

12-2013

# c-REL is a Transcriptional Target of Mesoderm Inducer in *Xenopus* Like 1 (MIXL1)

Aaron C. Raymond

Follow this and additional works at: [http://digitalcommons.library.tmc.edu/utgsbs\\_dissertations](http://digitalcommons.library.tmc.edu/utgsbs_dissertations)

 Part of the [Medicine and Health Sciences Commons](#), and the [Molecular Genetics Commons](#)

---

## Recommended Citation

Raymond, Aaron C., "c-REL is a Transcriptional Target of Mesoderm Inducer in *Xenopus* Like 1 (MIXL1)" (2013). *UT GSBS Dissertations and Theses (Open Access)*. 415.  
[http://digitalcommons.library.tmc.edu/utgsbs\\_dissertations/415](http://digitalcommons.library.tmc.edu/utgsbs_dissertations/415)

This Dissertation (PhD) is brought to you for free and open access by the Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in UT GSBS Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact [laurel.sanders@library.tmc.edu](mailto:laurel.sanders@library.tmc.edu).

***c-REL is a Transcriptional Target of Mesoderm Inducer in Xenopus Like 1 (MIXL1)***

by

Aaron Craig Raymond, M.S.

APPROVED:

---

Supervisory Professor, Lalitha Nagarajan, PhD

---

Richard Behringer, PhD

---

Yasuhide Furuta, PhD

---

Peng Huang, MD, PhD

---

Pierre McCrea, PhD

APPROVED:

---

Dean, The University of Texas  
Graduate School of Biomedical Sciences at Houston

***c-REL* is a Transcriptional Target of *Mesoderm Inducer in Xenopus Like 1*  
(*MIXL1*)**

A Dissertation

Presented to the Faculty of  
The University of Texas  
Health Science Center at Houston  
and  
The University of Texas  
M. D. Anderson Cancer Center  
Graduate School of Biomedical Sciences  
in Partial Fulfillment

of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

by

Aaron C Raymond, M.S.  
Houston, Texas

December, 2013

***c-REL is a Transcriptional Target of *Mesoderm Inducer in Xenopus Like 1* (MIXL1)***

Aaron Raymond, M.S.

Supervisory Professor: Lalitha Nagarajan, Ph.D.

MIXL1, an evolutionarily conserved, paired-type homeobox transcription factor induced by BMP4/TGF $\beta$  signaling, is a critical regulator of embryonic and adult hematopoiesis. Several lines of evidence implicate MIXL1 in hematopoietic transformation: (i) Aberrant MIXL1 expression is seen in human CML (Chronic Myelogenous Leukemia) in blast crisis, AML (Acute myelogenous leukemia), B cell lymphomas and pediatric ALL (Acute lymphocytic leukemia). (ii) Retroviral transduction of Mixl1 induces AML in murine models. Nonetheless, mechanisms underlying MIXL1 mediated proliferative, survival advantages are unknown.

The goal of my studies is to understand if and how aberrant MIXL1 expression contributes to leukemogenesis. As a first step, I sought to determine transcriptional targets of MIXL1. Using MIXL1 overexpression lines established in the human myelomonocytic leukemia cell line, U937, I performed global chromatin immunoprecipitation coupled sequencing (ChIP-Seq), expression

profiling studies I identified several putative targets. Of these, the proto-oncogene c-REL was an important target. Both ectopically expressed and endogenous MIXL1 bound *REL* promoters in myelogenous leukemia cells. Furthermore, cREL was induced under MIXL1 over-expression conditions in the leukemic cell line U937. Targeted shRNA mediated knockdown of endogenous MIXL1 in AML cell lines down regulated *REL* expression.

c-REL the cellular homolog of the viral oncogene v-rel encoded by the oncogenic reticuloendotheliosis retrovirus, is a member of the Nf- $\kappa$ B/Rel gene family. c-REL transcriptionally activates proliferation and immune response and represses apoptosis through upregulation of anti-apoptotic genes. c-REL plays a critical role in hematopoietic differentiation. It is required for B-cell differentiation, and dendritic cells to activate T cells. In the myeloid lineage, c-REL is critical for macrophage mediated innate immunity. My findings suggest that transcriptional up regulation of *REL* by MIXL1, therefore would promote survival or proliferation through activation of the Nf- $\kappa$ B pathway in AML cells. Consistent with these established activities of c-REL, expression of two anti-apoptotic genes BCL2A1 and BCL2L1 increased under MIXL1 over-expression conditions and decreased under conditions of MIXL1 knockdown by targeted shRNA in U937 cells. In summary, my studies identify c-REL to be a novel mediator of MIXL1-induced survival signals in leukemia.

## TABLE OF CONTENTS

ABBREVIATIONS – Genes and Proteins of Interest	xi
ABBREVIATIONS – Other Terms	xvi
Chapter 1. Introduction	1
1. MIXL1	1
2. MIXL1 is regulated by TGF- $\beta$ /BMP Pathways	2
3. MIXL1 in Hematopoiesis and Hematopoietic Neoplasms	3
4. Homeoboxes in Acute Myeloid Leukemia	4
5. NF- $\kappa$ B Pathway in Leukemia and Lymphoma	5
6. Clarifying the role of MIXL1 in Acute Myeloid leukemia	6
Chapter 2. Materials and Methods	7
1. Cell Culture	7
2. Establishment of Overexpression and Knockdown Cell Lines	7
3. Chromatin Immunoprecipitation	8
4. Global Expression Analysis	10
5. RT-QPCR	13
6. Co-Immunoprecipitation	16
7. Immunoblotting	16
8. Luciferase Reporter Assay	17
9. MTS Assay	20
10. Analysis and Data mining	21

Chapter 3. Results	22
1. Characterization of AML cell lines with overexpression of MIXL1	22
2. ChIP-Seq analysis of MIXL1 overexpression lines to identify transcriptional targets	27
3. Global Expression Profiling of MIXL1 overexpression to confirm transcriptional targets	39
4. c-REL as a transcriptional target of MIXL1	39
5. <i>MIXL1</i> knockdown in KG1 represses <i>c-REL</i> and <i>BCL2A1/L1</i> expression	49
6. MZF1 and MIXL1 interact with the c-REL promoter	49
7. MIXL1 is preferentially induced by BMP4 in human CD34+ cord blood hematopoietic stem progenitor cell	59
8. <i>MIXL1</i> expressing cells are sensitive to the BMP inhibitor LDN-193189	66
9. Endogenous MIXL1 in AML Patients	71
Chapter 4. Discussion	79
1. <i>MIXL1</i> is over-expressed in a subset of AML	79
2. MIXL1-positive AML lines are sensitive to LDN-193189	82
3. Transcriptional regulation by MIXL1 in hematopoietic cells	85
4. <i>c-REL</i> is a transcriptional target of MIXL1	87
5. Summary	88
Appendix: Peaks identified by all analyses of the MIXL-overexpression ChIP-Seq dataset.	89
References	96
Vita	112

## TABLE OF FIGURES

Figure 1 – Stable transfectants of U937 cells express MIXL1 at levels similar to endogenous protein in AML cell lines.	23
Figure 2 – MIXL1 expression reduces sensitivity of U937 cells to doxorubicin.	25
Figure 3 – Multiple analyses of MIXL1-ChIP-Seq identify 179 high-quality peaks.	28
Figure 4 – The majority of peaks for MIXL1 identified by ChIP-Seq localize to the Promoter region.	30
Figure 5 – Predicted MIXL1 motifs are not enriched in U937 overexpression MIXL1 ChIP-seq Peaks.	33
Figure 6 – The most commonly occurring motifs in the ChIP-seq peaks are Zinc-Finger binding motifs.	35
Figure 7 – ChIP-qPCR analysis of selected gene-associated peaks confirms enrichment.	37
Figure 8 – RT-qPCR confirms change of expression in selected genes as identified by Microarray..	40
Figure 9 – The NF- $\kappa$ B pathway is regulated by MIXL1.	42
Figure 10 – MIXL1 expressing clones show enhanced transcript levels for c-REL, BCL2A1, and BCL2L1.	45



Figure 11 – ChIP localizes endogenous MIXL1 to c-REL promoter in KG1 cells.	39
Figure 12 – Knockdown of MIXL1 decreased while enforced expression of c-REL increased c-REL, BCL2A1, and BCL2L1 transcript levels.	50
Figure 13 – c-REL over expression rescues of MIXL1 knockdown mediated growth arrest in KG1 cells.	52
Figure 14 – MIXL1 interacts with the center of the identified peak region of the REL promoter.	55
Figure 15 – Known Transcription Factor Binding Motifs in the REL Promoter.	57
Figure 16 – MZF1 is identified by co-immunoprecipitation of Flag-tagged MIXL1 in U937.	60
Figure 17 – MZF1 binds to the same locus as MIXL1 on c-REL promoter.	62
Figure 18 BMP4 induces MIXL1 in CD34 <sup>+</sup> .	64
Figure 19 – High MIXL1 expression cell lines are sensitive to 3 $\mu$ M LDN-193189.	67
Figure 20 – High MIXL1 expression cell lines are sensitive to consistent exposure of LDN-193189 for 4 days.	69
Figure 21 – MIXL1 overexpression marks a set of AML cases distinct from those expressing CDX2, HLX or HOXA9.	72
Figure 22 – MIXL1 expression co-occurs with REL and BCL2L1 expression, but not with HOXA9.	74

Figure 23 – High MIXL1 expression patient samples contain previously identified somatic mutations of AML.	77
Figure 24 – Proposed BMP/MIXL1/REL pathway.	80

## TABLE OF TABLES

Table 1 – Primer sets used for ChIP-qPCR	11
Table 2 – Probe sets used for RT-qPCR	14
Table 3 – Primers used for <i>REL</i> Promoter Luciferase construct construction	18
Appendix - Peaks identified by all analyses of the MIXL-overexpression	
ChIP-Seq dataset.	89

## **ABBREVIATIONS – Genes and Proteins of Interest**

**ABL1** – *c-ABL, Abelson murine leukemia viral oncogene homolog 1*. A proto-oncogene tyrosine kinase associated with CML.

**Activin A** – A cytokine belonging to the TGF- $\beta$  superfamily.

**ALK2 (ACVR1)** – Activin A receptor, type 1, a BMP-type member of the TGF- $\beta$  receptor superfamily. Can be activated by Activin and BMP.

**ALK3 (BMPR1A)** – Bone Morphogenic Protein receptor, type 1A, a BMP-type member of the TGF- $\beta$  receptor superfamily. Can be activated by BMP proteins.

**ALK6 (BMPR1B)** – Bone Morphogenic Protein receptor, type 1B, a BMP-type member of the TGF- $\beta$  receptor superfamily. Can be activated by BMP proteins.

**AVCR2A** – *Activin type 2 receptor A, a BMP-type member of the TGF- $\beta$  receptor superfamily*. Can be activated by Activin and BMP.

**ALX4** – *ALX Homeobox 4*, a paired-type homeobox transcription factor.

**APBB2** – *Amyloid beta A4 precursor protein-binding family B member 2*.

**BCL2** – *B-cell lymphoma 2*, an anti-apoptotic gene that regulates cytochrome-c mediated apoptosis.

**BCL2A1** – *BCL2-related protein A1*, a *BCL2* family member and an anti-apoptotic gene that regulates cytochrome-c mediated apoptosis.

**BCL2L1** – *BCL2-like protein 1*, a *BCL2* family member and an anti-apoptotic gene that regulates cytochrome-c mediated apoptosis.

**BMP4** – Bone Morphogenic Protein 4, a cytokine belonging to the TGF- $\beta$  superfamily.

**BMPR1** – Bone Morphogenic Protein receptor, type 2, a BMP-type member of the TGF- $\beta$  receptor superfamily. Can be activated by BMP proteins.

**BRCA1**– *Breast Cancer 1, early onset*, a RING zinc finger protein.

**CAMKK2** – *Calcium/calmodulin-dependent protein kinase kinase 2*, a calmodulin-dependant serine/threonine kinase.

**CDX2** – *Caudal type homeobox 2*, a caudal-related homeobox transcription factor.

**DLX1** –*Distal-less homeobox 1*, a homeobox transcription factor with similarity to the *Drosophila* distal-less gene.

**DLX4** –*Distal-less homeobox 4*, a homeobox transcription factor with similarity to the *Drosophila* distal-less gene.

**EGR1** – *Early growth response 1*, a C2H2-type zinc-finger transcription factor.

**EIF1** – *Eukaryotic translation initiation factor 1*.

**EOMES** – *Eukaryotic translation initiation factor 1*.

**GSC** – *Goosecoid homeobox*, a paired-type homeobox transcription factor.

**HHEX**– *Hematopoietically-expressed homeobox*, a homeobox transcription factor.

**HLX** – *H2.0-like homeobox*, a homeobox transcription factor.

**HOXA9** –*Homeobox A9*, a homeobox transcription factor of the cluster A subgroup.

**IL18R1** –*Interleukin 18 receptor 1*, an Interleukin cytokine receptor, that specifically binds to interleukin 18.

**MIXL1** – *Mesoderm inducer in Xenopus like 1*, a paired type homeobox transcription factor.

**MZF1** – *Myeloid zinc finger 1*, a C2H2-type zinc finger transcription factor.

**NEUROD1** – *Neuronal differentiation 1*, a helix-loop-helix transcription factor.

**NF-κB** – A ubiquitous transcription factor protein dimer that consists of either NF-κB1/p50 or NF-κB2/p49 in complex with REL, RELA, or RELB. The most prevalent form is NF-κB1/p50 in complex with RELA.

**NFKB1** – *Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1*, a NF-κB/REL-family transcription factor that produces the protein NF-κB1/p50, part of the most prevalent “canonical” NF-κB protein dimer alongside RELA.

**NFKB2** – *Nuclear factor of kappa light polypeptide gene enhancer in B-cells 2*, a NF-κB/REL-family transcription factor that produces the protein NF-κB2/p49.

**NFKBIA,B,E** – *Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha, beta, and epsilon*, a family of proteins that bind to NF-κB/REL dimers to inhibit their function.

**NKX2-1** – *NK2 homeobox 1*, a homeobox transcription factor.

**OTX2** – *orthodenticle homeobox 2*, a paired-type homeobox transcription factor.

**PCGF2** – *Polycomb group RING finger protein 2*, a RING finger containing protein.

**PML** – *Promyelocytic leukemia protein*, is a homeobox transcription factor.

**POU4F2** – *POU class 4 homeobox 2*, is a homeobox transcription factor.

**c-REL** – v-rel avian reticuloendotheliosis viral oncogene, a NF-κB/REL-family transcription factor and family member that shares the most sequence homolog to the viral oncogene.

**RELA** – v-rel avian reticuloendotheliosis viral oncogene homolog A (p65), a NF-κB/REL-family transcription factor and part of the most prevalent “canonical” NF-κB protein dimer alongside NF-κB1/p50.

**RELB** – v-rel avian reticuloendotheliosis viral oncogene homolog B, a NF-κB/REL-family transcription factor.

**RUNX1** – *RUNT related transcription factor 1*, a transcription factor that regulates hematopoietic stem cell differentiation.

**SLC39A13** – *Solute carrier family 39 (zinc transporter), member 13*, a transmembrane zinc transporter.

**SMAD** – A family of signal transduction proteins that transcriptionally regulate targeted genes in response to TGF-β family signaling pathways.

**SMYD5** – *SMYD family member 5*, a methyltransferase.

**Sp1** – *Sp1 transcription factor*, a SP/KLF family zinc finger transcription factor.

**T** – A gene that encodes the T box transcription factor Brachyury.

**TBX6** – *T-box 6*, a T box transcription factor.

**TBX20** – *T-box 20*, a T box transcription factor.

**TGF-β** – Transforming growth factor beta, a family of cytokines that regulate proliferation, differentiation, migration, and other pathways through interaction with TGF-β receptors. Part of the same superfamily of cytokines as BMP and Activin.

**VEG-FR** – Vascular endothelial growth factor Receptor, a family of receptors that interact with VEGF, and are important for vasculogenesis and angiogenesis.

**YES1** – *v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1*, src family tyrosine kinase and homolog to the Yamaguchi sarcoma virus oncogene.

**ZP3** – *Zona pellucida glycoprotein 3 (sperm receptor)*, an extracellular matrix protein.



## **ABBREVIATIONS – Other**

**6-FAM** – 6-Carboxyfluorescein, a common fluorescent dye used for attachment to oligonucleotides, especially in real time PCR techniques.

**ALL** – Acute lymphoblastic leukemia.

**AML** – Acute myelogenous leukemia.

**ChIP** – Chromatin Immunoprecipitation, A technique for identification of in vitro or in vivo interactions between proteins and chromosomal DNA by DNA/protein crosslinking, immunoprecipitation and then DNA extraction.

**ChIP-qPCR** – ChIP-coupled qPCR, a technique for quantifying and evaluating the prevalence of a DNA/protein interaction by using quantitative PCR against desired targets on DNA samples produced by ChIP.

**ChIP-Seq** – ChIP-coupled sequencing, a high throughput technique for identifying, quantifying, and evaluating the prevalence of DNA/protein interactions by using high-throughput sequencings techniques on DNA samples produced by ChIP.

**CML** – Chronic myelogenous leukemia.

**Ct (qPCR)** – Cycle Threshold, a raw data value representing the prevalence of nucleic acid sequencing in a quantitative PCR reaction. The Ct value is the cycle of PCR reaction where the fluorescence signal passed a set threshold, so lower Ct values mean higher abundance of DNA sequence.

**DMSO** – Dimethyl sulfoxide.

**Dorsomorphin** –Dorsomorphin dihydrochloride, a serine/threonine kinase inhibitor which can potently inhibit for BMP/Activin related receptor kinases and AMP-activated protein kinase (AMPK).

**Doxorubicin** – A cancer chemotherapy commonly used for leukemia and other cancers. It functions by intercalating DNA, blocking DNA replication.

**ECL Reagent** – Enhanced chemiluminescence reagent, a luminol-based detection reagent used for western blotting.

**FLAG** – A polypeptide epitope tag. Amino Acid Sequence: DYKDDDDK

**Fibrodysplasia ossificans progressiva** – A rare chronic disease that causes fibrous tissues to ossify, form bone tissue, when injured or spontaneously. Mutations in ACVR1 have been implicated in this disease.

**Gag-pol** – Group Antigens (ag)-Pol, A retroviral polyprotein that is cleaved to create the non-viral envelope proteins required for producing a retrovirus, including the reverse transcriptase (Pol) and retroviral core proteins (Gag). An expression construct containing this is used alongside an envelope-producing construct and the desired retroviral expression vector to create a replication-dead retrovirus.

**HA** –Human influenza hemagglutinin, commonly used as an epitope tag. The tag's Amino Acid Sequence: YPYDVPDYA

**Iowa Black FQ** – An IDT (Integrated DNA Technologies)-developed quencher with an absorbance range of 420 to 620 nm. The main use in this case is to attachment to oligonucleotides, to quench the activity of the nearby 6-FAM group.

**LDN-193189** – A serine/threonine kinase inhibitor derived from dorsomorphin to be a more specific inhibitor to BMP/Activin related ALKs.

**MTS** – 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, a tetrazole that, when present with phenazine methosulfate (PMS), produces a formazan product with the absorbance range of 490-500 nm, when reduced by cellular enzymes.

**qPCR** – Quantitative or Real-Time Polymerase Chain Reaction, a quantitative form of polymerase chain reaction where a fluorescence activity is tested for at the end of each cycle of PCR, with the activation of a fluorescent dye being directly linked with each successful polymerase reaction.

**SELEX**– systematic evolution of ligands by exponential enrichment. An approach for identifying the binding site of proteins on RNA or DNA in vitro.

**shRNA** – small hairpin RNA, a short sequence of RNA, usually artificially produced, to silence the targeted gene through RNA interference/post transcriptional gene silencing.

**TCGA** – The Cancer Genome Atlas, <http://cancergenome.nih.gov/>

**TSE** – Transcriptional Sequence End, the predicted endpoint for transcription of a gene locus. In this dissertation, TSE refers to a 5 kb untranscribed region downstream the endpoint for the purpose of defining the location of ChIP peaks in reference to nearby gene loci.

**TSS** – Transcriptional Start Site, the transcriptional start site of a gene locus. In this dissertation, the promoter region is defined as a 5 kb region upstream of the transcriptional start site.

## **Chapter 1. Introduction**

### **1. MIXL1**

*MIXL1*, *Mesoderm Inducer in Xenopus Like 1*, is the human ortholog of *Xenopus* Mix.1, required for mesoderm development. *MIXL1* is the sole human member of the Mix/Bix homeobox sub-family. While the Mix/Bix family is comprised of multiple genes in *Xenopus laevis* (*Mix.1-4*, *Bix.1-4*, and *Mixer*) and *Danio Rerio* (*Bon*, *Mtx1* and *Mtx2*); birds and mammals each contain only one ortholog for *Mix.1* and no pseudogenes (1-3). The Mix/Bix family has been well characterized as major regulators of mesoderm and endoderm specification in early embryonic development in *Xenopus*. Expression of *Mixl1*, the mouse ortholog of MIXL1 was detected at day 5.5 post coitum (dpc) in the visceral endoderm (3, 4), primitive streak and in the nascent mesoderm between days 6.5-8.0 dpc (3-5). *Mixl1*<sup>-/-</sup> embryos died before days 10.5 dpc with multiple defects (6). Enforced expression of *Mixl1* in mouse ES cells promoted mesodermal, hemangioblastic and hematopoietic progenitors (7).

As a paired-type subfamily, Mix/Bix family proteins are characterized by 60 amino acid segments homologous to that of the paired type homeobox transcription factors. Paired-type homeodomain-containing proteins preferentially bind to DNA at a sequence motif of "TAAT". Mix/Bix family are classified as Q50 paired-types, denoting both a glutamine in the 50<sup>th</sup> amino acid in the homeodomain, and a preference for a 3 nucleotide spacer of sequence "TGA" between both halves of the dimer motif i.e TAATTGAATTA (8, 9). The preference to a TAATNNNATTA sequence was confirmed for mouse Mixl1(10).

