UNIVERSITY OF CALIFORNIA

Los Angeles

The Role of Long Non-coding RNAs in B-acute Lymphoblastic Leukemia

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy in Cellular and Molecular Pathology

by

Norma Iris Rodríguez-Malavé

2015

ABSTRACT OF THE DISSERTATION

The Role of Long Non-coding RNAs in B-acute Lymphoblastic Leukemia

by

Norma Iris Rodríguez-Malavé

Doctor of Philosophy in Cellular and Molecular Pathology

University of California, Los Angeles, 2015

Professor Dinesh S. Rao, Chair

Non-coding RNAs play pivotal roles in a wide variety of molecular processes. The functions of long non-coding RNAs (lncRNAs), in particular, have deep implications for both development and oncogenesis. Dysregulated expression of lncRNAs has been found in various cancers, but had not been comprehensively described in B lymphoblastic leukemia (B acute lymphoblastic leukemia; B-ALL). We completed a gene expression profiling study in human B-ALL samples and found differential lncRNA expression in samples with particular cytogenetic abnormalities. We determined that lncRNA expression could discriminate between B-ALL with specific karyotype abnormalities as well as, predict patient survival. Two promising lncRNAs from our study, designated B-ALL associated long RNAs (BALRs), have the highest expression in B-ALL samples carrying the MLL rearrangement when compared non-MLL rearranged and normal CD19+ cells. MLL rearranged B-ALL cases have a very poor prognosis and occur frequently in infants, making them particularly difficult to treat. This thesis investigates the role of lncRNAs BALR-2 and BALR-6 in MLL translocated B-ALL.

In the first part of this thesis work, we found that high expression of BALR-2 was correlated with diminished response to prednisone treatment. Knockdown led to a reduction in proliferation, increased apoptosis, and increased sensitivity to prednisolone treatment. Conversely, overexpression of the lncRNA caused increased cell growth and resistance to prednisone treatment. Remarkably, BALR-2 expression was repressed by prednisolone treatment and its dysregulation led to changes in the glucocorticoid response pathway in human and mouse B-cells. These findings indicate an important role for BALR-2 in the pathogenesis of B-ALL.

Much like BALR-2, siRNA mediated knockdown of BALR-6 in human B-ALL cell lines caused decreased proliferation and increased apoptosis. Additionally, overexpression of BALR-6 isoforms caused a significant increase in progenitor populations in mice and increased proliferation in mammalian cell lines. To understand the functional role of BALR-6, differential expression analysis from cell lines with knockdown was carried out. The analysis indicated an enrichment of genes involved in leukemia. Among these genes were SP1 and its known target genes. Luciferase reporter assays uncovered a positive regulatory role for BALR-6 in SP1 mediated transcription. Together, these data elucidate a role for BALR-6 in transcriptional regulation.

Thus, this thesis identifies novel non-coding RNA transcripts that regulate gene expression, and thereby pathogenesis in B-ALL with MLL rearrangement. This work suggests novel diagnostic, prognostic, and therapeutic utility for lncRNAs in B-ALL.

The dissertation of Norma Iris Rodríguez-Malavé is approved.

Gay M. Crooks

David W. Dawson

Gregory S. Payne

Dinesh S. Rao, Committee Chair

University of California, Los Angeles

2015

DEDICATION

This dissertation in dedicated to my family: my parents, Carmelo and Norma; my brothers, Miguel and Richard; my beloved, Michael; and the love of my life, Jäger.

TABLE OF CONTENTS

FIGURES AND TABLES

CHAPTER I

CHAPTER III

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God for all the blessings in my life. I want to thank my advisor Dr. Dinesh Rao for giving me the opportunity to carry out my research studies and providing support, guidance, and mentorship though these years. I am sincerely grateful to have worked in his lab. I would like to thank my committee for their guidance and advice in science and in my career choices. Also, I want to thank all the members of the Rao Lab through the years that have taught me, guided me, supported me, and lent a hand. In particular, Thilini Fernando, Jayanth Palanichamy, Tiffany Tran, Jorge Contreras, Neha Goswami, Parth Patel, Jasmine Gajeton, Jaime Anguiano, Kim Pioli, and Mike Alberti. You are my family forever.

The work presented in this thesis was supported in part by the Eugene V. Cota-Robles Fellowship from UCLA and the Graduate Research Fellowship Program from the National Science Foundation (DGE-1144087). Additional support was given by a Career Development Award K08CA133521, the Sidney Kimmel Translational Scholar Award SKF-11-013, the Irving Feintech Family Foundation/Tower Cancer Research Foundation Research Grant, the University of California Cancer Research Coordinating Committee, the Stein-Oppenheimer Endowment Award, and the UCLA Broad Stem Cell Research Center to Dr. Dinesh Rao.

The work presented in Chapter I is a version of Rodriguez-Malave and Rao “Long non-coding RNAs in hematopoietic malignancies” that has been accepted for publication at Briefings in Functional Genomics. The work presented in Chapter II is a reprint of Fernando et al. “LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia” published in February 2015 with permission from Molecular Cancer Research. The research was carried out under the direction of PI Dinesh S. Rao and was primarily lead by Dr. Thilini R. Fernando and myself. We conceptualized, planned, carried out, and interpreted the experiments. Completion of the manuscript was made possible by the important experimental

contributions from Ella V. Waters, Weihong Yan, David Casero, Martina Pigazzi, and Giuseppe Basso.

The work presented in Chapter III is a version of Rodriguez-Malave et al. “BALR-6 regulates cell growth and cell survival in B-acute lymphoblastic leukemia” which is currently in revision at Molecular Cancer. The research was under the direction of PI Dinesh S. Rao. I lead, conceptualized, planned, executed, and interpreted the experiments. Additionally, I compiled and constructed the manuscript. Experimental contributions were made by, Thilini R. Fernando, Parth C. Patel, Jorge R. Contreras, Jayanth K. Palanichamy, Tiffany M. Tran, Jaime Anguiano, Kimanh T. Pioli, Michael O. Alberti, and Salemiz Sandoval. Contribution to result interpretation and construction of the manuscript was made by Michael J. Davoren, Gay M. Crooks, and Dinesh S. Rao. We would like to thank the UCLA Clinical Microarray Core for performing the microarray hybridization experiments, Alejandro Balazs at Caltech for lentiviral vector backbones, and Ken Dorshkind for helpful discussions. Also, we thank Neha Goswami, Ella Waters, Nolan Ung, Jennifer King, Jasmine Gajeton, and May Paing for their technical support. Very special thanks to Diana C Márquez-Garbán at Dr. Richard Pietras lab (UCLA) for use of their facilities for some of the experiments.

The works included in the appendices are reprints of Fernando et al. “MicroRNAs in B cell development and malignancy” and of Contreras et al. “MicroRNA-146a modulates B-cell oncogenesis by regulating Egr1” with permission from the Journal of Hematology and Oncology and Oncotarget, respectively. I contributed to the conceptualization and construction of the article in Appendix I. Notably, the work in Appendix II was mainly carried out by Jorge Contreras in collaboration with other members of the Rao Lab. I assisted in experimental design, execution, data analysis, and the overall construction of the article.

Finally, I would like to thank all my friends and family for their unconditional love and support. I could not have made it so far without all of you.

VITA

NORMA IRIS RODRÍGUEZ-MALAVÉ

EDUCATION

2010-2015 PhD Candidate University of California Los Angeles, CA

Cellular and Molecular Pathology Program, Cum Laude

2009-2010 Graduate Student University of Puerto Rico Mayagüez, PR

Eukaryote Molecular Genetics, and Human Population Genetics, Magna Cum Laude

2009 Bachelors in Science (B.S.) University of Puerto Rico Mayagüez, PR

Industrial Biotechnology, Magna Cum Laude

PUBLICATIONS

- **Rodríguez-Malavé NI**, Fernando TR, Patel PC, Contreras JR , Palanichamy JK , Tran TM , Anguiano J, Davoren MJ, Alberti MO, Pioli K, Sandoval S, Crooks GM, and Rao DS (**2015**) "BALR-6 regulates cell growth and cell survival in B-acute lymphoblastic leukemia" *Molecular Cancer*, submitted (addressing reviews) (MOLC-D-15-00459).
- **Rodríguez-Malavé NI**, and Rao DS (**2015**) "Long non-coding RNAs in hematopoietic malignancies" *Briefings in Functional Genomics* (accepted) (BFGP-15-0039.R1).
- Contreras JR , Palanichamy JK , Tran T , Fernando TR , **Rodríguez-Malavé NI**, Casero D, Arboleda V, and Rao DS (**2015**) "MicroRNA-146a modulates B-cell oncogenesis by targeting Egr1." *Oncotarget*, Apr 13 (PMID: 25906746).
- **Rodríguez-Malavé NI***, Fernando TR*, Waters EV, Yan W, Casero D, Basso G, Pigazzi M, and Rao DS (**2015**) "LncRNA expression discriminates karyotype and regulates cell survival in B-lymphoblastic leukemia." *Molecular Cancer Research*, May, **13**:5, 839-851. ***Co-authorship**.
- Fernando TR, **Rodríguez-Malavé NI**, and Rao DS (**2012**) "microRNAs in B-cell development and malignancy." *Journal of Hematology and Oncology*, Mar 8, **5**:7.

