buffer, cells were resuspended in FACS buffer containing 2μl propidium iodide (20ug/ml). Samples were analyzed on an Accuri FACS machine and FACS data were processed using the CFow plus software from Accuri.

**Western blot**

Proteins were resolved by SDS poly-acrylamide gel electrophoresis and electro-transferred onto Amersham Hybond membrane using a semi-dry transfer cell. Membranes were blocked in PBST (0.05 % Tween 20 in PBS) containing 5 % milk for one hour at room temperature, washed with PBST, and incubated with specific primary antibody overnight at 4 °C. After washing three times with PBST, HRP-conjugated 2nd antibodies (Jackson immunoresearch) were incubated at room temperature for one hour. Bound antibodies were visualized by ECL Plus Western detection system (Amersham) and analyzed with a FluorChemQ gel imaging system (Alpha Innotech).

**Immunofluorescence**

The human B lineage EU12 cells or murine 2A pre-B cells were cytopspun onto superfrost glass slides at 800 rpm for 5min. For analysis of Pax5 distribution in whole cells, the whole cell slides were fixed in 3% formaldehyde in PBS at room temperature for 25 min and permeabilized with 0.1 % Triton X-100. For preparation of NM on the slides, slides with cells were air dried overnight, extracted with CSK buffer, digested with DNase I (200 unit / 1 ml), and washed with the digestion buffer containing 0.25 M
ammonium sulfate. The prepared NM slides were fixed with 3% formaldehyde in PBS at room temperature for 25 min. For immunofluorescence staining, the slides were washed twice with PBS and blocked with 5% BSA in PBS at RT for 1 hour. Specific anti-Pax5, anti-TBP, and anti-RNA polymerase II antibodies were diluted at 1:100 in PBST containing 1% BSA and incubated at 4°C overnight. After four times washing with PBST, secondary antibodies diluted at 1:1000 in PBS were added and incubated for 1 hour at room temperature. After extensive washing with PBST four times, the slides were briefly counter-stained with DAPI and analyzed with Olympus X81 fluorescence microscope.

**Chromatin immunoprecipitation (ChIP) assays and NM-precipitation**

2A (Pax5+/+) or G5 (Pax5−/−) murine pro B cells were used for the ChIP assays. 2A or G5 cells (50^7^ cells) were cross-linked for 10 min at room temperature with 1% formaldehyde solution in growth media followed by two washes with PBS. Crosslinked cells were lysed with lysis buffer 1 (50 mM HEPES pH7.9, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100) and washed with lysis buffer 2 (10 mM Tris-HCl pH8.0, 200 mM NaCl, 1 mM EDTA pH 8.0, 0.5 mM EGTA pH 8.0). Cells were resuspended and sonicated in sonication buffer (50 mM Tris-HCl pH 7.5, 140 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS) for 20 cycles at 30 sec each on ice (16W) with 60 sec on ice between cycles. Sonicated lysates were cleared by centrifuge at 14000 rpm for 10 minute and incubated at 4°C overnight with primary antibodies. The beads were added and incubated the following day. Beads were washed three times with sonication buffer, one time with sonication
buffer with 500 mM NaCl, one time with LiCl wash buffer (20 mM Tris pH8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5 % Na-deoxycholate), and one time wash with TE (10 mM Tris pH 8.0, 1 mM EDTA). DNA was eluted in elution buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH8.0, 1% SDS) at 65°C for 15 min with occasional vortexing. Crosslinking were reversed by overnight incubation at 65°C. RNA and protein were digested using RNase A and Proteinase K, respectively, and DNA was purified with phenol chloroform extraction and ethanol precipitation. The Specific primers (8 pairs) spanning the CD19 loci were used to analyze the ChIP DNA samples. Real time PCR was performed on the ABI PRISM 9700HT DNA sequence detection system using the ABI SBYR green PCR master mix. For NM-precipitation assays, NM bound DNA was purified after extensive digestion with DNAse and RNase. The remaining DNA is analyzed by realtime PCR.

Electrophoretic mobility shift assay (EMSA)

EMSA was performed using fluorescence dye labeled probes. Briefly, the top strand DNA was conjugated with the Alexa-488 or Cy5 fluorescence dye and annealed with complementary single strand in the annealing buffer (10 µl of 100 mM stock solution of each oligomer, 10 µl of 1 M NaCl, 70 µl of H2O) by heating to 100°C for 5 min and slowing cooling down to room temperature. The 5X binding reaction buffer contains 50 µl of 1M Hepes pH 7.9, 250 µl of 1M KCl, 500 µl of 100% glycerol, 90 µl of H2O, 100 µl of 10 mg/ml BSA, 5 µl of 1M DTT, and 10 µl of 0.1 M PMSF. Pax5 probe was incubated with nuclear extract expressing various mutated Pax5 constructs in 1X binding buffer (4 µl of 5X binding buffer, 2 µl of nuclear extract, 1 µl of polydIdC (1
mg/ml)). Samples were separated on 5-6% non-denature PAGE gel. The Alexa 488 or Cy5 signals were analyzed on a FlourchemQ gel imaging system (Alpha Innotech). To confirm that Pax5 specifically binds to the probe, supershift assay was performed by adding anti-Pax5 antibody into the binding reaction prior to the adding of fluorescence probe.

**Microarray analysis**

G5 (Pax5−/−) pro B cells were reconstituted with wild type Pax5, the K67A, or K87A/89A mutant Pax5. Transduced GFP+ cells were purified by FACS sorting. Total RNA was extracted using Trizol reagent from GFP+ cells (10⁶ cells) and used for Microarray analysis (Affymetrix Mouse V1 Microarray chip). The results were analyzed using Expression console (Affymetrix). Detailed analysis was performed using Gene Spring software ver 10.0 and BRB program.
Reference List


34. van Wijnen, A. J., Bidwell, J. P., Fey, E. G., Penman, S., Lian, J. B., Stein, J. L.,


36. Zeng, C., van Wijnen, A. J., Stein, J. L., Meyers, S., Sun, W., Shopland, L.,

37. Choi, J. Y., Pratap, J., Javed, A., Zaidi, S. K., Xing, L., Balint, E., Dalamangas, S.,

38. Zhang, Z., Espinoza, C. R., Yu, Z., Stephan, R., He, T., Williams, G. S., Burrows,

19, 21-31


42. Gerner, C., Holzmann, K., Meissner, M., Gotzmann, J., Grimm, R., and
Figure 1

A

EU12 cells

<table>
<thead>
<tr>
<th></th>
<th>CK</th>
<th>CH</th>
<th>2M</th>
<th>NM</th>
<th>α-</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Pax5
- Pol II
- E12
- E47
- Satb2
- Lamin

B

ZA cells

<table>
<thead>
<tr>
<th></th>
<th>CK</th>
<th>CH</th>
<th>2M</th>
<th>NM</th>
<th>α-</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Pax5
- Pol II
- E12
- E47
- Satb2
- Lamin

C

<table>
<thead>
<tr>
<th>Cell</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Pax5
- Lamin
- Merge
- Hoechst

D

<table>
<thead>
<tr>
<th>Cell</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Pax5
- Pol II
- Merge
- Hoechst

E

<table>
<thead>
<tr>
<th>Cell</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Pax5
- TBP
- Merge
- Hoechst
Figure 1. Pax5 is associated with the NM in B lineage cells.

(A, B) Western blot analyses of the distribution of Pax5, E12, and E47 in different cellular fractions prepared from human EU12 B lineage cells (A) or murine 2A pro B cells (B). SO, soluble, CK, cytoskeleton, CH, chromatin, 2M, 2M NaCl wash, and NM, nuclear matrix. After dialysis of the NM fraction dissolved in 8M urea, the pellet (P) and soluble proteins (S) are separated by centrifugation. The distributions of the RNA polymerase II (Pol II) and NM associated proteins Satb1, Lamin A/C were analyzed as positive controls. (C, D, E) Immunofluorescence analyses of the distribution of fixed and permeabilized whole cell (Cell) or NM prepared on slides (NM) of EU12 cells. (C) Pax5 (green) and Lamin (Red), (D) Pax5 (red) and Pol II (green), (E) Pax5 (red) and TBP (green). Genomic DNA was visualized with Hoechst blue staining.
Figure 2. The N-terminal of Pax5 is required for the association with the NM. (A) Western blot analyses of the distribution of a series of Pax5 truncation constructs in different cellular fractions prepared from transiently transfected HEK293 cells. The diagram shows different Pax5 constructs (B) Direct fluorescence analyses of the localization of GFP-Pax5 truncation constructs in whole cell (Cell) or NM prepared on slides (NM) of transiently transfected HEK293 cells. Genomic DNA was visualized by DAPI blue staining. All experiments were repeated more than three times.
Figure 3. The K67 and K87/K89 within Pax5 PRD domain are required for association with NM. (A) Identification of potential Pax5 NMTS within PRD domain using the motif base search (NEME) with Runx1 NMTS. (B) Sequence alignment of different Pax members in human. (C) Sequence alignments of Pax5 from different species. (D) Potential Pax5 NMTS within PRD domain. The mutation site is highlighted in red. (E) Western blot analyses of the distribution of different Pax5 constructs in different cellular fractions prepared from transiently transfected HEK293 cells. (F) Direct fluorescence analysis of localization of different GFP-Pax5 constructs in fixed whole cells (Cell) or the NM prepared on slides (NM) of transiently transfected HEK293 cells. Genomic DNA was visualized by DAPI blue staining. All experiments were repeated more than three times.
Figure 4
Figure 4.

