April 2013

The Role of PU.1 and Spi-B in B Cell Acute Lymphoblastic Leukemia

Shereen Turkistany
*The University of Western Ontario*

Supervisor
Rodney DeKoter
*The University of Western Ontario*

Graduate Program in Microbiology and Immunology

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

© Shereen Turkistany 2013

Follow this and additional works at: [https://ir.lib.uwo.ca/etd](https://ir.lib.uwo.ca/etd)

Part of the [Medical Immunology Commons](https://ir.lib.uwo.ca/etd)

**Recommended Citation**


[https://ir.lib.uwo.ca/etd/1225](https://ir.lib.uwo.ca/etd/1225)

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlswadmin@uwo.ca.
THE ROLE OF PU.1 AND Spi-B IN B CELL ACUTE LYMPHOBLASTIC LEUKEMIA

Thesis format: Monograph
By:
Shereen A. Turkistany

Graduate program in Microbiology and Immunology

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Shereen A. Turkistany 2013
ABSTRACT

ETV6-RUNX1 is the most common chromosomal alteration in pediatric B cell acute lymphoblastic leukemia. ETV6-RUNX1 represses RUNX1 target genes. However, little is known about the target genes of ETV6-RUNX1 that are involved in promoting leukemogenesis. A recent study with two human leukemia cell lines AT-2 and REH, which express ETV6-RUNX1, suggested that SPIB was one of the top of the genes that were upregulated after knocking down the ETV6-RUNX1 fusion protein. In addition, our lab showed that deletion of PU.1 and Spi-B in B cells in mice resulted in the development of B cell acute lymphoblastic leukemia with 100% incidence. It is still not clear what pathways and target genes are affected by the loss of PU.1 and Spi-B and contribute to leukemogenesis. We hypothesized that ETV6-RUNX1 function as an oncogenic driver by repressing SPIB transcription, leading to impaired B cell receptor signaling. This hypothesis was divided into two parts. First, examining how PU.1 and Spi-B could act as tumor suppressor genes in B cells in our mouse model using gene expression profiling. The results of the microarray specified impairment in the expression of genes involved in BCR signaling pathways. Second, we aimed to determine the molecular mechanism of how SPIB is repressed by ETV6-RUNX1. ChIP results suggested that ETV6-RUNX1 directly interacts with SPIB. This study is expected to lead to a deeper understanding of the underlying biology of leukemogenesis caused by the loss of SPIB in ETV6-RUNX1 leukemia patients, hence allowing for the development of new molecular targeted therapies for B-cell leukemia.

Keywords: B cell acute lymphoblastic leukemia, ETV6-RUNX1 B-ALL, PU.1, SPIB, and B cell receptor signaling (BCR).
To my Father’s soul,
ACKNOWLEDGMENTS

I would like to acknowledge all those people who helped me in this journey to complete my master’s degree. I would like to thank my supervisor Dr. DeKoter for all his patience, guidance, and support through out my project. I would also like to thank my advisory committee Dr. Anargyros Xenocostas and Dr. Sung Kim for their encouragements and insightful suggestions throughout my thesis project.

Many thanks to all past and present lab members of the DeKoter lab for providing help when I need it, and making my research time full of joy.

I would like to thank my mom for her constant support, love, and blessings. I would also like to thank my brothers and sisters for their endless love and encouragements.

I would like to thank Ola Ismail, Najwa Zebian, and Bodour Al-Khamees for being nice and helpful.

Finally, I would like to thank all the people I met in the department of Microbiology and Immunology especially the professors who taught me in the undergrad and the grad courses. They all were great and helpful.
TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................... ii

ACKNOWLEDGMENTS ......................................................................................................................... iv

TABLE OF CONTENTS ........................................................................................................................ v

LIST OF FIGURES ................................................................................................................................. vii

LIST OF TABLES .................................................................................................................................. ix

LIST OF ABBREVIATIONS ..................................................................................................................... x

Chapter 1: Introduction ........................................................................................................................... 1

1.1 Overview: ........................................................................................................................................ 1

1.1.1 Stages of B cell development: ................................................................................................. 1

1.1.2 Transcriptional regulation of B cell commitment: ................................................................. 2

1.1.3 Interleukin 7 (IL-7) and Interleukin 7 receptor (IL-7R) in B cell development: ................. 3

1.1.4 Pre-B cell receptor signaling in B cell development: .............................................................. 5

1.2 B cell acute lymphoblastic leukemia (B-ALL): ............................................................................. 6

1.3 Purine-Rich Box1 (PU.1): ............................................................................................................. 7

1.3.1 The Expression and Regulation of PU.1 in B Cell development: ....................................... 8

1.3.2 The Role of PU.1 in B cell development and its target genes: ............................................. 8

1.4 Spleen Virus Integration Site (Spi-B): ......................................................................................... 9

1.4.1 The Expression and the Role of Spi-B in B Cell Development: .......................................... 9

1.5 The role of PU.1 and Spi-B in B-ALL: ....................................................................................... 9

1.6 ETV6-RUNX1 acute lymphoblastic leukemia: .......................................................................... 10

1.7 Hypothesis and objectives: ......................................................................................................... 11

CHAPTER 2: Materials and Methods .................................................................................................... 12

2.1 Mouse strains: ............................................................................................................................... 12

2.2 Cell sorting: .................................................................................................................................. 12

2.3 Affymetrix Mouse Exon Array hybridization: ........................................................................... 12

2.3.1 RNA isolation and evaluation: .............................................................................................. 12

2.3.2 RNA processing: ................................................................................................................... 13

2.4 Data pre-processing and batch effect removal: ........................................................................ 13

2.4.1 Data pre-processing: ............................................................................................................. 13

2.4.2 Batch effect removal: ............................................................................................................. 14

2.5 Data analysis: ............................................................................................................................... 15

2.5.1 Differentially expressed genes: ........................................................................................... 15

2.5.2 Partek pathway: .................................................................................................................... 15

2.5.3 Heat map generation: ........................................................................................................... 15

2.6 Real-Time quantitative PCR (qPCR): ......................................................................................... 15

2.7 Bioinformatics analysis: ............................................................................................................. 16

2.7.1 Matinspector: ...................................................................................................................... 16

2.7.2 Macvector: ........................................................................................................................... 16

2.7.3 UCSC Genome Browser website: ......................................................................................... 16
CHAPTER 3: Gene Expression Analysis of B cell Acute Lymphoblastic Leukemia

Caused by the Deletion of Genes Encoding PU.1 and Spi-B

3.1 Introduction: .................................................................................................................. 21
3.2 Sorting of B cell populations for gene expression analysis: ........................................ 21
3.2 Improved correlations between the replicates within each population after batch effect removal ........................................................................................................................................ 26
3.3 Removing batch effect from the data set increases the number of significantly differentially expressed genes ........................................................................................................... 27
3.4 High fold changes in differentially expressed genes in ΔBP LEUK cells compared to ΔB cells ........................................................................................................................................... 27
3.5 44 genes known to be activated by PU.1 were reduced in our data set .......................... 30
3.6 Signaling pathways predicted to be the most affected by decreased gene expression ...... 36
3.7 Cell cycle regulation pathways predicted to be the most affected by increased gene expression ........................................................................................................................................... 36
3.8 Confirmation of target genes potentially involved in leukemia ....................................... 42

CHAPTER 4: The role of SPIB in ETV6-RUNX1 acute lymphoblastic leukemia

4.1 Introduction: .................................................................................................................. 45
4.2 SPIB was one of the top genes that were up regulated after knocking down the ETV6-RUNX1 fusion protein ............................................................................................................. 47
4.3 SPIB expression level was reduced in ETV6-RUNX1 pediatric patient samples compared to other subgroups of B cell acute lymphoblastic leukemia patient samples .......................... 49
4.4 Five predicted RUNX1 binding sites were located in transcriptionally active sites of the SPIB gene ........................................................................................................................................... 49
4.5 P3 is highly conserved through evolution ....................................................................... 54
4.6 P3 is highly enriched compared to the negative controls (NC, CRORFII) and the other RUNX1 predicted binding sites (P1, P2, P4, P5) ........................................................................................................................................... 54

CHAPTER 5: Discussion ......................................................................................................... 58

5.1 Gene expression analysis of ΔBP LEUK cells suggested increase in cell proliferation, and block in B cell differentiation caused by disruption of key pathways ........................................................................................................... 58
5.2 Reduced levels of SPIB in childhood ETV6-RUNX1 B cell acute lymphoblastic leukemia: .................................................................................................................................................. 62
5.3 The ETV6-RUNX1 fusion protein could repress SPIB transcription ............................... 63
5.4 Future directions: ........................................................................................................... 65
5.5 Summary and conclusions: ............................................................................................ 66

Bibliography ......................................................................................................................... 68

APPENDIX ............................................................................................................................ 74

Curriculum Vitae .................................................................................................................... 75
LIST OF FIGURES

Figure 1.1 stages of B cell development 4
Figure 3.1 Identification of control and leukemic B cell population 23
Figure 3.2 Overview of work flow for Affymetrix Mouse Exon 1.0 ST Arrays 25
Figure 3.3 Improved correlation between replicates for each population after removing the batch effect 28
Figure 3.4 Removing the batch effect from the dataset increases the number of significantly differentially expressed genes 29
Figure 3.5 Heat Maps of the top 100 genes differentially expressed in ΔBP LEUK cells compared to ΔB cells. 35
Figure 3.6 The identities of genes directly activated by PU.1 37
Figure 3.7 44 genes known to be activated by PU.1 were reduced in our dataset 38
Figure 3.8 The B cell receptor signaling pathway was predicted to be the most affected pathway from the decreased gene expression seen in ΔBP LEUK compared to ΔB cells 41
Fig 3.9 confirmation of target genes potentially involved in leukemia 43
Figure 3.10 “Volcano plot” of changes in gene expression in splenic ΔBP LEUK cells compared to DB cells from the microarray experiment. 44
Figure 4.1 hypothetical model of ETV6-RUNX1 function in leukemogenesis. 46
Figure 4.2: SPIB and PU.1 expression levels in ETV6-RUNX1 knocked down cells (E/R KD) compared to the control in AT-2 and REH cell lines. 48
Figure 4.3: SPIB expression levels in ETV6-RUNX1 cells compared to other sub types of B-ALL in pediatric patients 51
Figure 4.4: Cross-referenced RUNX1 predicted binding sites on SPIB and POLD1 gene to human ENCODE data. 53
Figure 4.5: sequence alignment of P3 RUNX1 site in intron one of the SPIB gene in five mammalian species 55
Figure 4.6: Representative ChIP 56
Figure 4.7: ChIP analysis. 57
Figure 5.1 hypothetical model of ETV6-RUNX1 function in leukemogenesis 67
LIST OF TABLES

Table 2.1 qPCR primer sequences and efficiencies 18

Table 2.2 ChIP primer sequences 20

Table 3.1 The number of replicates for each cell population and time processed 24

Table 3.2 The top 30 genes with decreased in expression in ΔBP LEUK cells compared to ΔB cells 31

Table 3.3 The top 30 genes with increased in expression in ΔBP LEUK cells compared to ΔB cells 33

Table 3.4 The top 20 pathways generated from the gene list that were decreased in expression in ΔBP LEUK cells compared to ΔB cells 39

Table 3.5 The top 20 pathways generated from the gene list that were increased in expression in ΔBP LEUK cells compared to ΔB cells 40

Table 4.1: Pediatric acute lymphoblastic patient subgroups 50
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSC</td>
<td>Hematopoietic stem cells</td>
</tr>
<tr>
<td>BCR</td>
<td>B cell receptor signaling</td>
</tr>
<tr>
<td>B-ALL</td>
<td>B cell acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>MPP</td>
<td>Multipotent progenitor cells</td>
</tr>
<tr>
<td>CMP</td>
<td>Common myeloid progenitor</td>
</tr>
<tr>
<td>ELP</td>
<td>Early lymphoid progenitor</td>
</tr>
<tr>
<td>IL-7R</td>
<td>Interleukin-7 receptor</td>
</tr>
<tr>
<td>CLP</td>
<td>Common lymphoid progenitor</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IgHC</td>
<td>Immunoglobulin heavy chain</td>
</tr>
<tr>
<td>IgLC</td>
<td>Immunoglobulin light chain</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin dependent kinase</td>
</tr>
<tr>
<td>SLC</td>
<td>Surrogate light chain</td>
</tr>
<tr>
<td>BLNK</td>
<td>B cell linker protein</td>
</tr>
<tr>
<td>Btk</td>
<td>Bruton’s tyrosine kinase</td>
</tr>
<tr>
<td>PU.1</td>
<td>Purine-Rich Box1</td>
</tr>
<tr>
<td>Sfpi1</td>
<td>SFFV proviral integration site 1</td>
</tr>
<tr>
<td>Spi-B</td>
<td>Spleen Virus Integration Site</td>
</tr>
<tr>
<td>LPS</td>
<td>lipopolysaccharide</td>
</tr>
<tr>
<td>RMA</td>
<td>Robust Multiarray Analysis</td>
</tr>
<tr>
<td>ComBat</td>
<td>Combating Batch Effects</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>KEGG</td>
<td>Kyoto Encyclopedia of Genes and Genomes</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>Lyn</td>
<td>Lck/yes-related novel</td>
</tr>
<tr>
<td>Vav2</td>
<td>Guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Lef1</td>
<td>Lymphoid enhancer binding factor 1</td>
</tr>
<tr>
<td>Tcf7l2</td>
<td>transcription factor 7 like 2</td>
</tr>
<tr>
<td>Lgr5</td>
<td>leucine rich repeat containing G protein coupled receptor 5</td>
</tr>
<tr>
<td>Ptprc</td>
<td>Protein tyrosine phosphatase, receptor type, C</td>
</tr>
<tr>
<td>Enpep</td>
<td>glutamyl aminopeptidase</td>
</tr>
<tr>
<td>BM</td>
<td>Bone marrow</td>
</tr>
<tr>
<td>PB</td>
<td>Peripheral blood</td>
</tr>
<tr>
<td>ENCODE</td>
<td>Encyclopedia of DNA Elements</td>
</tr>
</tbody>
</table>
1.1 Overview:

B cell development from hematopoietic stem cells (HSC) to mature B cells is a highly regulated process. This multi-step process is tightly controlled by the interaction of the extrinsic cytokines and B cell receptor signaling (BCR) pathways and intrinsic transcriptional programming [1]. Disruption of this balanced interaction could lead in most cases to a block in B cell development and death. However, in certain circumstances the disruption might reprogram the B cell fate into a new developmental fate. This reprogramming occurs in vivo in human B cell acute lymphoblastic leukemia (B-ALL). In general, the molecular players leading to these aberrant fates have key roles in the normal B cell development [1].