RESEARCH EXPERIENCE

July 2011-Dec 2015 **University of California** **Los Angeles, CA**

Graduate Student Researcher Principal Investigator**: Dinesh S. Rao, MD, PhD.**

Primary Project: Elucidating the role of long intervening non-coding RNA in B-lymphoblastic leukemia (B-acute lymphoblastic leukemia, B-ALL).

Aug 2009-Jul 2010 **University of Puerto Rico** **Mayagüez, PR**

Graduate Student Researcher, Principal Investigators: **Taras K. Oleksyk PhD. & Juan Carlos Martinez-Cruzado, PhD.**

ADDITIONAL EXPERIENCE

Sept 2013-Dec 2013 **University of California** **Los Angeles, CA**

Teaching Assistant, PATH 222

Jan 2012-Mar 2012 **University of California** **Los Angeles, CA**

Teaching Assistant, MCDB CM156

AWARDS AND FELLOWSHIPS

Dec 2013 **American Society of Hematology** **Washington, DC**

Minority Graduate Student Abstract Achievement Award

Apr 2011–June 2015 **National Science Foundation** **Arlington, VA**

Graduate Research Fellowship Program

Sep 2010–June2012 **University of California** **Los Angeles, CA**

Eugene Cota Robles Fellowship

Sep 2010 **University of California** **Los Angeles, CA**

Paul D. Boyer Award

CHAPTER I:

Introduction:

“Long Non-Coding RNAs in Hematopoietic Malignancies”

Abstract

Recent years have witnessed the discovery of several classes of non-coding RNAs (ncRNAs), which are indispensable for the regulation of various cellular processes. Many of these RNAs are regulatory in nature with functions in gene expression regulation such as piRNAs, siRNAs and miRNAs. Long non-coding RNAs (lncRNAs) comprise the most recently characterized class. LncRNAs are involved in transcriptional regulation, chromatin remodeling, imprinting, splicing, and translation, among other critical functions in the cell. Recent studies have elucidated the importance of lncRNAs in hematopoietic development. Dysregulation of lncRNA expression is a feature of various diseases and cancers, and is also seen in hematopoietic malignancies. This article focuses on lncRNAs that have been implicated in the pathogenesis of hematopoietic malignancies.

Introduction

Blood cell development is a complex and highly ordered process that occurs as successive waves of fetal and embryonic hematopoiesis prior to definitive adult hematopoiesis in mammals. The bone marrow is the site of adult hematopoiesis in mammals, leading to the genesis of all the major lineages in the blood including lymphoid cells, myeloid cells, as well as anucleated red blood cells and platelets. This complex developmental process is intimately connected with the regulation of gene expression, as hematopoietic stem cells become successively committed to one or the other developmental lineage [1]. This system has been extensively studied in model organisms, including mice, with several important insights. Hierarchical models of development have been established, with loss of developmental potential characterizing each successive step. Several models of development exist and specify different relationships between progenitor cells in the bone marrow [1-7]. Most models include separate hierarchies for developing myeloid and lymphoid cells, with the former giving rise to

erythrocytes, platelets, granulocytes, and monocytes, and the latter giving rise to B- and T-lymphocytes [2, 6]. Importantly, specific transcriptional regulators, which tightly control gene expression, are known to function at different branch points in the hematopoietic hierarchy [8-10]. In addition, changes in chromosomal accessibility and epigenetic regulation play highly important roles. New discoveries showing that non-coding elements of the genome can also control gene expression may help us better understand both hematopoietic development and cancer.

Long non-coding RNAs (lncRNAs) are a recently described class of non-coding transcripts, the importance of which is still being recognized [11-15]. These RNAs are >200 nucleotides in length, transcribed by RNA polymerase II or III, undergo splicing, and are usually polyadenylated. They generally lack open reading frames (ORFs) of significant length, although a few may contain small ORFs within one or more alternative splice forms [16, 17]. These transcripts show functional and genomic conservation in vertebrates, despite rapid sequence evolution [18]. Over 10,000 lncRNAs have been identified so far; some of these are well described, while the majority requires further characterization [15, 18-23]. These transcripts can be classified by their position relative to neighboring protein coding genes: sense, antisense, intronic, exonic, or intergenic/intervening [24, 25]. LncRNAs are important regulators of a wide variety of cellular processes with distinct functions that differ based on the cellular compartment in which they are found. Nuclear and cytoplasmic transcripts can regulate gene expression at the transcriptional or translational level, respectively (Figure 1). They can recruit chromatin-remodeling complexes, regulate transcription, promote mRNA translation, and stabilize mRNA transcripts to prevent or induce decay (Figure 1A-D) [26-30]. Additionally, they can influence changes in the spatial conformation of chromosomes, imprinting, and splicing (Figure 1E) [17, 31-34]. A few loci that produce lncRNAs can also produce small biologically active peptides, and the notation of such loci as “non-coding” can therefore be debated [17]. Nonetheless, a variety

of non-coding functions can be ascribed to the majority of lncRNAs; some can serve as sponges competing for microRNAs, and some even harbor microRNAs within their transcripts (Figure 1F) [35-37]. The fact that lncRNAs are intrinsically involved with a diverse array of cellular mechanisms attests to potentially pivotal roles in development in general, and to hematopoiesis in particular.

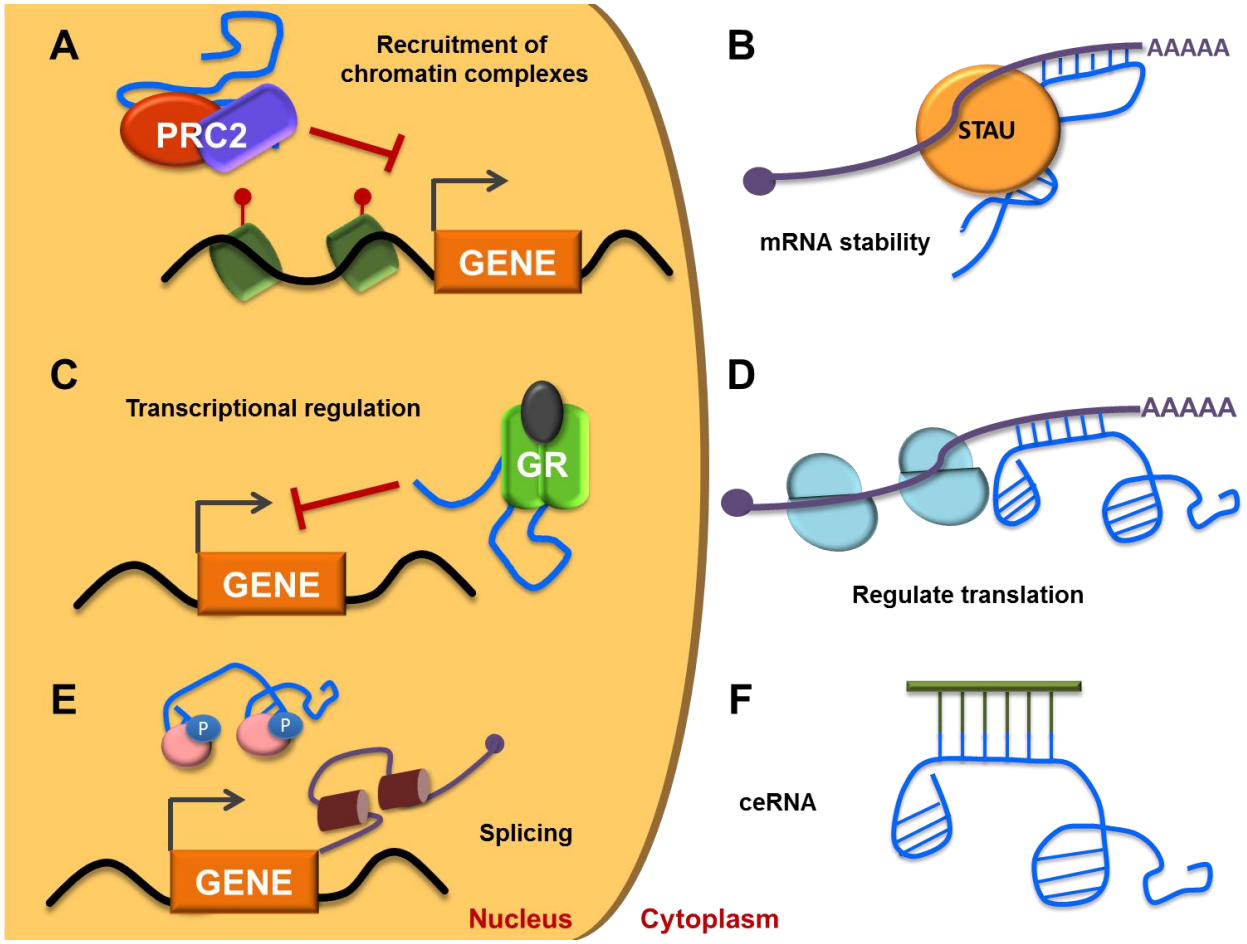


Figure 1: Schematic of lncRNA mechanisms of action. A) Epigenetic silencing of target loci by recruitment of chromatin-remodeling complexes, such as PRC2. **B)** mRNA stability is regulated by the interaction of between lncRNAs and STAU proteins. **C)** lncRNAs can regulate transcription by acting as decoys for transcription factors, like the glucocorticoid receptor (GR), inhibiting their binding to target promoter sequences. **D)** Translation of mRNAs can be inhibited or promoted by lncRNAs. **E)** Alternative mRNA splicing modulated by lncRNAs that regulate SR splicing factor phosphorylation. **F)** lncRNAs can act as competitive endogenous RNA (ceRNA) by binding to target microRNAs.