<table>
<thead>
<tr>
<th>Gene Accession</th>
<th>Gene Symbol</th>
<th>Pax5</th>
<th>K67A</th>
<th>K87R/K98A</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_008644</td>
<td>Cd19</td>
<td>9.99</td>
<td>1.41</td>
<td>1.57</td>
</tr>
<tr>
<td>NM_008528</td>
<td>Blnk</td>
<td>9.52</td>
<td>2.98</td>
<td>2.00</td>
</tr>
<tr>
<td>NM_163162</td>
<td>Tnrd3</td>
<td>8.60</td>
<td>1.43</td>
<td>1.47</td>
</tr>
<tr>
<td>NM_067891</td>
<td>Pcp4</td>
<td>5.81</td>
<td>3.62</td>
<td>1.74</td>
</tr>
<tr>
<td>NM_011927</td>
<td>Cea3amb</td>
<td>4.87</td>
<td>1.12</td>
<td>1.13</td>
</tr>
<tr>
<td>NM_002021</td>
<td>Stc1s5</td>
<td>4.63</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>NM_175260</td>
<td>Myh10</td>
<td>4.29</td>
<td>1.82</td>
<td>1.10</td>
</tr>
<tr>
<td>NM_001002896</td>
<td>Bfsp2</td>
<td>3.83</td>
<td>1.41</td>
<td>1.59</td>
</tr>
<tr>
<td>NM_130450</td>
<td>Elovl6</td>
<td>3.71</td>
<td>0.63</td>
<td>1.09</td>
</tr>
<tr>
<td>NM_175778</td>
<td>Scal</td>
<td>3.33</td>
<td>2.05</td>
<td>1.20</td>
</tr>
<tr>
<td>NM_011771</td>
<td>Ikatf3</td>
<td>3.04</td>
<td>1.23</td>
<td>1.08</td>
</tr>
<tr>
<td>NM_007521</td>
<td>Bach2</td>
<td>2.95</td>
<td>1.51</td>
<td>1.54</td>
</tr>
<tr>
<td>NM_009721</td>
<td>Atph1f</td>
<td>2.86</td>
<td>1.36</td>
<td>1.15</td>
</tr>
<tr>
<td>BC147127</td>
<td>4950420K17Rik</td>
<td>2.62</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>NM_198613</td>
<td>Apst1</td>
<td>2.76</td>
<td>1.10</td>
<td>0.94</td>
</tr>
<tr>
<td>NM_172930</td>
<td>Fam70a</td>
<td>2.69</td>
<td>1.29</td>
<td>1.27</td>
</tr>
<tr>
<td>NM_013674</td>
<td>Ift4</td>
<td>2.69</td>
<td>1.85</td>
<td>1.30</td>
</tr>
<tr>
<td>NM_009946</td>
<td>Cplx2</td>
<td>2.67</td>
<td>1.01</td>
<td>1.18</td>
</tr>
<tr>
<td>NM_009020</td>
<td>Rag2</td>
<td>2.63</td>
<td>0.83</td>
<td>1.57</td>
</tr>
<tr>
<td>NM_018615</td>
<td>Nup210</td>
<td>2.61</td>
<td>1.42</td>
<td>1.22</td>
</tr>
<tr>
<td>NM_019814</td>
<td>Prp3</td>
<td>2.59</td>
<td>1.37</td>
<td>1.36</td>
</tr>
<tr>
<td>NM_001162917</td>
<td>Dnnd4a</td>
<td>2.55</td>
<td>1.51</td>
<td>1.12</td>
</tr>
<tr>
<td>NM_013044</td>
<td>Gsr</td>
<td>2.52</td>
<td>0.65</td>
<td>0.87</td>
</tr>
<tr>
<td>A148411</td>
<td>El000240C6Rik</td>
<td>2.40</td>
<td>1.34</td>
<td>1.19</td>
</tr>
<tr>
<td>NM_008149</td>
<td>Gpn4</td>
<td>2.39</td>
<td>1.23</td>
<td>1.06</td>
</tr>
<tr>
<td>NM_178900</td>
<td>Prkd2</td>
<td>2.38</td>
<td>1.30</td>
<td>0.96</td>
</tr>
<tr>
<td>NM_001034337</td>
<td>Cds2</td>
<td>2.36</td>
<td>1.12</td>
<td>1.19</td>
</tr>
<tr>
<td>NM_013706</td>
<td>Cds2</td>
<td>2.34</td>
<td>0.97</td>
<td>1.18</td>
</tr>
<tr>
<td>NM_021609</td>
<td>Cc2b3b</td>
<td>2.33</td>
<td>0.90</td>
<td>1.22</td>
</tr>
<tr>
<td>NM_001033270</td>
<td>Sdc4a</td>
<td>2.32</td>
<td>1.11</td>
<td>1.34</td>
</tr>
<tr>
<td>NM_139200</td>
<td>Cytip</td>
<td>2.30</td>
<td>1.20</td>
<td>0.99</td>
</tr>
<tr>
<td>NM_001162917</td>
<td>Dnnd4a</td>
<td>2.27</td>
<td>1.11</td>
<td>0.92</td>
</tr>
<tr>
<td>BC062840</td>
<td>D200013E6Rik</td>
<td>2.24</td>
<td>1.12</td>
<td>1.13</td>
</tr>
<tr>
<td>NM_198163</td>
<td>Rab35</td>
<td>2.22</td>
<td>1.31</td>
<td>1.13</td>
</tr>
<tr>
<td>NM_001162917</td>
<td>Dnnd4a</td>
<td>2.22</td>
<td>1.27</td>
<td>1.08</td>
</tr>
<tr>
<td>NM_018784</td>
<td>Stgap1</td>
<td>2.21</td>
<td>1.18</td>
<td>1.13</td>
</tr>
<tr>
<td>NM_001162917</td>
<td>Dnnd4a</td>
<td>2.17</td>
<td>1.24</td>
<td>1.06</td>
</tr>
<tr>
<td>NM_030715</td>
<td>Polh</td>
<td>2.14</td>
<td>0.94</td>
<td>0.88</td>
</tr>
<tr>
<td>NM_198408</td>
<td>Crihbo</td>
<td>2.11</td>
<td>1.18</td>
<td>1.95</td>
</tr>
<tr>
<td>NM_009117</td>
<td>Saaf1</td>
<td>2.10</td>
<td>1.26</td>
<td>1.26</td>
</tr>
<tr>
<td>NM_001162917</td>
<td>Dnnd4a</td>
<td>2.06</td>
<td>1.24</td>
<td>0.92</td>
</tr>
<tr>
<td>NM_011521</td>
<td>Sdc4</td>
<td>2.03</td>
<td>1.15</td>
<td>1.10</td>
</tr>
<tr>
<td>NM_175380</td>
<td>Gpdt1</td>
<td>1.98</td>
<td>0.93</td>
<td>1.04</td>
</tr>
<tr>
<td>NM_008540</td>
<td>Man2af</td>
<td>1.97</td>
<td>0.93</td>
<td>1.12</td>
</tr>
<tr>
<td>NM_023154</td>
<td>Eth1</td>
<td>1.96</td>
<td>1.14</td>
<td>1.09</td>
</tr>
<tr>
<td>NM_027844</td>
<td>Krta5p-2</td>
<td>1.96</td>
<td>1.21</td>
<td>0.95</td>
</tr>
<tr>
<td>NM_133643</td>
<td>Edaradd</td>
<td>1.95</td>
<td>0.96</td>
<td>1.02</td>
</tr>
<tr>
<td>NM_183028</td>
<td>Pcmtd1t</td>
<td>1.94</td>
<td>1.15</td>
<td>1.06</td>
</tr>
<tr>
<td>NM_183031</td>
<td>Gpr183</td>
<td>1.91</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>NM_001142918</td>
<td>Tcf7l2</td>
<td>1.87</td>
<td>1.08</td>
<td>0.84</td>
</tr>
<tr>
<td>NM_178677</td>
<td>Scc2zc</td>
<td>1.86</td>
<td>0.90</td>
<td>0.98</td>
</tr>
<tr>
<td>NM_01162917</td>
<td>Dnnd4a</td>
<td>1.79</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>NM_007656</td>
<td>Cds2</td>
<td>1.75</td>
<td>0.93</td>
<td>0.86</td>
</tr>
<tr>
<td>NM_021280</td>
<td>Pirc1</td>
<td>1.74</td>
<td>1.06</td>
<td>0.85</td>
</tr>
<tr>
<td>NM_008737</td>
<td>Nr1f1</td>
<td>1.74</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>NM_020287</td>
<td>Clec24</td>
<td>1.73</td>
<td>1.05</td>
<td>0.72</td>
</tr>
<tr>
<td>NM_001162917</td>
<td>Dnnd4a</td>
<td>1.70</td>
<td>1.01</td>
<td>0.91</td>
</tr>
<tr>
<td>NM_172900</td>
<td>Siglec9</td>
<td>1.63</td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td>NM_178697</td>
<td>Tilt1</td>
<td>1.62</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>NM_022325</td>
<td>Ctsz</td>
<td>1.57</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>NM_025447</td>
<td>Dnmt1</td>
<td>1.53</td>
<td>0.86</td>
<td>0.89</td>
</tr>
</tbody>
</table>
Figure 4. Lost NM association compromises Pax5-mediated gene expression.