1.1.1 Stages of B cell development:

The multiple stages of B cell development that fall between HSC and mature B cells have been extensively studied [2, 3]. HSCs are identified by their cell surface markers; lineage negative (Lin<sup>−</sup>) c-kit<sup>+</sup> and Sca-1<sup>+</sup> [4]. HSCs are heterogeneous group of cells, which include long term HSC and short term HSC. The short term HSCs give rise to the multipotent progenitor cells (MPP) [5, 6]. MPP cells share the same phenotype of HSCs. In addition, they express Flt3 receptors. The MPP cells have the potential to give rise to two cell compartments, which represent two different lineages of cells; the common myeloid progenitor (CMP) and the early lymphoid progenitor (ELP) [7, 8]. As the name implies, CMP is the transitional compartments from which erythroid and myeloid cells are generated [7]. Lymphoid cells are generated from the ELP cells. These cells are characterized by
Lin\textsuperscript{neg} c-kit\textsuperscript{high} Sca-1\textsuperscript{pos} Flt3\textsuperscript{pos} IL-7R\textsuperscript{neg} [8]. Further differentiation of ELP cells to common lymphoid progenitor (CLP) cells is characterized by decreased expression of c-kit and Sca-1, the acquisition of additional cell surface markers that include the interleukin-7 receptor (IL-7R) and AA4.1 (CD93), and the increased expression of RAG1/2 [9, 10]. These markers are considered hallmarks of further maturation toward B-lymphocytes. CLPs are the first progenitor cells that specify B lineage development. CLP cells progeny consist of pre-pro-B, pro-B, pre-B, and finally, newly produced sIgM\textsuperscript{pos} B lymphocytes [2]. Pre-pro-B cells are characterized by Lin\textsuperscript{neg} CD45R\textsuperscript{pos} CD43\textsuperscript{pos} AA4.1\textsuperscript{pos} CD19\textsuperscript{neg} Ly6C\textsuperscript{neg} [11]. Pro-B cells are defined as CD45R\textsuperscript{pos} CD19\textsuperscript{pos} CD43\textsuperscript{pos} AA4.1\textsuperscript{pos}, and are considered the first progeny committed toward B cell development [8]. In addition to the cell surface markers, B cell stages are defined by the status of immunoglobulin (Ig) gene rearrangement [12, 13]. Immunoglobulin heavy chain (IgHC) gene rearrangements occur in CLP, pre-pro-B, and pro-B cells. If pro-B cells successfully rearrange their Igs, they mature to pre-B cells. In pre-B cells, the successful production of immunoglobulin light chain (IgLC) rearrangement results in the production of immature B cells (sIgM\textsuperscript{pos} B lymphocytes) that will migrate from the bone marrow to the spleen to complete the maturation process [2, 12].

1.1.2 Transcriptional regulation of B cell commitment:

Each stage of B cell development can be defined by the expression of key transcription factors [14, 15]. These transcription factors act in complex networks involving cross and auto regulation, which lead to the activation of specific target genes involved in B cell maturation [16]. The transcription factors PU.1, Ikaros, E2A, EBF, and PAX5 are essential in B cell specification and commitment [17-21]. PU.1 and Ikaros are required in the early stages of B cell development. They are both expressed in lymphoid primed multipotent progenitor cells and promote the expression of FLT3. FLT3 signaling acts with PU.1 to induce the expression
of IL-7R. IL-7 signaling acts to induce the transcription factor E2A in CLP. PU.1 and E2A, with IL-7 induce the expression of EBF, which is considered a critical event in specification of B cell development. CLP will give rise to the pro-B cell where the transcription factor PAX5 is expressed. PAX5 functions to promote B cell commitment by reinforcing B cell specific gene expression and repressing other lineage gene expression (Figure 1.1) [16].

1.1.3 Interleukin 7 (IL-7) and Interleukin 7 receptor (IL-7R) in B cell development:

IL-7 is a type 1 cytokine, which has a central role in early B cell development that includes commitment, survival, proliferation and maturation. IL-7 is mainly secreted by stromal cells in the lymphoid tissues [22-24]. The receptor for IL-7 is composed of two chains IL-7Rα chain and common γ chain (dimer receptor) [25].

Studies have provided evidence for the role of IL-7 in the commitment of the CLP to B cell lineage [26]. First, CLP without the presence of IL-7 cannot undergo B cell differentiation [26]. Second, as mentioned above, IL-7R signaling induces the expression of E2A. Then E2A with PU.1 and IL-7R induce the expression of EBF, which is an important transcription factor in B cell commitment [27]. Third, in the presence of IL-7, Pax5 −/− pro-B cells maintained their identity as early B cells, where without IL-7 these cells developed a macrophage phenotype [28].

In support of the notion that IL-7 is required for maturation and proliferation during early B cell development, studies with targeted deletions of genes encoding IL-7 or IL-7R in mice, showed a block in B cell differentiation at pro-B cell stage with only a few B cells maturing. In addition, pro-B cells stimulated with IL-7 undergo robust proliferation [24]. IL-7 signaling regulates many proteins involved in cell cycle regulation. N-myc and c-myc are expressed when pre-B cells are stimulated with IL-7. Enforced expression of c-myc enhances the activity of cyclin dependent kinase (CDK), which stimulates cell cycle progression [29, 30].
Figure 1.1 stages of B cell development

Sequential stages of early B cell differentiation from LMPP to Pre-B cells. Key transcription factors, growth factor receptors and cell surface markers of each stage are shown.
1.1.4 Pre-B cell receptor signaling in B cell development:

Pre-BCR signaling has an important role in proliferation and differentiation of the pre-B cells. Pre-BCR receptor expression serves as a checkpoint for successful immunoglobulin heavy chain gene rearrangement (IgHC). The successful recombination of the IgHC gene encodes for Ig\(\mu\) that associates with the germline-encoded surrogate light chain (SLC), and the signaling molecules Ig\(\alpha\) and Ig\(\beta\) to produce the pre-BCR receptor complex. Once the pre-BCR complex is on the cell surface, signaling through the receptor leads to robust proliferation [31, 32]. This proliferation is mediated through the activation of signaling molecules down stream of the pre-BCR receptor. The cytosolic kinase Syk is rapidly phosphorylated after the engagement of the pre-BCR receptor. Phosphorylation of Syk leads to the activation of the P13K pathway. P13K activates the adaptor protein BAM32 and other signaling proteins including PKB/AKT. These proteins stimulate cell proliferation and inhibit apoptosis [33-35]. In addition to promoting cell proliferation, Syk induces pre-B cell differentiation by activating the adaptor molecules B cell linker protein (BLNK) and Bruton’s tyrosine kinase (Btk) [36].

Once BLNK is activated, it provides docking sites for many signaling molecules such as Vav, Btk, Grb2, and PLC-\(\gamma\)2 [36]. Recent studies have shown the role of BLNK in transmitting differentiation signals and inhibiting proliferation signals. First, up-regulation of pre-BCR surface receptor expression and increased cell proliferation has been shown in Blnk deficient pre-B cells. In addition, forced expression of BLNK in these cells resulted in the down regulation of pre-BCR surface receptor and enhanced pre-B cell differentiation. Second, many studies have illustrated the negative effect of BLNK on IL-7R signaling through different mechanisms [37-39]. In summary, BLNK has a dual function in B cell development by promoting differentiation and down regulating proliferation.
Another signaling molecule with a potential role in B cell development is Btk. Studies have illustrated that Btk has an important role in the progression of pre-B cells to immature B cells. Also, Btk has a role in limiting pre-B cell proliferation, as Btk deficient pre-B cells exhibit increased proliferative response to IL-7 [40, 41].

In summary, B cell development is carried out through a balanced and tightly regulated interaction between proliferation and differentiation signaling pathways. Disruptions at any level (transcriptional, cytokine-cytokine receptor signaling, pre-BCR signaling) in B cell development can lead to B cell arrest and leukemia. Dysregulation of the proteins mentioned above by mutations, deletions, and repression has been reported in B cell acute lymphoblastic patients.

1.2 B cell acute lymphoblastic leukemia (B-ALL):

B cell acute lymphoblastic leukemia is a clonal malignant disease characterized by a block of B cell differentiation and the accumulation of these early B cells [42]. B-ALL represents 80% of the cases of childhood leukemia, and can be further divided into subgroups according to the associated genetic lesions [43]. These genetic lesions could be due to aneuploidy that is found in hyperdiploidy and hypodiploidy B-ALL subtypes. Also, genetic lesions could be due to chromosomal translocations that can lead to either deregulated expression of proto-oncogenes or the expression of a fusion genes with properties different from their wild-type counterparts [44]. Examples of chromosomal translocations in B-ALL subtypes include ETV6-RUNX1, BCR-ABL1, TCF3-PBX1, and many others. There is an accumulation of evidence that many of these chromosomal translocations occur prenatally, suggesting that these chromosomal translocations push the cells toward a pre-leukemic state. Cooperating mutations are required for the cells to become fully transformed [1, 45]. Recent studies of genome wide profiling of B-ALL have increased our understanding of the biology and the
molecular pathways affected in each subtype of B-ALL, and of the cooperating mutations occur in B-ALL [46-48]. First, microarray gene expression profiling of B-ALL has illustrated that known genetic lesion subtypes of B-ALL have distinct gene expression profiles that are different from one another. In addition, these studies led to the discovery of a novel group of ALL without known chromosomal translocations [48]. Second, studies with SNP arrays have identified novel altered genes in B-ALL. The two most important findings from the SNP array studies are: 1) in over two thirds of B-ALL cases, genes encoding transcriptional regulators involved in the regulation of B cell development were mutated one example is PU.1 [47, 49]; 2) the nature and the number of genetic lesions are different among B-ALL subtypes. For example, in ETV6-RUNX1 and BCR-ABL1 it has been reported that 6-8 genetic lesions are found per case. On the other hand, ALL cases with MLL rearrangements had much fewer numbers of genetic lesions [50].

In conclusion, each B-ALL subtype has a unique gene expression profile, which suggests that underlying biology and the affected pathways in each subtype are different among the subgroups. In addition, most of the genes affected by deletion, translocations and sequence mutations are genes encoding transcriptional regulators of B cell development. This may explain the block of B cell differentiation in B-ALL.

1.3 Purine-Rich Box1 (PU.1):

The transcription factor PU.1 belongs to the E26 transformation-specific (ETS) family. PU.1 was discovered in murine erythroleukemia, which was caused by the proviral integration of the spleen focus-forming virus SFFV [51]. PU.1 is encoded by the SFFV proviral integration site 1 (Sfpi1) in mice and SPI1 in humans [52]. It has three functional domains: the DNA binding domain (ETS domain), a PEST domain, and a transactivation domain. The ETS
domain recognizes the consensus core sequence “PU-box” GGAA or AGAA [53]. Studies of PU.1 have shown the critical role that PU.1 has in regulating the differentiation of different hematopoietic cells.

1.3.1 The Expression and Regulation of PU.1 in B Cell development:
PU.1 is expressed only in the hematopoietic system. PU.1 expression levels are a critical factor in determining the fate of differentiated hematopoietic cells. Studies have shown that PU.1 is expressed at low levels in early hematopoietic progenitor cells including CMP and CLP [54, 55]. It has been shown that PU.1 expression levels increase during myeloid differentiation, and decrease during B cell differentiation [54, 55]. Fine-tuning and tight regulation of PU.1 concentration is crucial for the development and function of B cells. Studies have shown that PU.1 concentration in macrophages is ~5 fold higher compared to B cells, and this is essential for the normal development of both cell types [56, 57].

1.3.2 The Role of PU.1 in B cell development and its target genes:
The importance of PU.1 in early B cell development has been demonstrated through gene knock out studies. Mice with homozygous null mutations of PU.1 have no B, T, or myeloid progenitor cells [58]. In addition, multipotential lymphoid-myeloid progenitors from the fetal liver of Sfpi1−/− mice failed to differentiate into pro-B cells in response to stromal cell contact and IL7 [59]. These findings were explained by the impaired expression of the IL-7R in fetal liver cells. Therefore, PU.1 regulates early B cell development by the regulation of IL-7R [21]. However, PU.1 does not have a major role in B cell development after the commitment to B cell lineage. This was shown by the conditional deletion of PU.1 in B cells, which had minimal functional defects and resulted in normal B cell development [60], suggesting that the function of PU.1 in B cells could be complemented by the related transcription factor
1.4 Spleen Virus Integration Site (Spi-B):

The transcription factor Spi-B also belongs to the ETS family. PU.1 and Spi-B share 67% homology. In addition, Spi-B binds to the same consensus core sequence “PU-box”. This suggests that Spi-B could regulate PU.1 target genes [61, 62].

1.4.1 The Expression and the Role of Spi-B in B Cell Development:

Spi-B expression in B cells starts at the pro-B stage of B cell development. Spi-B has an overlapping pattern of expression with PU.1 in B cells, and binds to the same consensus core sequence. This suggests that PU.1 and Spi-B are functionally redundant in B cells [62]. In contrast to Sfpi1−/− mice, Spi-B−/− mice have normal hematopoietic development of all cell lineages. However, Spi-B−/− B cells have functional defects in the response to antigenic stimulation through the BCR and the amount of antibody produced after immunization. In addition, these B cells proliferate poorly in vitro in response to BCR stimulation and in response to lipopolysaccharide (LPS) stimulation [63]. These results suggested that Spib−/− B cells are defective in one or more BCR signaling components, which could be directly regulated by Spi-B. In addition, Sfpi1+/− Spi-B−/− B cells exhibit more extensive defects in BCR signaling [64]. Therefore, PU.1 and Spi-B have an important role in regulating BCR signaling.