Indeed, several recent studies have revealed that lncRNAs can regulate hematopoietic development, influencing differentiation, proliferation, and cell survival. They have important regulatory roles at many developmental stages including, the determination of hematopoietic stem cell (HSC) fate, and the differentiation of progenitor and precursor blood cells of lymphoid, myeloid, and erythroid lineages [38-43]. The relatively well-known lncRNA, H19, is involved in maternal imprinting, and was one of the first lncRNAs described to function in adult HSC quiescence. H19 is highly expressed in long-term HSCs, with gradual downregulation in short-term HSCs. This transcript promotes HSC quiescence by regulating the Igf2-Igf1r pathway [38]. Thymic (T-lymphocyte) specific lncRNAs have also been described. One example of this is thymus specific non-coding RNA (Thy-ncR1), a cytoplasmic riboregulator of MFAP degradation [39]. BIC, which harbors miR-155-5p and miR-155-3p, is critical in hematopoietic lineage differentiation and in activation of mature B-cells [40, 44]. In the myeloid lineage, HOX antisense intergenic RNA myeloid 1 (HOTAIRM1), induces the expression of myeloblast differentiation genes, including HOXA1 and HOXA4 [41]. Similarly, the Eosinophil Granule Ontogeny lncRNA (EGO) is involved in stimulating eosinophilic differentiation of CD34+ progenitor cells by regulating protein expression [42]. During murine erythroid differentiation, the Erythroid Pro-Survival lincRNA (lincRNA-EPS) promotes terminal differentiation of mature erythrocytes by inhibiting pro-apoptotic gene *Pycard* [43]. Hence, lncRNAs play a variety of roles in controlling different steps in hematopoietic differentiation, including the maintenance of HSCs and the differentiation of myeloid, erythroid, and lymphoid lineages.

The disruption of the highly ordered differentiation of hematopoietic elements by somatic mutations leads to cancers of the hematopoietic system. These diseases can broadly be divided into lymphoid and myeloid diseases. Also, these malignancies can be thought of as diseases derived from the bone marrow, which occur clinically as a primarily “hematologic” (blood and bone marrow) presentation, or as diseases of mature lymphoid organs, which present primarily

as solid tumors of lymph nodes and other tissues. In this review we will focus on bone marrow derived hematopoietic malignancies, which include diseases derived from both lymphoid and myeloid progenitor cells. These diseases demonstrate many different patient presentations, and current classification systems for the diseases are based both on clinical background and on insights gained from analyses of the molecular alterations. Modern studies increasingly emphasize the role of lncRNAs as important elements in hematopoietic malignant progression (Table 1). Their modulation of a vast amount of cellular processes, especially those involved in differentiation and cell fate, ties them closely to the pathogenesis of these diseases. In the sections below, we will discuss lncRNAs that are involved in myeloid derived diseases, including myelodysplastic syndromes (MDS), myeloproliferative neoplasms (MPN), and acute myeloid leukemia (AML). We will also discuss a disease of lymphoid origin in the context of acute lymphoblastic leukemia (ALL), which constitutes the most important bone marrow-derived lymphoid malignancy and the most common childhood malignancy.

Table 1: lncRNAs implicated in hematopoietic malignancies

Disease	Name	Cellular function	Pathways	Role in hematopoietic malignancies	References
MDS	XIST	Chromatin remodeling, imprinting	X chromosome inactivation	Tumor suppressor	[45, 46]
	MEG3	Transcriptional regulation, protein scaffold	PRC2 scaffold p53 binding enhancer	Unknown, possibly tumor suppressor	[16, 47, 48]
MPN (CML)	BGL3	Competitive endogenous RNA	PTEN and γ-globin expression regulation	Tumor suppressor	[49]
	H19	Transcriptional regulation, miR-675 precursor	Maternal imprinting Rb regulation	Tumor suppressor/Oncogene	[36, 50-53]
	XIST	Chromatin remodeling, imprinting	X chromosome inactivation	Tumor suppressor	[45, 46]
AML	ANRIL	Transcriptional regulation, protein scaffold	PRC1 and PRC2 scaffold	Oncogene	[54-56]
	MEG3	Transcriptional regulation, protein scaffold	PRC2 scaffold p53 binding enhancer	Unknown, possibly tumor suppressor	[16, 47, 48]
	H19	Transcriptional regulation, miR-675 precursor	Maternal imprinting HSC quiescence	Unknown, possibly tumor suppressor	[51]
	BIC	miR-155 precursor	Interleukin-1 signaling B cell activation	Oncogene	[46, 57, 58]
	IRAIN	Spatial conformation of chromatin, enhancer	IGF1R expression	Unknown, possibly tumor suppressor	[59, 60]
	UCA1	mRNA stability, translational regulation	$p21^{kip1}$ expression	Oncogene	[61-63]
	CRNDE	Chromatin remodeling, protein scaffold	PRC2 and COREST scaffold	Oncogene	[64-67]
	WT1-AS	Unknown	Possibly WT1 regulation	Unknown	[61, 68]
	vtRNA2-1	aka nc886, miR-886 precursor	PKR regulator	Unknown, possibly tumor suppressor	[69, 70]
	HOTAIR	Transcriptional repressor	PRC2 scaffold	Oncogene	[71, 72]
	PVT-1	Competitive endogenous RNA, protein stability	miR-200 sponge MYC regulator	Unknown, possibly oncogenic	[61, 73]
	MONC	miR-99a precursor, unknown as lncRNA	Cellular proliferation and differentiation	Oncogene	[74]
	MIR100HG	miR-100 precursor, unknown as lncRNA	Cellular proliferation and differentiation	Oncogene	[74]
ALL	ANRIL	Transcriptional regulation	PRC1 and PRC2 scaffold	Oncogene	[54-56, 75, 76]
	BALR-2	Unknown	Glucocorticoid response pathway	Oncogene	[22]
	WT1-AS	Unknown	Possibly WT1 regulation	Unknown	[68]

Myelodysplastic Syndromes

Myelodysplastic syndromes are a heterogeneous group of clonal hematopoietic disorders that are characterized by decreases in mature peripheral blood cells, or cytopenias. This is thought to result from a clonal set of somatic mutations that confer a selective advantage on early progenitor cells, but concurrently inhibit their ability to differentiate into mature blood cells. This expansion of “immature” cells contrasts with MPN and AML profiles, in which different stages of myeloid development are observed to be the origin of malignant transformation. Recent work has implicated a number of pathways that are involved in the pathogenesis of this group of diseases. Perhaps the best-characterized MDS is 5q- syndrome, which results from the loss of the long arm of chromosome 5 in humans [77]. It has been determined that several coding and non-coding genes that are located in this genomic area are important in the pathogenesis of 5q- syndrome. Indeed, microRNA-146a is downregulated as a consequence of this deletion, and experimentally induced deficiency of this microRNA results in myeloproliferative and myelodysplastic phenotypes [78-80]. However, several other forms of MDS are known to exhibit a variety of mutations based on recent high throughput sequencing studies. These include mutations in splicing factor genes, raising the idea that MDS results from global changes in RNA splicing [81, 82]. DNA methylation is also affected, causing global changes in gene expression regulation, which result in the observed phenotypes of hematopoietic progenitor cell dysplasia and peripheral cytopenias [83]. This work has now extended into analysis of long non-coding RNAs in MDS, with a few specific transcripts, such as XIST and MEG3, involved in disease pathogenesis.

Although X-inactive specific transcript (XIST) was one of the earliest described functional non-coding transcripts, it is only now beginning to be appreciated in hematopoietic homeostasis. XIST is located on the X chromosome along with multiple non-coding transcripts, which together are crucial for regulation of X chromosome inactivation by imprinting [46, 70, 84-86]. Inactivation

of the X chromosome occurs via recruitment of protein complexes for epigenetic modification of histones, ultimately silencing target loci (Figure 2) [87]. Importantly, several studies have shown that X chromosome aberrations are involved in human cancers [88-90]. Recently, Yildirim et al showed that conditional knockout of *Xist* in hematopoietic stem cells of female mice caused highly aggressive myeloid neoplasms with 100% penetrance. Features of myelodysplastic syndrome were observed in the bone marrow, including hypolobated megakaryocytes and binucleated erythroid precursors. In the myeloid lineage, they observed hypogranularity and abnormal lobation in neutrophils and myelocytes, reminiscent of additional morphologic changes in human MDS. Lastly, circulating erythroid precursors were present and showed nuclear irregularities [45]. However, this likely does not represent a pure MDS phenotype. Rather, mice presented hyperplasia of all splenic compartments. In the bone marrow, mice had abnormal mitotic figures including multipolar mitoses, myeloid hyperplasia, and increased fibroblasts. In the peripheral blood, white blood cell counts were elevated when compared to wild-type females, including increased numbers of immature myelomonocytic cells, reminiscent of manifestations seen in chronic myelomonocytic leukemia. Gene expression analyses demonstrated reactivation of several X-chromosome encoded genes; this abnormal increase in gene dosage was postulated to explain the formation of myeloid malignancies. These findings highlight a role for lncRNAs in the normal development of bone marrow progenitors and demonstrate, *in vivo*, the pathological effects of deletion of a lncRNA.