Affimexrix microarray analyses of global gene expression in Pax5−/− pro B cells (G5) reconstituted with wild type Pax5, or the K67A, K87A/89A mutant Pax5. Pax5-activated genes (A) and suppressed genes (B). (C) The list of genes strongly activated by Pax5 but not K67A and K87/89A mutant. (D) The list of genes strongly repressed by Pax5 but not K67A and K87/89A mutant. The expression level of listed genes is more than 40 % increased or decreased by Pax5 compared to K67A, K87/89A mutants
Figure 5

A. Mouse CD19 locus

B. Pax5 | RNA pol II

C. 2A | G5

Fold induction

Fold induction
Figure 5. Pax5 is essential to recruits CD19 locus to the NM bound RNA pol II complex. (A) Diagram of mouse CD19 locus. Arrows indicate the position of 8 sets of primers used to analyze the Cd19 locus. (B) ChIP assays were performed with anti-RNA polymerase II and anti-Pax5 antibodies in 2A cells (Pax5+/+), G5 cells (Pax5−/−), G5 cells reconstituted with the K67A mutant Pax5, or K87A/K89A mutant Pax5 constructs. (C) The real time PCR analyses of the CD19 locus in NM bound DNA in 2A (Pax5+/+) versus G5 (Pax5−/−) cells. Results shown are the fold enrichment at 8 sites along the CD19 locus. Error bars indicate standard deviation from triple experiments.
Figure 6

A

B
Figure 6. Pax5 recruits target gene promoter to NM bound transcription centers.

Diagram shows that Pax5 binds to its target gene promoter region initially localized in the chromatin loop (A) and recruits its target gene to the NM bound RNA polymerase complex to activate transcription (B).
PAX5 REORGANIZES THE NUCLEAR MATRIX TO FACILITATE B LINEAGE SPECIFIC GENE EXPRESSION PROGRAM

by

SANG YONG HONG, LIN HWANG, WEIHUA XUE, SHI-JIAN DING, KAIHONG SU, PHILIP W. TUCKER AND ZHIXIN ZHANG.

In preparation for *Journal of Biological Chemistry*
ABSTRACT

Pax5 is an essential regulator for B lineage cell development. Our recent studies showed that Pax5 controls the B lineage specific gene expression through association with the nuclear matrix (NM). Global proteomic comparison of the NM associated proteins in wild type pro B cells verses Pax5⁻/⁻ pro B cells revealed that wild type pro B cells have established a NM infrastructure to facilitate B lineage specific gene expression program, in which many important transcription factors for lymphocyte or B lineage development, including Ikzf3, Bcl11a-XL, EBF1, and Pax5 are enriched in the NM. In contrast, in Pax5⁻/⁻ pro B cell NM, many important transcription factors for non-B lineage development, such as members of the c/EBP, Smad, and Runx families are enriched. Restoration of Pax5 expression in Pax5⁻/⁻ pro B cells induced Bcl11a-XL. Interestingly, enforced expression of Bcl11a-XL in Pax5⁻/⁻ pro B cells induced 56 positive target genes and repressed 22 Pax5 negative target genes. Based on these results, we made a conclusion that Pax5 has the capacity to reorganize the NM through induction of other NM-associated lymphoid and B lineage specific transcription factors to fully control the expression of hundreds genes.
INTRODUCTION

The nuclear matrix (NM) refers to the fibril network inside of nucleus originally observed under electron microscopy after sequential extraction to remove soluble proteins, chromatin DNA, and RNA (1;2). Accumulating studies indicated that the NM provides a dynamic structural support for many biological reactions inside of the nucleus. For example, it has been shown DNA replication machinery, RNA transcription factory, RNA splicing complexes, and telomere complexes are associated with the NM (3-10). Recent study further showed that RNA polymerase complex are stabilized on the NM even without transcription initiation and elongation (11). On the other hand, chromatin is organized into various sized loops anchored on the NM through matrix attachment regions (MARs) or scaffold attachment regions (SARs) (12;13). Emerging studies showed that many transcription regulators are associated with the NM to execute their function in different types of cells or at different stages during cell development. Alteration of their NM associated proteins may provide an important mechanism to control global events during development and differentiation (14-16).

B cell commitment and development is controlled by the concert action of a group of lymphoid and B lineage specific transcription factors, including PU.1, Ikaros, Bcl11a, E2A, EBF1, and Pax5 (17;18). In particular, essential function of Pax5 in B lineage commitment and maintenance has been well documented (19-21). In $Pax5^{-/-}$ mice, B
lineage cell development is blocked at the pro B cell stage. Interestingly, *Pax5*−/− pro B cells have the capacity to be differentiated into other lineages, including macrophage, T cells, natural killer cells, osteoclast, and dendritic cells under appropriate conditions (21-23). Conditional inactivation of *Pax5* in mature B cells can also lead to dedifferentiation of mature B cells back to T cells *in vivo* (24). *Pax5* represses many lineage and differentiation inappropriate genes, including *M-CSFR, Notch1, and Flt3* to restrict B lineage development pathway. Meantime, *Pax5* activates more than 300 B lineage specific genes including *CD19, mb-1, Blink, and Ebf1* (25-27). Although it has been shown that *Pax5* can recruit transcriptional co-activators or co-repressors to control gene expression, it is far from clear how *Pax5* controls the global events to direct B cell development.

Our recent studies showed that *Pax5* is associated with the NM to control B lineage specific gene expression program. Our global proteomic comparison of the NM associated proteins in wild type pro B cells verse *Pax5*−/− pro B cells uncovered an interestingly lineage specific distribution of NM associated proteins. In wild type pro B cells, several lymphoid or B lineage specific factors, such as *Pax5*, *Ebf1*, *Ikzf3*, and *Bcl11a*, are highly enriched in the NM in wild type pro-B cells. In contrast, many non-B lineage factors, such as members of c/EBP, Runx3, Smad family are enriched in *Pax5*−/− pro-B cells. Among the B lineage specific NM associated factors, *Ikzf3* and *Ebf1* are all direct target genes of *Pax5*. Here, we also show that *Bcl11a-XL* is specifically induced by *Pax5*. Gene expression profile analyses showed that enforced expression of *Bcl11a-XL* in *Pax5*−/− pro-B cells controls a subgroup of *Pax5* positive or negative target genes. Taken together, these results uncover a role of *Pax5* in reorganization of the NM infrastructure.
through induction of a group of NM-associated lymphoid or B lineage specific transcription factors to fully control the B lineage specific gene expression program.
RESULTS

Proteomic analysis of NM associated proteins in 2A and G5 pro B cells.

Our recent studies showed that Pax5 is associated with the NM to control B lineage specific gene expression. To explore additional B lineage factors that are associated with the NM, we performed global proteomic analysis to compare NM associated proteins in 2A (\textit{Pax5}^{+/+}) verse G5 (\textit{Pax5}^{-/-}) murine pro B cells (Fig. 1A). Our proteomic analyses identified 44667 peptides from 2A cell NM and 55622 peptides from G5 cell NM, which are derived from 1007 proteins in 2A cells and 1481 proteins in G5 cells, respectively (Fig. 1B). There are 737 NM associated proteins that were found in both 2A and G5 cells. 270 proteins are enriched in 2A cell NM fraction and 744 proteins are enriched in G5 cell NM fraction. These results indicated that a significant difference of the NM associated proteins in 2A cells verse those in G5 cells (Fig.1C). As a control, multiple Pax5 peptides were detected in the 2A cell NM fraction but not G5 cell NM fraction. The shared NM associated proteins in 2A and G5 cells contain the essential factors for nuclear organization, DNA replication, RNA transcription, RNA splicing, and chromosomal remodeling, supporting the current theory that the NM provides a dynamic structural support for many biological processes inside of nuclei (Fig. 1D). Many important functional domains, such as SMC flexible hinge region, helicase domain, and RNA recognition motif, were identified in both 2A and G5 NM fraction.
Identification of NM associated proteins that are differently distributed in 2A or G5 cells.