1.5 The role of PU.1 and Spi-B in B-ALL:

Our lab has provided direct evidence that PU.1 and Spi-B act as complementary tumor suppressor genes in B cells. This was done by generating conditional PU.1 knockout mice on a Spib−/− background. These mice developed B cell acute lymphoblastic leukemia with 100% incidence [65]. In addition, a recent study with transgenic mice expressing the human ETV6-
RUNX1 revealed that SPIB transcript levels were reduced [66]. Another study with two human leukemia cell lines, AT-2 and REH, which express ETV6-RUNX1, suggested that SPIB was one of the top genes that were most upregulated after knocking down the ETV6-RUNX1 fusion protein [67]. PU.1 mutations have been reported in relapsed B-ALL patients [49]. In conclusion, these studies clearly show that loss of PU.1 and SPIB could play a role in the leukemogenesis of B-ALL.

1.6 ETV6-RUNX1 acute lymphoblastic leukemia:

The chromosomal translocation t(12;21) fuses the ETV6 (TEL) gene to the RUNX1 (AML1) gene. This is the most common chromosomal alteration in pediatric B cell acute lymphoblastic leukemia, representing 25% of the cases [68]. ETV6-RUNX1 fusions can arise prenatally as an initiating event, which can contribute to the persistence of pre-leukemic clones. However, these persistent pre-leukemic clone cells cannot become fully transformed unless additional genetic mutations are acquired. The molecular mechanism underlying ETV6-RUNX1 function is to interfere with normal RUNX1 function. ETV6-RUNX1 represses RUNX1 target genes by the recruitment of nuclear core repres sor/histone deacetylase complexes [69]. RUNX1 is a transcription factor that has crucial regulatory function in the development of multiple blood cells. Conditional knock out of RUNX1 in the bone marrow has revealed that RUNX1 is required for the terminal differentiation of B cells [70]. RUNX1 regulates and physically interacts with key transcription factors in B cell development such as PU.1 and PAX5 [71, 72]. In addition to dysregulation of RUNX1 target genes, ETV6-RUNX1 is proposed to repress the expression of tumor suppressor genes and to increase the expression of anti-apoptotic genes [73]. In summary, ETV6-RUNX1 fusion protein plays a key role in transcriptional deregulation that might make the cells more persistent and susceptible to secondary mutations.
1.7 Hypothesis and objectives:

The chromosomal translocation t(12;21) results in the fusion of the ETV6 (TEL) gene to the RUNX1 (AML1) gene. This is the most common chromosomal alteration in pediatric B cell acute lymphoblastic leukemia, representing 25% of the cases. ETV6-RUNX1 represses RUNX1 target genes by the recruitment of nuclear core repressor/histone deacetylase complexes. However, little is known about the target genes of ETV6-RUNX1 that are involved in promoting leukemogenesis. A recent study with transgenic mice carrying the human ETV6-RUNX1 revealed reduced SPIB transcript levels. Another study with two human leukemia cell lines AT-2 and REH, which express ETV6-RUNX1, suggested that SPIB was one of the top genes that were up regulated after knocking down the ETV6-RUNX1 fusion protein. These studies lead to our hypothesis that ETV6-RUNX1 functions as an oncogenic driver by repressing SPIB transcription, leading to impaired B cell receptor signaling.

To test this hypothesis, the following aims were conducted:

1. To determine what target genes and pathways are affected by the loss of PU.1 and Spi-B in B cell acute lymphoblastic leukemia (Chapter 3).
2. To determine the expression levels of the SPIB gene before and after knocking down ETV6-RUNX1, and the molecular mechanism of how SPIB is regulated by ETV6-RUNX1 (Chapter 4).
CHAPTER 2: Materials and Methods

2.1 Mouse strains:
Mice used in this study were $Spib^{-/-}$ (ΔB mice) and $CD19^{+cre} Sfpi1^{lox/lox} Spib^{-/-}$ (ΔBP LEUK mice). Mice mating and breeding was done in concordance with an approved animal care protocol by the University of Western Ontario Council on animal care. All mice were on the C57BL/6 background. $CD19^{+cre}$ mice were mated to $Sfpi1^{lox/lox}$ mice to generate $CD19^{+cre} Sfpi1^{+/-}lox$ mice. Then, these mice were mated to $Spib^{-/-}$ mice to generate $CD19^{+cre} Sfpi1^{+/-}lox Spib^{-/-}$ mice. Lastly, these mice were intercrossed to generate $CD19^{+cre} Sfpi1^{lox/lox} Spib^{-/-}$ (ΔBP LEUK mice) [65].

2.2 Cell sorting:
Single cell suspensions from the spleen of mice were stained with phycoerythrin (PE)–conjugated anti-CD19 (1D3), and allophycocyanin-conjugated anti-B220 (RA3-6B2) (BD Pharmingen, Franklin Lakes, NJ, USA). Sorting of antibody-stained cells was done using the FACSAnia II system.

2.3 Affymetrix Mouse Exon Array hybridization:

2.3.1 RNA isolation and evaluation:
RNA extraction was done using RNA Bee RNA Isolation Agent (Tel-Test Friendswood, Texas). RNA concentration and purity were measured by the Nanodrop ND-1000 Spectrophotometer Thermo Fisher Scientific Inc. All RNA samples had high purity, with a 260/280 ratio of 1.8-2.1. RNA integrity was analyzed using Agilent 2100 Bioanalyzer Scans. The integrity of RNA was determined by two factors. First, the RNA integrity number (RIN),...
which can vary from 0 (degraded RNA) to 10 (intact RNA). Second, the ratio of 28s/18s ribosomal RNA, which is 2.0 for intact RNA. All samples had a RIN of (8-9.3) and a ratio of 28s/18s of (1.9-2.7).

2.3.2 RNA processing:
RNA processing for hybridization was done at the London Regional Genomic Center, according to the manufacture’s instruction (Affymetrix). Probes for the exon arrays were designed in the anti-sense configuration. Therefore, RNA had to be treated to produce a labeled single stranded sequence in the sense orientation before hybridization. In brief, RNA was converted to cDNA, followed by *in vitro* transcription to amplify the cDNA and produce cRNA. Then, a second cycle of cDNA synthesis was applied using the cRNA to produce single strand DNA in the sense orientation. This DNA is enzymatically fragmented and fluorescently labeled, then hybridized on Affymetrix Mouse Exon 1.0 ST Arrays. A total of eight arrays were hybridized at three different time points representing three batches of samples. Arrays were scanned using Affymetrix GeneChip Scanner 3000 7G, and converted to processed image. Then .CEL files were generated, containing the raw signal intensities (one signal intensity value per probe).

2.4 Data pre-processing and batch effect removal:

2.4.1 Data pre-processing:
.CEL files were exported to Partek Genomic Suite software (Partek Inc., St. Louis, MO, USA), where all the steps of data pre-processing and data analysis were done. Data pre-processing for Affymetrix arrays included background correction, quintile normalization, log2 transformation, and summarization. All previous steps were done using the Robust Multiarray Analysis (RMA) technique, which is implanted in Partek software. First,
background correction for all exon arrays was done by predicting the hybridization properties of a given probe based on the GC-content, as probes with higher GC-content will hybridize better. Second, a quintile normalization algorithm was applied to all exon arrays at the probe level, which makes the assumption that all arrays have an equal intensity distribution. Third, all signal intensity values for the arrays were transformed to log2 scale, and all subsequent analysis was done on these log-transformed values. Log transformation makes the intensity distribution symmetrical and bell shaped. Finally, summarization was done to convert multiple probe measurements into a single value of a probe sets. Probe sets represent values at the exon level. Additional summarization was applied on these probe sets values to convert them in to single value Meta-probe sets. These Meta-probe sets represent values at the gene level. All summarization steps were done using the median polish method. After that, intensity values of probe sets for various controls were checked. All the arrays had their control values within the expected range. Scan date was one of the major sources of variation that contributed to the variance between the samples. ComBat algorithm was used to remove the scan date factor of variation [74].

2.4.2 Batch effect removal:

In order to remove batch effect (scan date) between the batches, an Empirical Bayes method called Combating Batch Effects (ComBat) [74] was used. All work with removing the batch effect was done using the R script. In brief, R script was downloaded on the computer used in the London regional genomic center. Then a working directory was created. In the working directory, the ComBat.R script file, which has the ComBat code, was saved. Then, the gene expression file and the information file were created and saved in the working directory. After that, script codes were typed to run ComBat in R. codes that can be found in (http://www.bu.edu/jlab/wp-assets/ComBat/Usage.html). An output excel file with adjusted
data was generated. Finally, the adjusted data were copied and pasted in a new spreadsheet in Partek for data analysis [74].

2.5 Data analysis:

2.5.1 Differentially expressed genes:

To identify differentially expressed genes between the three different populations, a one-way analysis of variance (ANOVA) was used. Then two lists were created from the ANOVA spreadsheet. First, genes increased in ΔBP LEUK cells compared to ΔB cells with > 1.5 fold change and with p < 0.05 was generated. Second, genes decreased in ΔBP LEUK cells compared to ΔB cells with > - 1.5 fold change and with p < 0.05 was generated.

2.5.2 Partek pathway:

The increased and decreased lists of genes were assessed for over representation analysis using Fisher exact test to identify significantly enriched biological pathways. These biological pathways were generated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

2.5.3 Heat map generation:

Heat maps were generated using Hierarchical Clustering Explorer. Log2 transformed values were used for the top 100 genes that were up/down regulated in ΔBP LEUK cells compared to ΔB cells.

2.6 Real-Time quantitative PCR (qPCR):

RNA extraction was done using RNA Bee RNA Isolation Agent (Tel-Test Friendswood, Texas). RNA was reverse transcribed into cDNA using the iScript DNA synthesis kit (Bio-Rad, Hercules, CA). qPCR was done using the iQ SYBR Green Supermix Kit (Bio-Rad). A
Rotor-Gene 6000 instrument was used to run the qPCR reactions. Beta-2 microglobulin (B2m) was chosen as a reference gene to normalize the data to correct for differences in quantities of cDNA used as template. B2m had consistent CT values between samples, which made it an optimal reference gene. Relative expression software tool (REST), which put in to consideration the differences in the efficiency between the primers, was used to measure the relative expression levels between the samples [75]. Each result was performed in triplicate and significance was determined using the Students t-test.

Efficiency for each primer was calculated using two fold serial dilution standard curves. Also, primer specificity was tested by gel electrophoresis. Primer sequences and efficiency are listed in Table 2.1.

2.7 Bioinformatics analysis:

2.7.1 Matinspector:

Matinspector [76] was used to identify predicted RUNX1 binding sites on the SPIB and POLDI locus. 0.8 was the matrix score cutoff for predicted RUNX1binding sites to be chosen for further analysis.

2.7.2 Macvector:

Multiple alignments of the intron one of the SPIB for five mammalian species (rat, mouse, human, cow, dog) were done using the CLUSTALW method in Macvector (Accelrys, San Diego, CA). Sequences were acquired from the Ensemble.org database.

2.7.3 UCSC Genome Browser website:

UCSC Genome Browser website was used to visualize the ENCODE data (DNase I hypersensitivity sites and histone acetylation marker (H3K27AC)), and to cross reference the
predicted RUNX1 binding sites to transcriptionally active sites in B cell line GM12878 [77].

2.8 Cell culture:

The REH (B –ALL) cell line was purchased from the American Type Culture Collection (ATCC), Manassas, USA. These cells were cultured in complete medium containing Roswell Park Memorial Institute (RPMI), 10% fetal bovine serum, 100 U/mL penicillin/streptomycin, 2mM L-glutamine, and $5 \times 10^{-5}$ M 2-mercaptoethanol. Cell cultures were maintained at 37 °C in 5% CO₂ atmosphere.
Table 2.1 qPCR primer sequences and efficiencies:

<table>
<thead>
<tr>
<th>GENE (mouse)</th>
<th>Primer sequence</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2M F</td>
<td>5' TGGCTCACACTGAATTCAACCCCA'3</td>
<td>0.97</td>
</tr>
<tr>
<td>B2M R</td>
<td>5' TCTCGATCCCAGTAGACGGTCTTGG'3</td>
<td></td>
</tr>
<tr>
<td>Tcf7l2 F</td>
<td>5' CCGCCCGAACCCTCCATTTTCA'3</td>
<td>0.98</td>
</tr>
<tr>
<td>Tcf7l2 R</td>
<td>5' TGTCTCTGGAGGCTTCTGCTTG'3</td>
<td></td>
</tr>
<tr>
<td>Lgr5 F</td>
<td>5' TCAAGTTCAAGATGAGCGGACCT 3'</td>
<td>0.98</td>
</tr>
<tr>
<td>Lgr5 R</td>
<td>5' AGCCAGCTACAAATAGGTGCTCA 3'</td>
<td></td>
</tr>
<tr>
<td>lef1 F</td>
<td>5' TGGCATCCCTCATCCAGCTATTGT 3'</td>
<td>1</td>
</tr>
<tr>
<td>lef1 R</td>
<td>5' TGAGGCTTCACGTGATTAGTCA 3'</td>
<td></td>
</tr>
<tr>
<td>Blnk F</td>
<td>5'TCCAAGTCTCTTGGCCTGCCC'3</td>
<td>1.05</td>
</tr>
<tr>
<td>Blnk R</td>
<td>5'TGCATTCTGGAGGAGGAC'3</td>
<td></td>
</tr>
<tr>
<td>Btk F</td>
<td>5' ACAGATTCCCGAGGAGGTGAGG3'</td>
<td>1.05</td>
</tr>
<tr>
<td>Btk R</td>
<td>5' GGTCCTTCATCATATACAACCTGGAATGG3'</td>
<td></td>
</tr>
<tr>
<td>Enpep F</td>
<td>5' ACCAGGTCTGTGAGCAGG3'</td>
<td>0.9</td>
</tr>
<tr>
<td>Enpep R</td>
<td>5' CAGACATCTTGCTGTAGGAGG3'</td>
<td></td>
</tr>
<tr>
<td>Il7ra F</td>
<td>5' AGCAAGGGGTTGAAAGCAACTCG3'</td>
<td>1.05</td>
</tr>
<tr>
<td>Il7ra R</td>
<td>5' TCAGACTTGGATTTCATACATGTTTGG3'</td>
<td></td>
</tr>
<tr>
<td>Ptprc F</td>
<td>5' AGAAACGCTAAGCTTAGTTGTG3'</td>
<td>1.07</td>
</tr>
<tr>
<td>Ptprc R</td>
<td>5' TGGGGTTTAGATGACACTCAGG3'</td>
<td></td>
</tr>
<tr>
<td>Lyn F</td>
<td>5' CAGCAAAAGCCAGCTTCTGA 3'</td>
<td>0.93</td>
</tr>
<tr>
<td>Lyn R</td>
<td>5' ACAAGTCTCTGGTAGTGC'3</td>
<td></td>
</tr>
<tr>
<td>Vav 2 F</td>
<td>5' GCAAAGGGATCAGGCCATTTTC 3'</td>
<td>0.95</td>
</tr>
<tr>
<td>Vav 2 R</td>
<td>5' ATCTTCACACGGGACACATG3'</td>
<td></td>
</tr>
</tbody>
</table>
2.9 Chromatin immunoprecipitation (ChIP):