While the Mix/Bix family is well characterized to bind as a dimer with other mesoderm-specific homeodomain-containing proteins, including other Mix/Bix family members, Gsc and Siamois (11), human MIXL1 was found to interact, in mammalian a high throughput two-hybrid screen, with homeodomain-containing proteins associated with a diverse set of lineages, including the neuronal-associated *POU4F2*, *OTX2*, and *DLX1*, thyroid-associated *NKX2-1*, and dermis-associated *ALX4* (12). In the same study, MIXL1 interacted with the basic helix-loop-helix transcription factor *NEUROD1*, implying that MIXL1 may have the potential to interact with non-homeobox proteins.

## **2. MIXL1 is regulated by TGF- $\beta$ /BMP Pathways**

One of the developmental roles of the BMP ligand BMP4 is the induction of mesoderm during early embryogenesis, and in *Xenopus*, Mix.1 is an necessary intermediate for BMP4 induction of the mesoderm (1). This translates to humans as well; when human embryonic stem cells are exposed to BMP4, mesoderm differentiation is induced (13). Indeed, both Activin A and BMP4 can induce *MIXL1* expression in human embryonic stem cells (14-16), and Mix.1 is an intermediate for BMP4-mediated dorsal-ventral patterning in *Xenopus* (17). The promoter regions for both human MIXL1 and mouse Mixl1 contain multiple SMAD binding elements (12).

TGF- $\beta$  ligands have also been implicated in the regulation of *MIXL1* expression. TGF- $\beta$  signaling can activate *Mixl1* expression in mouse in collaboration with the transcription factor FoxH1 (18), and FoxH1 can collaborate

with Gsc to suppresses *Mixl1* expression (19). MIXL1 may be transcriptionally regulated by the TGF- $\beta$  superfamily of cytokines through TGF- $\beta$ /BMP receptors to activate the SMAD family of transcription factors.

### **3. MIXL1 in Hematopoiesis and Hematopoietic Neoplasms**

While MIXL1 is well characterized in the role of mesoderm formation during embryogenesis, its roles in post natal homeostasis have gone uncharacterized. In adults, *MIXL1* is expressed in hematopoietic progenitor cells (2). This has implicated a potential role of MIXL1 in hematopoiesis, and indeed loss of *Mixl1* in mouse embryonic stem cells leads to significant hematopoietic defects (20). Additionally, *MIXL1* is highly expressed in Hodgkin's, Burkitt's and Diffuse Large B Cell Lymphoma (21), and leukemia and lymphoma cell lines (2), implying a role in the pathogenesis of hematopoietic neoplasms, leukemia and lymphoma progression.

In mouse, enforced expression of *Mixl1* in hematopoietic stem cells resulted in a transplantable acute myeloid leukemia (AML) with a 200 day latency in 100% of mice (22). Furthermore, enforced expression of *Mixl1* in hematopoietic stem cells conferred abnormal self-renewal potential to granulocytic precursors (23). Additionally, viral integration into the *Mixl1* locus in mice was identified as a potential collaborator to *p27Kip1* loss in inducing T-cell leukemia (24). MIXL1 therefore may play a significant role in a subset of hematopoietic neoplasms.

#### **4. Homeobox genes in Acute Myeloid Leukemia**

A large body of evidence implicates the deregulation of homeobox genes is important for the progression of AML and is often associated with the poorer outcomes. Aberrant expression of both type I clustered homeobox-containing proteins, including HOXA9 and HOXB3 (25-27), and type II non-clustered homeobox genes, including HLX and CDX2, have been reported (28-32). Non-random chromosomal translocations resulting in homeobox-containing proteins are also a common occurrence in AML, and further underscore the importance and leukemogenic potential of aberrant expression of homeobox proteins (33, 34). While homeobox-containing proteins expression is tightly controlled and each protein is only expressed at a restricted stage of differentiation during normal hematopoietic maturation, homeobox-containing proteins are constitutively expressed in AML. A recurrent functional consequence of all homeobox genes over expression is self-renewal, a limiting step in AML pathogenesis (35). While many homeobox genes identified are strongly associated with AML are considered markers of poor prognosis, whether there is overlap in expression between many of the identified homeobox proteins in AML is unknown. Similar hematopoietic-relevant homeobox genes may have a similar function in transformation of the remaining cases.

## 5. NF- $\kappa$ B Pathway in Leukemia and Lymphoma

The NF- $\kappa$ B family includes transcription factors that regulate stress response, proliferation, apoptosis, cell growth, and immune response. Each active NF- $\kappa$ B complex consists of a dimer between a NF- $\kappa$ B/REL protein including a transactivation domain (RELA, RELB, and c-REL), and one that does not (NF $\kappa$ B1 or NF $\kappa$ B2). Of the five NF $\kappa$ b genes, *c-REL* is a proto-oncogene as the chicken reticuloendotheliosis viral ortholog *v-rel* transforms hematopoietic cells. (36).

The activation of NF- $\kappa$ B complexes has been implicated in both hematopoietic progenitor self-renewal and leukemia and lymphoma progression. In hematopoietic stem cell transplantation, knockout of both *NF- $\kappa$ B2* and *RelB* impairs engraftment due to increased self-renewal capacity (37). In Diffuse large B-cell lymphoma, NF- $\kappa$ B complexes are commonly activated, and elevated *c-REL* expression in the germinal center subtype correlates with poor survival (38). NF- $\kappa$ B complexes are also constitutively active in leukemic progenitor cells and in acute myeloid leukemia (39).

One of the functional consequences of NF- $\kappa$ B activation is the transcriptional activation of the *BCL2* (*B-cell CLL/lymphoma 2*) family. The BCL2 family consists of three proteins (BCL2, BCL2A1, BCL2L1), all of which function as anti-apoptotic factors by inhibiting the release of Cytochrome C release during caspase-mediated apoptosis. Overexpression or amplification of *BCL2* are common pro-survival modifications in cancer, and recurrent chromosomal



translocations resulting in constitutive expression of *BCL2* is common of diffuse large B-cell lymphoma (40).

## **6. Clarifying the role of *MIXL1* in Acute Myeloid leukemia**

While *MIXL1* has been implicated in acute myeloid leukemia, little is known about its overall role in leukemia. I hypothesize that *MIXL1* is regulated by the BMP4 pathway in hematopoiesis and leukemia cell lines to promote proliferation through transcriptional activation of downstream targets. In the following pages, I identify a novel direct transcriptional target of *MIXL1* in leukemic cell lines, the proto-oncogene *c-REL*, by chromatin immunoprecipitation (ChIP) and mRNA expression analysis. I also identify that *MIXL1* is induced in normal hematopoietic stem progenitors by BMP4. Consequently, BMP pathway inhibitor LDN-193189 is sufficient to impair growth in *MIXL1*-expressing AML lines *in vitro*.

## **Chapter 2. Materials and Methods**

### **1. Cell Culture**

AML cell lines U937, HL-60, OCI-AML2, ML3, and chronic myeloid leukemia cell line K562, were grown under conditions of 5% CO<sub>2</sub> and 37°C in RPMI 1640 with 10% FBS. AML cell line KG-1 was grown in RPMI 1640 with 20% FBS. Human embryonic cell line HEK-293T was grown in DMEM with 10% FBS under 5% CO<sub>2</sub> and 37°C.

Human cord blood hematopoietic stem progenitors cells (CD34)<sup>+</sup> from three subjects denoted 47, 51, and 60, were provided by Drs. Lisa St. John and Jeffrey Molldrem.

### **2. Establishment of Overexpression and Knockdown Cell Lines**

Inducible expression cell lines for FLAG-HA-tagged MIXL1 were previously established using U937T cells containing pTET-VP16PURO by electroporation following a protocol previously described by the lab (41). Clones used are denoted U937.1MIXL and U937.2MIXL, while an empty vector control is denoted U937.Control.

Lentiviral Production: Viral particles were generated in HEK-293T by transient transfection of the lentiviral expression construct, envelope construct pCMV-VSV-G, encoding the viral envelop protein VSV-G, and a gag-pol-encoding construct, at a ratio of 2:1:1 respectively. Lentiviral constructs were purchased from OpenBiosystems (Pittsburg, PA) and designated as such: MIXL1 KD1 = TRCN0000019155, MIXL KD2 = TRCN0000019156, and c-REL

Expression = ccsbBroad304\_11094. The cells were incubated for 48 hours at 37°C before harvesting the media containing viral particles. Viral particles were purified by passage through a 45 µm filter, and stored in -80°C until use.

Lentiviral transductions were performed by resuspending  $2 \times 10^5$  cells in 1 ml of virus-containing conditioned medium with 8 µg/ml polybrene, and then incubated at 37°C for 24 hours. After the incubation, the cells were pelleted by centrifugation and resuspended in the growth medium. Growth assay and expression experiments were performed 48 hours after transduction.

### **3. Chromatin Immunoprecipitation**

Chromatin Immunoprecipitation was based on the one used by Chadee et al (42), with the following modifications. For each U937 cell line, a total of  $10^8$  cells were cross-linked with 1% formaldehyde in growth media at 37°C for 20 minutes. Cells were harvested by centrifugation at 3000 RPM for 10 minutes, and resuspended in 500 µL RIPA lysis buffer. After 10 minutes on ice, the cells were sonicated 20 times at 4-5 watts for 20 seconds, with a rest time of 40 seconds between each sonication. The samples were then centrifuged for 5 minutes at 4°C and precleared with 20 µL A/G agarose slurry for 1 hour at 4°C. Two aliquots from each lysate was processed as follows (i) Flag-IP: 240 µL of the lysate, 250 µL lysis buffer, and 8.4 µg mouse anti-flag antibody (anti-Flag-M2, Sigma), (ii) IgG-IP 240 µL of the lysate, 250 µL lysis buffer, and 8.4 µg mouse IgG. The samples were incubated overnight at 4°C with rotation, and then incubated for an additional hour with 20 µL A/G agarose slurry. The agarose

beads were recovered by centrifugation and washed for 15 minutes each in: RIPA lysis buffer, high salt RIPA buffer, lithium chloride RIPA buffer, and finally TE prior to Proteinase K and RNase treatment at 37°C overnight. The samples were incubated for 6 hours at 65°C to reverse crosslinking. DNA was precipitated overnight at -20°C in 75% ethanol, and then washed twice in 75% ethanol. For each sample DNA from a 20 µL aliquot of the pre immune precipitation lysate diluted with 180 µL TE buffer served as another control for target amplification. The DNAs were resuspended in 50 µL ddH<sub>2</sub>O and the Pre-IP sample was then diluted with 450 µL of ddH<sub>2</sub>O.

50 ng each of immuno precipitated DNAs from (1) U937 Control IgG-IP, (2) U937 Control Flag-IP, (3) U937 1MIXL Flag-IP, and (4) U937 2MIXL-IP were used to construct libraries by the Sequencing and Microarray Facility (SMF) at MD Anderson Cancer center using the Beckman SPRIworks system. Illumina analysis pipeline GAPIipeline-1.5.0 was used for base calling and alignment to human genome. Peak calling was done by MACS v1.3.7.1 at  $p\text{value} \leq 1e-5$ . Peaks were identified against the human genome (UCSC assembly hg18, NCBI36) using genome model-based analysis of ChIP-Seq [MACS](43), and were generated by normalizing to the two control samples in three combinations: Flag-1MIXL to IgG-Control and Flag-Control, Flag-2MIXL to IgG-Control and Flag-Control, and Flag-1MIXL and Flag-2MIXL combined to IgG-Control and Flag-Control. The primary dataset used for analysis was overlapping peaks in all three analyses. The combined dataset was tested for predicted Paired-Q9 binding motifs(8) using Motif alignment and Search Tool [MAST](44), and

enriched motifs were identified using Multiple EM for Motif Elicitation [MEME](45).

Peaks were allocated to genes within a 25kb region, designated as follows: upstream – 5 to 25 kb upstream of the transcription start site, promoter – 0 to 5 kb upstream of the transcription start site, body – between the transcription start and end, TSE – 0 to 5 kb downstream of the transcriptional end, downstream – 5 to 25 kb downstream of the transcriptional end, and distant – not allocated to a gene.

ChIP-qPCR confirmation was performed by SYBR green quantitative PCR using primersets for each peak region identified and an exonic *c-REL* region primer set for control, as listed in Table 1. For ChIP-qPCR against endogenous proteins, 5 µg anti-MIXL1-N and anti-MIXL1-C antibodies (2) and anti-MZF1 (sc-66991, Santa Cruz Biotechnologies, Dallas, TX) were used for immunoprecipitation.

#### **4. Global Expression Analysis**

The original microarray analysis was performed previously by Dr. Hong Liang. To identify potential targets of the MIXL1 transcription factor, Dr. Liang performed global expression profiling analysis on MIXL1-expressing cells by microarray. The cell lines 1MIXL and Control were cultured without TET for 24 hours at 5X10<sup>4</sup> cells/mL. RNA was extracted from 5x10<sup>6</sup> cells by RNEASY Minikit (QIAGEN, Valencia, CA). Extracted RNA was hybridized against a Human Genome Affymetrix HG133A Microarray (Affymetrix, Santa Clara, CA).

For analysis, dChip analyzer software (46) was used, normalizing the 1MIXL dataset to the Control dataset, and gene expression models were obtained through PM (Perfect Model)-only approach. Differentially expressed genes were defined as genes in which the difference between the detected expression levels was at least 100, and the ratio was at least 1.2.

The networks analyses were generated through the use of Ingenuity Pathway Analysis (Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com)).

## **5. RT-QPCR**

RNA was purified using RNeasy Mini Kit from  $5 \times 10^6$  cells for each line. 200 ng from each sample was then reverse transcribed and diluted 10-fold. The samples were assayed by qPCR in triplicate using Taqman primers obtained from Applied Biosystems/Life Technologies (Carlsbad, CA) and Integrated DNA Technologies (Coralville, Iowa), as denoted in Table 2, and then quantified by delta-delta-CT method.

For CD34<sup>+</sup> cells, RNA was harvested in triplicate from ~2500 cells by RNeasy Micro Kit + (QIAGEN, Valencia, CA) 2 hours after addition of either 50 ng/ml BMP4 (314-BP, R&D Systems, Minneapolis, MN) or 2 ng/ml TGF- $\beta$ 1 (Sigma-Aldrich) in X Vivo-15 medium (Lonza, Allendale, NJ), then 100 ng from each sample was reverse transcribed and assayed.

**Table 1. Primer sets used for ChIP-qPCR.** Listed are the primers used for quantitative PCR analysis of the Chromatin Immunoprecipitation techniques. Each primer set is will generate a 100-200 bp product encoded near the center of the ChIP-Seq identified gene locus peak, with exception of Rel-Int-F/-R, which covers a 100-200 bp region internal of the *REL* gene locus not enriched in the ChIP-Seq analysis.

<b>Primers used ChIP-qPCR</b>	
Rel-Int-F	5-TTACCAGGATTTTGGCAAGG-3
Rel-Int-R	5-CAGGCAGTTTGGGGATAAGA-3
Rel-F	5-GGAACCCACCTCTCGAAAACC-3
Rel-R	5-TCCAGGTTGTTCTTCCGAGT-3
EIF1-F	5-TGACTCCGTGGGTAGTAGGG-3
EIF1-R	5-CCTTCTTGACCCTGTTGCAT-3
SLC39A13-F	5-CCTGAGGTTCCCAGTGAAAA-3
SLC39A13-R	5-GAGGACTACTGTGCGCTCCT-3
SMYD5-F	5-TTCCCCCTTTCATGACTCTG-3
SMYD5-R	5-CTCAGCTCAGTCCCCAAGAG-3
ZP3-F	5-ACCTCAGCCTCCCCAGTAGT-3
ZP3-R	5-TTGATCCAAAAGCAGCTGAA-3



**Table 2. Probe-sets used for RT-QPCR analysis.** Listed are the probes used for reverse transcription-coupled quantitative PCR for each of the respective genes. Probes for *MIXL1*, *REL*, *BCL2L1*, and *BCL2A1* were supplied by Applied Biosystems (Carlsbad, CA), while probes for *APBB2*, *IL18R1*, and *PCGF2* were customized and ordered from Integrated Integrated DNA Technologies (Coralville, Iowa), to be fully compatible with TaqMan analysis.

Probes used for qPCR	
MIXL1	Hs00968440_m1 TaqMan, Applied Biosystems
REL	Hs00231279_m1 TaqMan, Applied Biosystems
BCL2L1	Hs99999146_m1 TaqMan, Applied Biosystems
BCL2A1	Hs00187845_m1 TaqMan, Applied Biosystems
EGR1	Hs00152928_m1 TaqMan, Applied Biosystems
APBB2	Probe 5-CTCCCAAATCCCGACTGTGTCT-3 5' 6-FAM, 3' Iowa Black FQ Primer1 5-GATGGGTAGAGATGGCAGAAG-3 Primer2 5-GGATCAGGTACATGTCTTTCCC-3 Integrated DNA Technologies
IL18R1	Probe 5-CCTCCAGGCACTACATCCCTTTCAA-3 5' 6-FAM, 3' Iowa Black FQ Primer1 5-CACCTTTGCTGTGGAGATTTTG-3 Primer2 5-TCTTCGGCTTTTCTCTATCAGTG-3 Integrated DNA Technologies
PCGF2	Probe 5-TGGACATCGCCTACATCTACCCCT-3 5' 6-FAM, 3' Iowa Black FQ Primer1 5-GACGAGCCCACTGAAGGAATAC-3 Primer2 5-GCTGGACACGGTACTTGAG-3 Integrated DNA Technologies

## 6. Co-Immunoprecipitation

$10^7$  cells of U937-1MIXL were lysed with 1 ml ice-cold RIPA buffer (150 mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris (pH 7.5), 5  $\mu$ g/ml leupeptin, 8  $\mu$ g/ml aprotinin, 1 mM PMSF, 0.7  $\mu$ g/ml pepstatin A and 1 mM EDTA (pH 8.0)). 900  $\mu$ l of the lysate was precleared by incubation with protein A/G-agarose beads for 1 hour at 4°C. 5  $\mu$ g M2 anti-flag monoclonal antibody and 5  $\mu$ g mouse IgG antibody were each added to separate aliquots of 450  $\mu$ l cleared lysate, and incubated for 2 hours at 4°C. 20  $\mu$ l of protein G-agarose beads was added to the immune complexes and incubated for 1 hour at 4°C. After four washes with NET-N buffer (100 mM NaCl, 1 mM EDTA (pH 8.0), 20 mM Tris (pH 8.0), 0.2% NP-40, 5  $\mu$ g/ml leupeptin, 8  $\mu$ g/ml aprotinin, 1 mM PMSF, 0.7  $\mu$ g/ml pepstatin A), the precipitated proteins were eluted in 20  $\mu$ l of 2 $\times$  SDS loading buffer (100 mM Tris (pH 6.8), 200 mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol) and then resolved by 10% NuPAGE gel (Invitrogen, Carlsbad, CA).

## 7. Immunoblotting

Cells were lysed in Whole Cell lysis buffer (20 mM Tris, 250 mM NaCl, 2 mM EDTA, 1% Triton X-100, 1 mM DTT, 2  $\mu$ g/ml Aprotinin, 2  $\mu$ g/ml Leupeptin, 2  $\mu$ g/ml Pepstatin A, 1 mM NaVO<sub>3</sub>, 1 mM PMSF), then denatured in 4x SDS buffer at 90°C for 10 minutes, then resolved on 10% NuPAGE gel (Life Technologies) using MOPS buffer, and transferred to a PVDF membrane (GE Healthcare, Pittsburgh, PA).

Membranes were first blocked by incubation in TBS-T (50 mM Tris, 150 mM NaCl, and 0.05% Tween 20 adjusted to pH 7.6) with 5% milk for one hour, after which they were washed in TBS-T two times for two minutes. The membrane was then incubated with the first antibody for 2 hours 2% milk TBS-T at room temperature then washed in TBS-T two times for 5 minutes, then once for 15 minutes. The blot was then incubated with the secondary antibody in 2% milk TBS-T for 1 hour, followed by another round of washes in TBS-T. Blots were visualized using Amersham ECL western blotting detection reagent (GE Healthcare) with 5 minutes of film exposure.

Primary antibodies used were anti-MIXL1-N at 1:1500, MZF1 (sc-66991, Santa Cruz Technologies, Dallas, TX) at 1:100, and beta-actin (Sigma Aldrich) at 1:5,000. The secondary antibodies used were anti-rabbit-HRP (GE Healthcare) at 1:10,000 for the majority of primaries and anti-mouse-HRP (GE Healthcare) at 1:7,000 for the beta-actin primary antibody.

## **8. Luciferase Reporter Assay**

Luciferase constructs were generated by PCR amplification of segments of the *c-REL* peak region identified by ChIP-Seq. Promoter segments were amplified from genomic DNA extracted from human placenta with the primers listed in Table 3 and sub-cloned into the pBV-Luc vector between Mlu1 and EcoR1 sites. Each luciferase reporter vector insert was then confirmed by sequencing. 293T cells were transfected with 200 ng of expression vector, 200 ng of luciferase vector, and 0.2 ng *Renilla* Luciferase vector by Lipofectamine

**Table 3. Primers used for REL Promoter Luciferase construct construction.**

Listed are the primer sets used to generate the *REL* promoter region inserts for construction of p-BV-Luc-based luciferase vectors for use in the luciferase reporter assay. Each forward primer contains a Mlu1 target sequence and each reverse primer contains an EcoR1 target sequence for restriction digest. Primers were used in pairs. First set: Rel-R-EcoR1 paired with Rel-FM-130, Rel-FM-500, Rel-FM-700, and Rel-FM-944, with the Forward primer name denoting length of product. Second set: Rel-FM-700 paired with Rel-RE-150, Rel-RE-200, Rel-RE-300, and Rel-RE-580, with the Reverse primer pair denoting length of product.