Detailed analyses of the NM associated proteins differently distributed in 2A or G5 cell NM fractions revealed an interesting lineage specific pattern. To our particular interests, lymphocyte or B lineage specific factors, such as Pax5, Ebf1, Bcl11a, and Ikzf3 proteins are only found or highly enriched in 2A cell NM but not in G5 cell NM (Fig. 2A). In contrast, many non-B lineage factors, including, members of the c/EBF, Smad, Runx3 are only found or enriched in G5 cell NM (Fig. 2B). The level of enrichment is quantified by ratio of the number of specific peptides verse the number of total peptides. Among the B cell specific factors identified in 2A NM, Ikzf3 and Ebf1 are well known Pax5 target genes. Bcl11a-XL is specifically expressed in wild type pro B cells. We focused on analyses of the Bcl11a gene. There are 4 major isoforms of Bcl11a by alternative splicing termed as Bcl11a-XL (longest form), Bcl11a-L (long form), Bcl11a-S (short form), and Bcl11a-XS (shortest form) (28). Detailed analyses of the mass spectrometry results obtained from 2A NM identified many peptides specific for Bcl11a-XL (Fig. 2C). Western blot analyses of different cellular fractions prepared from 2A and G5 cells confirmed that Pax5 and Bcl11a-XL are only expressed in2A cell NM, but not in G5 cell NM. Ikzf3 is enriched in the 2A cell NM (Fig. 2D). These results showed that Pax5 induced a specific group of NM associated B cell factors.

Pax5 induces Bcl11a-XL expression

Bcl11a was originally identified as a proto-oncogene for murine myeloid leukemia (29) or human B lymphocyte leukemia (30). Bcl11a was also known as COUP interacting proteins (31). Bcl11a deficient mice have no B cells and abnormal T cell
development (32). Early studies suggested that Bcl11a is responsible for induction of RAG, Pax5 expression. Here, our results showed that restoration of Pax5 expression in G5 cells induced Bcl11a mRNA by 3.6 fold on microarray analysis (Fig. 3A). As expected, the well known Pax5 target genes, such as CD19, Blnk, Ikzf3, and Ebf1 are induced after Pax5 expression (25). The Pax5-mediated induction of Bcl11a expression was further confirmed by qRT-PCR using primers specific for all isoforms of Bcl11a (Fig. 3B). When primers specific for the Bcl11a-XL isoform were used in RT-PCR analysis, the results showed the Bcl11a-XL isoform is specifically expressed in 2A cells or in G5 cells reconstituted with Pax5 while the expression of Bcl11a-S isoform is slightly higher in 2A cells or G5 cells reconstituted with Pax5 (Fig. 3C). Using a Tamoxifen inducible ER-Pax5 expression in G5 cells, we confirmed that Bcl11a-XL proteins are strongly induced after adding tamoxifen for 1, 2, and 3 days (Fig. 3D). The activation of the ER-Pax5 system is monitored by FACS analysis of CD19 expression. These results indicate that Bcl11a-XL is a novel Pax5 target gene.

Identification and characterization of Pax5 binding sites within the Bcl11a promoter

To determine whether Pax5 directly activates Bcl11a-XL transcript, we analyzed Bcl11a gene promoter region and identified two potential Pax5 binding sties to the translation start codon (Fig. 4A). Using a serial of luciferase reporter constructs containing different fragments of the Bcl11a promoter regions, we found that coexpression of Pax5 induced luciferase expression in a Pax5-dose dependent manner (Fig. 4B). The electrophoretic mobility shift assay (EMSA) results further showed that the site 2 Pax5 binding site within a Bcl11a promoter binds to purified Pax5 protein. The
association of Pax5 to site 2 was further confirmed by supershift assays using anti-Pax5 antibodies. In comparison with the positive control CD19 probe, the Bcl11a site 2 probe has low binding affinity to Pax5 (Fig. 4C). Moreover, chromatin immunoprecipitation (ChIP) assay results showed that Pax5 and RNA polymerase II occupied the Bcl11a promoter region in 2A cells but not in G5 cells (Fig. 4D). These results demonstrate that Pax5 induces Bcl11a-XL expression consistent with the proteomic analysis and Western blot results.

Forced expression of Bcl11a-XL controls a subset of Pax5 target genes.

Without Pax5, G5 cells express almost none detectable Bcl11a-XL transcript and protein. Artificially expression of Bcl11a-XL in G5 cells by retroviral transduction showed global changes of gene expression with induction of more than 500 genes and inhibition of 70 genes. In the same system, forced expression of Pax5 induces 245 genes and represses 71 genes. Interestingly, 56 genes are induced by both Bcl11a-XL and Pax5 (Fig. 5A), whereas 22 genes are inhibited by both Bcl11a-XL and Pax5 (Fig. 5B). For example, many B cell specific genes, including Rag1, Rag2, Foxo1 (33), Lef1, and CD24a are upregulated by both Pax5 and Bcl11a-XL (Fig. 5A) whereas non B lineage genes, including Gfiba and Fcerg1g, are suppressed by both Pax5 and Bcl11a-XL (Fig. 5B). These results suggested that Pax5 and Bcl11a-XL function cooperatively control the B lineage gene expression program. However, some genes activated by Pax5, such as CD79a and VpreB are inhibited by Bcl11a-XL, implying that Bcl11a-XL also has different roles for B cell development.
DISCUSSION

It has been proposed that the NM provides the structural framework for the fundamental biological processes including DNA replication, RNA transcription, and RNA splicing (10). Furthermore, the components of NM associated proteins are distinct in different types of cells or at different differentiation stage, implying that cell type or differentiation specific factors are associated with the NM to exert their functions (14). Pax5 is an essential regulator for B cell commitment by activating B lineage specific gene expression and repressing lineage or differentiation stage inappropriate gene expression. \( \textit{Pax5}^{+/−} \) pro B cells are arrest for further B lineage differentiation, but are capable of differentiation into other lineages.

Here, global comparison of NM associated proteins in wild type pro B cells (2A) verse Pax5/-/- pro B cells (G5) indicated the several B cell specific factors are distributed in 2A NM, whereas non-B lineage factors are distributed in G5 cell NM, suggesting that Pax5 plays a role in regulating the NM components. Among the NM associated factors, \( \text{iKzf3}, \text{Ebf1}, \) and \( \text{Bcl11a} \) are all induced by Pax5. Although Ebf1 can be expressed without Pax5 (34), Ebf1 was not detected in G5 NM. Previous study showed that Bcl11a deficient mice have no B cells and have defects in T cell development (32). The failure to detect \( \textit{RAG1}, \textit{Pax5}, \) and \( \textit{IL-7R} \) gene expression suggested that Bcl11a is necessary for early lymphoid development. In this study, we showed that the Bcl11a-XL isoform is
specifically induced by Pax5. The Bcl11a gene contains one Pax5 binding element. Previous study suggests that Bcl11a is required for induction of Pax5 expression in early B cell development. In our study, Pax5 can also induce Bcl11a-XL expression, suggesting that there is a positive feedback loop between Pax5 and Bcl11a to enhance their expression. Forced expression of Bcl11a-XL altered expression of many genes and interestingly, a subset of genes regulated by Bcl11a-XL are also controlled by Pax5. It is possible that Pax5 controls this subgroup of genes through induction of Bcl11a-XL. Such positive feedback loop has been shown previously between Ebf1 and Pax5 (27;35).

Although the initiation of Ebf1 expression does not require Pax5, Pax5 can enhance Ebf1 expression. In our proteomic analysis, Ebf1 is not seen in the \(\text{Pax5}^{-/-}\) G5 cell NM, but is identified in wild type pro B cell NM. The association of Ebf1 with NM is also different from that of Pax5 because only a small fraction of Ebf1 is distributed in the NM in wild type pro B cells. Previous studies have identified Ikzf3 as a Pax5 inducible gene. Here, our results showed that Pax5 further enhances Ikzf3 expression and NM association. One of interesting findings is that in \(\text{Pax5}^{-/-}\) pro B cells, many non-B lineage transcription factors, such as \(c/EBP_g\), \(c/EBP_z\), \(Samds\), and \(Runx3\) are distributed in the NM (36;37). However, none of these genes are significantly inhibited by Pax5. These results led us to propose a competition model that all these lineage specific regulatory factors are competing for the limited transcription centers on the NM to activate their target genes. These results indicate that Pax5 can reorganize the NM infrastructure through induction of B lineage specific NM-associated factors and prevention of association with NM of non-B lineage factors to control B lineage specific gene expression program.
MATERIAL AND METHOD

Cell lines

Abeloson virus transformed murine 2A and G5 cells were cultured in RPMI 1640 medium (Invitrogen) supplemented with 10% heat inactivated FBS (Invitrogen), 100 units/ml penicillin/streptomycin, 2 mM L-glutamate and 50 μM β-mercaptoethanol. HEK 293 cells were maintained in DMEM medium supplemented with 10% heat inactivated FBS, 100 unit/ml penicillin/streptomycin, 2 mM L-glutamate and 10 mM Hepes.