Maternally expressed 3 (MEG3) is a myeloid related lncRNA that has been implicated in multiple human malignancies [91]. MEG3 is important for MDM2 down-regulation and for enhanced p53 binding to promoter regions of genes, such as to the growth differentiation factor 15 (GDF15) promoter [48]. Additionally, it is believed that MEG3 regulates genes by recruiting Polycomb Repressive Complex 2 (PRC2) for epigenetic silencing of targets [92]. Its locus is tightly regulated by modifications in its differentially methylated regions (DMRs), and

suppression of MEG3 expression in solid tumors occurs by hypermethylation of the DMRs [93]. In a cohort of 85 patients with MDS or AML, it was shown that 35% of MDS patients had aberrant hypermethylation of the MEG3 promoter region. These modifications trended with decreased overall patient survival [47]. These data once again strongly suggest that loss of a lncRNA can contribute to MDS pathogenesis and/or disease severity.

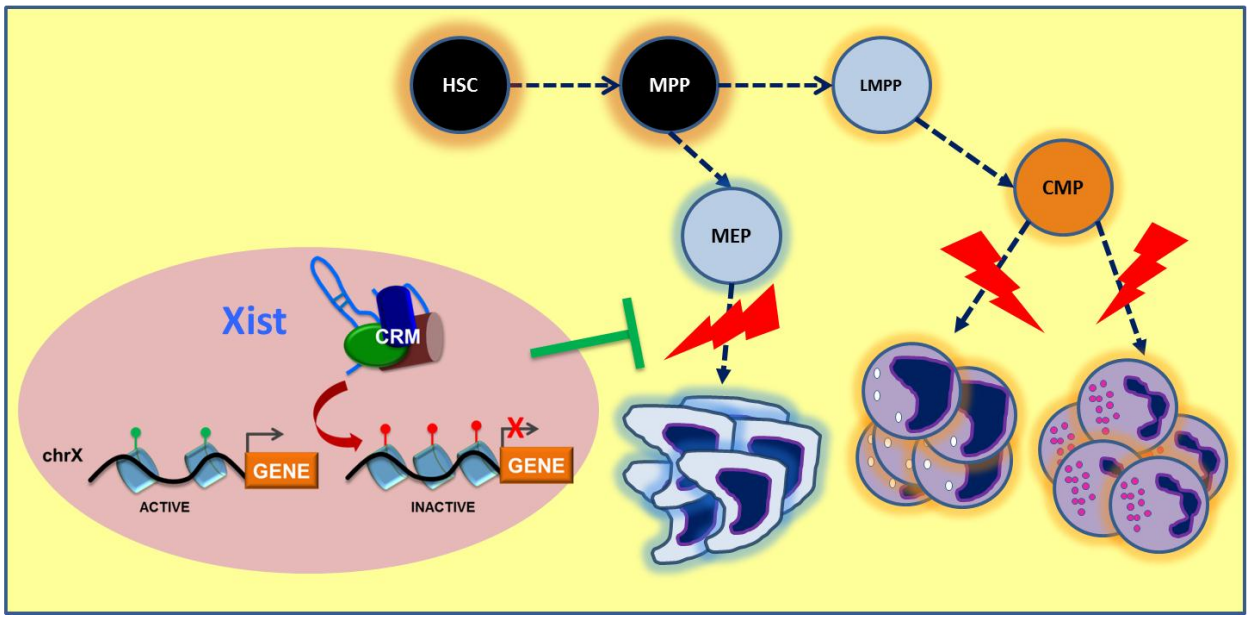


Figure 2: Xist in normal development of myeloid and erythroid progenitors. Knockout of *Xist* in the murine hematopoietic system causes alterations in the development of megakaryocytes and myeloid cells. Xist functions by directing chromatin remodeling complexes to target loci along the X chromosome. In the knockout mice, the loss of Xist led to reactivation of many genes on the X chromosome, and this dysregulated gene expression was thought to lead to cancer. HSC: hematopoietic stem cell, MPP: multipotent progenitor, MEP: megakaryocyte–erythroid progenitor cell, LMPP: lymphoid-primed multipotent progenitor, and CMP: common myeloid progenitor.

Although the number of studies is still small, these data provide compelling evidence for lncRNAs playing some role in MDS pathogenesis. Global profiling studies of lncRNA expression, and/or reanalysis of existing datasets with standardized algorithms designed to identify known and novel lncRNAs, may help us further understand the role of non-coding elements in this incompletely understood disease.

Myeloproliferative neoplasms

In contrast to MDS, myeloproliferative neoplasms are characterized by increased mature peripheral blood cells of one or more lineages, with the clonal proliferative advantage directing proliferation and differentiation of the erythroid, myeloid, or megakaryocytic lineages. The study of molecular pathogenesis and therapy in MPNs has been an area of rapid progress. Myeloproliferative neoplasms are classically divided into several subsets based on the lineage primarily affected. For example, the MPN Polycythemia Vera (PV) is a disease of red blood cell overproduction, while the MPN known as essential thrombocythemia (ET) is a disease of platelet overproduction. Many MPNs arise from mutations in tyrosine kinases, resulting in constitutive activation of pro-survival and proliferation signaling pathways, within certain primitive hematopoietic elements. The prime example in this category is chronic myelogenous leukemia (CML), a chronic form of leukemia in which the understanding of genetic abnormalities has led to striking new therapies, and improved patient outcome. CML is caused by a translocation between chromosomes 9 and 22, which results in the production of a chimeric fusion protein, BCR-ABL, a constitutively activated tyrosine kinase [94]. This fusion protein can cause a CML-like disease when overexpressed in mouse bone marrow. Remarkably, targeted therapies against BCR-ABL, such as treatment with the tyrosine kinase inhibitor Imatinib, have resulted in prolongation of disease free and overall survival in patients [95, 96]. The gene expression changes in CML are also an area of active investigation, and it is likely that lncRNAs are a part of the BCR-ABL-driven gene expression program.

Recently, Guo et al carried out a comprehensive study of lncRNA transcripts in human CML cells [49]. Specifically, they used a lncRNA cDNA microarray to identify dysregulated transcripts in K562 cell lines transduced with BCR-ABL shRNA or luciferase shRNA. The analysis revealed that 338 lncRNAs were upregulated and 108 were downregulated after knockdown of BCR-ABL. Among the most significantly dysregulated lncRNAs was Beta Globin

Locus 3 (BGL3), which may play a role in regulating γ-globin expression [97]. Expression of BGL3 was elevated in human cell lines with BCR-ABL inhibition. In the presence of imatinib, ectopic expression of the transcript reduced cell viability, while shRNA mediated knockdown increased it. In *BGL3* transgenic mice, bone marrow transformation by BCR-ABL was impaired. Additionally, the transcript acted as a competitive endogenous RNA for binding of microRNAs that regulate PTEN expression (Figure 3). Hence, it is likely that the dysregulation of this tumor suppressor lncRNA is involved in CML progression [49].

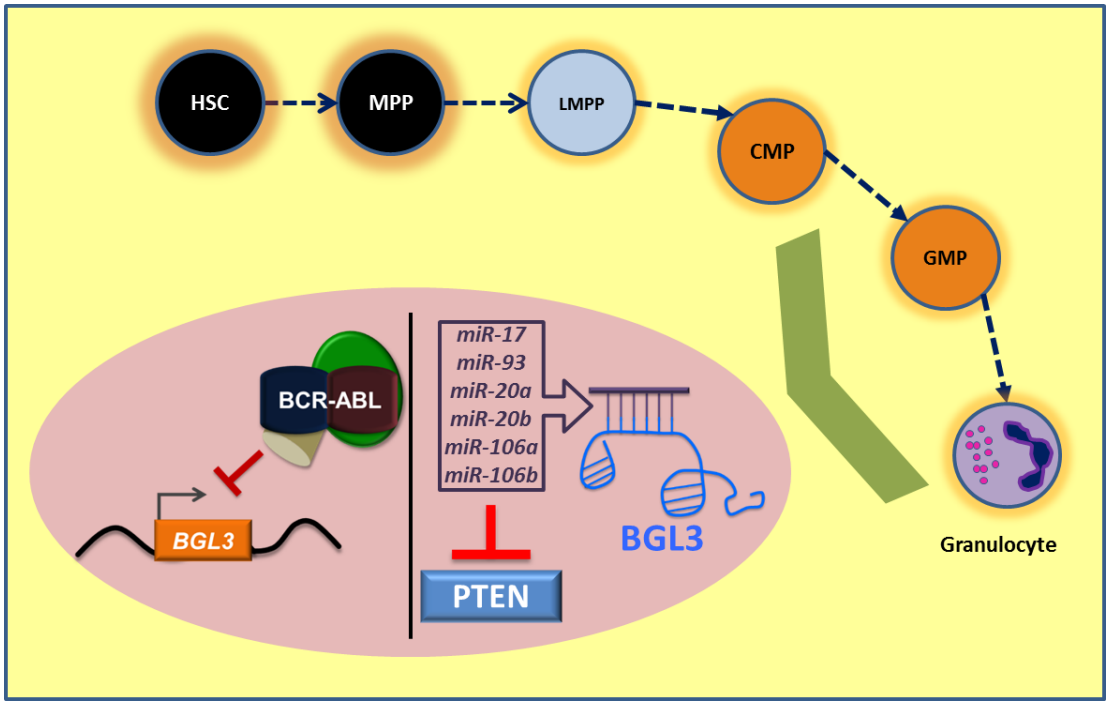


Figure 3: CML fusion protein, BCR-ABL, regulates BGL3 ceRNA expression. BGL3 is a competitive endogenous RNA that binds to several microRNAs known to target tumor suppressor PTEN. By doing so, it releases miRNA-mediated repression of PTEN. In CML, the BCR-ABL fusion protein represses BGL3 expression, resulting in increased repression of the tumor suppressor by miRNAs.