<b>Primers used for Luciferase Reporter Construction</b>	
Rel-R-EcoR1	5-ctgt-gaattc-CGCAGTCAGTCAGTCAGGAG-3
Rel-FM-130	5-ctgt-acgcgt-AGAATTCAGGGGTTGGGAAG-3
Rel-FM-500	5-ctgt-acgcgt-GGAAGAACAACCTGGAGGAG-3
Rel-FM-700	5-ctgt-acgcgt-GAACCACCTCTCGAAAACC-3
Rel-FM-944	5-ctgt-acgcgt-GGAGCTTTGGAGTCAGACAA-3
Rel-RE-150	5-ctgt-gaattc-CAGGTTGTTCTTCCGAGT-3
Rel-RE-200	5-ctgt-gaattc-GGCTAGCAGCGTGAGAAGG-3
Rel-RE-300	5-ctgt-gaattc-GACGCAGCAACCCTCACC-3
Rel-RE-580	5-ctgt-gaattc-AACCCCTGAATTCTTGAC-3

(Invitrogen, Grand Island, NY). The activity was then tested by Dual Luciferase Reporter Assay System (Promega, Madison, WI) 48 hours post-transfection.

## **9. MTS Assay**

The growth assay for comparison of knockdown cell lines were performed by plating  $3 \times 10^4$  cells/well for each line into a 96-well plate in triplicate, then testing the culture growth by MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay every 24 hours of incubation.

For comparative analysis of the effects of doxorubicin on U937 MIXL1-overexpressing lines, the three clones (Control, 1MIXL, and 2MIXL) were seeded to a 96-well plate at a density of  $3 \times 10^4$  cells/well. Doxorubicin (Sigma-Aldrich, St. Louis, MO) was added to each well at 0 – 2  $\mu$ M in triplicates, for 24 hours before MTS treatment.

To measure survival response to high concentrations of LDN-193189, each cell line was plated at  $3 \times 10^4$  cells/well on 96-well plates and treated with either 3  $\mu$ M LDN-193189 (Cellagen, San Diego, CA) or the equivalent concentration of DMSO in triplicate. Each line was then tested by MTS assay immediately, then again every 24 hours.

For the dose response experiments for LDN-193189, each line was plated at  $3 \times 10^3 \times 10^4$  cells/well on 96-well plates and treated with 0 – 0.7  $\mu$ M LDN-193189 in triplicate. The cells were then incubated for 4 days, with additional treatments of 0 – 0.7  $\mu$ M LDN-193189, before being tested by MTS assay.

The MTS assay was performed by the CellTiter 96 AQueous Non-Radioactive Cell Proliferation Assay kit (Promega) with an incubation time of 1 hour.

## **10. Analysis and Data mining**

TCGA AML database (47) was accessed and analyzed through the cBioPortal (48, 49) for cases either with a mutation, copy-number alteration, or expression change of 2-fold or over (1.0 threshold).

Identification of transcription factor binding motifs in the *c-REL* promoter was performed using TFSearch. (50)



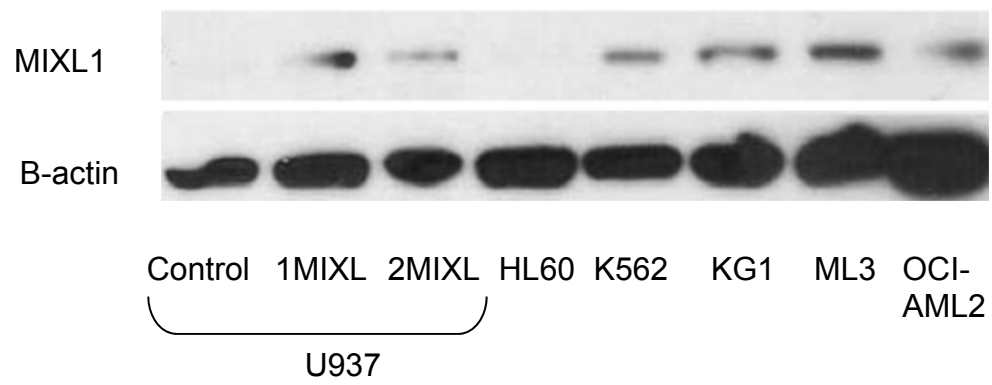
## **Chapter 3. Results**

### **1. Characterization of AML cell lines with overexpression of *MIXL1***

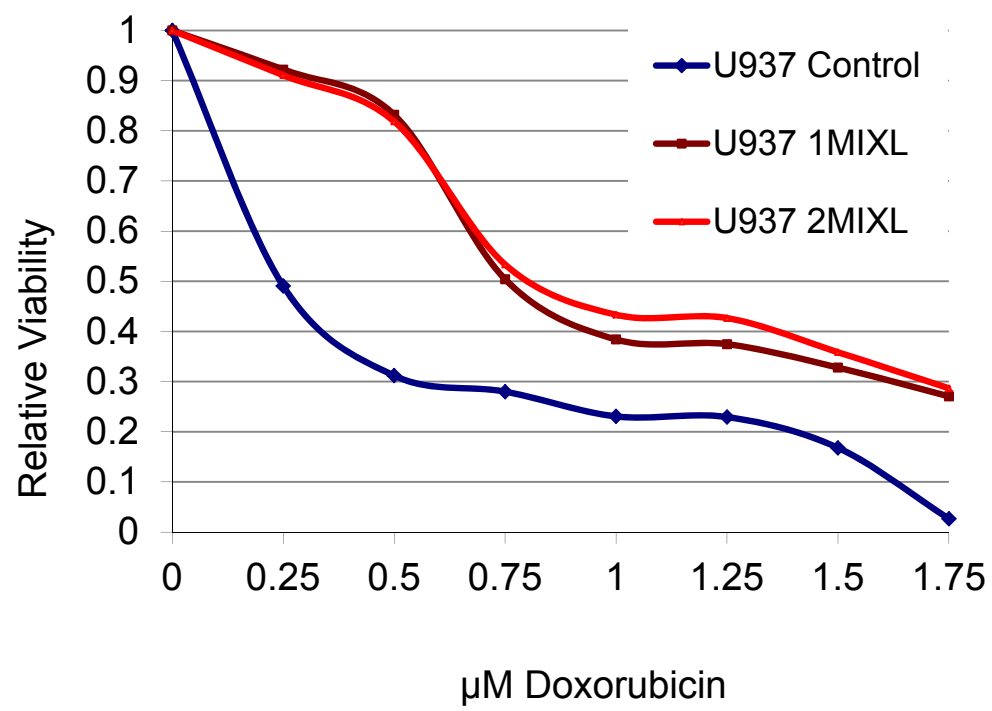
To identify the role of MIXL1 in AML, I used FLAG-HA-tagged MIXL1 inducible expression cell lines established in U937, a histiocytic lymphoma, previously created in the lab, denoted U937.1MIXL and U937.2MIXL with an empty vector control line U937.Control. While these clones were initially established with MIXL1 expression inducible by tetracycline withdrawal, after multiple passaging both clones exhibited enforced MIXL1 expression regardless of tetracycline induction, and were subsequently treated as overexpression lines. To determine if the levels of MIXL1 protein in the overexpression lines was comparable to endogenous protein levels in leukemia, I compared the protein levels of 1MIXL and 2MIXL to that of three AML lines, the high-expressing KG1 and ML3 and low-expressing HL60, where MIXL1 protein levels had already been tested (2), the previously untested AML OCI-AML2, and the CML line K562 (Figure 1). Protein expression in 1MIXL and 2MIXL was comparable to that of the expressing cell lines K562, KG1, ML3, and OCI-AML2, while U937.Control and HL60 had no detectable levels of MIXL1.

To test if MIXL1 infers a survival advantage in leukemic cell lines, I performed a MTS assay comparing the survival of the U937 subclones under 24 hour exposure to 0 – 2  $\mu\text{M}$  of the leukemia chemotherapy drug doxorubicin (Figure 2). After 24 hours of treatment, doxorubicin had a  $\text{LD}_{50}$  of 0.25  $\mu\text{M}$  for the U937.Control line and 0.75  $\mu\text{M}$  for 1MIXL and 2MIXL, with only the Control line nearly completely eliminated at the 1.75  $\mu\text{M}$  concentration of doxorubicin.

**Figure 1. Stable transfectants of U937 cells express MIXL1 at levels similar to endogenous protein in AML cell lines.** MIXL1 detected by probing 30ug whole cell lysates resolved on SDS-PAGE, transferred to PVDF membrane, with rabbit antibodies against N-terminal epitope of MIXL1 and  $\beta$  actin for loading control.



**Figure 2. MIXL1 expression reduces sensitivity of U937 cells to doxorubicin.** The cell lines were treated with 0  $\mu$ M to 1.75  $\mu$ M Doxorubicin on Day 0. Cell survival was measured at 24 hours by MTS assay. Absorbance of untreated cells was normalized to 1. Relative viability at varying concentrations of doxorubicin is denoted.



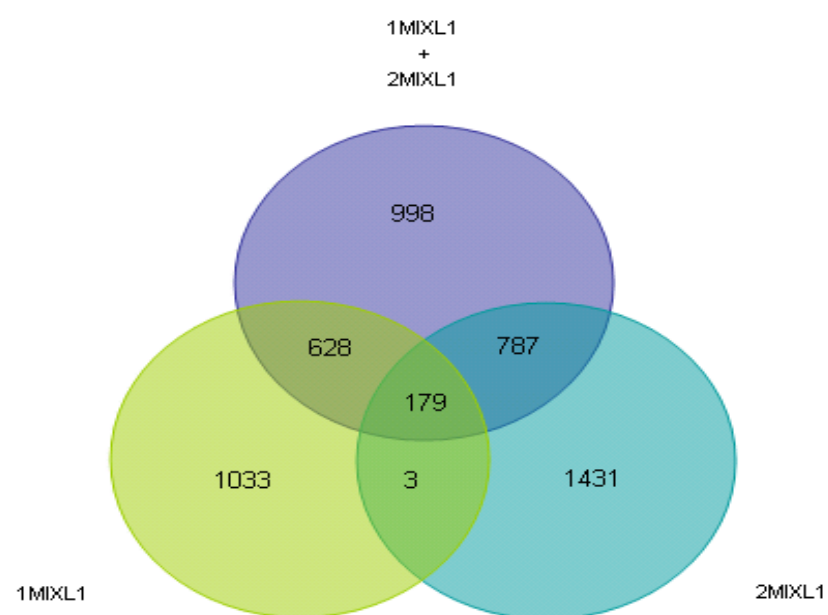
## **2. ChIP-Seq analysis of *MIXL1* overexpression lines to identify transcriptional targets**

As few transcriptional targets for MIXL1 in hematopoiesis have been identified, I performed ChIP-Seq in our overexpression clones to identify transcriptional targets of MIXL1. Flag-HA-MIXL1 and bound DNA segments were immunoprecipitated from both 1MIXL and 2MIXL by FLAG antibodies, alongside two control samples of FLAG-IP in U937.Control with no expected target, and IgG-IP in U937.Control. The sequenced samples were then aligned to the genome and analyzed in different pairwise combinations by Dr. Yue Lu and Dr. Shoudan Liang: 1MIXL-Flag normalized to Control-Flag and Control-IgG, 2MIXL-Flag normalized to Control-Flag and Control-IgG, 1MIXL-Flag and 2MIXL-Flag combined then normalized to Control-Flag and Control-IgG. Each of these sets identified over 1500 peaks, though only a small subset containing 179 peaks was present in all four groups (Figure 3). The overlap subset shared by all three groups was isolated and used in further studies (Appendix).

As a homeobox transcription factor, we expected MIXL1 to primarily interact with promoter regions of gene loci. To determine if this was the case, Yue Lu and Dr. Shoudan Liang identified where each peak in the 179 peak set was in relation to its nearest gene loci (Appendix and Figure 4). Out of the 179 peaks, 64% of the peaks localized to the promoter region (within 5 kb upstream of the transcription start site), 8% were localized to transcribed gene regions, 4% were between 25 kb and 5 kb upstream of the transcription start site, 1% were

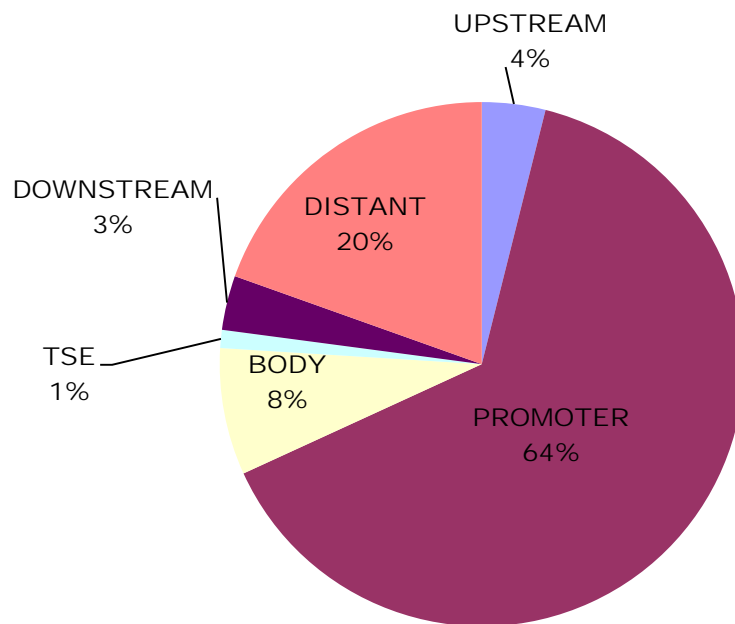
**Figure 3. Multiple analyses of MIXL1-ChIP-Seq identify 179 high-quality peaks.** Venn diagram of all the peaks identified using three different analyses of the two experimental (Flag immunoprecipitation in U937 1MIXL or 2MIXL) and two negative control (Flag or IgG immunoprecipitation in U937 Control) : 1MIXL vs Controls, 2MIXL vs Controls and 1MIXL+2MIXL vs Controls. Out of the peaks identified, 179 peaks were significantly enriched in all sets.

This figure was generated as part of the statistical analysis performed by Dr. Yue Lu and Dr. Shoudan Liang for ChIP-seq.





**Figure 4. The majority of peaks for MIXL1 identified by ChIP-Seq localize to the Promoter region.** Pie chart depicting the localization of MIXL1 to human genome. Peaks were classified based on distance to nearest transcribed gene locus, using the following criteria: upstream – 5 to 25 kb upstream of the transcription start site, promoter – 0 to 5 kb upstream of the transcription start site, body – between the transcription start and end, TSE – 0 to 5 kb downstream of the transcriptional end, downstream – 5 to 25 kb downstream of the transcriptional end, and distant – not allocated to a gene. Note that majority of peaks (67%) localized to gene promoters.



within 5 kb of the transcription sequence end, and 3% were between 25 kb and 5 kb of the transcription sequence end. 20% of the peaks were classified as not being within 25 kb of any known genes.

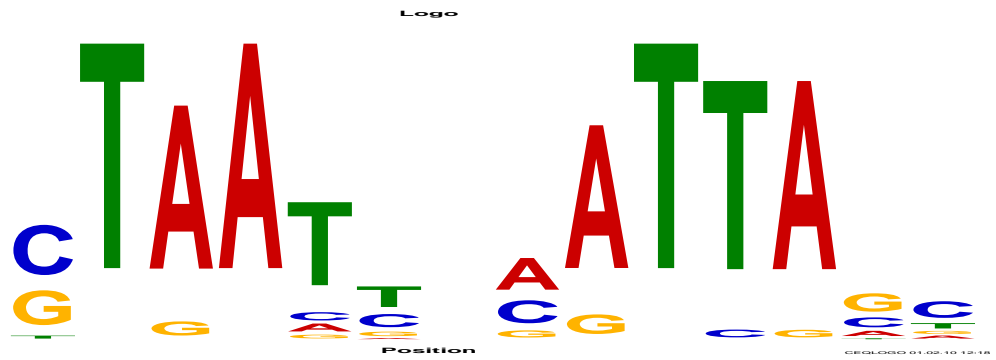
As stated in the introduction, MIXL1 is a paired type homeobox with a previously characterized predicted motif (8). To determine if this motif is how MIXL1 is bound to the identified motif locations, Dr. Yue Lu and Dr. Shoudan Liang analyzed the 179 peak set (Appendix) and a comparable randomized DNA fragment set for two different variants of the motif using the Motif alignment and Search Tool [MAST] (44) (Figure 5). Neither motif was significantly over-represented in the 179 peak set compared to the random set; therefore, the expected motif for MIXL1 is not enriched in our peak set. To characterize what motifs were enriched in our peaks, Dr. Yue Lu and Dr. Shoudan Liang next performed Multiple EM for Motif Elicitation [MEME] (45) for the sequences 200 base-pairs around each peak summit. The two highest peak motifs (Figure 6) are CG-rich and show no similarity to the AT-rich predicted homeobox motifs (TAAT); instead, they are more similar to zinc finger motifs.

To confirm the peak regions enriched in SET4 of the ChIP-seq analysis, I reanalyzed five of those regions, peaks near *EIF1*, *c-REL*, *SLC39A13*, *SMYD5*, and *ZP3*, by ChIP-qPCR in U937.Control, 1MIXL and 2MIXL by ChIP-qPCR using the Flag-antibody for immunoprecipitation (Figure 7). While each peak was enriched by ChIP-qPCR, only EIF1, SLC9A13, and SMYD5 had the scale of enrichment over the control samples that were seen in the ChIP-seq. This may be due to variability between the techniques.

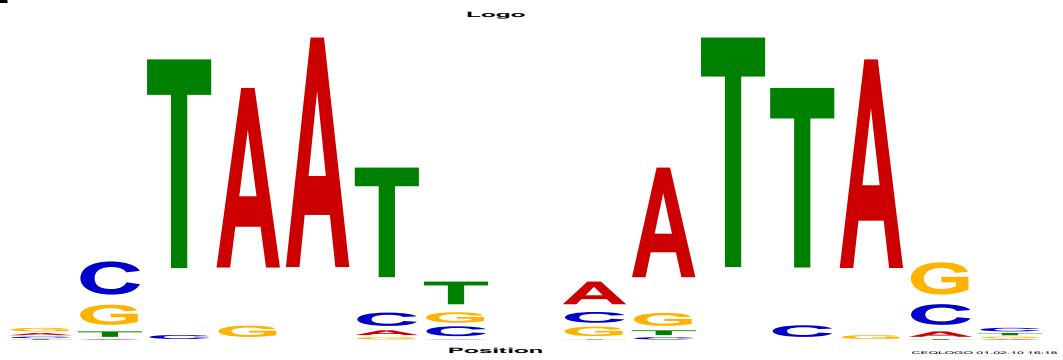
**Figure 5. Predicted MIXL1 motifs are not enriched in U937 overexpression MIXL1 ChIP-seq Peaks.** Motif1 and Motif2 were previously characterized putative MIXL binding motifs, and were searched against the stringent set of peaks (Appendix) at various distances (1000bp, 500bp, 200bp, or 100bp as listed) from the peak apex (top lines) and from the center of randomized sequences of the same length (bottom lines) , at either a p-value of  $1e^{-4}$ ,  $5e^{-4}$ , or  $1e^{-3}$ . In all cases, the likelihood of finding the MIXL1 binding motif was either not statistically significantly different or even lower than the likelihood of identifying the motif in randomized sequences, therefore the MIXL1 binding motif is not enriched in this dataset.

This figure was generated as part of the statistical analysis performed by Dr. Yue Lu and Dr. Shoudan Liang for ChIP-seq.

## Motif1



## Motif2



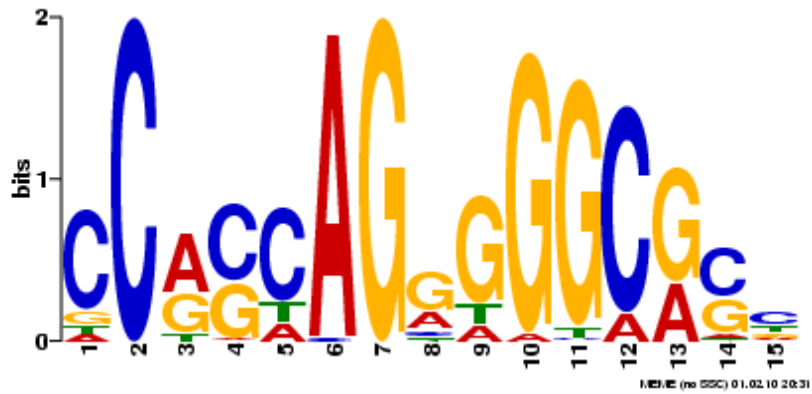
<b>Motif1</b>	(179 peaks)		
<i>p</i> value	1e-4	5e-4	1e-3
1000bp	3 4.97	31 44.6	59 77.75
500bp	1 1.65	9 18.2	18 35.65
200bp	0 0.43	2 5.76	3 12.06
100bp	0 0.18	0 2.75	1 5.62

<b>Motif2</b>	(179 peaks)		
<i>p</i> value	1e-4	5e-4	1e-3
1000bp	4 3	13 17.57	27 37.06
500bp	2 0.98	3 6.49	9 14.55
200bp	0 0.32	1 1.72	2 4.57
100bp	0 0.11	0 0.91	0 2.1

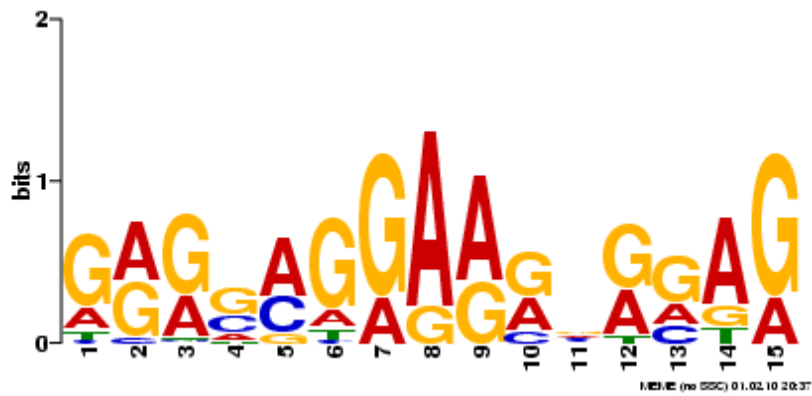
**Figure 6. The most commonly occurring motifs in the ChIP-seq peaks are Zinc-Finger binding motifs.** Motif1 and Motif2 were the two most statistically significant common motifs generated using the Multiple EM for Motif Elicitation [MEME] against the peak regions identified in the ChIP-seq analysis (Appendix). Both motifs are C/G-heavy regions of similarity to known zinc finger motifs.