Preparation of samples for Mass spectrometry

The NM fraction was prepared from one billion 2A or G5 cells. The NM fraction were incubated at 80°C for 30 min in the presence of 10 mM DTT, cooled down to room temperature, and incubated in dark with 55 mM iodoacetamide. Prior to digestion, the samples were diluted by 5 fold using 25 mM ammonium bicarbonate buffer. Calcium chloride was added to the final concentration of 5 mM. Trypsin was added at trypsin-to-protein ratio of 1:100 (w/w) and samples were digested at 37°C for 16 hours. The digested peptides were desalted using reversed-phase Sep-Pak cartridge (Waters, Milford, MA) and dried using Speed-Vac. The samples were dissolved in 10 mM NH4COOH, pH 3 for SCX Fractionation, which were performed on the Ultimate 3000 HPLC system, Dionex (Sunnyvale, CA) by SCX chromatography using a PolySulfoethyl A column (1.0 x 150
mm, 5 μm particle size; 300 Å pore size). The fractions were collected at 1 min intervals, a flow rate of 50 ul/ml, and gradient of 10-400 mM NH4COOH in 25% acetonitrile.

**Mass spectrometry analysis**

SCX fractions were analyzed by online LC MS/MS using an LTQ Orbitrap. All nano-HPLC-MS experiments were performed on an Eksigent Nano HPLC system, Eksigent (Dublin, CA) and a LTQ-Orbitrap mass spectrometer (Thermo Electron, Waltham, MA) with a nanoelectrospray ion source. The mass spectrometer was operated in the data-dependent mode to automatically switch between Orbitrap-MS and Orbitrap-MS/MS (MS²) acquisition. Survey full scan MS spectra (from m/z 375 to 1,575) were acquired in the Orbitrap with resolution R = 100,000 (after accumulation to a target value of 1,000,000 charges in the linear ion trap). The most intense ions (up to five, depending on signal intensity) were sequentially isolated for fragmentation in the linear ion trap using collisionally induced dissociation at a target value of 100,000 charges. The data analysis of MS² spectra were searched with the Mascot search engine (version 2.1.04, Matrix Science, London, UK). Mascot was used to match the MS² fragmentation spectra with sequences from the composite Mouse International Protein Index (IPI) protein sequence database (Version 3.52) containing the normal IPI mouse proteins, commonly observed contaminants, and the reverse sequences of all proteins. In the database search, carbamidomethylation of Cysteine was set as the fixed modification, while oxidation of methionine, acetylation of Lysine, carbamylation of Lysine, oxidation of methionine, dimethylation of Lysine, dimethylation at N-terminus, dimethylation of Arginine, methylation of Lysine and trimethylation of Lysine were set as variable modifications.
The precursor tolerance was first set as 10 ppm, MS$^2$ tolerance was 0.5 Da, final peptides was selected as mass accuracy of less than 5 ppm for the precursor mass, and peptide score higher than 15.

**Normalizing the data set of the identified nuclear matrix associated proteins**

Frequency = (the number of peptides / total number of peptides) X constant factor.

We normalized the data set by calculating the frequency and rank the detected peptides in 2A and G5 cells. The unique or common nuclear matrix associated proteins in 2A and G5 were grouped and compared using this frequency ratio. We characterize the NM associated proteins by bioinformatic approach through GeneCard and NCBI web site.

**Construction of Pax5 truncation mutants**

A series of Bcl11a truncation constructs were generated by PCR using specific primers. All the Bcl11a fragments were subcloned in pGL2 vector for luciferase assay, pMI vector for retroviral transduction, pCMV-GFP vector for direct fluorescence staining, or pcDNA3.1 vector for protein expression in mammalian cells.

**Transfection**

HEK 293 cells were transfected using Polyethylenimine (PEI) method. Breifly, HEK 293 cells were seeded one day prior to transfection in the 100 mm plates at about 50% confluence. 10µg DNA and 30µl PEI solution (1mg/ml) were added to each tube with 500 µl serum free DMEM media and then gently mixed together. After incubation at room temperature for 20 min, the DNA/PEI mixture mix was added into each plate. The
transfection efficiency is routinely monitored by flow cytometry analysis after transfection of pCMV-GFP plasmid.

**Retroviral transduction**

Bcl11a-XL or Pax5 cDNA fragments were subcloned into the pMI retroviral vector in front of the IRES-GFP expression cassette. Recombinant retrovirus was produced by co-transfection of the retroviral expression vectors together with the GP and ECO booster vectors into semi confluent ΦNX packaging cells. Recombinant retrovirus was collected from the culture supernatant 48 hours after transfection and filtered through a 0.45 μm filter. The resulting live virus was used to transduce Pax5−/− mouse pro B cells with the help of polybrene (4 mg/ml). The culture medium was changed to polybrene-free RPMI in the following day. The transduction efficiency was monitored by FACS analysis of the frequency of GFP positive cells. Samples were analyzed on an Accuri FACS machine and FACS data were processed using the CFlow plus software from Accuri.

**Microarray analysis**

G5 (Pax5−/−) pro B cells were reconstituted with Pax5, or Bcl11a-X. Transduced GFP+ cells were purified by FACS sorting. Total RNA was extracted using Trizol reagent from GFP+ cells (10^6 cells) and 100 ng of RNA was used for Microarray analysis (Affymetrix Mouse V1 Microarray chip). The results were analyzed using Expression console (Affymetrix). Detailed analysis was performed using Gene Spring software ver 10.0 and BRB program.
Reference List


Figure 1.

A

2A or G5 murine pro-B cell

Nuclear matrix fractionation

Trypsinization of protein samples

Fractionation of the peptides

Mass Spectrometer analysis

Data analysis

B

<table>
<thead>
<tr>
<th></th>
<th>2A</th>
<th>G5</th>
<th>mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hits</td>
<td>44667</td>
<td>55622</td>
<td>N/A</td>
</tr>
<tr>
<td>Proteins</td>
<td>1219</td>
<td>1844</td>
<td>59017</td>
</tr>
<tr>
<td>Genes</td>
<td>1007</td>
<td>1481</td>
<td>25614</td>
</tr>
</tbody>
</table>

C

D

<table>
<thead>
<tr>
<th>Function</th>
<th>Identified Nuclear Matrix Associates Proteins in both 2A and G5 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Replication</td>
<td>Chaf1b, Dnajc2, Hmgb1, Lig1, Mcm2, Mcm3, Mcm4, Mcm5, Mcm6, Mcm7, Npm1, Pdna. Phb, Pold1, Pold2, Rbbp4, Rbbp7, Rfc1, Rfc3, Rfc5, Rpa1, Rpa2, Ssrp1, Supt16h</td>
</tr>
<tr>
<td>Chromatin Remodeling</td>
<td>Actl6a, Baz1a, Baz1b, Mta2, Rbbp4, Rbbp7, Smarca1, Smarca4, Smarca5, Smarch1</td>
</tr>
<tr>
<td>mRNA Transcription</td>
<td>Polr2a</td>
</tr>
<tr>
<td>RNA Splicing</td>
<td>Bat1a, Ddx39, Ddx5, Ddx15, Eftud2, Hnmpa1, Hnmpa2b1, Hnmpa3, Hnmpc, Hnmpf, Hnmp1, Hnmp2, Hnmp3, Hnmpu, Khsrp, Nono, Npm1, Pabpc1, Pbn, Ppie, Prpf18, Prpf19, Prpf4, Prpf40a, Prpf6, Prtp1, Raly, Rbm25, Rbm39, Rbm43, Rbm46, Rpns1, Sf3a1, Sf3b1, Sf3b3, Sfpq, Sfrs7, Snmpa1, Snmpd1, Snmpd2, Snmpd3, Srm2, Syncr, Tardbp, Tra2a, U2af1, U2af2, 0610009D07Rik</td>
</tr>
</tbody>
</table>
Figure 1. Global proteomic analyses of the NM associated protein in wild type (2A) versus Pax5−/− (G5) pro-B cells. (A) The procedure of mass spectrometry analysis of the NM associated proteins from 2A and G5 cells (B) The total number of peptides, proteins, and genes identified from the NM fraction of 2A and G5 cells by mass spectrometry analyses. (C) The numbers of NM associated proteins specific for 2A or G5 cells or shared by 2A and G5. (D) The NM associated proteins shared by 2A and G5 cells are involved in fundamental biological reactions including DNA replication, RNA transcription, RNA splicing, and chromatin remodeling as classified using the Gene Ontology (GO) database.
Figure 2.