In addition to its aforementioned role in normal hematopoietic development, recent studies have shown a role for H19 in myeloid neoplasms. As noted before, H19 regulates the imprinting of its chromosomally adjacent gene, insulin-like growth factor 2 (*IGF2*), and regulates murine genes *Igfr2* and *Dlkl* in *trans* [98, 99]. This transcript has been implicated in various

cancers as an oncogene or tumor suppressor, depending on the context [52, 100]. Additionally, H19 harbors miR-675 in its first exon, which is known to regulate Igf1r [53]. Thus H19 acts as a lncRNA or a microRNA precursor in different cellular processes [36]. Bock et al examined expression levels of this lncRNA in normal bone marrow cells, as well as tissues from patients with different chronic myeloproliferative disorders, including PV, ET, CML, cellular phase primary myelofibrosis, fibrotic phase primary myelofibrosis, and chronic myelomonocytic leukemia (CMML). qRT-PCR revealed that all of these patient samples had reduced expression of H19 when compared to normal control cases [50]. Similarly, Tessema et al analyzed H19 expression in bone marrow biopsies and peripheral blood samples from normal donors, CMML patients, CML patients, and AML patients, finding reduced expression when compared to the healthy controls [51]. Conversely, studies by Guo and colleagues demonstrated that the expression of H19 is BCR-ABL kinase dependent in a BCR-ABL-positive cell line, and in primary CML cells derived from patients [101]. Knockdown of H19 in K562 cells induced apoptosis, and in mice it inhibited xenograft growth. Functional studies revealed that H19 expression is regulated by MYC in CML cell lines. These data suggest that H19 is required for tumorigenesis induced by BCR-ABL, acting in this context as an oncogene [101]. Further studies on H19 need to be carried out to fully understand its role in myeloproliferative neoplasms, and these studies highlight the importance of studying lncRNA function in a cell-type appropriate context.

Acute myeloid leukemia

In AML, a uniform population of blast cells replaces the normal heterogeneous population of maturing hematopoietic progenitor cells within the bone marrow. These cells are myeloid progenitor cells, and are thought to derive from multipotent myeloid-biased progenitor cells, or from committed myeloid progenitors such as the common myeloid progenitor (CMP).

AML is characterized by a bimodal incidence in terms of age, and is most common in older adults. Much work has gone into characterizing the molecular changes that are present in AML, and recurrent cytogenetic and molecular abnormalities are present in a large portion of AML patients. Many of these changes result in a dramatic proliferative advantage and/or an arrest in differentiation, contributing to the massive proliferation of blast cells that is seen in AML. Gene expression changes in AML are widespread and involve many different pathways. Therefore, it is likely that coding and non-coding elements play important roles in AML pathogenesis.

As a disease, AML is most devastating in older adults, who have a poor prognosis and low overall survival rates [102-104]. Although 50-55% of patients have chromosomal abnormalities, 45-50% of patient cases are cytogenetically normal (CN-AML) [103-105]. Nonetheless, CN-AML carries frequent somatic mutations in several oncogenes. Among these frequently mutated genes are nucleophosmin 1 (*NPM1*), Fms-related tyrosine-protein kinase-3 (*FLT3*), and CCAAT/enhancer-binding protein alpha (*CEBPA*) [105-107]. Garzon et al examined lncRNA expression in CN-AML patients to determine any correlations with known mutations. Indeed, lncRNA expression profiles segregated with mutations in the aforementioned genes. [61]. In *NPM1* mutated cases; they observed upregulation of antisense transcripts of HOX genes, and of the plasmacytoma variant translocation 1 (PVT1) lncRNA. It has been shown that PVT1 is required for high-level expression of MYC in human cancer cells, and hence may play a pathogenetic role in AML [108]. In addition, the Wilms tumor 1 antisense RNA (WT1-AS) was found exclusively in *FLT3-ITD* mutated cases. The WT1-AS transcript is alternatively spliced in AML, and ALL samples [68]. Lastly, among the upregulated lncRNAs in the *CEBPA* mutated cases was urothelial cancer associated 1 (UCA1), the role of which in AML has been recently studied [61, 62].

UCA1 lncRNA was first identified by a comprehensive expressed sequence tag analysis of candidate markers in patients with transitional cell carcinoma of the bladder [109].

Interestingly, this lncRNA can cause chemoresistance in bladder cancer by upregulating the Wnt signaling pathway [110]. UCA1 has also been implicated in a variety of malignancies, such as colorectaral cancer, and breast cancer [63, 111]. In AML, Hughes et al studied the effects of a dominant negative isoform of CEBPA, known as CEBPA-p30 [62, 112]. Genome wide transcriptome analysis of K562 cells with inducible *CEBPA-p30*, identified lncRNAs that are negatively regulated by the mutant protein [62]. Among the significantly dysregulated transcripts identified was UCA1. Both CEBPA and CEBPA-p30 bind to the UCA1 promoter; CEBPA repressed its expression while CEBPA-p30 induced its expression. UCA1 maintained proliferation in AML cells by repressing $p27^{kip1}$ (Figure 4). This function appeared to be dependent on interfering with the function of hnRNP1, which normally facilitates translation of $p27^{kip1}$. Cases with biallelic CEBPA mutations demonstrated an increase in UCA1 expression, further supporting the idea that UCA1 expression depends on mutant CEBPA. Taken together, these data suggest that UCA1 acts as an oncogenic lncRNA in CN-AML [62].

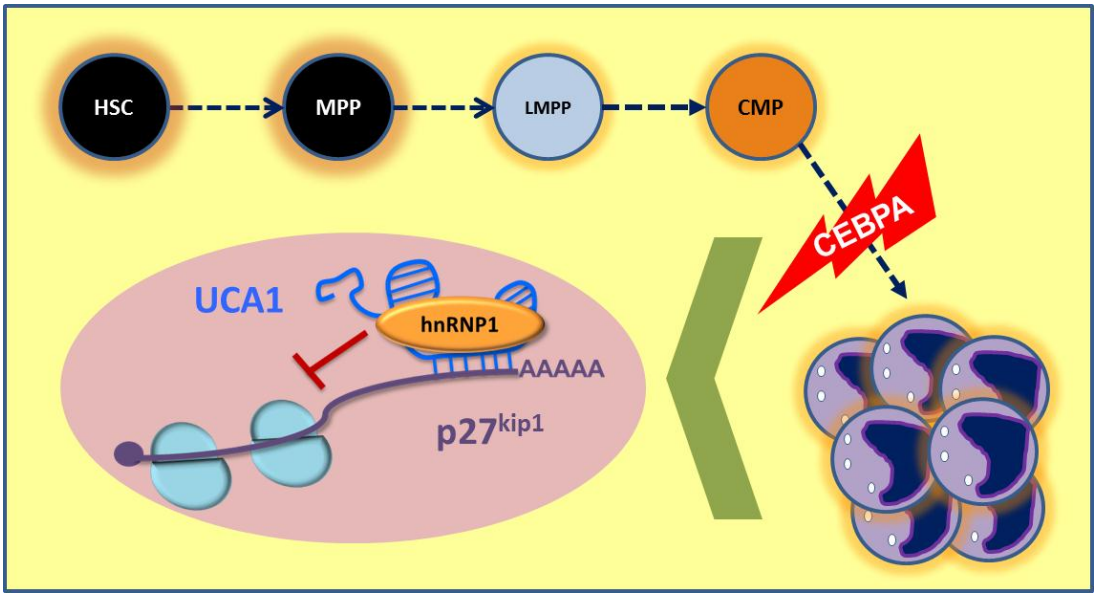


Figure 4: UCA1 represses $p27^{kip1}$ in CN-AML. hnRNP1 normally facilitates the translation of the tumor suppressor $p27^{kip1}$. UCA1 interacts with hnRNP1, which results in translational suppression of $p27^{kip1}$. Mutation of CEBPA in AML results in elevated UCA1 and decreased $p27^{kip1}$ expression.

Another recent addition to the AML-related lncRNA profile is the antisense non-coding RNA in the *INK4* locus (ANRIL), a polyadenylated cell-cycle related lncRNA that is transcribed antisense to p15INK4b [113]. This transcript has multiple isoforms that are transcribed in an antisense orientation to the *INK4* locus. One variant, p15AS, was isolated from two AML cell lines. p15AS was up-regulated in 11 of 16 AML and ALL primary samples [54]. This lncRNA was responsible for silencing of *p15INK4b* by regulating H3K9me2 and H3K4me2 levels at the promoter regions, prompting heterochromatin formation. It is likely that p15AS recruits PRC2, since EZH2 and SUZ12 (components of the PRC2 complex) were required for stable silencing of the locus. Kotake et al recently demonstrated that ANRIL does in fact recruit PRC2 by binding to SUZ12 to silence tumor suppressor genes in the *p15INK4b* locus, indicating its role as an oncogene (Figure 5) [55].