REH cells were cross-linked with formaldehyde for 10 minutes at room temperature. Cross-linked cells were neutralized with 125 mM glycine, and washed with cold phosphate buffered saline. Fixed cells were lysed with buffer containing protease inhibitors (50 mM Tris-HCl [pH 8.1], 10 mM EDTA, 1% SDS). Samples were sonicated using a Biorupture (Diagenode, Denville, USA) to an average DNA length of 200-1000 bp. Sonicated chromatin was incubated with rabbit anti-RUNX1, or control rabbit anti-IgG (ChIP grade, Abcam), which were coupled with protein G DynaBeads overnight at 4 °C. These beads were collected and washed five times with different washing buffers. Beads were washed once with low-salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-Cl, 150 mM NaCl). Then, beads were washed once with high-salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, 500 mM NaCl). After that, beads washed with LiCl buffer (0.25 M LiCl, 1% Nonidet P-40, 1% Na-deoxycholate, 1 mM EDTA, 10 mM Tris-Cl). Finally, beads were washed twice with Tris-EDTA buffer at pH 8. After washing, antibody, beads, chromatin complex were eluted with elution buffer (1% SDS, 0.1 M NaHCO3). Eluted solution was incubated at 65 °C overnight to reverse the cross-linking. Then, DNA was purified using the Qia-Quick PCR purification Kit (Qiagen). qPCR method was used to measure the enrichment of the immunoprecipitated DNA using specific primers designed to amplify the predicted RUNX1 binding sites (Table 2.2). Fold enrichment and percent input were calculated.
<table>
<thead>
<tr>
<th>GENE</th>
<th>RUNX1 binding position</th>
<th>primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIB</td>
<td><strong>P-2</strong> intron1</td>
<td>F 5' CCACCTGCACCTGCCCTC 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' TGGACCCCCCTCCCTCCAC 3'</td>
</tr>
<tr>
<td></td>
<td><strong>P-3</strong> intron1 conserved</td>
<td>F 5' GAGTCCGGTGAATGTGGTG 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' CTGGAGGGGAGAGACACAG 3'</td>
</tr>
<tr>
<td></td>
<td><strong>P-4</strong> intron 2</td>
<td>F 5' CGGAATACTATAACACACCCTTG 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' GCATAAACAGGGGCTTTTTCAA 3'</td>
</tr>
<tr>
<td></td>
<td><strong>P-5</strong> 3' down stream</td>
<td>F 5' AAAACAATTAGCCGGAGCGTGG 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' ACAGCTCACTACAGCTCCTCA 3'</td>
</tr>
<tr>
<td>POLD1</td>
<td><strong>P-1</strong> intron20</td>
<td>F 5' TGGTGAGCAGTGAAAGAGTC 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' AGACCAGGATGACTCTTGGGG 3'</td>
</tr>
<tr>
<td>C4ORFII</td>
<td>Negative control</td>
<td>F 5' GTATTCAGCCAGCCTTTCTTTG 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' ACACAGCTTATCTCAAGGTGACA 3'</td>
</tr>
<tr>
<td>SPIB</td>
<td><strong>NC</strong> intron5 8050 to 8198</td>
<td>F 5' AGGCATGATGGTGCACTCTGTG 3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R 5' TTTTTCAAGGGAGGTCTCGTT 3'</td>
</tr>
</tbody>
</table>
CHAPTER 3: Gene Expression Analysis of B cell Acute Lymphoblastic Leukemia

Caused by the Deletion of Genes Encoding PU.1 and Spi-B.

3.1 Introduction:

Our lab generated conditional PU.1 knockout mice on a Spib\textsuperscript{-/-} background, where the \textit{Sfpi1}\textsuperscript{lox} allele is excised under the control of the B cell specific \textit{CD19} locus (\textit{CD19}\textsuperscript{+/cre} \textit{Sfpi1}\textsuperscript{lox/lox} \textit{Spib}\textsuperscript{-/-} mice) henceforth called ΔBP LEUK mice. These mice developed B cell acute lymphoblastic leukemia (ALL) with 100\% incidence. Characterization of these B-ALL cells revealed that 100\% express BP-1 and CD93, which are markers of immature B cells. These results suggested that PU.1 and Spi-B are important in normal B cell differentiation [65]. To understand how the loss of PU.1 and Spi-B led to B-ALL, microarray gene expression profiling was used.

Microarray technology has the advantage of measuring the transcript levels of thousands of genes in one assay. It helps the scientist to determine which genes are associated with certain diseases and/or treatment, by comparing diseased/treated and control expression profiles [74]. In this study, Affymetrix mouse Exon Arrays were used to gain extensive insight of target genes and pathways affected by the loss of PU.1 and Spi-B in B-ALL cells compared to the control cells.

3.2 Sorting of B cell populations for gene expression analysis:

Two cell surface markers CD19 and B220 were selected to isolate the different B cell populations. CD19 is a known B cell specific marker, which is expressed once the lymphocyte progenitor is committed to B cell specification. It continues to be expressed on
mature B cells [78]. The second marker B220 is a known PU.1 target gene [79]. Absence of PU.1 expression will block the expression of B220, which serves as a good indicator to identify the leukemic cells knocked out for PU.1.

The three B cell populations were sorted including the control B cells, B-ALL cells, and B cells from ΔBP LEUK mice. Control B cells were sorted from spleens of mice lacking Spi-B in the germ line (CD19	exSfpi1	barSplib
) henceforth called ΔB mice. Flow cytometric analysis determined that 56% of the ΔB splenocytes were CD19*B220+ (Figure 3.1 left panel). Second, B-ALL cells were sorted from the spleens of ΔBP LEUK mice. These B-ALL cells were CD19*B220− (ΔBP LEUK cells), and represented 41% of the whole spleen cells (Figure 3.1 right panel). Lastly, we sorted an interesting small population of CD19*B220+ (ΔBP cells) from the spleens of ΔBP LEUK mice, which consisted of 8% of the total splenocytes (Figure 3.1 left panel). Eight samples were used (ΔB cells n = three, ΔBP LEUK cells n = 3 ΔBP cells n = 2). These replicates were processed in three time points, representing three batches/groups of samples (Table 3.1). RNA was extracted from these cells using RNA Bee RNA Isolation Agent and processed according to the manufactures guidelines (Affymetrix, Inc., Santa Clara, CA). The processed RNA was hybridized to the Affymetrix mouse Exon Arrays. Then, arrays were scanned using Affymetrix® Scanner 3000 7G (Figure 3.2).

The raw signal intensities generated from the microarray experiments need to be preprocessed before analysis. The purpose of the preprocessing step is to remove most of the technical sources of variation [80]. These systemic variations arise due to
Figure 3.1 Identification of control and leukemic B cell population

Three B cell populations were sorted based on the expression of CD19 and B220 surface markers. A representative flow cytometric panel is shown. The left panel represents sorted CD19⁺B220⁺ splenocytes from the control mice (ΔB). The right panel represents sorted splenocytes from leukemic mice (ΔBP). CD19⁺B220⁻ (ΔBP LEUK) and CD19⁺B220⁺ (ΔBP cells) were sorted from ΔBP mice.
Table 3.1 The number of replicates for each cell population and time processed

<table>
<thead>
<tr>
<th>Time point</th>
<th>Population</th>
<th>Mouse number</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 2011</td>
<td>ΔB cells</td>
<td>667</td>
<td>Spib&lt;sup&gt;-/-&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>ΔBP cells</td>
<td>669</td>
<td>CD19&lt;sup&gt;+&lt;/sup&gt;cre&lt;sup&gt;+&lt;/sup&gt;Sfpi1&lt;sup&gt;lox/lox&lt;/sup&gt; Spib&lt;sup&gt;-/-&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>ΔBP LEUK cells</td>
<td>668</td>
<td></td>
</tr>
<tr>
<td>June 2011</td>
<td>ΔBP cells</td>
<td>760</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ΔB cells</td>
<td>761</td>
<td>Spib&lt;sup&gt;-/-&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>ΔB cells</td>
<td>762</td>
<td></td>
</tr>
<tr>
<td>July 2011</td>
<td>ΔBP LEUK cells</td>
<td>782</td>
<td>CD19&lt;sup&gt;+&lt;/sup&gt;cre&lt;sup&gt;+&lt;/sup&gt;Sfpi1&lt;sup&gt;lox/lox&lt;/sup&gt; Spib&lt;sup&gt;-/-&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>ΔBP LEUK cells</td>
<td>783</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.2 Overview of work flow for Affymetrix Mouse Exon 1.0 ST Arrays
experiments being performed at different times, on different chip lots, with different technicians, or even with changing atmospheric ozone levels. Careful assessment and removal of systemic variations is important for reducing bias, which can be strong enough to mask true biological differences [81]. The first step taken to remove technical variation is background correction (e.g. caused by incomplete washing). Second step is normalization, which makes samples comparable to each other by homogenization of the signal intensity distribution (Figure 3.2) [80]. However, these normalization methods are not designed to remove systemic variations between two or more different batch/group of samples. As we had three batches of samples, an additional approach was needed for removing systemic variations between batches/groups [81].

3.2 Improved correlations between the replicates within each population after batch effect removal

The term ‘batch’ refers to microarray experiments processed simultaneously at one site and using the same platform (microarray chip). The term ‘batch effect’ will be used to describe systematic differences between the measurements of different batches [81].

To assess and adjust our data for batch effect, an Empirical Bayes Method called ComBat was used. This method is very robust in removing batch effect in small size batches, as compared to other methods that require at least twenty-five samples in each batch [74]. Principal Component Analysis (PCA) was used to evaluate the correlation between the replicates within each population before and after adjusting the data sets for batch effect. Results of PCA analysis are visualized in a three-dimensional plot. Eight dots were observed,
with each dot representing the gene expression of one sample. Before the removal of batch effect, replicates of each cell population were scattered (not clustered) indicating different gene expression profiles within the same population (Figure 3.3A). These differences in the gene expression profile within the same population were due to the batch effect, which confounded the signal values of the expression profiles. After the removal of batch effect, the replicates of each cell population were grouped together indicating similar gene expression profiles (Figure 3.3B). These results revealed the degree of batch effect within our dataset, and suggested that ComBat was an effective tool for removing batch effect.

3.3 Removing batch effect from the data set increases the number of significantly differentially expressed genes

To determine changes in gene expression, a one-way ANOVA test was used to generate two lists; increased and decreased in ΔBP LEUK cells compared to ΔB cells. After filtering genes with more than 1.5 fold change and \( P \geq 0.05 \), there were 2908 genes down-regulated in ΔBP LEUK cells after removal of batch effect compared to only 1161 genes prior (Figure 3.4A). Also, there were 3477 genes up-regulated in ΔBP LEUK cells after removal of batch effect compared to only 1636 genes prior (Figure 3.4B). This data illustrates that after removing batch effect from our data set, the total number of significantly altered genes were increased by almost double.
Figure 3.3 Improved correlation between replicates for each population after removing the batch effect

A. Before batch effect removal. ΔBP LEUK cell population (red dots) were far from each other, and ΔB cell population (green dots) are widely separated indicating variation in gene expression between the replicates of each population. B. After batch effect removal. Removing batch effect grouped ΔBP LEUK cell populations (red dots), and ΔB cell populations (green dots) together indicating similar gene expression within each population as expected.
Figure 3.4 Removing the batch effect from the dataset increases the number of significantly differentially expressed genes

A Venn diagram was used to illustrate the total number of significantly differentially expressed genes before (green) and after (light red) the batch effect removal in two lists. A. List of genes with decreased expression in ΔBP LEUK cells compared to ΔB cells. It was observed that after batch effect removal, the total number of significantly altered genes is 2908 compared to 1161 genes before batch effect removal. B. List of genes with increased expression in ΔBP LEUK cells compared to ΔB cells. It was observed that after batch effect removal, the total number of significantly altered genes is 3477 compared to 1636 genes before batch effect removal.
3.4 High fold changes in differentially expressed genes in ΔBP LEUK cells compared to ΔB cells