A

<table>
<thead>
<tr>
<th>GeneBank</th>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_016707</td>
<td>Bcl11a</td>
<td>B-cell CLL/lymphoma 11A (zinc finger protein)</td>
</tr>
<tr>
<td>NM_008782</td>
<td>Pax5</td>
<td>Paired box gene 5 (B-cell lineage specific activator protein)</td>
</tr>
<tr>
<td>ENSMUST00000109268</td>
<td>Ebf1</td>
<td>Early B-cell factor 1</td>
</tr>
<tr>
<td>NM_011771</td>
<td>Ikzf3</td>
<td>IKAROS family zinc finger 3 (Aiolos)</td>
</tr>
<tr>
<td>NM_001024917</td>
<td>N4bp2</td>
<td>NEDD4 binding protein 2</td>
</tr>
<tr>
<td>NM_009532</td>
<td>Xrc1</td>
<td>X-ray repair complementing defective repair in Chinese hamster cells 1</td>
</tr>
<tr>
<td>NM_020083</td>
<td>Chaf1b</td>
<td>chromatin assembly factor 1, subunit B (p60)</td>
</tr>
<tr>
<td>NM_016682</td>
<td>Uba2</td>
<td>ubiquitin-like modifier activating enzyme 2</td>
</tr>
<tr>
<td>NM_019748</td>
<td>Sae1</td>
<td>SUMO1 activating enzyme subunit 1</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>GeneBank</th>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_009884</td>
<td>Cebp9</td>
<td>CCAAT/enhancer binding protein (C/EBP), gamma</td>
</tr>
<tr>
<td>NM_001024806</td>
<td>Cebp2</td>
<td>CCAAT/enhancer binding protein (C/EBP), zeta</td>
</tr>
<tr>
<td>NM_010754</td>
<td>Smad2</td>
<td>SMAD family member 2</td>
</tr>
<tr>
<td>NM_016769</td>
<td>Smad3</td>
<td>SMAD family member 3</td>
</tr>
<tr>
<td>NM_008540</td>
<td>Smad4</td>
<td>SMAD family member 4</td>
</tr>
<tr>
<td>NM_008541</td>
<td>Smad5</td>
<td>SMAD family member 5</td>
</tr>
<tr>
<td>NM_019483</td>
<td>Smad9</td>
<td>SMAD family member 9</td>
</tr>
<tr>
<td>NM_019732</td>
<td>Runx3</td>
<td>runt-related transcription factor 3</td>
</tr>
<tr>
<td>NM_010858</td>
<td>Myl4</td>
<td>myosin, light chain 4, alkali</td>
</tr>
<tr>
<td>NM_019391</td>
<td>Lsp1</td>
<td>lymphocyte-specific protein 1</td>
</tr>
<tr>
<td>NM_175836</td>
<td>Spnb2</td>
<td>spectrin, beta, non-erythrocytic 1</td>
</tr>
<tr>
<td>NM_007415</td>
<td>Parp1</td>
<td>Poly (ADP-ribose) polymerase 1</td>
</tr>
<tr>
<td>BC14735</td>
<td>Prt6</td>
<td>Centromere protein V</td>
</tr>
<tr>
<td>NM_012015</td>
<td>H2afy</td>
<td>H2A histone family, member Y</td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Repeat</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPFSVYSTEKHMKK (XL)</td>
<td>5</td>
<td>HMKTHGQVGK (XL)</td>
</tr>
<tr>
<td>FSTPFGELDGGISR</td>
<td>3</td>
<td>RHMKTHMHK</td>
</tr>
<tr>
<td>QCNGSCLCLEKGVDKPPSPSPIEMK</td>
<td>3</td>
<td>RSHTGEKPYK</td>
</tr>
<tr>
<td>HMKTHMHSKSSPTVK</td>
<td>3</td>
<td>RKQGPQHLSK</td>
</tr>
<tr>
<td>LGAEEMALATHPSAFDR</td>
<td>3</td>
<td>THMHHKSSPTVK</td>
</tr>
<tr>
<td>KASPVEVGIQVTPEDDDCLSTSSR</td>
<td>3</td>
<td>IDDGTVNGRGCSPGESASGGLSK</td>
</tr>
<tr>
<td>IYLESEGHSPLTPR</td>
<td>2</td>
<td>IDDGTVNGRGCSPGESASGGLSKK</td>
</tr>
<tr>
<td>LTRHMKTHQVVK (XL)</td>
<td>2</td>
<td>GVDKPPSPSPIEMK</td>
</tr>
<tr>
<td>VGISGLGAECPQRLHGIHIADNNPFNLLR</td>
<td>2</td>
<td>GVDKPPSPSPIEMK</td>
</tr>
<tr>
<td>SKSCEFCGKTFK</td>
<td>2</td>
<td>THQVQGKDVYK (XL)</td>
</tr>
<tr>
<td>THQVQGKDVYKCEICK (XL)</td>
<td>1</td>
<td>HMKTHQVQGKDVYK (XL)</td>
</tr>
<tr>
<td>NCSNLVHRRSHTGERPK (XL)</td>
<td>1</td>
<td>SHTGERPYKCELNCYACAQSSKLTR (XL)</td>
</tr>
</tbody>
</table>

D

<table>
<thead>
<tr>
<th>2A cells</th>
<th>G5 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

- Bcl11a
- Ikzf3
- Pax5
- Lamin B
Figure 2. Distribution of lineage specific NM associated proteins in 2A and G5 cells. 

(A) The list of proteins which are enriched or unique in 2A cell NM. Lymphoid or B cell specific transcription factors are highlighted in red. The ratio of enrichment was decided by the number of peptides of specific gene verses the total peptide number. (B) The list of proteins which are enriched or unique in G5 cell NM. Non-B lineage transcription factors are highlighted in blue. (C) Identification of peptides from Bcl11a and Bcl11a-XL. Bcl11a-XL specific peptides are highlighted in red. (D) Bcl11a-XL and Ikzf3 are expressed and associated with the NM in wild type pro B cells. Western blot analyses of the expression and distribution of Bcl11a-XL, Ikzf3, and Pax5 in different cellular fractions from 2A and G5 cells. LaminB was also analyzed as a positive control.
Figure 3.

A

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Fold induction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cdl9</td>
<td>27.73</td>
</tr>
<tr>
<td>Blnk</td>
<td>24.52</td>
</tr>
<tr>
<td>Ikrz3</td>
<td>5.80</td>
</tr>
<tr>
<td>Cd79a</td>
<td>5.59</td>
</tr>
<tr>
<td>Irf4</td>
<td>4.90</td>
</tr>
<tr>
<td>Bcl11a</td>
<td>3.63</td>
</tr>
<tr>
<td>Ebf1</td>
<td>2.00</td>
</tr>
</tbody>
</table>

B

![Bar graph showing fold induction](image)

C

![Western blots showing Bcl11a XL and Bcl11a S](image)

D

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>1 day</th>
<th>2 day</th>
<th>3 day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SO</td>
<td>CK</td>
<td>CH</td>
<td>2M</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>CK</td>
<td>CH</td>
<td>2M</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>CK</td>
<td>CH</td>
<td>2M</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>CK</td>
<td>CH</td>
<td>2M</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>CK</td>
<td>CH</td>
<td>2M</td>
</tr>
</tbody>
</table>

![Western blots showing Pax5, Bcl11a XL, Matrin3, Lamin B](image)
Figure 3. Pax5 induces Bcl11a-XL expression. (A) Microarray results show that reconstitution of Pax5 expression in $Pax5^{-/-}$ pro-B G5 cells by retroviral transduction induces different Pax5 target genes, including Bcl11a. (B) Real time results show that the transcription level of Bcl11a is induced after reconstitution of Pax5 expression in $Pax5^{-/-}$ pro-B G5 cells by retroviral transduction. (C) Reverse transcription-PCR (RT-PCR) to detect the specific isoforms of Bcl11a. Bcl11a-S is the short isoform and Bcl11a-XL is the extra long isoform. For semi-quantification analyses, serial 1 to 4 diluted samples were used. Resulting PCR products were separated on agarose gel and analyzed using FluroQ imaging system. (D) Western blot analyses of Pax5 and Bcl11a-XL in different cellular fraction prepared from G5 cells expressing the inducible ER-Pax5 fusion protein. Following treatment with tamoxifen for 1, 2, and 3 days, the expression and distribution of Pax5 and Bcl11a-XL were analyzed. The distribution of the Matrin-3 and Lamin proteins were also analyzed to monitor the quality of NM fraction.
Figure 4.