In addition to being dysregulated in MDS, aberrant hypermethylation of the DMRs in the MEG3 promoter was seen in 48% of AML patients. These findings correlated significantly with decreased survival in AML [93]. The finding of hypermethylation of the MEG3 promoter was confirmed in a second study examining 40 patients with AML [114]. Further study of MEG3 in AML is imperative for greater understanding of its role in pathogenesis.

One of the best studied lncRNAs is HOX transcript antisense RNA (HOTAIR). This lncRNA is expressed from the HOXC locus, and is a trans-acting repressor of genes in the HOXD locus [26]. Specifically, it recruits PRC2 to the target loci, and is required for H3K27me3 of chromatin associated with the HOXD locus. This lncRNA is crucial for homeotic and skeletal development [115]. In addition, this transcript has been implicated in different cancers and contributes to breast cancer progression by “reprogramming” the chromatin state [116-119]. Recently, Xing et al showed increased expression of HOTAIR in leukemic cells lines and primary AML blasts [71]. Patients with high HOTAIR expression showed the worst clinical outcome. shRNA-mediated knockdown of HOTAIR inhibited cell growth, caused apoptosis, and

reduced the number of colony formation units. Notably, HOTAIR regulates c-KIT expression by acting as a competitive endogenous RNA. It binds to miR-193a, which targets c-KIT. These results suggest an oncogenic role for HOTAIR in AML, in addition to its known role in epithelial cancers.

While AML is not as common in children as it is in adults, inherited mutations can cause a predisposition to this malignancy. Particularly, children with Down Syndrome (DS) have an increased risk of developing acute megakaryoblastic leukemia (AMKL), a type of AML. Interestingly, patients with DS-AMKL have a much better prognosis than cases of non-DS-AMKL [120-122]. Klusmann et al reported that an oncogenic microRNA (oncomir), miR-125b-2 located on chromosome 21, is highly expressed in patients with DS-AMKL when compared to non-DS-AMKL cases. This microRNA is located within the same locus as two lncRNAs, MONC (also known as MIR199AHG) and MIR100HG [123]. The roles of these two lncRNAs was examined in AMKL pathogenesis [74]. These transcripts are mainly nuclear, are highly expressed in AMKL blasts, and their expression was correlated with their corresponding microRNA clusters. Knockdown of MONC and MIR100HG resulted in impeded cell growth in both cell lines and primary patient samples. Using a lentiviral lncRNA vector, which purportedly conserves RNA secondary structure, MONC was overexpressed in hematopoietic stem and progenitor cells obtained from cord blood. This resulted in arrested myeloid differentiation, and enhanced the proliferation of erythroid progenitor cells. Taken together, these data show that MONC and MIR100HG play important roles in AMKL leukemic growth, independent of the miRNA clusters that they harbor [74]. Together, these studies demonstrate functional roles for several lncRNAs in AML and highlight a variety of cellular functions for these lncRNAs.

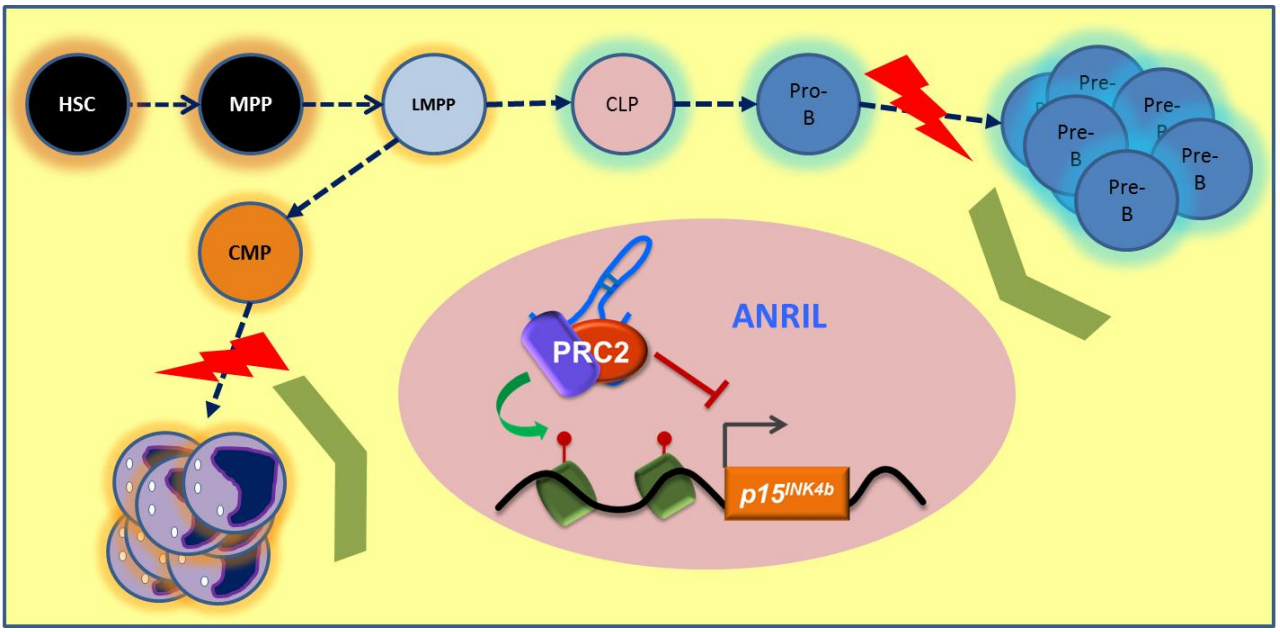


Figure 5: ANRIL represses cell cycle inhibitors in B-ALL. ANRIL recruits PRC2 to the $p15^{INK4b}$ locus, causing transcriptional repression of these cell cycle inhibitors. In B-ALL and AML leukemogenesis, ANRIL is upregulated, exacerbating the repression at the $p15^{INK4b}$ locus. CLP: common lymphoid progenitor, Pro-B: pro B cell, and Pre-B: pre B cell.

B-acute lymphoblastic leukemia (B-ALL)

Acute leukemia can be derived from lymphoid and myeloid progenitors, which constitute the traditional two branches in the hierarchy of hematopoietic development. Acute lymphoblastic leukemia can be derived from B- or T-cell progenitors, with the former being much more common. B-acute lymphoblastic leukemia (B lymphoblastic leukemia, B-ALL, ALL) is a disease with a bimodal distribution and is the most common malignancy in children. Years of research have identified recurrent cytogenetic abnormalities in B-ALL, including four common translocations, *ETV6-RUNX1*, *BCR-ABL*, *MLL* rearrangements, and *E2A-PBX1*, which account for about 30% of cases [124]. In addition, deletion of genes that are important in B-cell development, such as *PAX5* and *IKZF1*, have also been noted in many cases [125-128]. Together these findings implicate that abnormal expression of B-cell maturation genes, as well

as ectopically expressed genes, can contribute to the pathogenesis of B-ALL. Recently, these gene expression changes have shown to include non-coding RNA. Indeed certain microRNAs, such as miR-21, have now been recognized as capable of driving ALL in mouse models [129].

Very little is known about the role lncRNAs play in the pathogenesis of ALL. Recently two important profiling studies have given us insights on how these transcripts are involved in this important hematopoietic malignancy [22, 23]. Fang and colleagues carried out a genome wide lncRNA expression study on MLL-rearranged (*MLL-r*) ALL patient samples [23]. They identified 111 lncRNAs that were differentially expressed in *MLL*-r ALL samples when compared with normal bone marrow samples. Additionally, unique expression patterns between the different MLL translocations were observed. LncRNA expression correlated with certain broad categories of genes, including known MLL-fusion target proteins, such as HOXA9 and MEIS1. Functional studies with siRNAs demonstrated that particular lncRNAs could affect cellular proliferation and apoptosis.

Fernando et al expanded the scope of lncRNAs in B-ALL by looking at patient samples with several different translocations (*E2A-PBX*, *TEL-AML1*, *MLL-r*, *BCR-ABL*, and cytogenetically normal cases) [22]. Unbiased microarray profiling of human B-ALL samples revealed that lncRNA expression correlated with the specific cytogenetic abnormalities. This karyotype discrimination was confirmed by qRT-PCR of four transcripts, termed B-ALL associated long RNAs (BALRs). Amongst these lncRNAs, BALR-2 expression correlated significantly with poor overall survival and reduced patient response to prednisone treatment. Interestingly, functional studies carried out in human and murine B-ALL cell lines demonstrated that BALR-2 has an important role in regulating cell proliferation, apoptosis, and sensitivity to glucocorticoid treatment. Most importantly, a pivotal role for BALR-2 in the glucocorticoid response pathway was uncovered, with BALR-2 expression negatively regulating the expression of JUN and its pro-apoptotic target BIM. Notably, a different lncRNA, growth arrest-

specific 5 (GAS5) has been implicated in lymphoma, and acts as a decoy for the glucorticoid receptor, causing transcriptional silencing of target genes (Figure 1E) [130-135]. It is possible that BALR-2 may also downregulate glucocorticoid receptor mediated signaling; however, the mechanism has not yet been delineated. These results support the importance of BALR-2 in B-ALL leukemogenesis, prognosis, and treatment.