This figure was generated as part of the statistical analysis performed by Dr. Yue Lu and Dr. Shoudan Liang for ChIP-seq.

**MOTIF 1** width = 15 sites = 65 E-value = 1.1e-027

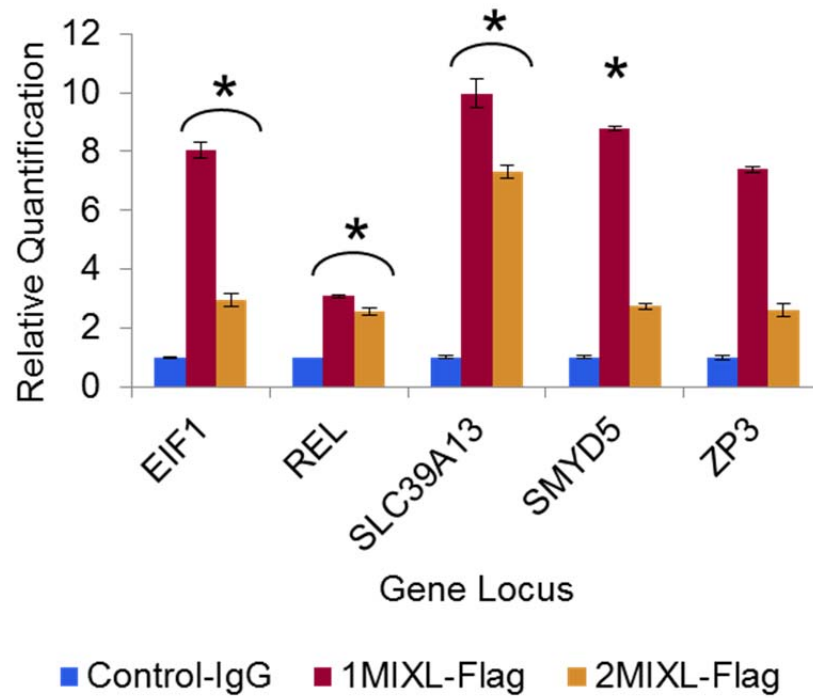


**MOTIF 2** width = 15 sites = 112 E-value = 3.6e-015



**Figure 7. ChIP-qPCR analysis of selected gene-associated peaks confirms enrichment.** ChIP of 5 candidate peaks with FLAG antibodies identified by ChIP-Seq (EIF1, c-REL, SLC39A13, SMYD5, ZP3) showed specific MIXL1 binding both 1MIXL and 2MIXL clonal ChIP with normal mouse IgG served as control. Error bars represent standard deviation between triplicates. \* denotes p value <0.05. The chart below is the ChIP-seq results for the constrained (179-peak) dataset for each of the peaks tested, for comparison.





ChIP-Seq Results for selected Peaks			
Gene	Peak Location	Enrichment	Proximity to Gene
EIF1	chr17:37097849-37099490	7.55	promoter
REL	chr2:60961441-60962652	9.79	promoter
SLC39A13	chr11:47385814-47387193	6.87	promoter
SMYD5	chr2:73293129-73293533	7.38	upstream
ZP3	chr7:75905120-75905486	20.74	body

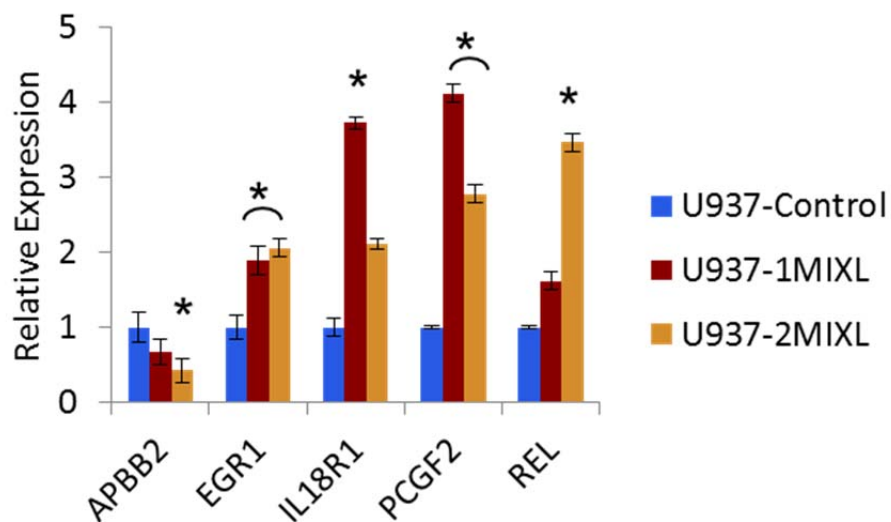
### **3. Global Expression Profiling of *MIXL1* overexpression to confirm transcriptional targets**

To better identify direct transcriptional targets of *MIXL1*, I took a previous global expression profiling experiment performed in the lab on U937.Control and 1*MIXL* using the U133 Plus 2.0 microarray by Dr. Hong Liang and reanalyzed it. To confirm the microarray analysis, a small subset of genes was tested by RT-qPCR in U937.Control 1*MIXL* and 2*MIXL* (Figure 8). While the expression of *APBB2* and *EGR1* were comparable between both methods at approximately -2-fold and +2-fold change, respectively, *IL18R1* expression was significantly higher in 1*MIXL* by qPCR at 3.5-fold in comparison to the 2.6 identified by microarray. Of specific note is *c-REL*, a gene also identified by ChIP-seq as a putative *MIXL1* transcriptional target, where both methods found a 1.66-fold increase in 1*MIXL*, and interestingly 2*MIXL1* expressed 3.5-fold over U937.Control by qPCR.

### **4. *c-REL* as a transcriptional target of *MIXL1***

Since *c-REL* was identified as a putative transcriptional target of *MIXL1* by both ChIP-Seq and global expression profiling, I decided to get a broader look at the expression of the overall NF- $\kappa$ B pathway (Figure 9). In the NF- $\kappa$ B pathway, NF $\kappa$ B1 or NF $\kappa$ B2 herterodimerize with RELA, RELB, or *c-REL* to transcriptionally regulate downstream targets, with the most common and best studied dimer pair being the RELA-NF $\kappa$ B1 dimer. One of the common transcriptional targets of the NF- $\kappa$ B dimers is the anti-apoptotic BCL2 family (BCL2, BCL2L1, and BCL2A1).

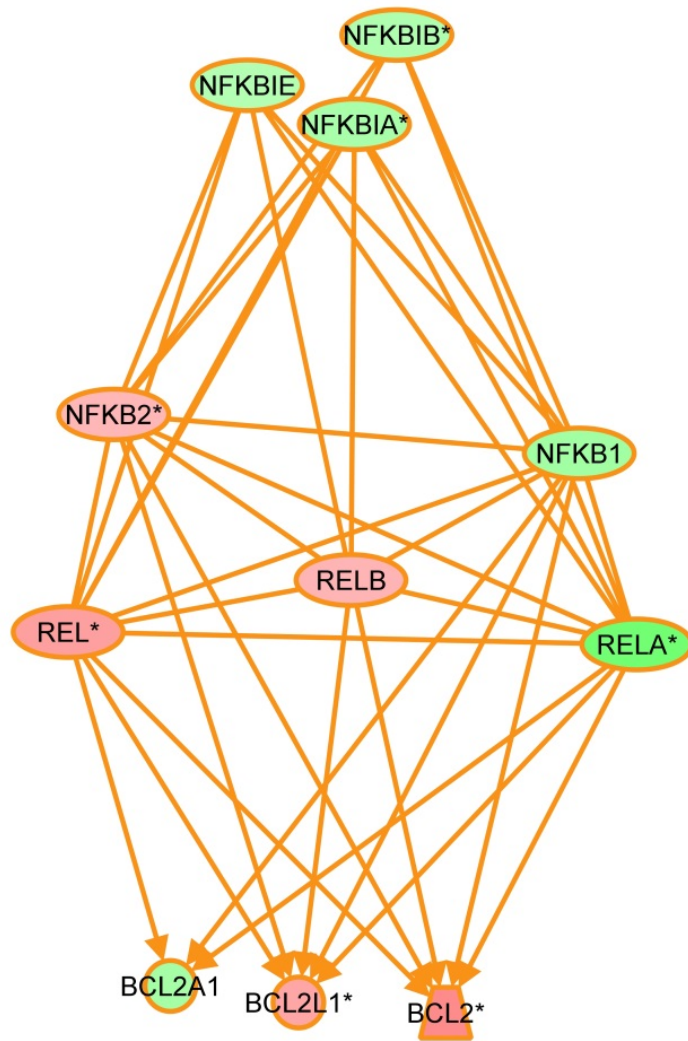
**Figure 8. RT-qPCR confirms change of expression in selected genes as identified by Microarray.** Quantitative RT-PCR confirms expression changes of selected genes identified by global expression profiling. Quantitative RT-PCR showing differences in gene expression between U937 control, 1MIXL and 2MIXL. Transcript levels were normalized to 18srRNA. Error bars represent standard deviation between triplicates. \* denotes p value <0.05. The chart below is the Micorarray results for the 5 genes tested, for comparison.



#### Microarray Results for selected genes

probe set	gene	Accession	EntrezGene	Fold Change
216747_at	APBB2	AK024871	323	-2.45
201694_s_at	EGR1	NM_001964	1958	2.78
206618_at	IL18R1	NM_003855	8809	2.63
240752_at	PCGF2	AW510760	7703	5.65
206035_at	REL	NM_002908	5966	1.66

**Figure 9. The NF- $\kappa$ B pathway is regulated by *MIXL1*.** The NF- $\kappa$ B pathway is regulated by MIXL1. Output from HG133A affymetrix microarray expression profiling was analyzed by Ingenuity Pathway analysis software (IPA). Green denotes decreased expression and red denotes increased expression relative to U937.control.



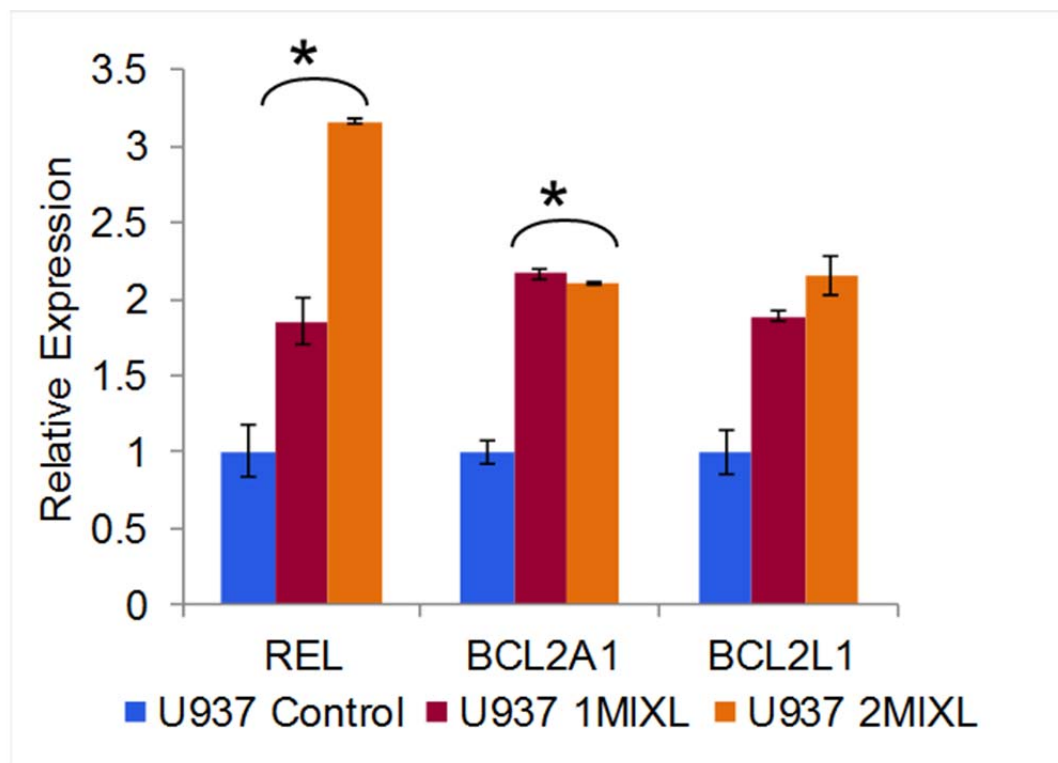
The activity of these NF- $\kappa$ B dimers can be inhibited by binding with I $\kappa$ B family proteins (NFKBIA/B/E), which block nuclear localization of the dimer. In the global expression profile under MIXL1 overexpression in 1MIXL, all these genes are differentially regulated: the I $\kappa$ B family members and both components of the abundant NF- $\kappa$ B dimer (*RELA* and *NF $\kappa$ B1*) are down-regulated, while the less abundant NF- $\kappa$ B members *c-REL* and *RELB* and the anti-apoptotic *BCL2L1* and *BCL2* are up-regulated. The anti-apoptotic *BCL2A1*, however, was down-regulated.

To confirm the change of expression characterized in the global expression profiling for *c-REL*, *BCL2A1* and *BCL2L1*, I analyzed RNA expression levels by RT-qPCR between U937.Control, 1MIXL and 2MIXL (Figure 10). Expression of *c-REL* increased approximately 1.8-fold in 1MIXL and 3-fold in 2MIXL, in agreement with the previous run (Figure 8); *BCL2L1* expression increased approximately 2-fold in each line in confirmation with the microarray results, while *BCL2A1* expression increased approximately 2-fold as well, raising the possibility of a potential technical difference between the two assays.

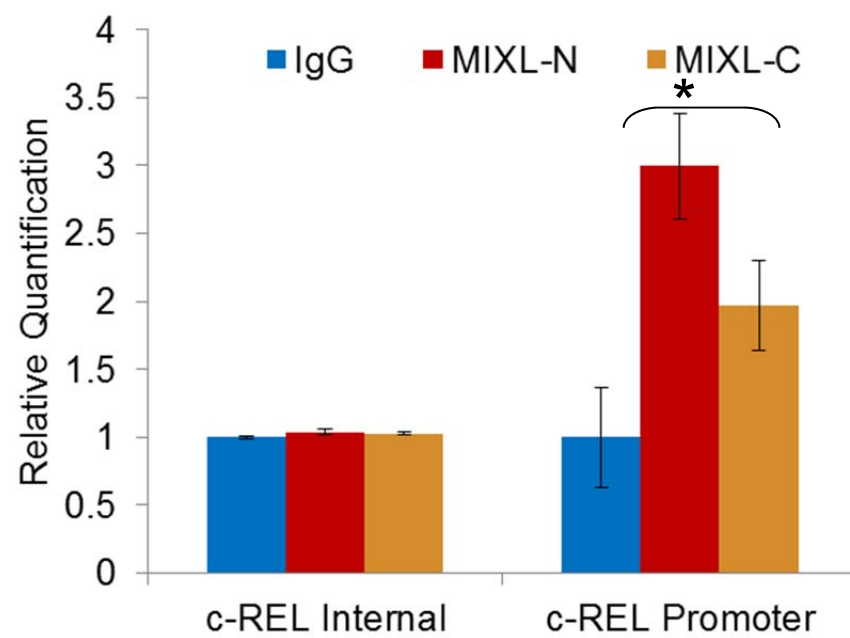
While I have identified *c-REL* promoter region as a target for MIXL1, it is possible that the perceived interaction is either an artifact of the Flag antibody or unique to U937 cells. To test both of these possibilities, I performed ChIP-qPCR on the endogenous C-REL promoter in the AML cell line KG1 using both our N-terminal and C-terminal MIXL1 antibodies (Figure 11). In both cases, the *c-REL* promoter was enriched: 3-fold for MIXL-N and 2-fold for MIXL-C, while the internal genomic locus was not.

**Figure 10. MIXL1 expressing clones show enhanced transcript levels for c-REL, BCL2A1, and BCL2L1.** Quantitative RT-PCR results showing the difference in expression between U937control, 1MIXL and 2MIXL for c-REL, BCL2A1, and BCL2L1. Expression was normalized to 18srRNA transcript levels. Error bars represent standard deviation between triplicates. \* denotes p value <0.05





**Figure 11. ChIP localizes endogenous MIXL1 to c-REL promoter in KG1 cells.** Quantitative genomic PCR analysis shows specific enrichment of endogenous MIXL1 immunoprecipitated with either N-terminal or C-terminal MIXL1 antibodies, on the c-REL promoter whereas an internal locus within the c-REL gene showed no MIXL1 occupancy. Error bars represent standard deviation between triplicates. \* denotes p value <0.05.



## **5. *MIXL1* knockdown in KG1 represses *c-REL* and *BCL2A1/L1* expression**

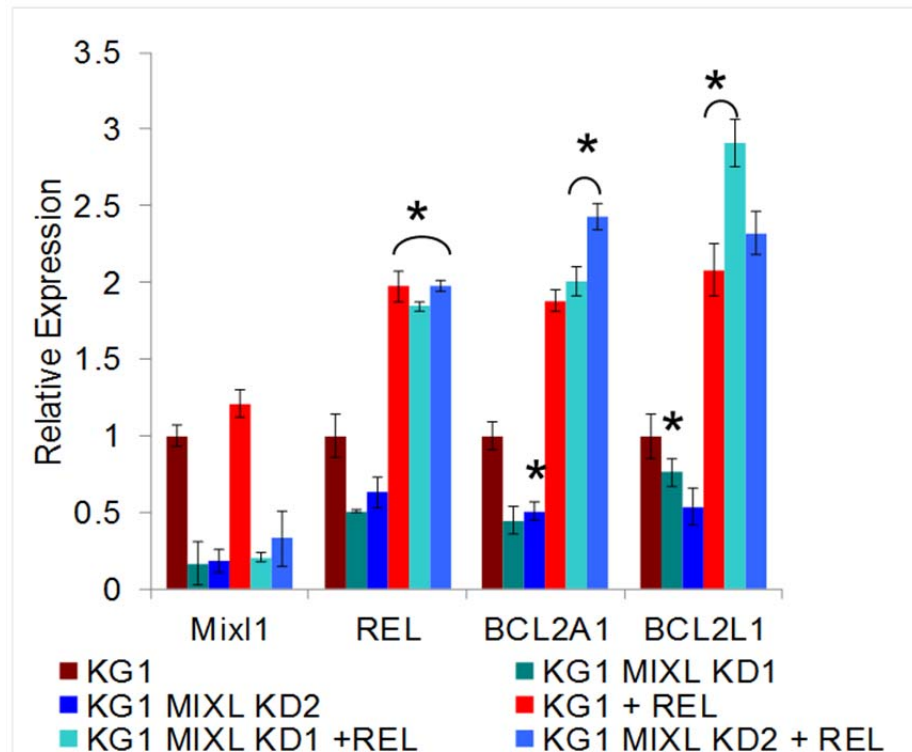
To confirm *MIXL1* regulation of *c-REL* and its downstream targets, I knocked down *MIXL1* expression by shRNA lentiviral transfection in the high-*MIXL1* expression AML cell line KG1, and then attempted to reverse the phenotypes by enforcing expression of *c-REL*. I first tested the expression of *MIXL1*, *c-REL*, *BCL2A1*, and *BCL2L1* under knockdown of *MIXL1* and enforced expression of *c-REL* by RT-qPCR (Figure 12). Under *MIXL1* knockdown, *MIXL1* expression is down to 0.25-fold, while expression of *c-REL*, *BCL2A1*, and *BCL2L1* down to about 0.6-fold, 0.5-fold, and 0.7-fold, respectively. With the enforced expression of *c-REL*, expression of *c-REL*, *BCL2A1*, and *BCL2L1* increase approximately 2-fold, regardless of *MIXL1* expression level.

To identify if loss of *MIXL1* affects the growth rate of KG1 cells, I assayed the knockdown lines over the course of four days for growth by MTS assay (Figure 13). While the control shRNA cell line and *REL*-enforced expression line grew at a stable rate over 4 days, the knockdown lineages had lower cell counts than the starting count at 24 hours before recovering to a stable growth rate by day four.