A

B

C

D
Figure 4. Identification and characterization of Pax5 binding sites within the Bcl11a gene promoter. (A) Identification of two potential Pax5 binding sites within the Bcl11a gene promoter region by DNA sequencing analyses. (B) Coexpression of Pax5 induces the Bcl11a promoter driven luciferase reporter gene expression. Three different length of Bcl11a promoter regions were amplified by PCR and subcloned into the promoterless pGL2 vector for luciferase assay. These constructs were transiently transfected into HEK293 cells with or without Pax5. Transfection efficiency was monitored by co-transfection of a CMV-luc2 reporter constructs. Results shown are fold induction above the basal level. Error bars indicate standard deviation from triplicated experiments (C) Electrophoretic mobility shift assay (EMSA) results show that purified GST-PAx5 proteins bind to the site 2 Pax5 binding sites within the Bcl11a promoter region. Supershift assays were performed using anti-Pax5 antibody. The CD19 probe was used as the positive control. (D) Chromatin immunoprecipitation (ChIP) assay results showed that Pax5 occupies the Bcl11a promoter region in wild type (2A) pro B cells but not Pax5<sup>−/−</sup> (G5) cells. ChIP samples were analyzed by real time PCR. Results shown are the fold enrichment of the Bcl11a promoter region in anti-Pax5 antibody or anti-RNA polymerase II ChIP samples verses control Ig ChIP samples.
Figure 5.

### A

<table>
<thead>
<tr>
<th>Gene Accession</th>
<th>Gene Symbol</th>
<th>Pax5 Bcl11a</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_009703</td>
<td>Bsr1</td>
<td>6.88</td>
</tr>
<tr>
<td>NM_008791</td>
<td>Pcp4</td>
<td>5.81</td>
</tr>
<tr>
<td>NM_133659</td>
<td>Erg</td>
<td>3.13</td>
</tr>
<tr>
<td>NM_016707</td>
<td>Bcl11a</td>
<td>3.04</td>
</tr>
<tr>
<td>NM_010703</td>
<td>Lef1</td>
<td>3.03</td>
</tr>
<tr>
<td>NM_011008</td>
<td>Ets1</td>
<td>2.96</td>
</tr>
<tr>
<td>NM_009946</td>
<td>Cdx2</td>
<td>2.96</td>
</tr>
<tr>
<td>NM_009721</td>
<td>Atp1b1</td>
<td>2.86</td>
</tr>
<tr>
<td>NM_016753</td>
<td>Lmx</td>
<td>2.73</td>
</tr>
<tr>
<td>NM_008538</td>
<td>Marcks</td>
<td>2.69</td>
</tr>
<tr>
<td>NM_009020</td>
<td>Rag2</td>
<td>2.63</td>
</tr>
<tr>
<td>NM_010344</td>
<td>Gsr</td>
<td>2.52</td>
</tr>
<tr>
<td>NM_175645</td>
<td>Xyst1</td>
<td>2.44</td>
</tr>
<tr>
<td>NM_006019</td>
<td>Rag1</td>
<td>2.42</td>
</tr>
<tr>
<td>AK149411</td>
<td>E4300204C06Rik</td>
<td>2.40</td>
</tr>
<tr>
<td>NM_024253</td>
<td>Nkb7</td>
<td>2.30</td>
</tr>
<tr>
<td>XP_030924</td>
<td>LOC100045972</td>
<td>2.21</td>
</tr>
<tr>
<td>NM_028127</td>
<td>Fmrn1</td>
<td>2.19</td>
</tr>
<tr>
<td>NM_011521</td>
<td>Dsc4</td>
<td>2.03</td>
</tr>
<tr>
<td>NM_030266</td>
<td>Inpd4b</td>
<td>1.99</td>
</tr>
<tr>
<td>NM_008548</td>
<td>Mar1a</td>
<td>1.97</td>
</tr>
<tr>
<td>NM_001035550</td>
<td>Lrcrb</td>
<td>1.96</td>
</tr>
<tr>
<td>NM_183028</td>
<td>Pcmtd1</td>
<td>1.94</td>
</tr>
<tr>
<td>NM_019730</td>
<td>Foxo1</td>
<td>1.92</td>
</tr>
<tr>
<td>NM_021987</td>
<td>Trps3tinp1</td>
<td>1.87</td>
</tr>
<tr>
<td>NM_021529</td>
<td>Lmo7</td>
<td>1.81</td>
</tr>
<tr>
<td>NM_020682</td>
<td>Stambp1</td>
<td>1.76</td>
</tr>
<tr>
<td>NM_004651</td>
<td>Btkl5</td>
<td>1.75</td>
</tr>
<tr>
<td>NM_029116</td>
<td>Htet11</td>
<td>1.75</td>
</tr>
<tr>
<td>NM_021280</td>
<td>Plg1</td>
<td>1.74</td>
</tr>
<tr>
<td>NM_020257</td>
<td>Clec2i</td>
<td>1.73</td>
</tr>
<tr>
<td>NM_00103782</td>
<td>Gm5592</td>
<td>1.73</td>
</tr>
<tr>
<td>NM_011158</td>
<td>Prkar2b</td>
<td>1.70</td>
</tr>
<tr>
<td>NM_001159963</td>
<td>Fbxl8</td>
<td>1.70</td>
</tr>
<tr>
<td>NM_003383</td>
<td>Hmg2</td>
<td>1.69</td>
</tr>
<tr>
<td>NM_023294</td>
<td>Ndc80</td>
<td>1.68</td>
</tr>
<tr>
<td>X16670</td>
<td>LOC290487</td>
<td>1.68</td>
</tr>
<tr>
<td>NM_026682</td>
<td>Prps2</td>
<td>1.66</td>
</tr>
<tr>
<td>NM_007332</td>
<td>Dck</td>
<td>1.65</td>
</tr>
<tr>
<td>NM_016769</td>
<td>Smad3</td>
<td>1.65</td>
</tr>
<tr>
<td>NM_001025261</td>
<td>Tpd52</td>
<td>1.65</td>
</tr>
<tr>
<td>NM_001081417</td>
<td>Cdh7</td>
<td>1.64</td>
</tr>
<tr>
<td>NM_015008</td>
<td>Krtap5-1</td>
<td>1.62</td>
</tr>
<tr>
<td>NM_018811</td>
<td>Dgka</td>
<td>1.62</td>
</tr>
<tr>
<td>NM_011487</td>
<td>Stat4</td>
<td>1.61</td>
</tr>
<tr>
<td>NM_001163026</td>
<td>Dnajc13</td>
<td>1.60</td>
</tr>
<tr>
<td>NM_177721</td>
<td>Ranbp6</td>
<td>1.57</td>
</tr>
<tr>
<td>NM_001439</td>
<td>Hmg1b1</td>
<td>1.56</td>
</tr>
<tr>
<td>NM_177733</td>
<td>Ez2</td>
<td>1.56</td>
</tr>
<tr>
<td>NM_178635</td>
<td>Uvrag</td>
<td>1.55</td>
</tr>
<tr>
<td>NM_010437</td>
<td>Hivc2</td>
<td>1.53</td>
</tr>
<tr>
<td>ENSMUST00000065104</td>
<td>4933439C10Rik</td>
<td>1.53</td>
</tr>
<tr>
<td>XP_030920</td>
<td>Gm9081</td>
<td>1.51</td>
</tr>
<tr>
<td>NM_172772</td>
<td>B230308D07Rik</td>
<td>1.51</td>
</tr>
<tr>
<td>NM_026441</td>
<td>Htt40</td>
<td>1.50</td>
</tr>
<tr>
<td>NM_172470</td>
<td>Wdr35</td>
<td>1.50</td>
</tr>
</tbody>
</table>
Figure 5. Pax5 and Bcl11a-XL control a subset of shared target gene expression. (A) Pax5 and Bcl11a-XL induce a shared group of target genes. A subset of genes are induced by both Pax5 and Bcl11a-XL upon reconstitution into $Pax5^{-/-}$ pro B cells. (B) List of a subset of genes induced by both Pax5 and Bcl11a-XL. (C) Pax5 and Bcl11a-XL inhibit a shared group of target genes after reconstitution into $Pax5^{-/-}$ pro B cells. (D) List of a subset of genes repressed by both Pax5 and Bcl11a-XL.
CONCLUSIONS

Pax5 is an essential regulator for B lineage cell development (Nutt et al., 1999). It not only activates B lineage specific gene expression, immunoglobulin gene $V_H$ to $DJ_H$ recombination, but also represses many lineage or developmental stage inappropriate genes to specify B lineage development pathway (Delogu et al., 2006; Zhang et al., 2006; Schebesta et al., 2007). The $Pax5^{-/-}$ pro-B cells can differentiate into other lineage cells, including macrophage, natural killer cells, T cells, dendritic cells, and osteoclast under appropriate conditions (Nutt et al., 1999; Rolink et al., 1999). Even mature B cells can dedifferentiate into T cells after conditional inactivation of Pax5 (Cobaleda et al., 2007). Despite of our understanding of the multiple functions of Pax5, it is not clear how a single transcription factor can control so many events.