The median of the top 100 genes that were decreased in expression in ΔBP LEUK cells compared to ΔB cells was -56 fold (Table 3.2 only showing the top 30 genes). The median of the top 100 genes that were increased in expression in ΔBP LEUK cells compared to ΔB cells was 27 fold (Table 3.3 only showing the top 30 genes). These are large changes and may reflect the fact that PU.1 is a key transcription factor in B cell development. Heat maps were generated from the top 100 genes increased and decreased in expression in ΔBP LEUK cells compared to ΔB cells using the signal intensity value for each replicate. Red and blue colors represent an increase and a decrease in the signal intensity respectively. In the heat map of the top 100 genes decreased in ΔBP LEUK, the ΔBP LEUK population columns were colored blue compared to ΔB and ΔBP populations which were colored red (Figure 3.5A). In the heat map of the top 100 genes increased in ΔBP LEUK cells compared to ΔB cells, an interesting trend was observed. From the right to the left, there was a gradual increase in the intensity from ΔB cell population (blue color) to ΔBP LEUK (red color) cell population, which suggested that ΔBP cell population was at an intermediate stage between the two (Figure 3.5B).
Table 3.2 The top 30 genes with decreased expression in ΔBP LEUK cells compared to ΔB cells

<table>
<thead>
<tr>
<th>Transcript ID</th>
<th>gene_assignment</th>
<th>Gene Symbol</th>
<th>p-value</th>
<th>Fold-Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>6868091</td>
<td>NM_029499 // Ms4a4c // membrane-spanning 4-domains, subfamily A, member 4C // 19</td>
<td>Ms4a4c</td>
<td>8.02E-08</td>
<td>-313.707</td>
</tr>
<tr>
<td>6932118</td>
<td>NM_019992 // Stap1 // signal transducing adaptor family member 1 // 5 E1l5 // 56</td>
<td>Stap1</td>
<td>8.09E-05</td>
<td>-201.617</td>
</tr>
<tr>
<td>6806245</td>
<td>NM_010745 // Ly86 // lymphocyte antigen 86 // 13 A3.3l13 // 17084 // NM_198610</td>
<td>Ly86</td>
<td>2.18E-07</td>
<td>-190.898</td>
</tr>
<tr>
<td>6850017</td>
<td>NM_010388 // H2-DMb2 // histocompatibility 2, class II, locus Mb2 // 17 B11l7 18</td>
<td>H2-DMb2</td>
<td>1.18E-05</td>
<td>-167.254</td>
</tr>
<tr>
<td>6798285</td>
<td>NM_178911 // Pld4 // phospholipase D family, member 4 // 12 Fl1l2 // 104759 ///</td>
<td>Pld4</td>
<td>1.79E-06</td>
<td>-160.007</td>
</tr>
<tr>
<td>6827944</td>
<td>NM_183031 // Gpr183 // G protein-coupled receptor 183 // 14 E5l14 // 321019 ///</td>
<td>Gpr183</td>
<td>8.06E-09</td>
<td>-155.01</td>
</tr>
<tr>
<td>6938678</td>
<td>NM_030682 // Tlr1 // toll-like receptor 1 // 5 C3.1l5 37.0 cM // 21897 /// ENSMU</td>
<td>Tlr1</td>
<td>1.35E-11</td>
<td>-153.518</td>
</tr>
<tr>
<td>6855018</td>
<td>NM_010378 // H2-Aa // histocompatibility 2, class II antigen A, alpha // 17 B111</td>
<td>H2-Aa</td>
<td>4.28E-07</td>
<td>-153.485</td>
</tr>
<tr>
<td>6765460</td>
<td>NM_007758 // Cr2 // complement receptor 2 // 1 H6l1 106.6 cM // 12902 /// ENSMUS</td>
<td>Cr2</td>
<td>4.61E-10</td>
<td>-152.625</td>
</tr>
<tr>
<td>6930383</td>
<td>NM_007646 // Cd38 // CD38 antigen // 5 B35 28.0 cM // 12494 /// ENSMUST00000030</td>
<td>Cd38</td>
<td>1.02E-05</td>
<td>-152.562</td>
</tr>
<tr>
<td>6890068</td>
<td>NM_011246 // Rasgrp1 // RAS guanyl releasing protein 1 // 2 E5l2 65.0 cM // 1941</td>
<td>Rasgrp1</td>
<td>8.99E-08</td>
<td>-145.315</td>
</tr>
<tr>
<td>6980089</td>
<td>NM_013517 // Fcer2a // Fc receptor, IgE, low affinity II, alpha polypeptide // 8</td>
<td>Fcer2a</td>
<td>1.28E-11</td>
<td>-143.523</td>
</tr>
<tr>
<td>6848579</td>
<td>NM_009835 // Ccr6 // chemokine (C-C motif) receptor 6 // 17 A11l7 // 12458 /// N</td>
<td>Ccr6</td>
<td>6.93E-11</td>
<td>-139.087</td>
</tr>
<tr>
<td>6850012</td>
<td>NM_010386 // H2-DMa // histocompatibility 2, class II, locus DMa // 17 B11l7 18</td>
<td>H2-DMa</td>
<td>3.02E-06</td>
<td>-133.578</td>
</tr>
<tr>
<td>Chromosome</td>
<td>Gene ID</td>
<td>Gene Name</td>
<td>Description</td>
<td>Expression (log2²FC)</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>-----------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>1</td>
<td>NM_133809 // Kmo // kynurenine 3-monooxygenase (kynurenine 3-hydroxylase) // 1 H</td>
<td>Kmo</td>
<td>8.37E-11</td>
<td>-127.329</td>
</tr>
<tr>
<td>19</td>
<td>NM_001013779 // Aim2 // absent in melanoma 2 // 1 H31 // 383619 // ENSMUST0000</td>
<td>Aim2</td>
<td>2.01E-05</td>
<td>-119.605</td>
</tr>
<tr>
<td>11</td>
<td>NM_019549 // Plek // pleckstrin // 11 A211 6.5 cM // 56193 // ENSMUST00001028</td>
<td>Plek</td>
<td>6.85E-07</td>
<td>-114.815</td>
</tr>
<tr>
<td>11</td>
<td>NM_001111136 // Ptprc // protein tyrosine phosphatase, receptor type, C // 1 E41</td>
<td>Ptprc</td>
<td>5.13E-07</td>
<td>-108.988</td>
</tr>
<tr>
<td>17</td>
<td>NM_008052 // Dtx1 // deltex 1 homolog (Drosophila) // 5 F15 65.0 cM // 14357 //</td>
<td>Dtx1</td>
<td>8.65E-09</td>
<td>-104.755</td>
</tr>
<tr>
<td>11</td>
<td>NM_021887 // II21r // interleukin 21 receptor // 7</td>
<td>7 F4 // 60504 // ENSMUST0000</td>
<td>II21r</td>
<td>1.35E-08</td>
</tr>
<tr>
<td>1</td>
<td>NM_007807 // Cybb // cytochrome b-245, beta polypeptide // X A1.11X // 13058 //</td>
<td>Cybb</td>
<td>6.82E-05</td>
<td>-94.7675</td>
</tr>
<tr>
<td>1</td>
<td>NM_011487 // Stat4 // signal transducer and activator of transcription 4 // 1 C1</td>
<td>Stat4</td>
<td>2.59E-08</td>
<td>-94.7443</td>
</tr>
<tr>
<td>14</td>
<td>NM_030098 // Rnase6 // ribonuclease, RNase A family, 6 // 14 C1114 // 78416 //</td>
<td>Rnase6</td>
<td>8.58E-06</td>
<td>-93.5864</td>
</tr>
<tr>
<td>14</td>
<td>NM_145968 // Tagap // T-cell activation Rho GTPase-activating protein // 17 A111</td>
<td>Tagap</td>
<td>2.50E-08</td>
<td>-93.0906</td>
</tr>
<tr>
<td>17</td>
<td>NM_182806 // Gpr18 // G protein-coupled receptor 18 // 14 E5114 68.0 cM // 11016</td>
<td>Gpr18</td>
<td>3.47E-09</td>
<td>-89.0148</td>
</tr>
<tr>
<td>10</td>
<td>NM_174990 // Gimap4 // GTPase, IMAP family member 4 // 6 B2.3l6 // 107526 // NM</td>
<td>Gimap4</td>
<td>1.69E-06</td>
<td>-86.3424</td>
</tr>
<tr>
<td>16</td>
<td>NM_007575 // Ciita // class II transactivator // 16 A1116 // 12265 // ENSMUST00</td>
<td>Ciita</td>
<td>2.77E-06</td>
<td>-82.8095</td>
</tr>
</tbody>
</table>
Table 3.3 The top 30 genes with increased expression in ΔBP LEUK cells compared to ΔB cells

<table>
<thead>
<tr>
<th>Transcript ID</th>
<th>gene_assignment</th>
<th>Gene Symbol</th>
<th>p-value</th>
<th>Fold-Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>6909278</td>
<td>NM_007934 // Enpep // glutamyl aminopeptidase // 3 G3I3 65.2 cM // 13809 // ENS</td>
<td>Enpep</td>
<td>0.000531666</td>
<td>149.683</td>
</tr>
<tr>
<td>6776185</td>
<td>NM_007706 // Socs2 // suppressor of cytokine signaling 2 // 10 C2I10 52.0 cM // ENSMUST0000066882</td>
<td>Socs2</td>
<td>6.43E-05</td>
<td>145.08</td>
</tr>
<tr>
<td>6905321</td>
<td>NM_019410 // Pfn2 // profilin 2 // 3 D3 29.3 cM // 18645 // ENSMUST0000066882</td>
<td>Pfn2</td>
<td>2.30E-05</td>
<td>74.9968</td>
</tr>
<tr>
<td>6812386</td>
<td>NM_028784 // F13a1 // coagulation factor XIII, A1 subunit // 13 A3.3I3 // 74145</td>
<td>F13a1</td>
<td>5.32E-06</td>
<td>62.5917</td>
</tr>
<tr>
<td>6970840</td>
<td>NM_172476 // Tmc7 // transmembrane channel-like gene family 7 // 7 F2I7 // 20976</td>
<td>Tmc7</td>
<td>1.00E-05</td>
<td>59.7736</td>
</tr>
<tr>
<td>6874057</td>
<td>NM_010279 // Gfra1 // glial cell line derived neurotrophic factor family receptor</td>
<td>Gfra1</td>
<td>6.16E-05</td>
<td>57.3056</td>
</tr>
<tr>
<td>6883110</td>
<td>NM_001012704 // Wfdc13 // WAP four-disulfide core domain 13 // 2H3I2 // 408190 //</td>
<td>Wfdc13</td>
<td>0.000842222</td>
<td>46.8146</td>
</tr>
<tr>
<td>6777190</td>
<td>NM_010195 // Lgr5 // leucine rich repeat containing G protein coupled receptor 5</td>
<td>Lgr5</td>
<td>0.000190008</td>
<td>46.3768</td>
</tr>
<tr>
<td>6905746</td>
<td>NM_016753 // Lxn // latexin // 3 E1I3 31.6 cM // 17035 // ENSMUST0000058981 //</td>
<td>Lxn</td>
<td>7.57E-05</td>
<td>45.9732</td>
</tr>
<tr>
<td>6938217</td>
<td>NM_172710 // Sel1I3 // sel-1 suppressor of lin-12-like 3 (C. elegans) // 5 C1I5</td>
<td>Sel1I3</td>
<td>1.43E-05</td>
<td>44.3316</td>
</tr>
<tr>
<td>7014110</td>
<td>NM_023270 // Rnf128 // ring finger protein 128 // X F1IX // 66889 // ENSMUST0000058981 //</td>
<td>Rnf128</td>
<td>0.000142421</td>
<td>42.1642</td>
</tr>
<tr>
<td>6768014</td>
<td>NM_010288 // Gja1 // gap junction protein, alpha 1 // 10 B4I10 29.0 cM // 14609</td>
<td>Gja1</td>
<td>7.95E-06</td>
<td>41.822</td>
</tr>
<tr>
<td>6901316</td>
<td>NM_130450 // Elovl6 // ELOVL family member 6, elongation of long chain fatty aci</td>
<td>Elovl6</td>
<td>1.77E-05</td>
<td>41.5706</td>
</tr>
<tr>
<td>6963854</td>
<td>NM_028930 // Tmc5 // transmembrane channel-like gene family 5 // 7I7 F3 // 74424</td>
<td>Tmc5</td>
<td>1.48E-05</td>
<td>41.2887</td>
</tr>
<tr>
<td>6815259</td>
<td>NM_198408 // Crhbp // corticotropin releasing hormone binding protein // 13</td>
<td>Crhbp</td>
<td>0.000134405</td>
<td>41.2017</td>
</tr>
<tr>
<td>Chromosome</td>
<td>GA</td>
<td>Symbol</td>
<td>Description</td>
<td>Position</td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>11</td>
<td>6888340</td>
<td>NM_145526 // P2rx3 // purinergic receptor P2X, ligand-gated ion channel, 3 // 2</td>
<td>P2rx3</td>
<td>1.60E-07</td>
</tr>
<tr>
<td>11</td>
<td>6978354</td>
<td>NM_018882 // Gpr56 // G protein-coupled receptor 56 // 8 D1I8 45.0 cM // 14766 /</td>
<td>Gpr56</td>
<td>7.95E-06</td>
</tr>
<tr>
<td>11</td>
<td>6921929</td>
<td>NM_019427 // Epb4.1l4b // erythrocyte protein band 4.1-like 4b // 4 B3I4 26.0 cM</td>
<td>Epb4.1l4b</td>
<td>1.10E-05</td>
</tr>
<tr>
<td>11</td>
<td>6904066</td>
<td>NM_013531 // Gnb4 // guanine nucleotide binding protein (G protein), beta 4 // 3</td>
<td>Gnb4</td>
<td>0.000122039</td>
</tr>
<tr>
<td>11</td>
<td>6820113</td>
<td>NM_008115 // Gfra2 // glial cell line derived neurotrophic factor family recepto</td>
<td>Gfra2</td>
<td>0.00118214</td>
</tr>
<tr>
<td>11</td>
<td>6824743</td>
<td>NM_010590 // Jub // ajuba // 14 C3I4 // 16475 /// ENSMUST00000054487 // Jub //</td>
<td>Jub</td>
<td>1.68E-06</td>
</tr>
<tr>
<td>11</td>
<td>6855659</td>
<td>NM_001025250 // Vegfa // vascular endothelial growth factor A // 17 CI7 24.2 cM</td>
<td>Vegfa</td>
<td>0.000456204</td>
</tr>
<tr>
<td>11</td>
<td>6922026</td>
<td>NM_025968 // Ptgr1 // prostaglandin reductase 1 // 4 R4 C1 // 67103 /// ENSMUST00</td>
<td>Ptgr1</td>
<td>3.71E-05</td>
</tr>
<tr>
<td>11</td>
<td>6784062</td>
<td>NM_010517 // Igfbp4 // insulin-like growth factor binding protein 4 // 11 D1I1 //</td>
<td>Igfbp4</td>
<td>1.02E-06</td>
</tr>
<tr>
<td>11</td>
<td>6861173</td>
<td>NM_001001979 // Megf10 // multiple EGF-like-domains 10 // 18 D3I8 // 70417 ///</td>
<td>Megf10</td>
<td>0.000119756</td>
</tr>
<tr>
<td>11</td>
<td>6912517</td>
<td>NM_027491 // Rragd // Ras-related GTP binding D // 4 A5I4 11.4 cM // 52187 /// E</td>
<td>Rragd</td>
<td>5.78E-06</td>
</tr>
<tr>
<td>11</td>
<td>6839836</td>
<td>NM_033474 // Arvcf // armadillo repeat gene deleted in velo-cardio-facial syndro</td>
<td>Arvcf</td>
<td>3.53E-09</td>
</tr>
<tr>
<td>11</td>
<td>6791229</td>
<td>NM_001163608 // Plxdc1 // plexin domain containing 1 // 11 D1I1 // 72324 /// NM_</td>
<td>Plxdc1</td>
<td>5.99E-07</td>
</tr>
<tr>
<td>11</td>
<td>6871471</td>
<td>NM_019699 // Fads2 // fatty acid desaturase 2 // 19I19 B // 56473 /// ENSMUST000</td>
<td>Fads2</td>
<td>0.000455211</td>
</tr>
</tbody>
</table>
Figure 3.5 Heat Maps of the top 100 genes differentially expressed in ΔBP LEUK cells compared to ΔB cells.