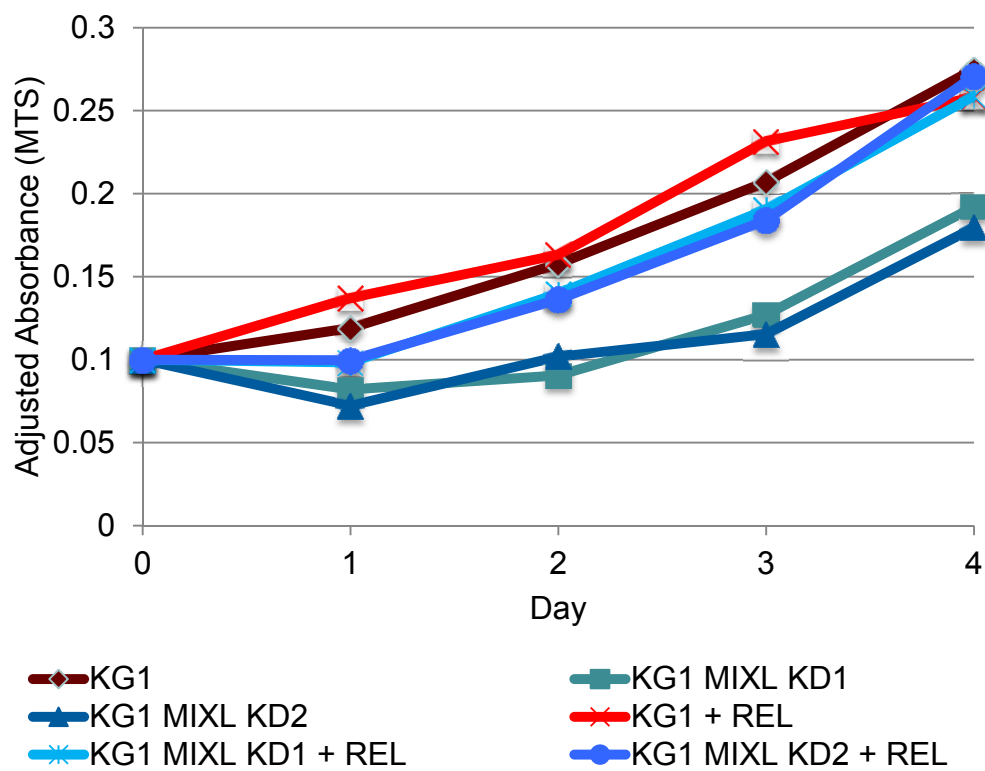
## **6. *MZF1* and *MIXL1* interact with the *c-REL* promoter**

From the ChIP-seq analysis, I have identified the *c-REL* promoter as a binding target for *MIXL1*; however the peak region identified is over 1000 base-pairs long, and is lacking in putative *MIXL1* motifs. These details pose the

**Figure 12. Knockdown of MIXL1 decreased while enforced expression of c-REL increased c-REL, BCL2A1, and BCL2L1 transcript levels.** MIXL1 shRNA lentivirus, c-REL retrovirus were transduced into KG1 cells. RT-qPCR in triplicates was performed on RNAs isolated 48 hours after transduction. Expression was normalized to 18srRNA levels, and error bars represent standard deviation between triplicates. \* denotes p value <0.05



**Figure 13. c-REL over expression rescues of MIXL1 knockdown mediated growth arrest in KG1 cells.** Growth was measured by MTS assay every 24 hours over a 4 day period in KG1 cells transduced with MIXL1 shRNA lentivirus and c-REL retrovirus. Absorbance was normalized to that of a non-transfected control sample.

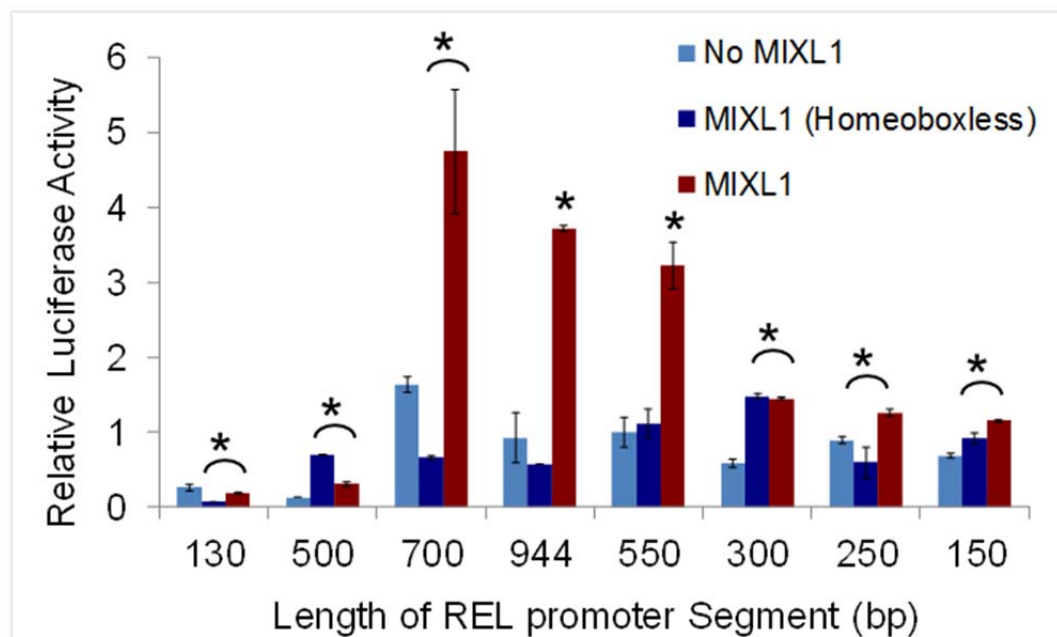
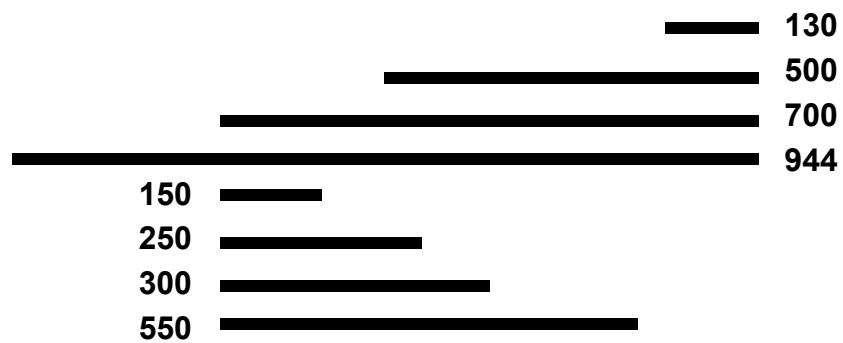
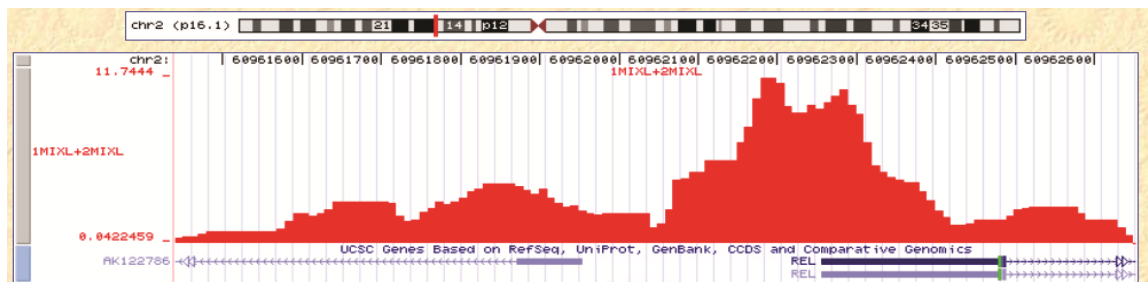




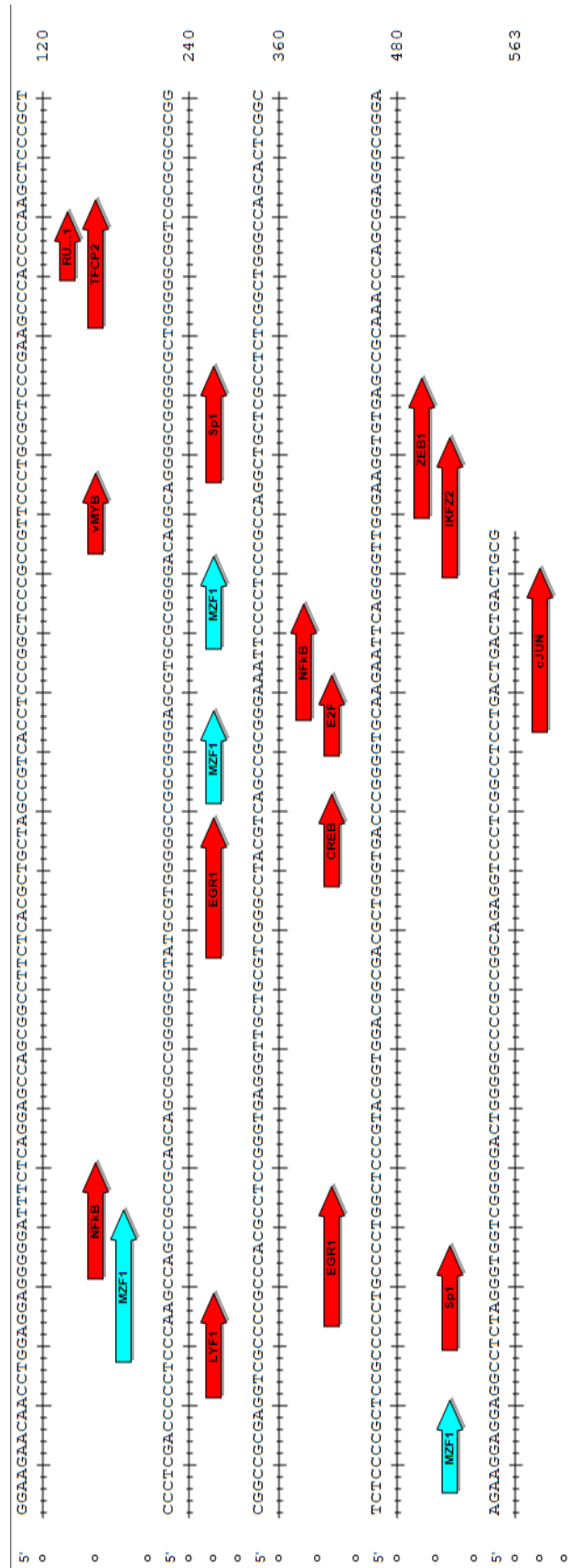
question: what part of this region is required for MIXL1 interaction and what else is required for it? To answer this question, I first tried to identify which region of the peak MIXL1 needs to regulate c-REL expression by performing a luciferase reporter assay with nested deletions of the c-REL promoter. By subcloning genomic fragments of the c-REL promoter generated by PCR (Figure 14, 130 to 944), I generated luciferase constructs in the backbone pBV-Luc. The luciferase constructs were then co-transfected into HEK293T cells with a *MIXL1* expression construct, a homeoboxless truncated MIXL1 expression construct, or an empty expression vector; the cells were tested for luciferase activity. Of these first sets, only the 700 base-pair and 944 base-pair promoters were significantly induced by full-length MIXL1. To narrow down the region further, I took the N-terminus region of the 700 base-pair promoter insert and extracted progressively shorter segments of the promoter in it (Figure 14, 550 to 150). The only promoter segment that was significantly induced by full-length MIXL1 was the 550 base-pair segment.

To identify potential binding partners for MIXL1 that are required for c-REL transcriptional regulation, I searched the 550 base-pair region of the c-REL promoter for known transcription factor binding motifs (Figure 15). The motifs identified included 2 NF- $\kappa$ B motifs, a RUNX1 motif, 2 Sp1 motifs, and 4 MZF1 motifs. As *MZF1* (*myeloid zinc finger 1*) is a transcription factor associated with the myeloid lineage and implicated in myeloid leukemia (51, 52), it may be expressed in *MIXL1* expressing cells and may in fact have a synergistic role. To determine if MIXL1 and MZF1 are actually interacting, I performed co-

**Figure 14. MIXL1 interacts with the center of the identified peak region of the *REL* promoter.** (Above) c-REL promoter peak region identified by ChIP-Seq, as generated by UC Santa Cruz genome browser is shown. The location and size of each promoter fragment used for the luciferase reporter assay is displayed underneath. (Below) MIXL1 binds to a 550 bp region within the c-REL promoter. Regions of the DNA depicted in 5A were cloned into the reporter vector pBV-Luc luciferase, transiently co-transfected into HEK293T with MIXL1, MIXL1 Homeobox-less, or empty expression vector. Equal amount of Renalia Luciferase co-transfected with the reporter constructs allowed normalization. Luciferase activity of each combination was tested after 48 hours in triplicate. Error bars represent standard deviation between triplicates. \* denotes p value <0.05



**Figure 15. Known Transcription Factor Binding Motifs in the *REL* Promoter.** A visual representation of the location of identified transcription factor binding motifs by TFSearch in the 550 bp *REL* promoter identified by luciferase reporter assay. Of note, the most prevalent motif in this region is the MZF1 zinc finger binding motif (marked in blue), which is present in four locations.



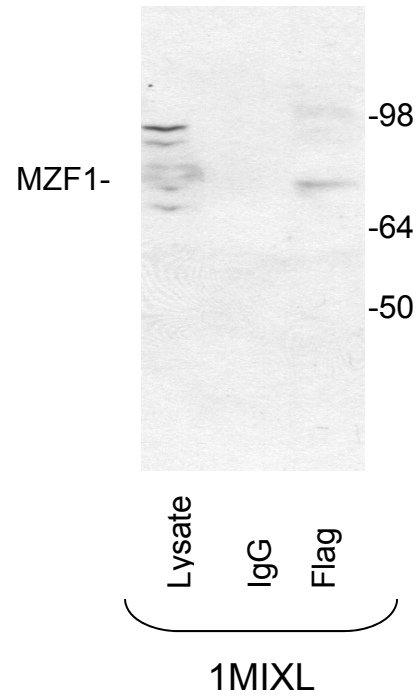
immunoprecipitation by anti-Flag antibodies in 1MIXL, using a mouse IgG as a control then blotted it with a MZF1 antibody (Figure 16). The expected MZF1 band (~82 kDa) was present in the 1MIXL lysate and Flag-immunoprecipitation but not the IgG-immunoprecipitation, implicating that MIXL1 and MZF1 may form a complex.

To determine if MZF1 binds to the c-REL promoter in the expected region, I performed ChIP-qPCR analysis using the *c-REL* promoter and intron primersets to probe MZF1-immunoprecipitated genomic DNA from cell line 1MIXL (Figure 17). The promoter region was enriched approximately 2.5-fold in the MZF1-immunoprecipitated fraction, while the intronic control region was not enriched.

## **7. MIXL1 is preferentially induced by BMP4 in human CD34+ cord blood hematopoietic stem progenitor cell**

*MIXL1* has previously been described as a transcriptional target of SMAD proteins (53) through the canonical TGF- $\beta$  and BMP signaling pathways. To identify whether the TGF- $\beta$  or BMP is the relevant ligand in hematopoiesis, I treated CD34+ hematopoietic progenitor cells with either 2 ng/ml of TGF- $\beta$ 1 or 50 ng/ml BMP4, and tested the relative mRNA levels of *MIXL1* after 2 hours (Figure 18). BMP4 was able to significantly activate MIXL1 expression relative to the un-induced CD34<sup>+</sup> samples by about 2-fold, while TGF- $\beta$  had no effect on *MIXL1* expression. This implicates BMP4 to be a preferred regulator of *MIXL1* expression in hematopoiesis.

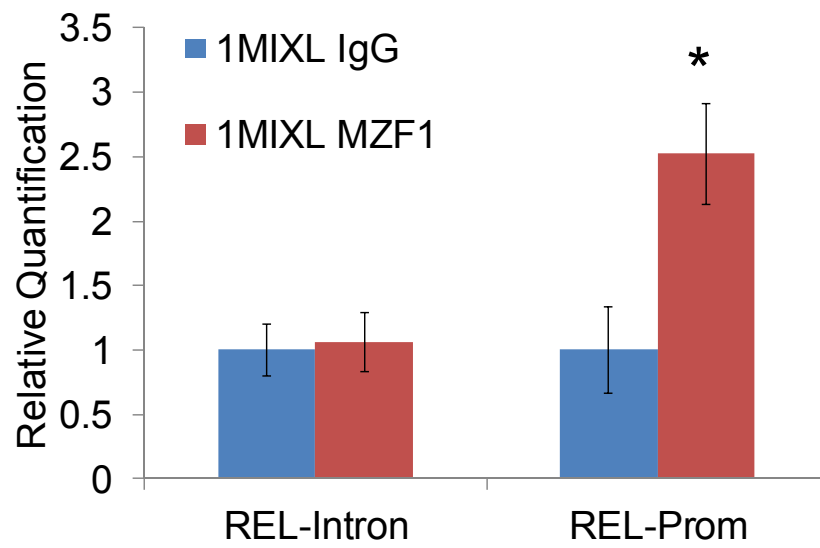
**Figure 16. MZF1 is identified by co-immunoprecipitation of Flag-tagged MIXL1 in U937.** Co-immunoprecipitation assay against Flag-HA-MIXL1 using an anti-Flag antibody in 1MIXL, and using an anti-MZF1 antibody to probe the immunoblot, with preclear sample as the positive control and IgG-immunoprecipitated sample as the negative control. MZF1 was detected in the preclear lysate and Flag-IP fraction, but not the IgG sample.



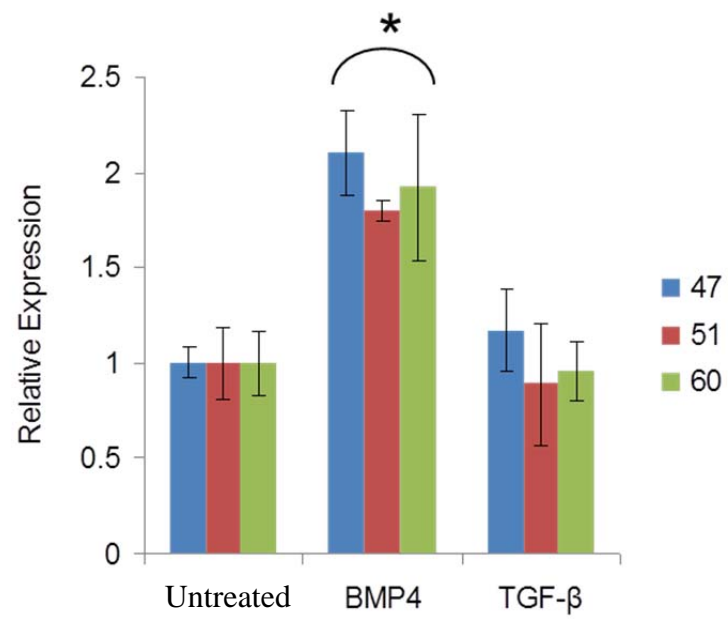


**Figure 17. MZF1 binds to the same locus as MIXL1 on *c-REL* promoter.**

Quantitative PCR analysis of the identified *c-REL* promoter region and *c-REL* intron control region, comparing the abundance of each genomic locus immunoprecipitated by either IgG or MZF1 antibodies, normalized to a standard curve. Error bars represent standard deviation between triplicates. \* denotes p value <0.05



**Figure 18. BMP4 induces MIXL1 in CD34+.** MIXL1 expression increased 2-fold in CD34+ cells treated with BMP4, than untreated or cells treated with TGF- $\beta$ . CD34+ HSPCs from three cord blood donors (47, 51, 60) were treated with either 50 ng/ml BMP4 or 2 ng/ml TGF- $\beta$ 1 for 2 hours. MIXL1 transcript levels were quantified by RT-qPCR using 18srRNA as normalization control. Error bars represent standard deviation between triplicates and \* denotes p value <0.05.

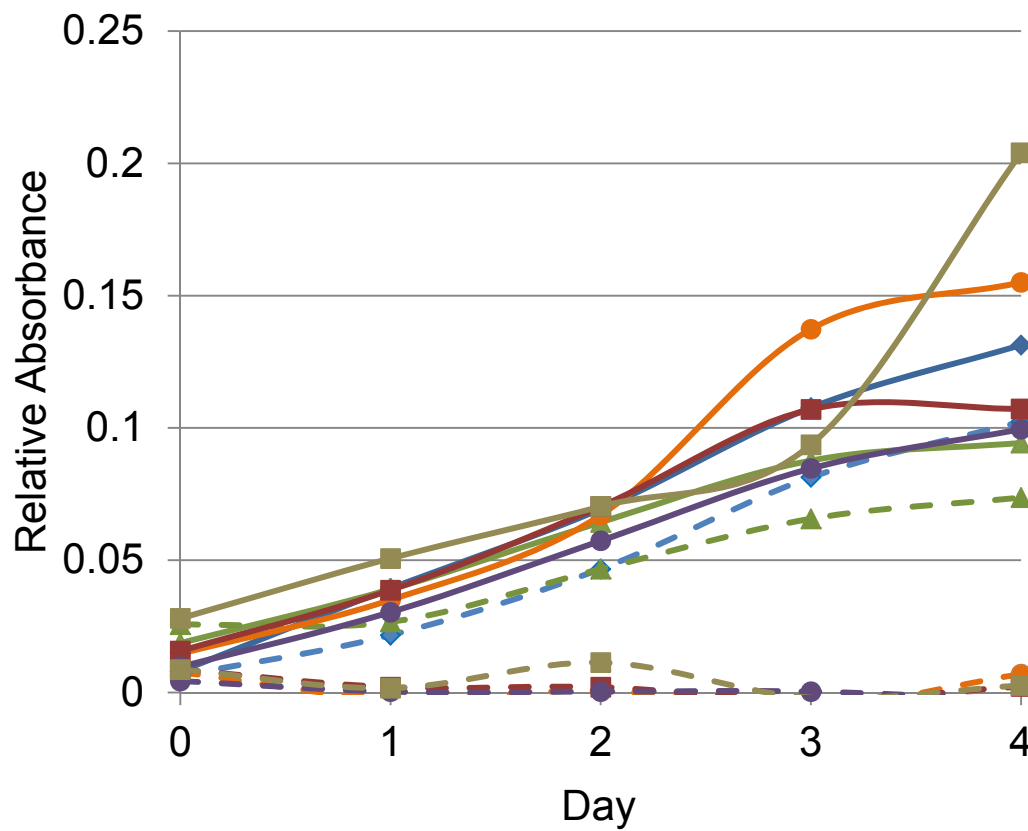


## **8. *MIXL1* expressing cells are sensitive to the BMP inhibitor LDN-193189**

As *MIXL1* is regulated by the BMP pathway, I postulated that *MIXL1*-expressing lines may have increased sensitivity to BMP pathway inhibition. To inhibit BMP pathway activation, I used the ALK2/3/6 (ACVR1, BMPR1A, and BMPR1B) inhibitor LDN-193189. Four high *MIXL1*-expression cell lines (OCI-AML2, KG1, ML3 and K562), and two low *MIXL1*-expression lines (U937 and HL60) were exposed to a single treatment of 3  $\mu$ M LDN-193189 before a 4 day incubation, with the relative survival assayed by the MTS analysis (Figure 19). By day 1, the high expression lines were no survival was detected in the LDN-193189 treated aliquots, while the growth of the low-expression lines was diminished compared to the untreated aliquots. However, on the following days, HL60 and U937 cell cultures have begun to recover, while the high-expressing lines stay undetectable.