Accumulating studies indicate that the nuclear matrix (NM) provides a dynamic structural support for various biological reactions inside the nuclei, including DNA replication, RNA transcription, and RNA splicing (Nickerson, 2001). Recent studies further showed that the transcription machinery exists as preassembled complex anchored on the NM independently of transcription initiation and elongation (Mitchell and Fraser, 2008). Correspondingly, chromatin DNA is packed into different sized loops attached to the NM (Mirkovitch et al., 1984; Cockerill and Garrard, 1986) and many transcription factors are found to be associated with the NM to execute their regulatory function (van Wijnen et al., 1993).
Based on these information and reasoning, I studied whether Pax5 is associated with the NM to execute its multiple functions. First, I analyzed the distribution of Pax5 in different cellular fractions. Western blot results showed that the majority of endogenous Pax5 protein was distributed in the NM in human and mouse B lineage cells together with RNA polymerase II and the well known NM-associated proteins SATB2 and LaminA/C. By contrast, other transcription factors which are also important for B cell development, such as E47 and E12 (Bain et al., 1994), are not distributed in the NM fraction, suggesting that association with the NM is a specific feature for Pax5. Currently, it is not clear why there are about half of the Pax5 proteins is associated with the NM and the other half is in the soluble fractions. One of the intriguing possibilities is that Pax5 may have the potential to translocate from the soluble fraction to the NM-bound RNA polymerase complex to regulate its target gene expression. Previous studied showed that association with the nuclear matrix can be regulated by posttranslational modification in Rb, NF-κB, p73, and CDK9 protein (Mancini et al., 1994; Ben-Yehoyada et al., 2003; Tanaka et al., 2007; Sabo et al., 2008). Pax5 can also be acetylated by Histone acetyltransferase (Emelyanov et al., 2002; Barlev et al., 2003; He et al., 2011). It will be interesting and important to determine if acetylation of Pax5 affects its association with the NM. Immunofluorescence analyses of NM prepared on slides showed that Pax5 is localized closely to many transcription foci marked by RNA Polymerase II and TBP. These results are consistent with previous observation that Pax5 interacts with TBP (Eberhard and Busslinger, 1999). It should be pointed out that the transcription machinery, including RNA polymerase II and TBP are anchored on the NM (Cook, 1999).
Such finding led us to further test if Pax5 association with the NM is essential for induction of its target genes.

Analyses of a serial N-and C-terminal truncation Pax5 constructs indicated that N-terminus of the PRD domain is required for NM targeting, although deletion of C-terminus of Pax5 also affects NM association. Currently, the only well characterized NMTS is from Runx family members (Zeng et al., 1997). To investigate the NMTS within PRD domain, we used the motif base search (MEME) program and found that the Pax5 65AA-98AA region contain structure similar to the Runx NMTS. This region forms an extended loop to connect two DNA binding domains within the PRD domain.

Interestingly, this region is highly conserved in all Pax5 protein from all species and in all the Pax members in human. We specifically focused on this region and tested a series of site-directed mutants. Among them, the K67A, P80R, and K87A/89A mutants showed reduced or diminished NM association and reduced activity on the induction of CD19 expression when reconstituted into Pax5<sup>−/−</sup> pro-B cells. The analyses of global gene expression profiles showed that the Pax5 K67A or K87A/K89A mutants failed to activate or inhibit many Pax5 target genes, indicating that the association with the NM is essential for Pax5 to control the B lineage specific gene expression. It should be pointed out that the P80R mutant was identified previously from childhood acute lymphoblastic leukemia (ALL) patient (Mullighan et al., 2007). Such results suggested that alteration of Pax5 association with the NM may cause abnormal B cell development and leukemia. It will be interesting to analyze that other Pax5 mutants identified from human B cell leukemia patients affect the association with the NM. Moreover, developing peptides or small
molecules that interfere with Pax5 association with NM may offer a novel strategy to modify Pax5 function.

Now, it becomes clear that RNA Pol II and the basal transcription machinery are physically anchored on the NM scaffold, rather than moving around in the nucleus to search for promoters (Jackson and Cook, 1985; Jackson et al., 1998). Having evidence that Pax5 is associated with the NM and distributed closely to transcription centers, I further analyzed if Pax5 binding to its target gene promoter region recruits this promoter to the NM-bound RNA pol II to initiate transcription. The conventional ChIP results showed that RNA Pol II occupied the CD19 promoter, which is dependent on Pax5 and its association with the NM. Our NM precipitation results provided the direct evidence that the CD19 promoter is associated with the NM in wild type pro-B cells (2A), but not in Pax5<sup>−/−</sup> pro-B cells (G5). So, we propose that Pax5 binding to the CD19 promoter region recruits the CD19 promoter to the NM-bound RNA Pol II. It will be interesting to perform global sequence analyses of NM-associated DNA to determine if association with the NM correlates with the gene expression status in 2A and G5 cells. Taken together, these results demonstrate that Pax5 association with the NM is essential for controlling its B lineage specific gene expression.

The NM plays an important role in supporting many biological processes inside of the nuclei. It has been shown that the composition of NM-associated proteins is distinct in different types of cells or cells at different developmental stage (Fey and Penman, 1988; Dworetzky et al., 1990; Partin et al., 1993). To investigate the role of Pax5 in NM organization, I performed global proteomic analyses to compare NM associated proteins in wild type pro-B cells (2A) verse Pax5<sup>−/−</sup> pro-B cells (G5). Interestingly, the results
showed that the wild type pro B cells have established a NM infrastructure to facilitate the B lineage specific gene expression program, in which many important factors for lymphocyte or B lineage development, including Ikzf3, Bcl11a-XL, EBF1, and Pax5 are enriched in the NM. By contrast, in Pax5−/− pro-B cells, transcription factors that are important for non-B lineage cell development, such as members of c/EBP, Smad, and Runx families (Woolf et al., 2003; Xie et al., 2004) are enriched in the NM fraction. Among the NM-associated lymphoid or B lineage specific factors, Ikzf3, Ebf1, and Bcl11a are all induced by Pax5. Although EBF can be expressed without Pax5 (Lin and Grosschedl, 1995), the level of EBF was not low to be detected in the G5 NM. Previous studies showed that Bcl11a deficient mice had no B cells and had defects in T cell development. The lack of Pax5 gene expression in Bcl11a−/− fetal liver cells suggested that Bcl11a was required for induction of Pax5 expression during early B cell development (Liu et al., 2003). Here, we showed that the Bcl11a-XL isoform is specifically induced by Pax5, implying that there is a positive feedback loop between Pax5 and Bcl11a-XL to enhance their expression. Interestingly, forced expression of Bcl11a-XL in Pax5−/− pro-B cells controls a subset of Pax5 target genes. It is possible that Pax5 controls this subgroup of genes through induction of Bcl11a-XL. One of the interesting findings is that in Pax5−/− pro-B cells, many non-B lineage transcription factors, such as c/EBPγ, c/EBPz, Smads, and Runx3 were distributed in the NM. However, none of these genes were significantly inhibited by Pax5. These results led us to propose a competition model that these lineage specific regulatory factors are competing for the limited foci on the NM to conduct their functions. A previous study showed that overexpression of Pax5 in hematopoietic progenitor cells promoted B lineage cell commitment at the expense of myeloid lineage
development (Schaniel et al., 2002a). Conversely, overexpression of c/EBPα enhanced myeloid lineage development (Akashi et al., 2000). It will be interesting to determine if reconstitution of Pax5 expression in Pax5−/− pro-B cells (G5) recruits B lineage specific factors to the NM and prevent non-B lineage factors from associating with the NM.

Taken together, works presented in this dissertation showed that Pax5 controls the B lineage specific expression program through association with the NM. In particular, Pax5 binding to its target gene promoter or regulatory regions recruits its target gene to the NM-bound transcription centers. Pax5 induces a group of NM-associated lymphoid or B lineage specific transcription factors to reorganize the NM infrastructure to fully control the B lineage specific gene expression program.
Reference List


60. Dickinson, P., Cook, P.R., and Jackson, D.A. (1990). Active RNA polymerase I is fixed within the nucleus of HeLa cells. EMBO J 9, 2207-2214.


63. Domen, J. and Weissman, I.L. (2000). Hematopoietic stem cells need two signals to prevent apoptosis; BCL-2 can provide one of these, Kitl/c-Kit signaling the other. J Exp. Med. 192, 1707-1718.


mammalian nuclei. Cell 71, 865-873.


EMBO J 30, 2388-2404.


135