Heat maps were generated using the signal intensity value (red color = increased, blue color = decreased) for each replicate. Each column represents one replicate. Replicates are aligned together for each cell population. A. Heat map of the top 100 genes decreased in ΔBP LEUK cells compared to ΔB cells. B. Heat map of the top 100 genes increased in ΔBP LEUK cells compared to ΔB cells.
3.5 44 genes known to be activated by PU.1 were reduced in our data set

There are at least 110 genes known to be activated by PU.1 published in the literature (Figure 3.6) [82]. From the 110 target genes 44 had decreased expression in ΔBP LEUK cells compared to ΔB cells (Figure 3.7A). 11 of the 110 target genes had increased expression in ΔBP LEUK cells compared to ΔB cells (Figure 3.7B). 55 of the 110 target genes did not meet the cutoff of 1.5 fold change or the significance of P ≤ 0.05. These results show that 40% of known PU.1 target genes were reduced in our data set.

3.6 Signaling pathways predicted to be the most affected by decreased gene expression

To determine which pathways were affected due to the loss of PU.1 and Spi-B in ΔBP LEUK cells compared to ΔB cells, we used the Partek Pathway tool to map the genes in the increased and decreased lists on to biological pathways. An enrichment score with P values was generated for each pathway using Fisher exact test or chi-square test.

Most of the genes that were decreased in ΔBP LEUK cells compared to ΔB cells were mapped to signaling pathways (Table 3.4). The top category affected in the list was the B cell receptor-signaling pathway (Figure 3.8). As shown, most of the genes in this pathway were reduced (blue colored). This data suggests the importance of PU.1 and Spi-B transcription factors in B cell receptor signaling pathway, and correlated with an arrest in cell differentiation in ΔBP LEUK cells.

3.7 Cell cycle regulation pathways predicted to be the most affected by increased gene expression

Most of the genes that were increased in expression in ΔBP LEUK cells compared to ΔB cells were mapped to cell cycle regulation, DNA replication, and mismatch repair pathways (Table 3.5). These results correlated with an increase in cell proliferation.
**Figure 3.6 The identities of genes directly activated by PU.1**

Genes activated by the transcription factor PU.1 are indicated in alphabetical order using the human nomenclature. Genes are grouped according to the sub cellular location of the proteins that are the products of the indicated genes. Sub cellular location is indicated on the left side.
Figure 3.7 44 genes known to be activated by PU.1 were reduced in our data set

A. Intersect between the list of PU.1 known target genes (110) represented by 136 probe sets (green) and the list of genes were decreased in expression in ΔBP LEUK cells compared to ΔB (pink). 44 genes known to be activated by PU.1 represented by 52 probe sets were reduced in expression in ΔBP LEUK cells compared to ΔB.

B. Intersect between the list of PU.1 known target genes (110) represented by 136 probe sets (green) and the list of genes were increased in expression in ΔBP LEUK cells compared to ΔB (pink). 11 genes known to be activated by PU.1 were increased in expression in ΔBP LEUK cells compared to ΔB.
Table 3.4 The top 20 pathways generated from the gene list that were decreased in expression in ΔBP LEUK cells compared to ΔB cells

<table>
<thead>
<tr>
<th>Pathway Name</th>
<th>Enrichment Score</th>
<th>Enrichment p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>B cell receptor signaling pathway</td>
<td>56.9403</td>
<td>1.87E-25</td>
</tr>
<tr>
<td>Olfactory transduction</td>
<td>55.859</td>
<td>5.51E-25</td>
</tr>
<tr>
<td>Osteoclast differentiation</td>
<td>41.7813</td>
<td>7.16E-19</td>
</tr>
<tr>
<td>Antigen processing and presentation</td>
<td>34.4911</td>
<td>1.05E-15</td>
</tr>
<tr>
<td>T cell receptor signaling pathway</td>
<td>32.5841</td>
<td>7.06E-15</td>
</tr>
<tr>
<td>Leishmaniasis</td>
<td>29.2515</td>
<td>1.98E-13</td>
</tr>
<tr>
<td>Influenza A</td>
<td>28.0184</td>
<td>6.79E-13</td>
</tr>
<tr>
<td>Toll-like receptor signaling pathway</td>
<td>25.7857</td>
<td>6.33E-12</td>
</tr>
<tr>
<td>Leukocyte transendothelial migration</td>
<td>24.2586</td>
<td>2.91E-11</td>
</tr>
<tr>
<td>Viral myocarditis</td>
<td>23.889</td>
<td>4.22E-11</td>
</tr>
<tr>
<td>NOD-like receptor signaling pathway</td>
<td>23.6298</td>
<td>5.47E-11</td>
</tr>
<tr>
<td>Herpes simplex infection</td>
<td>22.3261</td>
<td>2.01E-10</td>
</tr>
<tr>
<td>Legionellosis</td>
<td>22.0244</td>
<td>2.72E-10</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>19.6472</td>
<td>2.93E-09</td>
</tr>
<tr>
<td>Toxoplasmosis</td>
<td>18.2024</td>
<td>1.24E-08</td>
</tr>
<tr>
<td>Allograft rejection</td>
<td>17.5291</td>
<td>2.44E-08</td>
</tr>
<tr>
<td>Type 1 diabetes mellitus</td>
<td>16.7667</td>
<td>5.23E-08</td>
</tr>
<tr>
<td>Intestinal immune network for IgA production</td>
<td>16.5588</td>
<td>6.44E-08</td>
</tr>
<tr>
<td>Lysosome</td>
<td>16.2362</td>
<td>8.89E-08</td>
</tr>
</tbody>
</table>
Table 3.5 The top 20 pathways generated from the gene list that were increased in expression in ΔBP LEUK cells compared to ΔB cells

<table>
<thead>
<tr>
<th>Pathway Name</th>
<th>Enrichment Score</th>
<th>Enrichment p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olfactory transduction</td>
<td>145.106</td>
<td>9.58E-64</td>
</tr>
<tr>
<td>Cell cycle</td>
<td>51.3958</td>
<td>4.78E-23</td>
</tr>
<tr>
<td>DNA replication</td>
<td>36.5039</td>
<td>1.40E-16</td>
</tr>
<tr>
<td>Neuroactive ligand-receptor interaction</td>
<td>30.9574</td>
<td>3.59E-14</td>
</tr>
<tr>
<td>RNA transport</td>
<td>23.5215</td>
<td>6.09E-11</td>
</tr>
<tr>
<td>Cytokine-cytokine receptor interaction</td>
<td>20.7679</td>
<td>9.56E-10</td>
</tr>
<tr>
<td>Homologous recombination</td>
<td>16.7049</td>
<td>5.56E-08</td>
</tr>
<tr>
<td>Ribosome biogenesis in eukaryotes</td>
<td>15.0612</td>
<td>2.88E-07</td>
</tr>
<tr>
<td>Spliceosome</td>
<td>14.6238</td>
<td>4.46E-07</td>
</tr>
<tr>
<td>Oocyte meiosis</td>
<td>14.3249</td>
<td>6.01E-07</td>
</tr>
<tr>
<td>Pyrimidine metabolism</td>
<td>13.8491</td>
<td>9.67E-07</td>
</tr>
<tr>
<td>Base excision repair</td>
<td>13.1767</td>
<td>1.89E-06</td>
</tr>
<tr>
<td>Nucleotide excision repair</td>
<td>12.2301</td>
<td>4.88E-06</td>
</tr>
<tr>
<td>Mismatch repair</td>
<td>11.6785</td>
<td>8.47E-06</td>
</tr>
<tr>
<td>Chemokine signaling pathway</td>
<td>11.4977</td>
<td>1.02E-05</td>
</tr>
<tr>
<td>Cell adhesion molecules (CAMs)</td>
<td>10.9548</td>
<td>1.75E-05</td>
</tr>
<tr>
<td>Arachidonic acid metabolism</td>
<td>10.4704</td>
<td>2.84E-05</td>
</tr>
<tr>
<td>mRNA surveillance pathway</td>
<td>10.0653</td>
<td>4.25E-05</td>
</tr>
<tr>
<td>Serotonergic synapse</td>
<td>9.03487</td>
<td>0.000119181</td>
</tr>
<tr>
<td>Drug metabolism - cytochrome P450</td>
<td>8.6927</td>
<td>0.000167806</td>
</tr>
</tbody>
</table>
Figure 3.8 The B cell receptor signaling pathway was predicted to be the most affected pathway from the decreased gene expression seen in ΔBP LEUK compared to ΔB cells:

The B cell receptor signaling pathway was identified using Partek software from the decreased expression of gene in ΔBP LEUK compared to ΔB cells. Genes in blue had reduced expression levels in ΔBP LEUK compared to ΔB cells. Genes in black had no change in ΔBP LEUK compared to ΔB cells. Notably, most of the genes in this pathway were reduced.
3.8 Confirmation of target genes potentially involved in leukemia

RT-qPCR is commonly used to validate microarray data sets to support findings from microarray data [83]. RNA from sorted ΔB cells and ΔBP LEUK cells was used in RT-qPCR to confirm the microarray data. Ten genes were selected in total for validation, which were involved in different pathways that could promote leukemogenesis in our model (see discussion). Four of the selected genes were involved in the B cell receptor signaling pathway; Bruton’s tyrosine kinase Btk, B cell linker protein Blnk, Lck/yes-related novel Lyn, and guanine nucleotide exchange factor Vav2. Three of the selected genes were from the Wnt signaling pathway; lymphoid enhancer binding factor1 Lef1, transcription factor 7 like 2 Tcf7l2 and leucine rich repeat containing G protein coupled receptor 5 Lgr5. The remaining three genes were for receptors; interleukin seven receptor Il7r, protein tyrosine phosphatase, receptor type, C Ptprc which encodes B220, and glutamyl aminopeptidase Enpep which encodes Bp-1. Five genes demonstrated an increased expression in ΔBP LEUK cells compared to ΔB cells by microarray were also increased by q-PCR; Lef1, Tcf7l2, Lgr5, Il7r, Pb-1 (Figure 3.9). Four genes whose expression was decreased in ΔBP LEUK cells compared to ΔB cells by microarray were also decreased by q-PCR; Btk, Blnk, Lyn, B220 (Figure 3.9). Vav2 had no significant change by q-PCR (data not shown). The q-PCR levels of these transcripts concurred with those of the microarray results (Figure 3.10).
**Fig 3.9 confirmation of target genes potentially involved in leukemia**

RT-qPCR was performed on RNA prepared from sorted ΔB cells and ΔBP LEUK cells. ΔBP LEUK cells showed an increase in expression of *Lef1, Tcf7l2, Lgr5, Il7r*, and *Enpep* compared to the ΔB cells. ΔBP LEUK cells showed a decrease in expression of *Blnk, Btk, Ptprc*, and *Lyn* compared to the ΔB cells. Each result was performed in triplicate and significance with $P < 0.05$ was determined using the Students t-test.
Figure 3.10 “Volcano plot” of changes in gene expression in splenic ΔBP LEUK cells compared to ΔB cells from the microarray experiment. The nine genes that were significantly changed by qPCR and were concurred with those of the microarray are shown.
CHAPTER 4: The role of SPIB in ETV6-RUNX1 acute lymphoblastic leukemia

4.1 Introduction:

The chromosomal translocation t(12;21) fuses the ETV6 (TEL) gene to the RUNX1 (AML1) gene. This is the most common chromosomal alteration in pediatric B cell acute lymphoblastic leukemia representing 25% of the cases [68]. ETV6-RUNX1 fusions arise prenatally as an initiating event linked to pre-leukemia clonal expansion [68]. The molecular mechanism underlying ETV6-RUNX1 function is to interfere with normal RUNX1 function. ETV6-RUNX1 represses RUNX1 target genes by the recruitment of nuclear core repressor/histone deacetylase complexes [69]. However, little is known about the target genes of ETV6-RUNX1 that are involved in leukemogenesis. A recent study with transgenic mice carrying the human ETV6-RUNX1 revealed that SPIB had reduced transcript levels [66]. Another study with two human leukemia cell lines AT-2 and REH, which express ETV6-RUNX1, found that SPIB was one of the top genes that were up regulated after knocking down the ETV6-RUNX1 fusion protein [67]. These studies lead to our central hypothesis that ETV6-RUNX1 functions as an oncogenic driver by repressing SPIB transcription, leading to impaired B cell receptor signaling which could cause the block of B cell differentiation (Figure 4.1)
Figure 4.1 Hypothetical model of ETV6-RUNX1 function in leukemogenesis.