To further characterize the sensitivity to LDN-193189 in high *MIXL1*-expression cell lines, I performed a dose response analysis on each of the 6 lines over a 0-700 nM range of LDN-193189. Due to the short half-life of LDN-193189, the medium was changed every 24 hours, and relative survival was assayed them by MTS after four days (Figure 20). The number of cells in the low-expression lines U937 and HL60 were constant across all concentrations of LDN-193189, suggesting no significant change in growth rate, whereas in the high-*MIXL1* expression lines, there is a significant decrease between 100 nM and 200 nM in culture viability.

**Figure 19. High MIXL1 expression cell lines are sensitive to 3  $\mu$ M LDN-193189.** 3  $\mu$ M of LDN-193189 was cytotoxic to OCI-AML2, ML3, KG1 and K562 in contrast to U937 and HL60 cells. Each cell line was treated with vehicle or 3  $\mu$ M LDN-193189 on Day 0, and viability measured every 24 hours by MTS assay, in triplicate. Absorbance was normalized to that of a vehicle-only sample.



LDN-193189

0 μM

3 μM

—◆— U937

—◆—

—▲— HL60

—▲—

—●— OCI-AML2

—●—

—■— ML3

—■—

—●— KG1

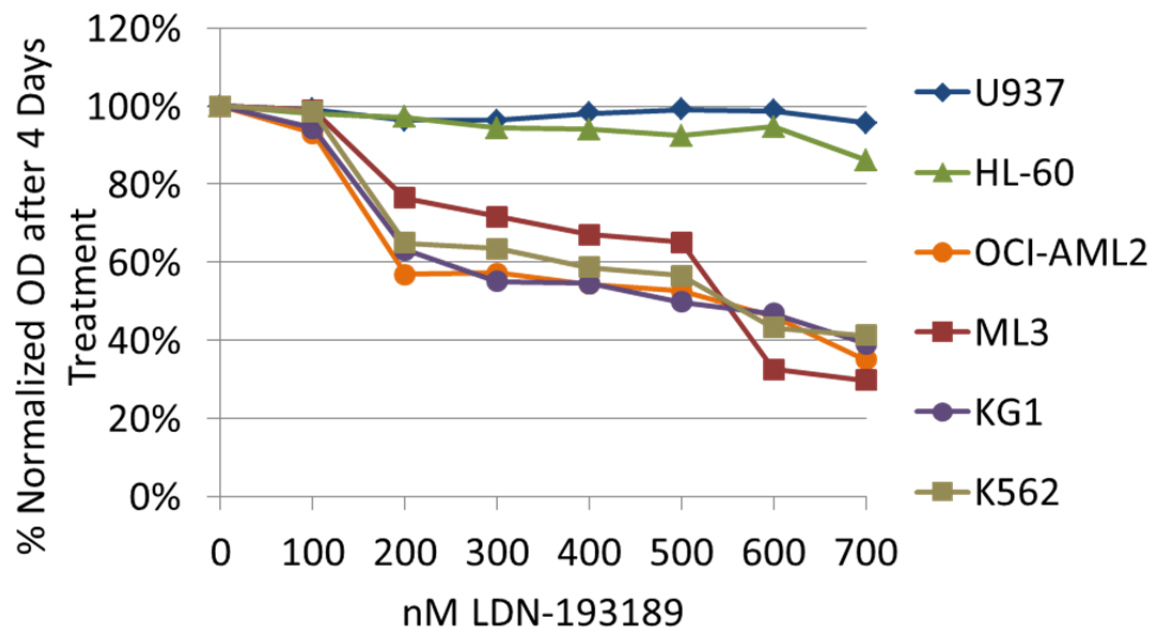
—●—

—■— K562

—■—

**Figure 20. High MIXL1 expression cell lines are sensitive to consistent exposure of LDN-193189 for 4 days.** OCI-AML2, ML3, KG1 and K562 were sensitive to 200 nM LDN-193189, while low MIXL1 expression lines U937 and HL60 were unaffected. Each cell line was treated with 0 nM to 700 nM LDN-193189, replenished every 24 hours with fresh drug or control medium for 4 days. Viability on day 4 was assayed then tested by MTS assay in triplicate. Absorbance was normalized to that of control vehicle-only treated samples.





## 9. Endogenous *MIXL1* in AML Patients

To determine if *MIXL1* is overexpressed in patient samples, I analyzed the TCGA AML patient sample dataset for at least 2-fold change of expression, chromosomal modification, or mutation of *MIXL1* by the cBioPortal, AML-associated homeobox transcription factors *CDX2* (30), *HLX*(28), *HOXA9* (27), and NF- $\kappa$ B pathway members *c-REL*, *BCL2*, *BCL2A1*, and *BCL2L1* (Figure 21). *MIXL1* was amplified or up-regulated in 13% of the total AML cases, while only 12% of the patients had *HOXA9* modifications and 6% had *CDX2* modifications. *HLX* was down regulated in 10% of patients and amplified or upregulated in 17%. Only a few patient samples contained expression changes in two of these genes, implicating *MIXL1* expression may be a marker of a subgroup of AML cases distinct from *HOXA9* and *CDX2* expressing AMLs.

To determine if *MIXL1*, *HOXA9*, and *CDX2* characterize distinct subgroups, I analyzed the genes for mutual exclusivity between the patient samples (Figure 22). While *MIXL1* and *HOXA9* had a tendency for mutual exclusivity, *CDX2* had no strong association with either of the two homeoboxes. *HLX* had a tendency to be mutually exclusive with *HOXA9*, but had a minor tendency to co-occur with *MIXL1* or *CDX2*.

I also performed the same mutual exclusivity analysis with NF- $\kappa$ B pathway members *c-REL*, *BCL2*, *BCL2L1*, and *BCL2A1*. While both *MIXL1* and *CDX2* had a tendency for co-occurrence with *c-REL*, *MIXL1* had a stronger tendency to co-occur with *BCL2L1*, while *CDX2* had a stronger tendency to co-occur with *BCL2*.

**Figure 21. *MIXL1* overexpression marks a set of AML cases distinct from those expressing *CDX2*, *HLX* or *HOXA9*.** The TCGA AML patient dataset was queried through the cBIOPortal database for >2 fold expression alterations as determined by RNA-Seq, amplifications, or homozygous deletions, across 166 AML patient cases. Each column represents a patient. *MIXL1* is amplified or up-regulated in 13% of the total AML cases.

Customize

Case Set: All Complete Tumors: All tumor samples that have mRNA, CNA and sequencing data (163 samples)

Altered in 123 (74%) of cases



**Figure 22. *MIXL1* expression co-occurs with *REL* and *BCL2L1* expression, but not with *HOXA9*.** A mutual exclusivity analysis from cBioPortal using the TCGA AML database as depicted in Figure 21. Expression of *MIXL1* has a tendency to co-occur alongside expression of *REL* and *BCL2L1*, while expression and modification in *HOXA9* tend to be excluded from cell lines expressing *MIXL1* or *BCL2L1*.

Gene	REL	HLX	BCL2	BCL2L1	MIXL1	BCL2A1	HOXA9	CDX2
REL	---	0.168697	0.366682	0.246175	0.107604	0.181816	0.452275	0.125329
HLX		---	0.026975	0.071816	0.073667	0.529387	0.150686	0.098407
BCL2			---	0.205304	0.456966	0.204341	0.330588	0.208945
BCL2L1				---	0.016194	0.419566	0.174625	0.560916
MIXL1					---	0.519936	0.241739	0.631472
BCL2A1						---	0.555332	0.266459
HOXA9							---	0.655451
CDX2								---

p-values <0.05, as derived via Fisher's Exact test are outlined in red.  
p-values are not adjusted for FDR.

#### Legend

Strong tendency towards mutual exclusivity ( $0 < \text{Odds Ratio} < 0.1$ )

Some tendency towards mutual exclusivity ( $0.1 < \text{Odds Ratio} < 0.5$ )

No association ( $0.5 < \text{Odds Ratio} < 2$ )

Tendency toward co-occurrence ( $2 < \text{Odds Ratio} < 10$ )

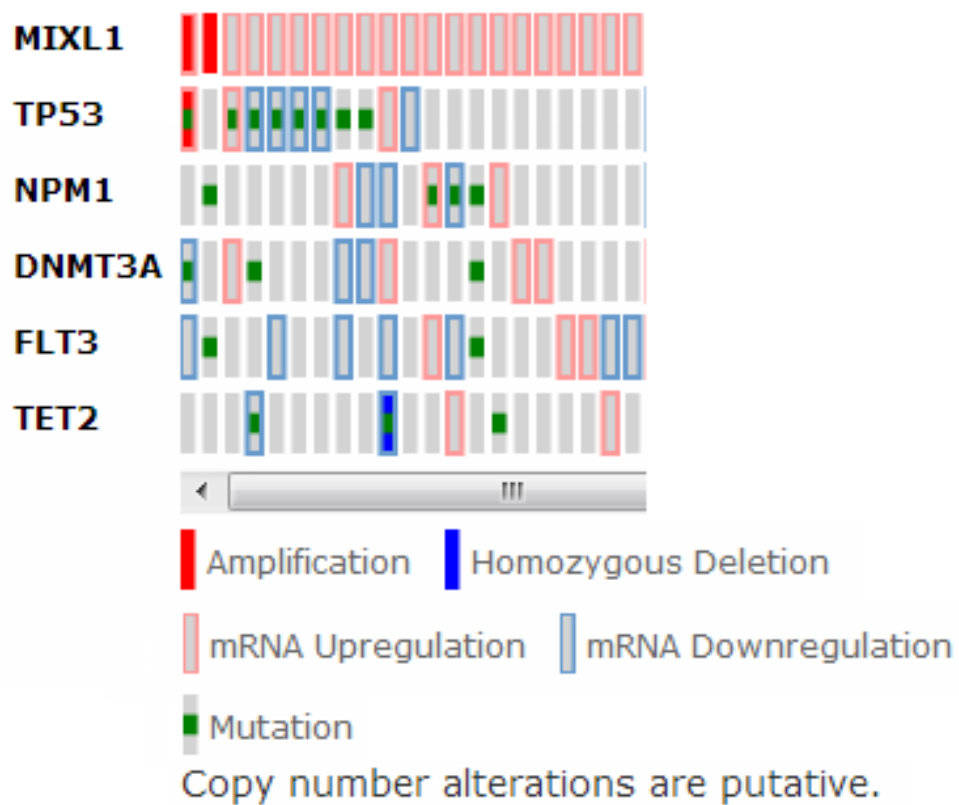
Strong tendency towards co-occurrence ( $\text{Odds Ratio} > 10$ )

No events recorded for one or both genes

I next analyzed TCGA AML patient sample dataset for genes with any of the commonly identified somatic mutations (*NPM1*, *FLT3*, *DNMT3A*, *TET2* and *TP53*) to determine with which mutations *MIXL1* co-occurred or co-operated with (Figure 23). The most common alteration was mutations in *TP53* seen in 8 cases (38%) of *MIXL1*<sup>+</sup> cases. Interestingly, this constitutes 66% of all the cases with *TP53* mutations (n=12) in the TCGA population. All of the other recurrent mutations were also seen, but less frequently. Four patients had mutations in *NPM1*; there were three cases each for *DNMT3A* and *TET2*. Two patients had mutations in *FLT3*.

**Figure 23. High MIXL1 expression patient samples contain previously identified somatic mutations of AML.** 76% (16/21 cases) of high MIXL1 patients in the TCGA AML patient dataset contained previously identified common somatic mutations for AML (*NPM1*, *FLT3*, *DNMT3A*, *FLT*, *JAK3* and *TP53*). The TCGA AML patient dataset was queried through the cBIOPortal database for >2 fold expression alterations, mutations, amplifications, or homozygous deletions, across 166 AML patient cases. Each column represents a patient.





## **Chapter 4. Discussion**

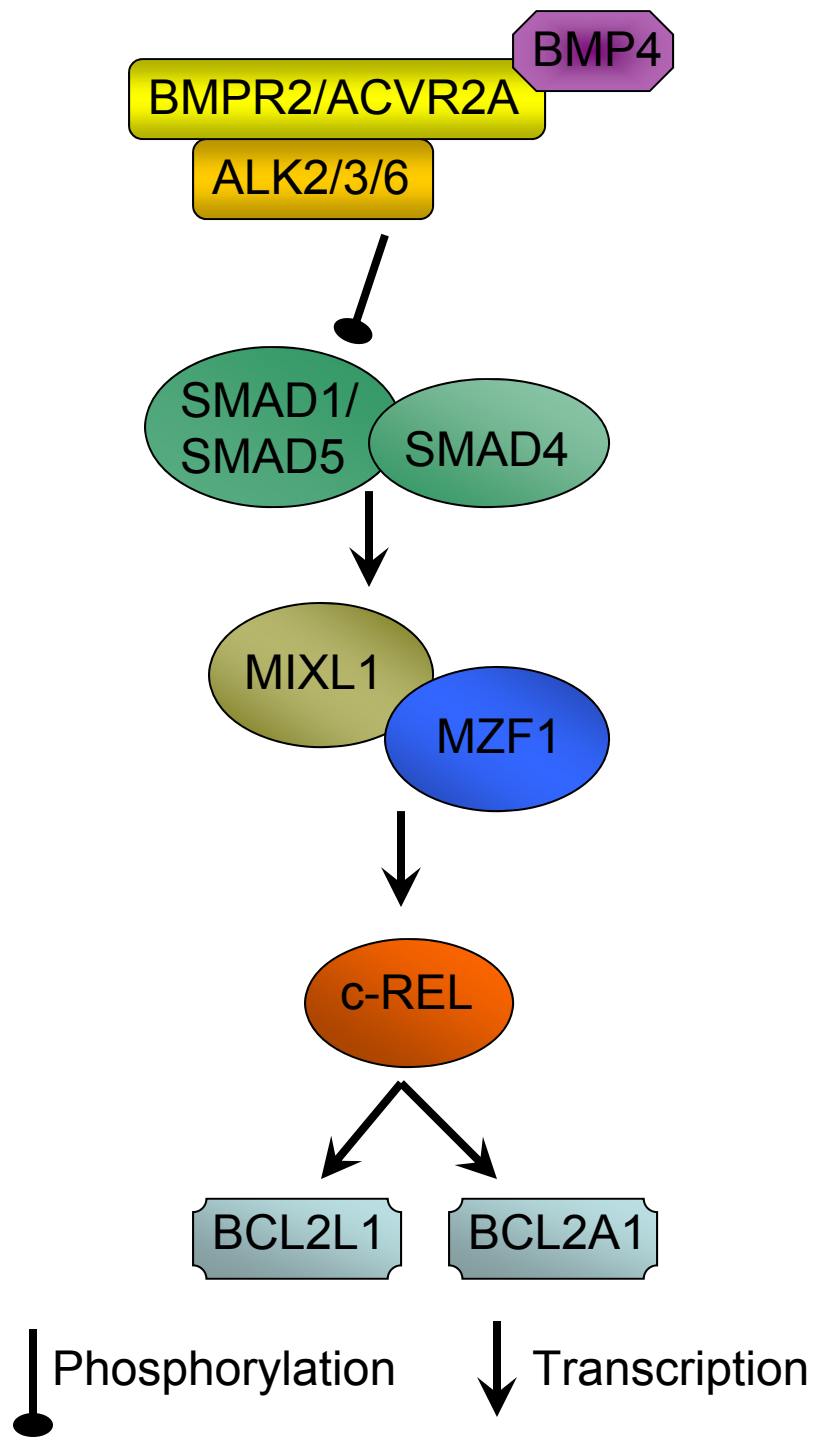
In this study I have identified a BMP4-MIXL1-c-REL pathway which may have an important role in normal hematopoietic development (Figure 24).

Mix.1 is induced in *Xenopus* embryos by BMP4 or Activin A in a SMAD5 dependent manner (17) although *MIXL1* can be induced by TGF- $\beta$  in human hepatocellular carcinoma (54) and TGF- $\beta$  and activin mouse ES cells (55). TGF- $\beta$  from the glial cells within the bone marrow niche is thought to maintain HSC quiescence (56). In humans, BMP2, BMP4 and BMP7 regulate maintenance, proliferation and repopulating activities of progenitors (57-59). Our studies demonstrate for the first time that MIXL1 induction in human HSPCs by BMP4. At present the distinction between the growth inhibitory versus mitogenic response to TGF- $\beta$  family ligands is dependent on biological assays. *MIXL1* expression may be a new marker that correlates with a pro-growth/ survival response.

### **1. *MIXL1* is over-expressed in a subset of AML**

Analysis of the TCGA AML patient sample set identified a potentially unique subset of high MIXL1 expressing cancers making up 13% of the cancer set, a group both distinct and larger than the CDX2 and HOXA9 subtype. A distinct aspect of the MIXL1 expressing cancer set is the prevalence of patients with mutated p53: 38% of high-MIXL1 patients had p53 mutations and 67% of patients with p53 mutations had high expression of MIXL1. p53 mutations are associated with a strong resistance to chemotherapy (60); if MIXL1 expression is

**Figure 24. Proposed BMP/MIXL1/REL pathway.** BMP4 binds to the complex between BMP type 2 receptor (BMPR2 or AVCR2A) and BMP type 1 receptor kinase (ALK2, ALK3, or ALK6). In response to activation, the kinase phosphorylates SMAD1 or SMAD5 which complexes with SMAD4 to activate the transcription of *MIXL1*. *MIXL1* complexes with MZF1 and activates the transcription of *c-REL*. C-REL in turn activates the transcription of BCL2 family members.



prevalent in the p53 mutant subset of AML, then it may identify a new potential avenue to treat these untreatable patients.

While this implicates *MIXL1* expression as a potentially strong diagnostic marker in leukemia, it will be important to show that this can be replicated in a much larger set of patients. The TCGA dataset contains complete patient data for 166 patients of diverse types of AML; however this may not be an accurate representation of the distribution of total AML cases. In addition, the more patient information gathered for further analysis, the more reliable and statistically significant any correlations determined will be.

A number of technical reasons preclude *MIXL1* detection in array based global expression profiling studies. In addition to weak probes, the long 3' UTR includes *Alu* elements which can result in quenched signals. Thus availability of the RNA-Seq data for the TCGA patients provided a valuable resource to establish the importance of *MIXL1*.

## **2. *MIXL1*-positive AML lines are sensitive to LDN-193189**

Expression of *MIXL1* appears to identify a distinct subset of AML, which is of interest if *MIXL1* expression also identified a potential therapeutic target. It would be quite relevant, then if the difference in sensitivity to kinase inhibitor LDN-193189 seen between *MIXL1*-expressing lines and non-expressing lines is also identified in patient samples. The differential cytotoxicity to 3  $\mu$ M LDN-193189, and growth suppression specific to those lines in lower expression

levels, implies that this defined subtype of AML may be responsive to Type 1 BMP receptor kinase suppression.

In the experiments presented, I treated a variety of leukemic lines with LDN-193189 at concentrations up to 3  $\mu$ M. While there was a distinct difference between how MIXL1-expressing and non-expressing lines responded below this concentration, any concentration at or above 4  $\mu$ M was lethal to all cell lines tested. One possible explanation for this is that activation of the BMP pathway may still be critical to the survival of U937 and HL60, yet this pathway does not include *MIXL1* activation in these cell lines. Indeed, the BMP intermediate SMAD1 has a different set of targets between U937 and K562 when induced by BMP4 (61). In particular, my survey of their datasets identified that SMAD1 binds to a unique peak in the MIXL1 locus in K562 but not U937.

Another likely explanation above 4  $\mu$ M for LDN-193189 are lethal to U937 and HL60 despite the lack of *MIXL1* activation is that LDN-193189 may be inhibiting other kinases critical to the survival of these lines. LDN-193189 functions in BMP signaling suppression by acting as an ATP-antagonist in the BMP/Activin-regulated kinases ALK2/3/6 (ACVR1, BMPR1A, and BMPR1B)(62), but it can function as an ATP-antagonist in other kinases, including ABL, VEGFR, YES1, and CAMKK2 (63), many of which have nearly as high sensitivity to the inhibitor as the BMP pathway kinases. It would therefore be likely that even cells not reliant on BMP pathway for survival would have some level of sensitivity to the inhibitor.

Since LDN-193189 can inhibit a large number of kinases beyond the BMP-regulated kinases, it is possible that the sensitivity seen in cell lines with high *MIXL1* expression is either due to a combination of BMP-regulated kinase inhibition and off-target, or off-target effects entirely. While I have shown that in CD34+ hematopoietic precursor cells, *MIXL1* can be activated by BMP4, I have yet to show that blocking this pathway is critical to the survival of the *MIXL1*-expressing cell lines. Additionally, even within the BMP-regulated kinases ALK2/3/6 (ACVR1, BMPR1A, and BMPR1B), I have yet to distinguish what role each has and which is most important as a target of LDN-193189 in each cell line. To better identify what role each potential target of LDN-193189 is playing in the overall survival of these cell lines, I propose inhibiting each relevant kinase through another means, specifically shRNA knockdown, and testing for both expression changes in the downstream pathways and changes in growth and survival of the cells.

While LDN-193189 targets a large number of kinases beyond the BMP-regulated kinases, this may not preclude its use in therapeutics. Indeed, many of the other kinase targets, including ABL1 and VEG-FR, are implicated in cancer and are either currently targeted by therapeutics or considered likely targets for therapeutics in the future. LDN-193189 has been already used in rodent models of hepcidin induced chronic anemia (64) and fibrodysplasia ossificans progressiva with constitutive activation of ACVR1 (65), and the side effects exhibited from drug treatment were minimal. LDN-193189 may have significant potential as a generally safe and therapeutic treatment for AML.