As mentioned previously, Yu et al showed that ANRIL was upregulated in both AML and ALL samples. In this study, ANRIL acted as a tumor promoting lncRNA, in particular, regulating the expression of p15INK4b and p16INK4a cell cycle inhibitors (Figure 5) [54, 55]. Iacobucci et al carried out a study using ALL peripheral blood cells (PBCs), compared to normal nonmalignant PBCs, which showed a clear association between ANRIL and a known BCR-ABL-related ALL nucleotide polymorphism [75]. Similar to what is seen in AML; these data suggest a role as a tumor promoting lncRNA for ANRIL by epigenetic regulation of cell cycle inhibitor genes.

Conclusions

In a remarkably short time, there has been much progress into understanding the molecular and cellular function of lncRNAs, as well as their involvement in various physiological and pathological processes. This review has focused mainly on the importance of a select few lncRNAs that have been studied in the pathogenesis of hematopoietic malignancies (Table 1). For the most part, the data is largely correlative and clinically derived, but there have been a few important mechanistic studies. It seems fair to say that, like other molecular abnormalities seen in hematolymphoid malignancies, abnormalities in lncRNA expression do not seem highly specific for a particular disease state. Rather, they may contribute to cellular proliferation and/or quiescence, and these functions may be shared across different cell types. We did not review the literature on other malignancies derived from the hematolymphoid system, including

lymphoma and chronic lymphocytic leukemia, but there are a few profiling and functional studies in these other diseases as well [136-138].

There are further studies underway that seek to delineate global lncRNA alterations during malignant hematopoietic development. In this vein, it will be extremely important to facilitate the development and implementation of bioinformatics pipelines to standardize lncRNA discovery and profiling in various disease states. As RNA species that are generally expressed at lower levels than protein-coding mRNAs, highly accurate methods of discovery are called for [18]. It will also be of great interest to define how dysregulation of certain lncRNAs seen in two or more diseases specifically contributes to the pathogenesis of a particular disease. For example, ANRIL, which is dysregulated in both AML and ALL, may function differently in these different contexts. The use of appropriate *in vitro* and *in vivo* disease models is therefore of paramount importance.

In terms of molecular mechanisms, additional work remains to be completed. The diversity of cellular functions ascribed to lncRNAs makes this a daunting task (Figure 1). However, the studies to date provide important clues as to how to proceed on this front. First, the subcellular localization of a lncRNA seems to be an important predictor of function. The majority of lncRNAs that function as modifiers of transcriptional and epigenetic mechanisms seem to be localized to the nucleus. Cytoplasmic lncRNAs may function in post-transcriptional gene expression regulation, and mechanistic studies should be guided by such knowledge of putative functions. In addition, work that relates lncRNA expression to a particular dysregulated transcription factor will help characterize upstream regulatory mechanisms and help develop an integrated picture of lncRNA function.

Future areas of clinical development include prospective trials to validate the use of lncRNAs as diagnostic and prognostic aids. Most of the studies to date have been performed on archival tissues. In this regard, it is important to consider testing platforms and the use of

appropriate control samples. For example, recent advances in RNA sequencing may render this a better platform for high-throughput testing than traditional microarrays. For small numbers of lncRNAs, it may be possible to perform qRT-PCR to characterize expression levels. In addition to the diagnostic/prognostic area, it would seem that lncRNAs are an excellent area for novel therapeutic development. Targeting via siRNAs has been shown in a number of studies to date, and seems to be a viable method of knocking down lncRNAs. Improvements in delivery technologies promise to pave the way for lncRNA-interfering therapeutics as a novel method of fighting hematopoietic malignancies in the future.

References

1. Silverstein AM. A History of Immunology. San Diego, CA,: Academic Press, 1989.
2. von Boehmer H. The developmental biology of T lymphocytes, Annu Rev Immunol 1988;6:309-326.
3. Pontvert-Delucq S, Breton-Gorius J, Schmitt C et al. Characterization and functional analysis of adult human bone marrow cell subsets in relation to B-lymphoid development, Blood 1993;82:417-429.
4. Liebermann DA, Hoffman-Liebermann B. Genetic programs of myeloid cell differentiation, Curr Opin Hematol 1994;1:24-32.
5. Akashi K, Kondo M, Cheshier S et al. Lymphoid development from stem cells and the common lymphocyte progenitors, Cold Spring Harb Symp Quant Biol 1999;64:1-12.
6. Hardy RR, Hayakawa K. B cell development pathways, Annu Rev Immunol 2001;19:595-621.
7. Shaw AC, Bassing CH, Swat W et al. Signal Transduction and Regulation of Antigen Receptor Gene Rearrangement during Early Lymphoid Development. In: Zon L. I. (ed)

Hematopoiesis: A Developmental Approach. New York, NY: Oxford University Press, 2001, 544-564.

8. Hayday AC, Barber DF, Douglas N et al. Signals involved in gamma/delta T cell versus alpha/beta T cell lineage commitment, Sem in Immu 1999;11:239-249.

9. Kee BL, Murre C. Transcription factor regulation of B lineage commitment, Curr Opin Immunol 2001;13:180-185.

10. Skalnik DG. Transcriptional mechanisms regulating myeloid-specific genes, Gene 2002;284:1-21.

11. Kapranov P, Cheng J, Dike S et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription, Science 2007;316:1484-1488.

12. Kapranov P, St Laurent G, Raz T et al. The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' un-annotated RNA, BMC Biol 2010;8:149.

13. Harrow J, Frankish A, Gonzalez JM et al. GENCODE: the reference human genome annotation for The ENCODE Project, Genome Res 2012;22:1760-1774.

14. Fernando TR, Rodriguez-Malave NI, Rao DS. MicroRNAs in B cell development and malignancy, J Hematol Oncol 2012;5:7.

15. Derrien T, Johnson R, Bussotti G et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression, Genome Res 2012;22:1775-1789.

16. Zhou Y, Zhong Y, Wang Y et al. Activation of p53 by MEG3 non-coding RNA, J Biol Chem 2007;282:24731-24742.

17. Bazzini AA, Johnstone TG, Christiano R et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation, EMBO J 2014;33:981-993.

18. Ulitsky I, Shkumatava A, Jan CH et al. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution, Cell 2011;147:1537-1550.

19. Guttman M, Amit I, Garber M et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals, Nature 2009;458:223-227.

20. Pang KC, Dinger ME, Mercer TR et al. Genome-wide identification of long noncoding RNAs in CD8+ T cells, J Immunol 2009;182:7738-7748.

21. Alvarez-Dominguez JR, Hu W, Yuan B et al. Global discovery of erythroid long noncoding RNAs reveals novel regulators of red cell maturation, Blood 2014;123:570-581.

22. Fernando TR, Rodriguez-Malave NI, Waters EV et al. LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia, Mol Cancer Res 2015;13:839-851.

23. Fang K, Han BW, Chen ZH et al. A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: biology and epigenetic target, Hum Mol Genet 2014;23:3278-3288.

24. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions, Nat Rev Genet 2009;10:155-159.

25. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs, Cell 2009;136:629-641.

26. Rinn JL, Kertesz M, Wang JK et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs, Cell 2007;129:1311-1323.

27. Carrieri C, Cimatti L, Biagioli M et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat, Nature 2012;491:454-457.

28. Huarte M, Guttman M, Feldser D et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response, Cell 2010;142:409-419.

29. Yoon JH, Abdelmohsen K, Srikantan S et al. LincRNA-p21 suppresses target mRNA translation, Mol Cell 2012;47:648-655.

30. Kretz M, Siprashvili Z, Chu C et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR, Nature 2013;493:231-235.

31. Ding DQ, Okamasa K, Yamane M et al. Meiosis-specific noncoding RNA mediates robust pairing of homologous chromosomes in meiosis, Science 2012;336:732-736.

32. Gabory A, Jammes H, Dandolo L. The H19 locus: role of an imprinted non-coding RNA in growth and development, BioEssays 2010;32:473-480.

33. Tripathi V, Ellis JD, Shen Z et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation, Mol Cell 2010;39:925-938.

34. Migeon BR, Chowdhury AK, Dunston JA et al. Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation, Am J Hum Genet 2001;69:951-960.

35. Wang Y, Xu Z, Jiang J et al. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal, Dev Cell 2013;25:69-80.

36. Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor, RNA 2007;13:313-316.

37. Lerner M, Harada M, Loven J et al. DLEU2, frequently deleted in malignancy, functions as a critical host gene of the cell cycle inhibitory microRNAs miR-15a and miR-16-1, Exp Cell Res 2009;315:2941-2952.

38. Venkatraman A, He XC, Thorvaldsen JL et al. Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence, Nature 2013;500:345-349.

39. Aoki K, Harashima A, Sano M et al. A thymus-specific noncoding RNA, Thy-ncR1, is a cytoplasmic riboregulator of MFAP4 mRNA in immature T-cell lines, BMC Mol Biol 2010;11:99.

40. Eis PS, Tam W, Sun L et al. Accumulation of miR-155 and BIC RNA in human B cell lymphomas, Proc Natl Acad Sci U S A 2005;102:3627-3632.

41. Zhang X, Lian Z, Padden C et al. A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster, Blood 2009;113:2526-2534.

42. Wagner LA, Christensen CJ, Dunn DM et al. EGO, a novel, noncoding RNA gene, regulates eosinophil granule protein transcript expression, Blood 2007;109:5191-5198.

43. Hu W, Yuan B, Flygare J et al. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation, Genes Dev 2011;25:2573-2578.