Boxes represent ETV6, RUNX1, and SPIB genes. Arrowheads indicate transcriptional activation. Blunt ends indicate transcriptional repression.
4.2 *SPIB* was one of the top genes that were up regulated after knocking down the ETV6-RUNX1 fusion protein.

A recent study published by the Grumayer group [67] showed that knocking down the ETV6-RUNX1 fusion protein in two human acute lymphoblastic cell lines (AT-2, REH), led to the up-regulation of the *SPIB* transcript levels by 1.6 and 3.59 in REH and AT-2, respectively (Figure 4.2). In brief, they used lentiviral shRNA vectors to knock down the ETV6-RUNX1 fusion protein in the REH and AT-2 cell lines. The cells that were selected for viral integration and stable fusion protein repression had 50-80% reduction of the ETV6-RUNX1 fusion protein between different experiments. These cells were used for expression profiling; three and two biological replicates were used from independent knockdown experiments of the REH and AT-2, respectively. Controls, which were transduced with a non-targeting shRNA vector, were used. Raw data of the Microarray (CEL files) are publically available at GEO (www.ncbi.nlm.nih.gov/geo/, accession number GSE29639) [4].

To reproduce the microarray data, CEL files were downloaded and reanalyzed using Partek software. The results confirmed that *SPIB* was one of the top genes that was up regulated after knocking down ETV6-RUNX1 in both cell lines. The mean of *SPIB* expression was higher in the reanalyzed data compared to the original study (Figure 4.2), which could be explained by the different software used to analyze the data in the original study (R program). Published data were replicated and suggested up-regulation of the *SPIB* transcript levels after knocking down the ETV6-RUNX1 fusion protein.
**Figure 4.2: SPIB and PU.1 expression levels in ETV6-RUNX1 knocked down cells (E/R KD) compared to the control in AT-2 and REH cell lines.**  

**A.** Data generated from the original study showed that *SPIB* expression was up-regulated by 1.6 and 3.59 in REH and AT-2 respectively. There was no significant change in the expression of PU.1.  

**B.** Re-analyzed data showed that *SPIB* expression level was up regulated by 2.27 and 7.3 in REH and AT-2 respectively. There was no significant change in the expression of PU.1.
4.3 *SPIB* expression level was reduced in ETV6-RUNX1 pediatric patient samples compared to other subgroups of B cell acute lymphoblastic leukemia patient samples.

To investigate whether SPIB expression levels were reduced in human pediatric patients who have ETV6-RUNX1 leukemia, microarray data published by Downing’s group was reanalyzed [48]. The goals of this study were to use gene expression profiling to accurately diagnose and sub-classify pediatric acute leukemias, and to enhance the understanding of the underlying biology of each sub-type of acute leukemias. The group collected 132 diagnostic bone marrow (BM) aspirates or peripheral blood (PB) samples from pediatric patients with different types of acute leukemia (Table 4.1). RNA was extracted from those samples, and hybridized to Affymetrix HG-U133A and HG-U133B oligonucleotide microarrays.

CEL files of the subgroups of B cell acute lymphoblastic leukemia patient samples were downloaded and reanalyzed using Partek software. One-way ANOVA was used to detect the differentially expressed genes between the ETV6-RUNX1 pediatric patients and the other subgroups of B cell acute lymphoblastic leukemia patient samples (BCR-ABL1, E2A-PBX1, MLL and Hyperdiploid >50). The results suggested that *SPIB* expression was reduced in ETV6-RUNX1 compared to MLL, E2A-PBX1, and Hyperdiploid patient samples (Figure 4.3). The only exception was BCR-ABL, as there was no difference in the *SPIB* expression level between the two groups (data not shown). These results were consistent with the idea that *SPIB* is repressed by ETV6-RUNX1.

4.4 Five predicted RUNX1 binding sites were located in transcriptionally active sites of the *SPIB* gene

Studies illustrated that ETV6-RUNX1 fusion protein can directly bind RUNX1 consensus
Table 4.1: Pediatric acute lymphoblastic patient subgroups [48]

Samples were collected from acute lymphoblastic leukemia pediatric patients. Patients were divided according to the genetic lesions associated with it.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Sample number</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR-ABL</td>
<td>15</td>
</tr>
<tr>
<td>MLL</td>
<td>20</td>
</tr>
<tr>
<td>E2A-PBX1</td>
<td>18</td>
</tr>
<tr>
<td>Hyperdiploid &gt;50</td>
<td>17</td>
</tr>
<tr>
<td>ETV6-RUNX1</td>
<td>20</td>
</tr>
<tr>
<td>OTHER</td>
<td>28</td>
</tr>
<tr>
<td>T-ALL</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>132</td>
</tr>
</tbody>
</table>
Figure 4.3: SPIB expression levels in ETV6-RUNX1 cells compared to other sub types of B-ALL in pediatric patients. SPIB expression levels were reduced by 5.9, 3.5, and 3.2 fold in ETV6- RUNX1 compared to MLL, E2A-PBX1, and hyperdiploidy respectively.
sequence in a target gene, and repress its activation. To determine if SPIB is directly regulated by ETV6-RUNX1, computational analysis using MatInspector was done on the SPIB gene to predict the RUNX1 binding sites [76]. There were ten predicted RUNX1 binding sites in the SPIB gene that met a cutoff of .8 of the matrix score. To determine whether these binding sites were located in transcriptionally active sites, we crossed referenced them to human Encyclopedia of DNA Elements (ENCODE) data [77]. We used two data sets from ENCODE to define transcriptionally active sites in the B cell line GM12878; DNase I hypersensitivity sites and histone acetylation marker (H3K27AC). Only four predicted RUNX1 binding sites were located in transcriptionally active sites (P = position; P2, P3, P4, P5) (Figure 4.4). NC is a silent site of the SPIB gene that did not have predicted RUNX1 binding sequence, and was used as a negative control for the chromatin immunoprecipitation experiments (ChIP) (Figure 4.7).

As SPIB is considered to be evolutionary duplicated gene from SPII, one would expect the same regulation pattern to exist. SPII gene is 20 kb apart from its neighboring gene, and has upstream regulatory elements [72]. However, this is not true for the SPIB gene. The neighboring gene POLD1 is only 800 base pairs apart from the SPIB gene. In addition, the POLD1 locus had an intron (intron 20), where blood specific transcription factors were bound (e.g. PU.1 and PAX-5) (data not shown). These data suggested that intron 20 of the POLD1 gene could serve as a distal enhancer that regulates the SPIB gene. There was only one predicted RUNX1 binding site in intron 20 of the POLD1 gene, which was located in a transcriptionally active site, henceforth called P1 (Figure 4.4). In summary, the five predicted RUNX1 binding sites (P1, P2, P3, P4, P5) that were located in transcriptionally active sites could be used to investigate the binding of RUNX1 by ChIP.
Figure 4.4: Cross-referenced RUNX1 predicted binding sites on SPIB and POLD1 gene to human ENCODE data. The pink peaks represent H3K27AC, which is a histone acetylation marker, found near active regulatory elements. The red peaks represent the signal of DNase I hypersensitivity at the site. For all the predicted RUNX1 binding sites P1, P2, P3, P4 and P5 have high peaks of DNase I and high peaks of H3K27AC markers except for P5. NC was chosen in a site that is transcriptionally silent (no H3K27AC peak and DNase I peak), and without a RUNX1 binding site.
4.5 P3 is highly conserved through evolution.

It is accepted that sequences conserved through evolutionary distance, are likely to have critical biological function [84]. To determine whether the sequence of the five predicted RUNX1 binding sites (P1, P2, P3, P4, P5) are conserved through evolution, a phylogenetic analysis was done. Multiple alignments of five mammalian species (rat, mouse, human, cow, dog) using Macvector revealed that only P3 was highly conserved between species. P3 is located in the first intron of the SPIB gene (Figure 4.5).

4.6 P3 is highly enriched compared to the negative controls (NC, CRORFII) and the other RUNX1 predicted binding sites (P1, P2, P4, P5).

To confirm that RUNX1 bound to the five predicted RUNX1 binding sites (P1, P2, P3, P4, P5), which were located in transcriptionally active sites, we performed chromatin immunoprecipitation (ChIP) using anti-RUNX1 antibody that could detect both RUNX1 and ETV6-RUNX1 proteins. The REH cell line was used, which was established from a relapsed acute leukemia patient. REH expresses both the fusion protein ETV6-RUNX1 and the wild type RUNX1 [4]. qPCR was used as the method of choice to analyze ChIP experiments. qPCR primers were designed to recognize the RUNX1 predicted binding sites and the negative controls. After that, the primers were tested by input DNA prepared from the REH cell line to ensure specificity. Results revealed that P3 was highly enriched with RUNX1/ETV6-RUNX1 compared to the negative controls (NC, CRORFII) (Figure 4.6) and the other RUNX1 binding sites (P1, P2, P4, P5) (Figure 4.7). Collectively, these data suggest that RUNX/ETV6-RUNX1 could regulate SPIB expression in pre-B cells by interaction with the P3.
Figure 4.5: sequence alignment of P3 RUNX1 site in intron one of the SPIB gene in five mammalian species. The location of the RUNX1 (aGTGGT) site is highlighted inside a square.
**Figure 4.6: Representative ChIP.** Chromatin was prepared from REH cell line and immunoprecipitated with anti-RUNX1. Immunoprecipitated DNA was quantified by relative qPCR using primers recognizing the RUNX1 binding sites P3 and the negative control C4ORFII. Amounts of immunoprecipitated DNA are expressed as percentage of input.
Figure 4.7: ChIP analysis. Chromatin was prepared from the REH cell line and immunoprecipitated with anti-RUNX1. Immunoprecipitated DNA was quantified by relative qPCR using primers recognizing the RUNX1 binding sites P1, P2, P3, P4, P5 and the negative controls NC and C4ORFII. Amounts of immunoprecipitated DNA are expressed as fold enrichment. Fold enrichment means enrichment using anti-RUNX1 antibody compared to nonspecific IgG antibody. The mean ± SD of data from three independent experiments with separate chromatin preparations are shown. P < 0.05 was determined using the Students t-test.
CHAPTER 5: Discussion

5.1 Gene expression analysis of ΔBP LEUK cells suggested increase in cell proliferation, and block in B cell differentiation caused by disruption of key pathways.

Our lab showed that reduced levels of PU.1 and Spi-B are enough to induce B-ALL in mice with 100% incidence [65]. This suggested that PU.1 and Spi-B act as complementary tumor suppressors in the B cell lineage [65]. The first aim of this study was to explore what target genes and pathways were affected by the loss of PU.1 and Spi-B that could explain leukemogenesis. This was done by gene expression profiling.

Microarray results in general demonstrated high fold changes in differential expressed genes (increased-decreased) between ΔBP LEUK cells and ΔB cells. This suggests that PU.1 is a key transcription factor in leukemogenesis, as using the Spib knock out control ΔB eliminated changes due to Spi-B. Mapping those genes to functional groups (pathways), showed an increase in the expression of genes involved in cell cycle, cytokine-cytokine receptor interaction, and DNA replication. These results may suggest enhanced cell proliferation in ΔBP LEUK cells. Also, block of B cell differentiation in ΔBP LEUK cells might be explained with a decrease in expression of genes involved in B cell receptor signaling. Key player genes that could explain leukemogenesis from those pathways include interleukin 7 receptor *Il7r* which was highly expressed in ΔBP LEUK cells, Bruton’s tyrosine kinase *Btk*, and B cell linker protein *Blnk*. Both latter genes were decreased in ΔBP LEUK cells and involved in B cell receptor signaling (BCR signaling).
Btk transcripts were reduced by 3.7 fold in ΔBP LEUK cells compared to ΔB cells. However, confirmation of this result by q-PCR showed a reduction in Btk expression by almost 100-fold change (Figure 3.9, 3.10). In addition, Btk was highly reduced at the protein level in ΔBP LEUK cells compared to ΔB cells, which was shown by western blot analysis done by other members of the lab (unpublished data). These results confirm what has been previously published in the literature that Btk expression requires PU.1 [85]. Studies of Btk deficient mice showed that Btk has a central role in B cell differentiation [40, 41]. Also, studies have illustrated the role of Btk in the progression of pre-B cells to immature B cells. Btk – deficient pre-B cells had increased proliferation capacity in response to IL7 [40]. In addition, during transition of these Pre-B cells to small resting Pre-B cells, they failed to down regulate the metallopeptidase BP-1 and sialoglycoprotein CD43, and had impaired up regulation of the CD2 adhesion molecule and IL2R [40]. Interestingly, ΔBP LEUK cells also failed to down regulate the BP-1 and CD43, and had increased CD2, which could be explained at least partially by the loss of Btk expression. Btk mutations alone do not cause B-ALL [86]. However, mice with mutations in both Btk and Blnk had pre-B cell leukemia with 75% incidence rate as compared to 5% in Blnk only deficient mice [86]. In addition, it has been reported that BTK had reduced expression levels in childhood acute lymphoblastic leukemia [87]. Therefore, loss of PU.1 expression can result in reduced Btk transcript levels that may play a role in the block or impairment of B cell differentiation in ΔBP LEUK cells in cooperation with other genes such as Blnk.

Blnk transcript levels were also reduced by 2.7 fold in ΔBP LEUK cells compared to ΔB cells. q-PCR results were concordant with the microarray data (Figure 3.9, 3.10). Recently, our lab established that PU.1 and Spi-B are direct activators of Blnk [88]. Blnk is an adaptor protein that has a dual function in B cells. It promotes differentiation by transmitting the
differentiation signals of the Pre-BCR, and it has negative effects on proliferation by blocking IL-7 signaling [89, 90]. There are several mechanisms of how Blnk might down regulate IL-7 signaling. First, Blnk can directly bind to JAK3 resulting in uncoupling JAK3 and IL-7R [91]. Second, Blnk can inhibit the activation of the PI3K-AKT pathway downstream of IL-7R signaling [90, 92]. Blnk mutations or reduced expression are associated with human B-ALL [47, 93]. Therefore, reduction in the expression of PU.1 and Spi-B will result in the reduction of Blnk in ΔBP LEUK cells. This will lead to decreased differentiation signals, and increased IL-7 signaling pathways. This might explain the impairment of B cell differentiation and enhanced proliferation through increased IL-7R signaling in ΔBP LEUK cells.