### 3. Transcriptional regulation by MIXL1 in hematopoietic cells

I identified several important transcriptional targets of *MIXL1* in this study. While I have focused on *c-REL*, other critical genes identified by the ChIP-Seq study may play a significant role in the leukemogenic activities of *MIXL1*. Future research will better characterize other downstream targets and elucidate which ones are responsible for the function of MIXL1.

The expected motif for MIXL1 binding is two TAATTGAATTA. This motif was originally identified *Gsc* (*gooseoid*) promoter and was confirmed by SELEX and reporter assays in ES cells and fibroblasts in vitro SELEX screens (66). The expected motif, however, was less abundant in the targeted sequences identified by ChIP-Seq than in randomized sequences. Indeed, while the homeobox binding motif is prevalent in AT-rich regions, CG-rich were very abundant in our sequences identified by ChIP-Seq. This is of stark contrast to an analysis of *Mixl1* targets in mouse embryonic stem cells which had found the homeobox binding motif to be considerably enriched (10).

Several possible reasons underlie the lack of this established DNA binding motif in our target peaks. One possibility is that the MIXL1 is binding directly to DNA in a way never previously characterized, which could be tested by purifying the MIXL1 protein and testing the binding affinity with different DNA sequences. A more likely explanation is that MIXL1 requires interactions with other proteins to bind to DNA in this sequence. This is supported by the abundance of other characterized transcription factor binding motifs in the peak regions and the



length of the *c-REL* promoter segment required for MIXL1 regulation, which implies a larger multiprotein complex is necessary. While MIXL1 has been shown previously to interact with other homeobox-containing proteins, it has also been shown to interact with NEUROD1 (12), which does not contain a homeobox domain, in a high throughput two-hybrid screen. As the most prevalent binding motifs identified in the ChIP-Seq were zinc finger motifs, and MZF1 motifs were prevalent in the *c-REL* promoter peak, MZF1 was a potential putative binding partner.

MZF1 (myeloid zinc finger 1) is a zinc finger transcription factor associated with the myeloid lineage (67, 68). Expression of MZF1 inhibits apoptosis and differentiation in myeloid leukemia cells (51, 52) and induces proliferation, migration, invasion, and metastasis in colorectal cancer, cervical cancer, and breast cancer (69-71).

Interactions between zinc fingers and homeobox proteins have been identified previously. In breast cancer, an isoform of distal-less homeobox DLX4 binds to the zinc finger BRCA1 (72). In hematopoietic cells, zinc finger PML binds to the homeobox HHEX (73). Additionally, a multitude of other potential interactions between homeoboxes and zinc fingers have been identified through two-hybrid assays (12).

Recent studies suggest that Mixl1 can interact with T box factors T, Eomes, Tbx6 and Tbx20 directly to form protein complexes which function as transcriptional repressors (74). Additionally, in a high throughput mammalian two hybrid screen, an interaction between TBX20 and MZF1 was identified (12).

MIXL1, TBX20, and MZF1 may bind to targets as a multiprotein complex during hematopoiesis, similar to the multiprotein complex between paired-type homeobox protein NKX2-5, T box protein TBX3, and zinc finger protein GATA4 in cardiac development (75).

To better identify how MIXL1 is interacting with these gene loci, we would need to determine what the subunit composition of the MIXL1-containing complexes are. I would need to repeat co-immunoprecipitation and luciferase assay on each combination of potential complex members. Next, I would want to over-express each potential member in every combination to see which combinations best activate the transcriptional targets. Finally, I would knock out each member from an expressing AML line to determine the minimal complex necessary for transcriptional activation.

#### **4. *c-REL* is a transcriptional target of MIXL1**

c-REL is a promising target for MIXL1 transcriptional activation due to its strong sequence similarity to the established viral oncogene v-REL and established role in the activation of anti-apoptotic genes BCL2L1 and BCL2A1 (76). Amplification and overexpression of c-REL is implicated in diffuse large B cell lymphoma (38). c-REL overexpression is seen in a wide array of other cancer types, including colorectal cancer (77), breast cancer (78), lung carcinoma (79), squamous cell carcinoma (80, 81), retinoblastoma (82), endometrial carcinoma (83), and pancreatic cancers (84).

Activity of the NF- $\kappa$ B DNA binding factors comprising of 5 proteins which heterodimerize to bind DNA, is regulated both at the transcriptional and post-transcriptional level. Cytoplasmic sequestration of the five proteins by inhibitors NFKBIA, B and E regulates their activity post-transcriptionally (85). Expression profiling suggests down regulation of all the NFKBI members and preferential upregulation of *c-REL*, *RELB* and *NFKB2*. Furthermore, recent evidence suggests potential functional differences in the activity depending upon the subunit composition. Thus *c-REL* homodimers or heterodimer with NFKB2 or RELB may have varied responses (37, 86). Notably, there is an absolute requirement for *c-REL* in B lymphoid development (87), though the factors surrounding both the regulation of *c-REL* and the downstream effects distinct from the other family members is currently unknown. Overall, this would implicate that MIXL1 and *c-REL* may activate preferentially a “non-canonical” NF- $\kappa$ B pathway distinct from the one regulated by the RELA/NFKB1 complex.

## **5. Summary**

In summation, we have identified MIXL1 to mark a new subset of AML and *c-REL* as a direct transcriptional target of MIXL1. Targeting the BMP receptor upstream of MIXL1 may be a novel therapeutic avenue to explore in AML.

**Appendix: Peaks identified by all analyses of the *MIXL*-overexpression**

**ChIP-Seq dataset.** The following set of peaks was generated by normalizing to the two control samples in three combinations and then finding the overlapping set (as visualized in Figure 5): Flag-1MIXL to IgG-Control and Flag-Control, Flag-2MIXL to IgG-Control and Flag-Control, and Flag-1MIXL and Flag-2MIXL combined to IgG-Control and Flag-Control. A total of 179 peaks were identified in this stringent set. This table was generated as part of the analysis performed by Dr. Yue Lu and Dr. Shoudan Liang for ChIP-seq. Columns are as followed:

Site position: The location of the peak on Human genome assembly Hg18/NCBI36.

Gene name: Gene loci within 25 kb of the peak.

Location: Relative location of the peak to the nearest gene loci. Designations are: Upstream – 5 to 25 kb upstream of the transcription start site (TSS), Promoter – 0 kb to 5 kb upstream of the TSS, Body – Between the TSS and transcription sequence end (TSE), TSE – 0 to 5 kb downstream of the transcription sequence end, Downstream – 5 to 25 kb downstream of the TSE. Unlabelled have no nearby loci.

Fold enrichment: Fold enrichment of the peak region in experimental sets over control sets.

Number of Tags: The number of sequence tags that contributed to this peak.

-10Log (pvalue): Negative Log base ten of the p-value for each peak.

<b>site_position</b>	<b>gene_name</b>	<b>Location</b>	<b>Fold enrichment</b>	<b>Number of tags</b>	<b>-10LOG (pvalue)</b>
chr1:100370463-100371655	SASS6,CCDC76	promoter	5.45	70	67.94
chr1:1082711-1084387			6.72	132	139.37
chr1:153411692-153412694	KRTCAP2,TRIM46	promoter	7.65	63	85.3
chr1:153480474-153481740	GBA	promoter	4.55	88	62.49
chr1:154452528-154453384	SLC25A44	downstream	3.99	296	151.79
chr1:154986531-154989082	HDGF	promoter	6.32	124	66.28
chr1:167341593-167342441	ATP1B1	promoter	7.31	48	124.79
chr1:178117582-178119503	TOR1AIP1	promoter	6.01	147	130.37
chr1:221054670-221055313			8.93	27	107.73
chr1:230832149-230833445			7.72	62	146.53
chr1:234753905-234754746	LGALS8	promoter	6.83	35	86.28
chr1:39228761-39230181	AKIRIN1	promoter	7.52	72	69.02
chr1:8799585-8800833	RERE	promoter	7.29	66	119.64
chr1:9473155-9473998			8.65	56	121.24
chr1:9609526-9610602			10.55	79	142.8
chr1:9633942-9634866	PIK3CD	promoter	5.73	60	78.06
chr10:121291414-121293247	RGS10	promoter	5.2	85	101.97
chr10:124757753-124759210	IKZF5,ACADSB	promoter	7.4	82	105.68
chr10:38731545-38732467	CDC10L	promoter	9.58	29	92.31
chr10:43163960-43164521			10.25	38	108.47
chr10:70937383-70938053	TSPAN15	TSE	7.21	47	95.72
chr10:76639460-76641469	VDAC2	promoter	4.83	98	114.52
chr10:88126705-88127529			12.11	39	106.9
chr10:93547695-93548906	TNKS2	promoter	10.42	60	148.58
chr10:98469350-98471263	PIK3AP1	promoter	6.66	94	101.59
chr11:119886935-119887385			7.25	25	69.29
chr11:127547489-127547696			16.58	14	134.56
chr11:18753771-18754240	PTPN5	body	16.01	17	90.98
chr11:2405808-2406392	TRPM5	upstream	10.64	25	96.64
chr11:34029256-34031806	CAPRIN1	promoter	5.15	95	66.1
chr11:47385814-47387193	SLC39A13	promoter	6.87	62	73.23



chr17:37097849-37099490	EIF1	promoter	7.55	92	156.7
chr17:39543640-39544659	HDAC5	body	5.29	61	75.92
chr17:41626353-41627609			5.79	90	148.39
chr17:53282123-53283069	MRPS23	promoter	10.49	66	159.6
chr17:7124493-7125013	SLC2A4	promoter	8.31	36	82
chr18:148098-149269	USP14	promoter	5.85	70	98.42
chr18:58341020-58341483	ZCCHC2	promoter	11.9	22	104.59
chr18:9092283-9093337	NDUFV2	promoter	7.67	82	124.29
chr19:10666975-10667833	ILF3	downstream	7.26	53	100.56
chr19:14377476-14378551	DDX39	downstream	5.47	57	105.59
chr19:2770179-2771331	ZNF554	promoter	10.66	56	98.46
chr19:4320061-4321948	SH3GL1	body	6.73	91	81.82
chr19:55182096-55182785	VRK3	body	8.54	56	112.82
chr19:60363117-60364407	TNNI3	upstream	10.66	47	94.92
chr19:63422082-63422569			15.31	18	99.28
chr19:8551405-8552171	MYO1F	upstream	12.34	59	155
chr2:109029691-109030422			6.93	27	52.19
chr2:144805719-144807166	GTDC1	promoter	6.17	61	62.76
chr2:183288716-183289819	DNAJC10	promoter	11.73	56	154.25
chr2:208197446-208199695	FAM119A	promoter	4.52	138	104.45
chr2:218328982-218329926	DIRC3	promoter	11.05	30	91.93
chr2:231284904-231286412	CAB39	promoter	6.16	81	58.89
chr2:240613319-240614157	NDUFA10	promoter	12.65	35	117.06
chr2:39517179-39518454	MAP4K3	promoter	6.32	73	83.87
chr2:60961441-60962652	REL	promoter	9.79	73	142.2
chr2:64568721-64569193			11.9	22	87.09
chr2:85696451-85697304	USP39	promoter	9.09	58	136.51
chr2:86495597-86495902			14.88	23	128.97
chr2:96294363-96296376	TMEM127,CIAO1	promoter	5.78	104	77.6
chr20:1046903-1047656	PSMF1	promoter	11.15	39	97.43
chr20:18066017-18066962	CSRP2BP	promoter	5.92	49	100.85
chr20:26136641-26137576			4.51	171	163.46
chr20:29663539-29664931			8.15	177	100.29

chr20:30019791-30020960	XKR7	promoter	12.96	156	173.76
chr20:34005247-34006579	SCAND1	promoter	7.11	86	145.41
chr20:43424790-43425893	SYS1	promoter	5.37	65	111.67
chr20:43995718-43997880	PCIF1	promoter	6.94	101	110.54
chr20:6051064-6051910	FERMT1	promoter	7.11	26	73.6
chr21:45183817-45185018	C21orf67	promoter	9.57	75	94.97
chr21:46166957-46168073	PCBP3	body	12.34	40	100.56
chr22:30215674-30216941	EIF4ENIF1	promoter	5.83	66	112.48
chr22:44890951-44891648	LOC400931	downstream	10.99	36	115.41
chr3:129881882-129883049			6.69	76	123.03
chr3:132704139-132704713	MRPL3	promoter	9.47	27	101.08
chr3:143779916-143780809	ATR	promoter	6.92	53	125.46
chr3:157874292-157875969	TIPARP	promoter	6.04	109	152.82
chr3:158288967-158290241			8.12	77	150.42
chr3:158439581-158439957			16.58	15	85.54
chr3:159844820-159845455	GFM1	promoter	11.03	46	182.96
chr3:180271815-180273139	ZMAT3	promoter	6.46	61	116.69
chr3:40325257-40326747	EIF1B	promoter	7.23	89	93.84
chr3:49951972-49953275	RBM6	promoter	5.67	88	195.31
chr3:53081401-53082439			11.9	64	143.45
chr3:88189857-88191895	CGGBP1	promoter	7.58	109	149.87
chr3:9412642-9414529	SETD5	promoter	4.53	161	164.66
chr4:140436020-140436664	NDUFC1	promoter	14.1	51	205.52
chr4:144653627-144655034	SMARCA5	promoter	4.74	83	79.27
chr4:15265187-15267009	FBXL5	promoter	7.4	94	126.9
chr4:172011274-172011568			27.78	15	133.58
chr4:185186291-185187153			9.53	41	148.02
chr4:186629373-186629928	CCDC110	promoter	9.65	19	78.06
chr4:2506433-2508320			4.89	133	140.79
chr4:40910100-40911927	APBB2	promoter	5.15	76	74.24
chr4:41631637-41632502	TMEM33	promoter	6.3	53	99.03
chr4:6627106-6629004	MAN2B2	promoter	6.01	86	83.36
chr4:6990416-6991886	TBC1D14	body	6.04	86	110.19



chr4:77287964-77289097	NUP54	promoter	10.39	98	284.03
chr4:84595427-84596455	HELQ,MRPS18C	promoter	9.17	76	156.45
chr5:126140153-126142456	LMNB1	promoter	5.04	162	145.67
chr5:131853280-131855024	IRF1	promoter	5.69	99	103.51
chr5:134268021-134269311	PCBD2	promoter	11.9	98	219.06
chr5:139906921-139908326	EIF4EBP3	promoter	5.49	89	131.77
chr5:150807080-150808183	SLC36A1	promoter	9.41	62	111.54
chr5:171546448-171548631	STK10	promoter	5.81	145	133.58
chr5:36341637-36341890	RANBP3L	upstream	20.13	17	134.56
chr5:36911650-36913606	NIPBL	promoter	4.74	103	106.91
chr5:43027710-43029352			4.36	105	90.45
chr5:74842885-74843826	COL4A3BP,POLK	promoter	6.19	51	99.08
chr6:10909536-10909757	MAK	body	20.83	14	140.47
chr6:109910129-109911249	ZBTB24	promoter	5.6	41	56.78
chr6:114397754-114399542	HDAC2	promoter	6.49	95	137.56
chr6:137155129-137156222	MAP3K5	promoter	7.09	84	178.43
chr6:142509768-142510478	VTA1	promoter	11.58	49	136.13
chr6:163754408-163755513	QKI	promoter	6.14	37	69.47
chr6:41861658-41863865	TOMM6	promoter	6.46	132	96.4
	SNHG5,SNORD50A, SNORD50B				
chr6:86444545-86445994	AKIRIN2,NCRNA001 20	promoter	7.66	109	197.73
chr6:88467636-88469003		promoter	7.83	90	175.61
chr7:105581441-105581841			12	20	78.09
chr7:116289481-116290748	CAPZA2	promoter	6.3	66	82.89
chr7:1480097-1480815	INTS1	body	12.83	39	132.6
chr7:4889432-4890279	RADIL	promoter	10.44	36	92.26
chr7:5402506-5404226	TNRC18	body	6.13	119	130.6
chr7:55400456-55401923	LANCL2	promoter	5.96	74	103.72
chr7:6278588-6279343	CYTH3	promoter	6.94	47	98.07
chr7:65783909-65785122	LOC100289135	TSE	6.92	65	120.47
chr7:75461351-75462413	TMEM120A	promoter	7.31	44	74.47
chr7:75905120-75905486	ZP3	body	20.74	21	141.62

chr7:87342930-87344460	SLC25A40,DBF4	promoter	8.38	79	130.92
chr7:99516745-99517741	ZNF3	promoter	10.31	87	208.49
chr7:99904614-99905602	LOC402573	upstream	9.97	51	111.46
chr8:142795117-142795846			15.31	30	104.36
chr8:144983269-144984286	PUF60	promoter	7.08	78	170.73
chr8:146589-147330	RPL23AP53	downstream	7.02	39	97.43
chr8:27187763-27188526			8.67	44	94.3
chr8:28535547-28536566			8.05	57	131.34
chr8:90983090-90984625	OSGIN2	promoter	7.27	68	115.28
chr8:91726570-91728158	TMEM64	promoter	5.76	78	129.92
chr9:114134880-114136159	ROD1	promoter	5.79	48	89.5
chr9:133367851-133368844	POMT1	promoter	6.34	38	89.75
chr9:138804959-138806804	TMEM141	promoter	6.13	88	131.77
chr9:139632021-139633388	C9orf37,EHMT1	promoter	5.73	56	85.15
chr9:36125703-36127420	GLIPR2	promoter	4.73	65	84.86
chrX:153371303-153372803	SLC10A3	promoter	7.63	74	104.54
chrX:69426163-69427181	PDZD11,KIF4A	promoter	9.34	53	164.21

## **References**

1. Peale, F. V., Jr., L. Sugden, and M. Bothwell. 1998. Characterization of CMIX, a chicken homeobox gene related to the *Xenopus* gene mix.1. *Mech Dev* 75:167-170.
2. Guo, W., A. P. Chan, H. Liang, E. D. Wieder, J. J. Molldrem, L. D. Etkin, and L. Nagarajan. 2002. A human Mix-like homeobox gene MIXL shows functional similarity to *Xenopus* Mix.1. *Blood* 100:89-95.
3. Pearce, J. J., and M. J. Evans. 1999. Mml, a mouse Mix-like gene expressed in the primitive streak. *Mech Dev* 87:189-192.
4. Robb, L., L. Hartley, C. G. Begley, T. C. Brodnicki, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, and A. G. Elefanty. 2000. Cloning, expression analysis, and chromosomal localization of murine and human homologues of a *Xenopus* mix gene. *Dev Dyn* 219:497-504.
5. Mohn, D., S. W. Chen, D. C. Dias, D. C. Weinstein, M. A. Dyer, K. Sahr, C. E. Ducker, E. Zahradka, G. Keller, K. S. Zaret, L. J. Gudas, and M. H. Baron. 2003. Mouse Mix gene is activated early during differentiation of ES and F9 stem cells and induces endoderm in frog embryos. *Dev Dyn* 226:446-459.
6. Hart, A. H., L. Hartley, K. Sourris, E. S. Stadler, R. Li, E. G. Stanley, P. P. Tam, A. G. Elefanty, and L. Robb. 2002. Mixl1 is required for axial

mesendoderm morphogenesis and patterning in the murine embryo. *Development* 129:3597-3608.

7. Willey, S., A. Ayuso-Sacido, H. Zhang, S. T. Fraser, K. E. Sahr, M. J. Adlam, M. Kyba, G. Q. Daley, G. Keller, and M. H. Baron. 2006. Acceleration of mesoderm development and expansion of hematopoietic progenitors in differentiating ES cells by the mouse Mix-like homeodomain transcription factor. *Blood* 107:3122-3130.
8. Wilson, D., G. Sheng, T. Lecuit, N. Dostatni, and C. Desplan. 1993. Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev* 7:2120-2134.
9. Galliot, B., C. de Vargas, and D. Miller. 1999. Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class. *Dev Genes Evol* 209:186-197.
10. Zhang, H., S. T. Fraser, C. Papazoglu, M. E. Hoatlin, and M. H. Baron. 2009. Transcriptional activation by the Mixl1 homeodomain protein in differentiating mouse embryonic stem cells. *Stem Cells* 27:2884-2895.
11. Luu, O., M. Nagel, S. Wacker, P. Lemaire, and R. Winklbauer. 2008. Control of gastrula cell motility by the Goosecoid/Mix.1/ Siamois network: basic patterns and paradoxical effects. *Dev Dyn* 237:1307-1320.
12. Ravasi, T., H. Suzuki, C. V. Cannistraci, S. Katayama, V. B. Bajic, K. Tan, A. Akalin, S. Schmeier, M. Kanamori-Katayama, N. Bertin, P. Carninci, C. O. Daub, A. R. Forrest, J. Gough, S. Grimmond, J. H. Han, T. Hashimoto, W. Hide, O. Hofmann, A. Kamburov, M. Kaur, H. Kawaji, A. Kubosaki, T.

- Lassmann, E. van Nimwegen, C. R. MacPherson, C. Ogawa, A. Radovanovic, A. Schwartz, R. D. Teasdale, J. Tegner, B. Lenhard, S. A. Teichmann, T. Arakawa, N. Ninomiya, K. Murakami, M. Tagami, S. Fukuda, K. Imamura, C. Kai, R. Ishihara, Y. Kitazume, J. Kawai, D. A. Hume, T. Ideker, and Y. Hayashizaki. 2010. An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* 140:744-752.
13. Zhang, P., J. Li, Z. Tan, C. Wang, T. Liu, L. Chen, J. Yong, W. Jiang, X. Sun, L. Du, M. Ding, and H. Deng. 2008. Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. *Blood* 111:1933-1941.
  14. Xu, X., V. L. Browning, and J. S. Odorico. 2011. Activin, BMP and FGF pathways cooperate to promote endoderm and pancreatic lineage cell differentiation from human embryonic stem cells. *Mech Dev* 128:412-427.
  15. Pick, M., L. Azzola, A. Mossman, E. G. Stanley, and A. G. Elefanty. 2007. Differentiation of human embryonic stem cells in serum-free medium reveals distinct roles for bone morphogenetic protein 4, vascular endothelial growth factor, stem cell factor, and fibroblast growth factor 2 in hematopoiesis. *Stem Cells* 25:2206-2214.
  16. Jackson, S. A., J. Schiesser, E. G. Stanley, and A. G. Elefanty. 2010. Differentiating embryonic stem cells pass through 'temporal windows' that mark responsiveness to exogenous and paracrine mesendoderm inducing signals. *PLoS One* 5:e10706.

17. Mead, P. E., I. H. Brivanlou, C. M. Kelley, and L. I. Zon. 1996. BMP-4-responsive regulation of dorsal-ventral patterning by the homeobox protein Mixl1. *Nature* 382:357-360.
18. Hart, A. H., T. A. Willson, M. Wong, K. Parker, and L. Robb. 2005. Transcriptional regulation of the homeobox gene Mixl1 by TGF-beta and FoxH1. *Biochem Biophys Res Commun* 333:1361-1369.
19. Izzi, L., C. Silvestri, I. von Both, E. Labbe, L. Zakin, J. L. Wrana, and L. Attisano. 2007. Foxh1 recruits Gsc to negatively regulate Mixl1 expression during early mouse development. *EMBO J* 26:3132-3143.
20. Ng, E. S., L. Azzola, K. Sourris, L. Robb, E. G. Stanley, and A. G. Elefanty. 2005. The primitive streak gene Mixl1 is required for efficient haematopoiesis and BMP4-induced ventral mesoderm patterning in differentiating ES cells. *Development* 132:873-884.
21. Drakos, E., G. Z. Rassidakis, V. Leventaki, W. Guo, L. J. Medeiros, and L. Nagarajan. 2007. Differential expression of the human MIXL1 gene product in non-Hodgkin and Hodgkin lymphomas. *Hum Pathol* 38:500-507.
22. Glaser, S., D. Metcalf, L. Wu, A. H. Hart, L. DiRago, S. Mifsud, A. D'Amico, S. Dagger, C. Campo, A. C. Chan, D. J. Izon, and L. Robb. 2006. Enforced expression of the homeobox gene Mixl1 impairs hematopoietic differentiation and results in acute myeloid leukemia. *Proc Natl Acad Sci U S A* 103:16460-16465.

23. Metcalf, D., S. Glaser, S. Mifsud, L. Di Rago, and L. Robb. 2007. The preleukemic state of mice reconstituted with Mixl1-transduced marrow cells. *Proc Natl Acad Sci U S A* 104:20013-20018.
24. Hwang, H. C., C. P. Martins, Y. Bronkhorst, E. Randel, A. Berns, M. Fero, and B. E. Clurman. 2002. Identification of oncogenes collaborating with p27Kip1 loss by insertional mutagenesis and high-throughput insertion site analysis. *Proc Natl Acad Sci U S A* 99:11293-11298.
25. Lawrence, H. J., C. D. Helgason, G. Sauvageau, S. Fong, D. J. Izon, R. K. Humphries, and C. Largman. 1997. Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis. *Blood* 89:1922-1930.
26. Andreeff, M., V. Ruvolo, S. Gadgil, C. Zeng, K. Coombes, W. Chen, S. Kornblau, A. E. Barón, and H. A. Drabkin. 2008. HOX expression patterns identify a common signature for favorable AML. *Leukemia* 22:2041-2047.
27. Drabkin, H. A., C. Parsy, K. Ferguson, F. Guilhot, L. Lacotte, L. Roy, C. Zeng, A. Baron, S. P. Hunger, M. Varella-Garcia, R. Gemmill, F. Brizard, A. Brizard, and J. Roche. 2002. Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. *Leukemia* 16:186-195.
28. Pandolfi, A., and U. Steidi. 2012. HLX in AML: novel prognostic and therapeutic target. *Oncotarget*.
29. Kawahara, M., A. Pandolfi, B. Bartholdy, L. Barreyro, B. Will, M. Roth, U. C. Okoye-Okafor, T. I. Todorova, M. E. Figueroa, A. Melnick, C. S.

- Mitsiades, and U. Steidl. 2012. H2.0-like homeobox regulates early hematopoiesis and promotes acute myeloid leukemia. *Cancer Cell* 22:194-208.
30. Scholl, C., D. Bansal, K. Döhner, K. Eiwen, B. J. Huntly, B. H. Lee, F. G. Rücker, R. F. Schlenk, L. Bullinger, H. Döhner, D. G. Gilliland, and S. Fröhling. 2007. The homeobox gene CDX2 is aberrantly expressed in most cases of acute myeloid leukemia and promotes leukemogenesis. *J Clin Invest* 117:1037-1048.
31. Rawat, V. P., S. Thoene, V. M. Naidu, N. Arseni, B. Heilmeier, K. Metzeler, K. Petropoulos, A. Deshpande, L. Quintanilla-Martinez, S. K. Bohlander, K. Spiekermann, W. Hiddemann, M. Feuring-Buske, and C. Buske. 2008. Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood* 111:309-319.
32. Riedt, T., M. Ebinger, H. R. Salih, J. Tomiuk, R. Handgretinger, L. Kanz, F. Grunebach, and C. Lengerke. 2009. Aberrant expression of the homeobox gene CDX2 in pediatric acute lymphoblastic leukemia. *Blood* 113:4049-4051.
33. Choi, C. W., Y. J. Chung, C. Slape, and P. D. Aplan. 2009. A NUP98-HOXD13 fusion gene impairs differentiation of B and T lymphocytes and leads to expansion of thymocytes with partial TCRB gene rearrangement. *J Immunol* 183:6227-6235.



34. Chung, K. Y., G. Morrone, J. J. Schuringa, M. Plasilova, J. H. Shieh, Y. Zhang, P. Zhou, and M. A. Moore. 2006. Enforced expression of NUP98-HOXA9 in human CD34(+) cells enhances stem cell proliferation. *Cancer Res* 66:11781-11791.
35. Argiropoulos, B., and R. K. Humphries. 2007. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* 26:6766-6776.
36. Gilmore, T. D., D. Kalaitzidis, M. C. Liang, and D. T. Starczynowski. 2004. The c-Rel transcription factor and B-cell proliferation: a deal with the devil. *Oncogene* 23:2275-2286.
37. Zhao, C., Y. Xiu, J. Ashton, L. Xing, Y. Morita, C. T. Jordan, and B. F. Boyce. 2012. Noncanonical NF- $\kappa$ B signaling regulates hematopoietic stem cell self-renewal and microenvironment interactions. *Stem Cells* 30:709-718.
38. Curry, C. V., A. A. Ewton, R. J. Olsen, B. R. Logan, H. A. Preti, Y. C. Liu, S. L. Perkins, and C. C. Chang. 2009. Prognostic impact of C-REL expression in diffuse large B-cell lymphoma. *J Hematop* 2:20-26.
39. Birkenkamp, K. U., M. Geugien, H. Schepers, J. Westra, H. H. Lemmink, and E. Vellenga. 2004. Constitutive NF-kappaB DNA-binding activity in AML is frequently mediated by a Ras/PI3-K/PKB-dependent pathway. *Leukemia* 18:103-112.
40. Iqbal, J., W. G. Sanger, D. E. Horsman, A. Rosenwald, D. L. Pickering, B. Dave, S. Dave, L. Xiao, K. Cao, Q. Zhu, S. Sherman, C. P. Hans, D. D. Weisenburger, T. C. Greiner, R. D. Gascoyne, G. Ott, H. K. Müller-

- Hermelink, J. Delabie, R. M. Braziel, E. S. Jaffe, E. Campo, J. C. Lynch, J. M. Connors, J. M. Vose, J. O. Armitage, T. M. Grogan, L. M. Staudt, and W. C. Chan. 2004. BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. *Am J Pathol* 165:159-166.
41. Liang, H., S. Samanta, and L. Nagarajan. 2005. SSBP2, a candidate tumor suppressor gene, induces growth arrest and differentiation of myeloid leukemia cells. *Oncogene* 24:2625-2634.
42. Chadee, D. N., M. J. Hendzel, C. P. Tylicski, C. D. Allis, D. P. Bazett-Jones, J. A. Wright, and J. R. Davie. 1999. Increased Ser-10 phosphorylation of histone H3 in mitogen-stimulated and oncogene-transformed mouse fibroblasts. *J Biol Chem* 274:24914-24920.
43. Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, and X. S. Liu. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9:R137.
44. Bailey, T. L., and M. Gribskov. 1998. Methods and statistics for combining motif match scores. *J Comput Biol* 5:211-221.
45. Bailey, T. L., and C. Elkan. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2:28-36.
46. Li, C., and W. H. Wong. 2001. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A* 98:31-36.

47. Network, C. G. A. R. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 368:2059-2074.
48. Cerami, E., J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, and N. Schultz. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2:401-404.
49. Gao, J., B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, and N. Schultz. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6:pl1.
50. Akiyama, Y. 1995. TFSEARCH: Searching Transcription Factor Binding Sites.
51. Hromas, R., S. Boswell, R. N. Shen, G. Burgess, A. Davidson, K. Cornetta, J. Sutton, and K. Robertson. 1996. Forced over-expression of the myeloid zinc finger gene MZF-1 inhibits apoptosis and promotes oncogenesis in interleukin-3-dependent FDCP.1 cells. *Leukemia* 10:1049-1050.
52. Robertson, K. A., D. P. Hill, M. R. Kelley, R. Tritt, B. Crum, S. Van Epps, E. Srour, S. Rice, and R. Hromas. 1998. The myeloid zinc finger gene (MZF-1) delays retinoic acid-induced apoptosis and differentiation in myeloid leukemia cells. *Leukemia* 12:690-698.

53. Sahr, K., D. C. Dias, R. Sanchez, D. Chen, S. W. Chen, L. J. Gudas, and M. H. Baron. 2002. Structure, upstream promoter region, and functional domains of a mouse and human Mix paired-like homeobox gene. *Gene* 291:135-147.
54. Mizutani, A., D. Koinuma, S. Tsutsumi, N. Kamimura, M. Morikawa, H. I. Suzuki, T. Imamura, K. Miyazono, and H. Aburatani. 2011. Cell type-specific target selection by combinatorial binding of Smad2/3 proteins and hepatocyte nuclear factor 4alpha in HepG2 cells. *J Biol Chem* 286:29848-29860.
55. Xi, Q., Z. Wang, A. I. Zaromytidou, X. H. Zhang, L. F. Chow-Tsang, J. X. Liu, H. Kim, A. Barlas, K. Manova-Todorova, V. Kaartinen, L. Studer, W. Mark, D. J. Patel, and J. Massague. 2011. A poised chromatin platform for TGF-beta access to master regulators. *Cell* 147:1511-1524.
56. Xu, X., V. L. Browning, and J. S. Odorico. 2011. Activin, BMP and FGF pathways cooperate to promote endoderm and pancreatic lineage cell differentiation from human embryonic stem cells. *Mechanisms of development* 128:412-427.
57. Bhatia, M., D. Bonnet, D. Wu, B. Murdoch, J. Wrana, L. Gallacher, and J. E. Dick. 1999. Bone morphogenetic proteins regulate the developmental program of human hematopoietic stem cells. *J Exp Med* 189:1139-1148.
58. Hutton, J. F., V. Rozenkov, F. S. Khor, R. J. D'Andrea, and I. D. Lewis. 2006. Bone morphogenetic protein 4 contributes to the maintenance of

- primitive cord blood hematopoietic progenitors in an ex vivo stroma-noncontact co-culture system. *Stem Cells Dev* 15:805-813.
59. Khurana, S., S. Buckley, S. Schouteden, S. Ekker, A. Petryk, M. Delforge, A. Zwijsen, and C. M. Verfaillie. 2013. A novel role of BMP4 in adult hematopoietic stem and progenitor cell homing via Smad independent regulation of integrin- $\alpha$ 4 expression. *Blood* 121:781-790.
  60. Wattel, E., C. Preudhomme, B. Hecquet, M. Vanrumbeke, B. Quesnel, I. Dervite, P. Morel, and P. Fenaux. 1994. p53 mutations are associated with resistance to chemotherapy and short survival in hematologic malignancies. *Blood* 84:3148-3157.
  61. Trompouki, E., T. V. Bowman, L. N. Lawton, Z. P. Fan, D. C. Wu, A. DiBiase, C. S. Martin, J. N. Cech, A. K. Sessa, J. L. Leblanc, P. Li, E. M. Durand, C. Mosimann, G. C. Heffner, G. Q. Daley, R. F. Paulson, R. A. Young, and L. I. Zon. 2011. Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell* 147:577-589.
  62. Cuny, G. D., P. B. Yu, J. K. Laha, X. Xing, J. F. Liu, C. S. Lai, D. Y. Deng, C. Sachidanandan, K. D. Bloch, and R. T. Peterson. 2008. Structure-activity relationship study of bone morphogenetic protein (BMP) signaling inhibitors. *Bioorg Med Chem Lett* 18:4388-4392.
  63. Vogt, J., R. Traynor, and G. P. Sapkota. 2011. The specificities of small molecule inhibitors of the TGF $\beta$ s and BMP pathways. *Cell Signal* 23:1831-1842.

64. Steinbicker, A. U., C. Sachidanandan, A. J. Vonner, R. Z. Yusuf, D. Y. Deng, C. S. Lai, K. M. Rauwerdink, J. C. Winn, B. Saez, C. M. Cook, B. A. Szekely, C. N. Roy, J. S. Seehra, G. D. Cuny, D. T. Scadden, R. T. Peterson, K. D. Bloch, and P. B. Yu. 2011. Inhibition of bone morphogenetic protein signaling attenuates anemia associated with inflammation. *Blood* 117:4915-4923.
65. Yu, P. B., D. Y. Deng, C. S. Lai, C. C. Hong, G. D. Cuny, M. L. Bouxsein, D. W. Hong, P. M. McManus, T. Katagiri, C. Sachidanandan, N. Kamiya, T. Fukuda, Y. Mishina, R. T. Peterson, and K. D. Bloch. 2008. BMP type I receptor inhibition reduces heterotopic [corrected] ossification. *Nat Med* 14:1363-1369.
66. Jolma, A., J. Yan, T. Whittington, J. Toivonen, K. R. Nitta, P. Rastas, E. Morgunova, M. Enge, M. Taipale, G. Wei, K. Palin, J. M. Vaquerizas, R. Vincentelli, N. M. Luscombe, T. R. Hughes, P. Lemaire, E. Ukkonen, T. Kivioja, and J. Taipale. 2013. DNA-binding specificities of human transcription factors. *Cell* 152:327-339.
67. Hromas, R., B. Davis, F. J. Rauscher, M. Klemsz, D. Tenen, S. Hoffman, D. Xu, and J. F. Morris. 1996. Hematopoietic transcriptional regulation by the myeloid zinc finger gene, MZF-1. *Curr Top Microbiol Immunol* 211:159-164.
68. Hui, P., X. Guo, and P. G. Bradford. 1995. Isolation and functional characterization of the human gene encoding the myeloid zinc finger protein MZF-1. *Biochemistry* 34:16493-16502.

69. Mudduluru, G., P. Vajkoczy, and H. Allgayer. 2010. Myeloid zinc finger 1 induces migration, invasion, and in vivo metastasis through Axl gene expression in solid cancer. *Mol Cancer Res* 8:159-169.
70. Rafn, B., C. F. Nielsen, S. H. Andersen, P. Szyniarowski, E. Corcelle-Termeau, E. Valo, N. Fehrenbacher, C. J. Olsen, M. Daugaard, C. Egebjerg, T. Bøttzauw, P. Kohonen, J. Nylandsted, S. Hautaniemi, J. Moreira, M. Jäättelä, and T. Kallunki. 2012. ErbB2-driven breast cancer cell invasion depends on a complex signaling network activating myeloid zinc finger-1-dependent cathepsin B expression. *Mol Cell* 45:764-776.
71. Yue, C. H., Y. W. Chiu, J. N. Tung, B. S. Tzang, J. J. Shiu, W. H. Huang, J. Y. Liu, and J. M. Hwang. 2012. Expression of protein kinase C  $\alpha$  and the MZF-1 and Elk-1 transcription factors in human breast cancer cells. *Chin J Physiol* 55:31-36.
72. Kluk, B. J., Y. Fu, T. A. Formolo, L. Zhang, A. K. Hindle, Y. G. Man, R. S. Siegel, P. E. Berg, C. Deng, T. A. McCaffrey, and S. W. Fu. 2010. BP1, an isoform of DLX4 homeoprotein, negatively regulates BRCA1 in sporadic breast cancer. *Int J Biol Sci* 6:513-524.
73. Topcu, Z., D. L. Mack, R. A. Hromas, and K. L. Borden. 1999. The promyelocytic leukemia protein PML interacts with the proline-rich homeodomain protein PRH: a RING may link hematopoiesis and growth control. *Oncogene* 18:7091-7100.
74. Pereira, L. A., M. S. Wong, S. M. Lim, A. Sides, E. G. Stanley, and A. G. Elefanty. 2011. Brachyury and related Tbx proteins interact with the Mixl1

- homeodomain protein and negatively regulate Mixl1 transcriptional activity. PLoS One 6:e28394.
75. Garg, V., I. S. Kathiriya, R. Barnes, M. K. Schluterman, I. N. King, C. A. Butler, C. R. Rothrock, R. S. Eapen, K. Hirayama-Yamada, K. Joo, R. Matsuoka, J. C. Cohen, and D. Srivastava. 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424:443-447.
  76. Banerjee, A., R. Grumont, R. Gugasyan, C. White, A. Strasser, and S. Gerondakis. 2008. NF-kappaB1 and c-Rel cooperate to promote the survival of TLR4-activated B cells by neutralizing Bim via distinct mechanisms. *Blood* 112:5063-5073.
  77. Fu, T., P. Li, H. Wang, Y. He, D. Luo, A. Zhang, W. Tong, L. Zhang, B. Liu, and C. Hu. 2009. c-Rel is a transcriptional repressor of EPHB2 in colorectal cancer. *J Pathol* 219:103-113.
  78. Sovak, M. A., R. E. Bellas, D. W. Kim, G. J. Zanieski, A. E. Rogers, A. M. Traish, and G. E. Sonenshein. 1997. Aberrant nuclear factor-kappaB/Rel expression and the pathogenesis of breast cancer. *J Clin Invest* 100:2952-2960.
  79. Mukhopadhyay, T., J. A. Roth, and S. A. Maxwell. 1995. Altered expression of the p50 subunit of the NF-kappa B transcription factor complex in non-small cell lung carcinoma. *Oncogene* 11:999-1003.
  80. Liu, C. J., S. C. Lin, Y. J. Chen, K. M. Chang, and K. W. Chang. 2006. Array-comparative genomic hybridization to detect genomewide changes



in microdissected primary and metastatic oral squamous cell carcinomas. *Mol Carcinog* 45:721-731.

81. Yang, X., H. Lu, B. Yan, R. A. Romano, Y. Bian, J. Friedman, P. Duggal, C. Allen, R. Chuang, R. Ehsanian, H. Si, S. Sinha, C. Van Waes, and Z. Chen. 2011.  $\Delta$ Np63 versatilely regulates a Broad NF- $\kappa$ B gene program and promotes squamous epithelial proliferation, migration, and inflammation. *Cancer Res* 71:3688-3700.
82. Qu, Y., F. Zhou, X. Dai, H. Wang, J. Shi, X. Zhang, Y. Wang, and W. Wei. 2011. Clinicopathologic significances of nuclear expression of nuclear factor- $\kappa$ B transcription factors in retinoblastoma. *J Clin Pathol* 64:695-700.
83. Pallares, J., J. L. Martínez-Guitarte, X. Dolcet, D. Llobet, M. Rue, J. Palacios, J. Prat, and X. Matias-Guiu. 2004. Abnormalities in the NF-kappaB family and related proteins in endometrial carcinoma. *J Pathol* 204:569-577.
84. Kallifatidis, G., S. Labsch, V. Rausch, J. Mattern, J. Gladkikh, G. Moldenhauer, M. W. Büchler, A. V. Salnikov, and I. Herr. 2011. Sulforaphane increases drug-mediated cytotoxicity toward cancer stem-like cells of pancreas and prostate. *Mol Ther* 19:188-195.
85. Le Beau, M. M., C. Ito, P. Cogswell, R. Espinosa, A. A. Fernald, and A. S. Baldwin. 1992. Chromosomal localization of the genes encoding the p50/p105 subunits of NF-kappa B (NFKB2) and the I kappa B/MAD-3 (NFKBI) inhibitor of NF-kappa B to 4q24 and 14q13, respectively. *Genomics* 14:529-531.

86. Stankovic, S., R. Gugasyan, K. Kyparissoudis, R. Grumont, A. Banerjee, P. Tschlis, S. Gerondakis, and D. I. Godfrey. 2011. Distinct roles in NKT cell maturation and function for the different transcription factors in the classical NF- $\kappa$ B pathway. *Immunol Cell Biol* 89:294-303.
87. Gilmore, T. D. 1992. Role of rel family genes in normal and malignant lymphoid cell growth. *Cancer Surv* 15:69-87.

Aaron Craig Raymond was born to of Nancy E. Witto Raymond and Craig R. Raymond in Winston-Salem, North Carolina on November 18, 1983. Aaron graduated high school at Simsbury High School in Simsbury Connecticut in 2002, and began his undergraduate education at Rensselaer Polytechnic Institute in Troy, NY later that year. He received a Bachelor of Science with a major in “Bioinformatics and Molecular Biology” and a minor in Music in May 2006. That August he entered the UT Graduate School of Biomedical Sciences at Houston. December 2009, Aaron received his Masters of Science from the GSBS Genes and Development program, with a thesis entitled “FZD6 AND MRS2: TWO PUTATIVE MIXL1 TRANSCRIPTIONAL TARGETS”, as a member of the Dr. Lalitha Nagarajan lab in the department of Genetics of UT MD Anderson Cancer Center.

Permanent address:

1711 Old Spanish Trail, Unit 413

Houston, Texas 77054