It has been shown that over expression of IL-7 can lead to leukemia in mice [94]. Also, IL-7R gain of function mutations is associated with pediatric acute lymphoblastic leukemia [95]. Il7r was highly expressed in ΔBP LEUK cells compared to ΔB cells. q-PCR results were in agreement with the microarray results (Figure 3.9, 3.10). In addition, our lab has demonstrated that IL-7R signaling pathway through PI3K-AKT is activated, and ΔBP LEUK cells can grew indefinitely in response to IL7 cytokine in vitro [88]. These results suggested that IL-7 and IL-7R are important in the proliferation of ΔBP LEUK cells.

In summary, these results suggest that PU.1 and Spi-B are important transcription factors that play a role in the balance between the proliferation and differentiation of B cells by activating genes required for BCR signaling that promote differentiation and inhibit proliferation.

Other candidate genes that could explain the disruption of B cell differentiation in ΔBP LEUK cells belong to the Wnt-signaling pathway. Although the Wnt-signaling pathway was
not significantly enriched, Lef1 was one of the top 100 genes that was increased by 20 fold in ΔBP LEUK cells (Figure 3.9, 3.10). Knock out studies of Lef1 showed that is required for pro-B cell proliferation [96]. On the other hand, overexpression of Lef1 caused a severe disruption of normal hematopoietic differentiation leading to B lymphoblastic and acute myeloid leukemias [97]. In addition, LEF1 mutations were associated with human ALL [46]. Therefore, increased Lef1 expression may contribute to the disruption of B cell development in ΔBP LEUK cells. Other candidate genes from the Wnt-signaling pathway included Tcf7l2 and Lgr5 that were also up regulated in ΔBP LEUK cells, and their expression levels were confirmed by q-PCR (Figure 3.9,3.10).

Overall, microarray results were consistent with what has been published in the literature about the PU.1 target genes[82]. The comparison of the 110 known target genes of PU.1 revealed that 40% of those genes were reduced in ΔBP LEUK cells (Figure 3.7). There were exceptions of 11 genes known to be activated by PU.1 that were up regulated in ΔBP LEUK cells. There are two potential explanations for these findings. First, the activation of certain genes such as Cdk6 and Irf4 by PU.1 were shown only in erythroid and myeloid cell types respectively, which may refer to the difference in the context of gene regulation between the cell types [98, 99]. Second, some genes may require PU.1 for activation but not for continual expression, such as Il7r [21]. The Il7r activated during the common lymphoid progenitor stage, which is before cre- mediated PU.1 deletion in ΔBP LEUK cells [24, 65]. In addition to consistency with the published literature, validation of the expression of selected genes by q-PCR showed an agreement in the direction of the fold change except for vav2, which did not change significantly (Figure 3.9,3.10). This agreement in the direction of fold change is considered enough as microarray technique differs from q-PCR in many aspects such as the region of the gene targeted by the microarray probes and the q-PCR primers, and the
normalization methods. These all will affect how the results will correlate [83].

In summary, ΔBP LEUK cells have reduced expression of key genes involved in BCR signaling, which also have been reported in human B-ALL. Although our data does not show the effect of loss of Spi-B on gene expression, it is reasonable to conclude that Spi-B will affect BCR signaling for the following reasons. First, PU.1 and Spi-B have identical binding sites, which suggests that they can trans-activate the same target genes [65]. Second, Spib \( \rightarrow \) B cells are defective in one or more BCR signaling components, which could be directly regulated by Spib [64]. This may indicate that reduced expression or mutations in PU.1 and SPIB could play a role in the pathogenesis of human B-ALL through the disruption of BCR signaling.

5.2 Reduced levels of SPIB in childhood ETV6-RUNX1 B cell acute lymphoblastic leukemia:

Recently, studies have demonstrated that impaired expression of PU.1 and SPIB is associated with human B-ALL. Mutations of SPI1 have been reported in relapsed acute lymphoblastic leukemia patients [49]. Also, a recent study by the Grumayer group with two human leukemia cell lines AT-2 and REH, which express ETV6-RUNX1, suggested that SPIB was one of the top genes that were up regulated after knocking down the ETV6-RUNX1 fusion protein [67]. These studies triggered us to question if ETV6-RUNX1 function as an oncogenic driver by repressing SPIB transcription. To confirm what the Grumayer group published, re-analysis of the published microarray data was done. The re-analyzed results were consistent with what was published in the original paper. SPIB was at the top of the gene list that was up regulated after knocking down ETV-RUNX1 protein. PU.1 transcript levels did not change between the knocked down and the control in either cell lines
(Figure 4.2). These results suggested that SPIB but not PU.1 has impaired expression in ETV6-RUNX1 leukemia. This led us to question if this phenomena will be true in actual patient samples from pediatric ETV6-RUNX1 leukemia. Re-analysis of the published microarray data by the Downing group confirmed the trend of reduced SPIB transcript levels in pediatric patients with ETV6-RUNX1 leukemia compared to the other sub-types of B-ALL (Figure 4.3)[48]. The exception was with BCR-ABL1 leukemia where there was no difference in the SPIB expression compared to ETV6-RUNX1 leukemia. This suggests that SPIB transcript levels are also reduced in BCR-ABL1 leukemia. This finding is in agreement with what was published by the Muschen group where they showed that SPIB transcript levels were reduced in BCR-ABL1 leukemia patients [100]. In summary, these results suggest that SPIB is suppressed in ETV6-RUNX1 human leukemia cell lines and pediatric patients with this subtype of B-ALL. The next step of this study was to investigate the mechanism by which SPIB expression is reduced in ETV6-RUNX1 leukemia.

5.3 The ETV6-RUNX1 fusion protein could repress SPIB transcription:

Previous studies have shown that ETV6-RUNX1 fusion proteins could interact with RUNX1 binding sites, which will cause the repression of RUNX1 target genes via the recruitment of nuclear core repressor/histone deacetylase complexes [69, 101]. We hypothesized that ETV6-RUNX1 would repress SPIB expression by interacting with RUNX1 binding sites on SPIB. To test this hypothesis we first used computational software (MATINSPECTOR) to predict RUNX1 binding sites within the SPIB gene [76]. These sites were scanned for transcriptional activation using ENCODE data [77]. Two markers were used to define transcriptional active sites in the B cell line GM12878. The first marker was DNase I hypersensitivity, and the second marker was histone acetylation modification (H3K27AC), which has been shown to distinguish active enhancers from inactive enhancer
elements [77, 102]. Only four predicted RUNX1 sites were located in transcriptionally active sites (P2, P3, P4, P5) (Figure 4.4). P3 is highly conserved through evolution, suggesting a critical biological function (Figure 4.5). P3 was located in the first intron below the transcription start site that has been shown to be active in early stages of B cell development [103]. This data indicate that we have identified an important enhancer/promoter that could be negatively regulated by ETV6-RUNX1 in pre-B ALL cells. In addition to the SPIB locus, intron 20 of the POLD1 gene could serve as a distal enhancer that regulates the SPIB gene (Figure 4.4). The rationale behind this is that SPIB is thought to be evolutionary duplicated from the SPIII gene[104]. Interestingly, unlike SPIB, SPIII gene is 20 KB apart from its neighboring gene, and has upstream regulatory elements [72]. POLD1 is only 800 base pairs apart from the SPIB gene. ChIP experiments showed that P3 was highly enriched with RUNX1/ETV6-RUNX1. Therefore, these results indicate that RUNX1/ETV6-RUNX1 could regulate SPIB. Future work needs to be done in order to determine whether it is suppressing or activating the SPIB gene.
5.4 Future directions:

This project provided insight into the role of PU.1 and Spi-B as tumor suppressor genes in B-ALL in mice. In addition, it specified a mechanism for the impairment of expression of \textit{SPIB} in ETV6-RUNX1 B-ALL. Yet, more work needs to be done to confirm the repression function of ETV6-RUNX1 on \textit{SPIB} and to address the downstream consequences of ETV6-RUNX1 inhibition of \textit{SPIB} expression along with the changes induced by other cooperating mutations. First, to confirm the repressor function of ETV6-RUNX1, two approaches can be used. Knocking down ETV6-RUNX1 can be done to illustrate the up-regulation of \textit{SPIB} in the REH cell line. In addition, direct examination of the repression function of ETV6-RUNX1 can be done by cloning the region around P3 and ligating it into luciferase reporter vector. If the conclusion from these assays were that ETV6-RUNX1 represses \textit{SPIB} expression, then it would be interesting if we can show that enforced expression of \textit{SPIB} in the REH cell line will promote differentiation and change cell cycle status in the REH cell line. In addition, this assay will help in identifying novel target genes that could be regulated by \textit{SPIB}. Finally, it would be interesting if the findings from the REH cell line could be validated in primary ETV6-RUNX1 patient samples. This could be done using standard curves to measure the transcript levels of \textit{SPII} and \textit{SPIB} using qPCR. Other techniques could be used. This study is expected to lead to a deeper understanding of the underlying biology of leukemogenesis caused by the loss of \textit{SPIB} in ETV6-RUNX1 leukemia patients, hence allow for the development new molecular targeted therapies for B-cell leukemia.
5.5 Summary and conclusions:

In this study, we provide mechanistic evidence to explain how PU.1 and Spi-B could play a role in the leukemogenesis of ETV6-RUNX1 B cell leukemia (Figure 5.1). We tackled this question using three approaches. The first approach was to understand how PU.1 and Spi-B function as tumor suppressor genes in B cells. A mouse model that conditionally knocked out PU.1 in B cells on a Spib−/− background was used. These mice develop B-ALL with 100% incidence. Gene expression analysis was used to explore what pathways and target genes were affected by the loss of PU.1 and Spi-B. This work identified the B cell receptor-signaling pathway as the most affected pathway where most of the involved genes had reduced expression in ΔBP LEUK cells, which might explain the block in B cell differentiation. The second approach was to measure the SPI1 and SPIB transcript levels by re-analyzing published microarray data from the human REH cell line and pediatric patient samples, which express the ETV6-RUNX1 fusion. Results from the microarray analysis suggested that SPIB is repressed by ETV6-RUNX1. The third approach was to address the mechanism of how SPIB is repressed in ETV6-RUNX1 leukemia patients and the human REH cell line. Chromatin immunoprecipitation experiments were done using the REH cell line to identify putative binding sites of RUNX1. Our findings strongly suggest that SPIB expression is reduced in ETV6-RUNX1 leukemia, and it could be repressed by ETV6-RUNX1 fusion protein.
Figure 5.1 hypothetical model of ETV6-RUNX1 function in leukemogenesis.
Bibliography

71.


68.


65.


67.


68.


69.


70.


71.


72.


73.


74.


75.


76.


77.


78.


79.

97.

96.

95.

94.

93.

92.

91.

90.

88.

87.

86.

85.

84.

83.

82.

81.

80.


APPENDIX

Animal use protocol

From: eSiriusWebServer
To: Rodney DeKoter
CC: auspc@uwoc.ca
Date: 7/20/2012 3:43 PM
Subject: eSirius Notification - New Protocol Modification Has Been APPROVED2009-010::4

Western

AUP Number: 2009-010
PI Name: DeKoter, Rodney
AUP Title: Regulation of Myeloid Versus Lymphoid Cell Fates by PU.1

Official Notification of AUS Approval: A MODIFICATION to Animal Use Protocol 2009-010 has been approved.
The holder of this Animal Use Protocol is responsible to ensure that all associated safety components
(biosafety, radiation safety, general laboratory safety) comply with institutional safety standards and have
received all necessary approvals. Please consult directly with your institutional safety officers.
Submitted by: Kirchlea, Will D
on behalf of the Animal Use Subcommittee
Curriculum Vitae

Name: Shereen Turkistany

Education:
University of Western Ontario (UWO), London, ON
Master of Science (M.Sc.), 2011- present
Department of Microbiology and Immunology

King Abdulaziz University, Jeddah, Saudi Arabia
Bachelor of Medical Technology, 2003-2007
Faculty of Applied Medical Science, Specialization: Medical Technology

Awards:
King Abdullah Scholarship Program; 2009- 2013

Canadian Society for Immunology Annual Conference 2012, St. John’s, Newfoundland; Winner of poster session; Awarded $ 300 and certificate.


Publications:
*The Transcription Factor PU.1 is a Critical Regulator of Cellular Communication in the Immune System (2011), Shereen A. Turkistany, Rodney P.DeKoter.*

*Deletion of genes encoding PU.1 and Spib in developing B cells Unlinks B cell receptor and interleukin-7-receptor signaling,* (manuscript in preparation), Darah A. Christie, Shereen A.Turkistany, Li S. Xu, Heather C. Broughton, Gillian I.Bell, David A.Hess, and Rodney P.DeKoter.
RELVANT TECHNICAL SKILLS:

Analyze microarray data using Partek software or the R program.

Excellent in using bioinformatics tools (ComBat, Ingenuity, Matinspecto, MacVector, USCS genome browser)

Real-Time PCR (relative and absolute measurements of gene expression)

Qualified in other research techniques such as Chromatin immunoprecipitation, Western blot, Cloning, cell culture and conventional PCR

Full access to Microsoft application and other Graphic software (Adobe Illustrator, InDesign, Photoshop)

Volunteer Service:

Let’s Talk Science, 2011 – Present
Planned new activities to increase youths exposure and involvement In science

Graduate Society Social Club, 2011 – present
Assisted with planning of numerous social events for the Department of Microbiology and Immunology

Infection and Immunity Research Forum, 2011- present
Assisted in organizing the event in November 2011.