44. Elton TS, Selemon H, Elton SM et al. Regulation of the MIR155 host gene in physiological and pathological processes, Gene 2013;532:1-12.

45. Yildirim E, Kirby JE, Brown DE et al. Xist RNA is a potent suppressor of hematologic cancer in mice, Cell 2013;152:727-742.

46. Penny GD, Kay GF, Sheardown SA et al. Requirement for Xist in X chromosome inactivation, Nature 1996;379:131-137.

47. Benetatos L, Hatzimichael E, Dasoula A et al. CpG methylation analysis of the MEG3 and SNRPN imprinted genes in acute myeloid leukemia and myelodysplastic syndromes, Leuk Res 2010;34:148-153.

48. Zhang X, Rice K, Wang Y et al. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions, Endocrinology 2010;151:939-947.

49. Guo G, Kang Q, Zhu X et al. A long noncoding RNA critically regulates Bcr-Abl-mediated cellular transformation by acting as a competitive endogenous RNA, Oncogene 2015;34:1768-1779.

50. Bock O, Schlue J, Kreipe H. Reduced expression of H19 in bone marrow cells from chronic myeloproliferative disorders, Leukemia 2005;17:815-816.

51. Tessema M, Langer F, Bock O et al. Down-regulation of the IGF-2/H19 locus during normal and malignant hematopoiesis is independent of the imprinting pattern, International J of Onc 2005;26:499-507.

52. Iizuka N, Oka M, Tamesa T et al. Imbalance in expression levels of insulin-like growth factor 2 and H19 transcripts linked to progression of hepatocellular carcinoma, Anticancer Res 2004;24:4085-4089.

53. Keniry A, Oxley D, Monnier P et al. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r, Nat Cell Biol 2012;14:659-665.

54. Yu W, Gius D, Onyango P et al. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA, Nature 2008;451:202-206.

55. Kotake Y, Nakagawa T, Kitagawa K et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15INK4B tumor suppressor gene, Oncogene 2011;30:1956-1962.

56. Yap KL, Li S, Munoz-Cabello AM et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a, Mol Cell 2010;38:662-674.

57. O'Connell RM, Rao DS, Chaudhuri AA et al. Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder, J Exp Med 2008;205:585-594.

58. Tam W, Ben-Yehuda D, Hayward WS. bic, a novel gene activated by proviral insertions in avian leukosis virus-induced lymphomas, is likely to function through its noncoding RNA, Mol Cell Biol 1997;17:1490-1502.

59. Sun J, Li W, Sun Y et al. A novel antisense long noncoding RNA within the IGF1R gene locus is imprinted in hematopoietic malignancies, Nucleic Acids Res 2014;42:9588-9601.

60. Kang L, Sun J, Wen X et al. Aberrant allele-switch imprinting of a novel IGF1R intragenic antisense non-coding RNA in breast cancers, Eur J Cancer 2015;51:260-270.

61. Garzon R, Volinia S, Papaioannou D et al. Expression and prognostic impact of lncRNAs in acute myeloid leukemia, Proc Natl Acad Sci U S A 2014;111:18679-18684.

62. Hughes JM, Legnini I, Salvatori B et al. C/EBPalpha-p30 protein induces expression of the oncogenic long non-coding RNA UCA1 in acute myeloid leukemia, Oncotarget 2015;6:18534-18544.

63. Huang J, Zhou N, Watabe K et al. Long non-coding RNA UCA1 promotes breast tumor growth by suppression of p27 (Kip1), Cell Death Dis 2014;5:e1008.

64. Payton JE, Grieselhuber NR, Chang LW et al. High throughput digital quantification of mRNA abundance in primary human acute myeloid leukemia samples, J Clin Invest 2009;119:1714-1726.

65. Le Dieu R, Taussig DC, Ramsay AG et al. Peripheral blood T cells in acute myeloid leukemia (AML) patients at diagnosis have abnormal phenotype and genotype and form defective immune synapses with AML blasts, Blood 2009;114:3909-3916.

66. Khalil AM, Guttman M, Huarte M et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression, Proc Natl Acad Sci U S A 2009;106:11667-11672.

67. Ellis BC, Molloy PL, Graham LD. CRNDE: A Long Non-Coding RNA Involved in CanceR, Neurobiology, and DEvelopment, Front Genet 2012;3:270.

68. Dallosso AR, Hancock AL, Malik S et al. Alternately spliced WT1 antisense transcripts interact with WT1 sense RNA and show epigenetic and splicing defects in cancer, RNA 2007;13:2287-2299.

69. Treppendahl MB, Qiu X, Sogaard A et al. Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML, Blood 2012;119:206-216.

70. Lee JT, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-inactivation centre, Nat Genet 1999;21:400-404.

71. Xing CY, Hu XQ, Xie FY et al. Long non-coding RNA HOTAIR modulates c-KIT expression through sponging miR-193a in acute myeloid leukemia, FEBS Lett 2015;589:1981-1987.

72. Hao SF, Shao ZH. HOTAIR is upregulated in acute myeloid leukemia and that indicates a poor prognosis, Int J of Clin & Exp Path 2015;8:7223-7228.

73. Storlazzi CT, Fioretos T, Paulsson K et al. Identification of a commonly amplified 4.3 Mb region with overexpression of C8FW, but not MYC in MYC-containing double minutes in myeloid malignancies, Hum Mol Genet 2004;13:1479-1485.

74. Emmrich S, Streltsov A, Schmidt F et al. LincRNAs MONC and MIR100HG act as oncogenes in acute megakaryoblastic leukemia, Mol Cancer 2014;13:171.

75. Iacobucci I, Sazzini M, Garagnani P et al. A polymorphism in the chromosome 9p21 ANRIL locus is associated to Philadelphia positive acute lymphoblastic leukemia, Leuk Res 2011;35:1052-1059.

76. Williams RT, Sherr CJ. The INK4-ARF (CDKN2A/B) locus in hematopoiesis and BCR-ABL-induced leukemias, Cold Spring Harb Symp Quant Biol 2008;73:461-467.

77. Nilsson L, Astrand-Grundstrom I, Arvidsson I et al. Isolation and characterization of hematopoietic progenitor/stem cells in 5q-deleted myelodysplastic syndromes: evidence for involvement at the hematopoietic stem cell level, Blood 2000;96:2012-2021.

78. Starczynowski DT, Kuchenbauer F, Argiropoulos B et al. Identification of miR-145 and miR-146a as mediators of the 5q- syndrome phenotype, Nat Med 2010;16:49-58.

79. Zhao JL, Rao DS, Boldin MP et al. NF-kappaB dysregulation in microRNA-146a-deficient mice drives the development of myeloid malignancies, Proc Natl Acad Sci U S A 2011;108:9184-9189.

80. Zhao JL, Rao DS, O'Connell RM et al. MicroRNA-146a acts as a guardian of the quality and longevity of hematopoietic stem cells in mice, eLife 2013;2:e00537.

81. Graubert TA, Shen D, Ding L et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes, Nat Genet 2012;44:53-57.

82. Papaemmanuil E, Cazzola M, Boultwood J et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts, N Engl J Med 2011;365:1384-1395.

83. Zhang L, Padron E, Lancet J. The molecular basis and clinical significance of genetic mutations identified in myelodysplastic syndromes, Leuk Res 2015;39:6-17.

84. Brown CJ, Hendrich BD, Rupert JL et al. The Human Xist Gene - Analysis of a 17 Kb Inactive X-Specific Rna That Contains Conserved Repeats and Is Highly Localized within the Nucleus, Cell 1992;71:527-542.

85. Sado T, Hoki Y, Sasaki H. Tsix silences Xist through modification of chromatin structure, Dev Cell 2005;9:159-165.

86. Lee JT. Epigenetic regulation by long noncoding RNAs, Science 2012;338:1435-1439.

87. Zhao J, Sun BK, Erwin JA et al. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome, Science 2008;322:750-756.

88. McDonald HL, Gascoyne RD, Horsman D et al. Involvement of the X chromosome in non-Hodgkin lymphoma, Genes Chro Cancer 2000;28:246-257.

89. Sirchia SM, Ramoscelli L, Grati FR et al. Loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells, Cancer Res 2005;65:2139-2146.

90. Yao Y, Ma J, Xue Y et al. Knockdown of long non-coding RNA XIST exerts tumor-suppressive functions in human glioblastoma stem cells by up-regulating miR-152, Cancer Lett 2015;359:75-86.

91. Zhang X, Zhou Y, Mehta KR et al. A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells, J Clin Endocrinol Metab 2003;88:5119-5126.

92. Zhao J, Ohsumi TK, Kung JT et al. Genome-wide identification of polycomb-associated RNAs by RIP-seq, Mol Cell 2010;40:939-953.

93. Kagami M, O'Sullivan MJ, Green AJ et al. The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: hierarchical interaction and distinct functional properties as imprinting control centers, PLoS Genet 2010;6:e1000992.

94. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining, Nature 1973;243:290-293.

95. Ben-Neriah Y, Daley GQ, Mes-Masson AM et al. The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene, Science 1986;233:212-214.

96. Kantarjian H, Sawyers C, Hochhaus A et al. Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia, N Engl J Med 2002;346:645-652.

97. Kiefer CM, Lee J, Hou C et al. Distinct Ldb1/NLI complexes orchestrate gamma-globin repression and reactivation through ETO2 in human adult erythroid cells, Blood 2011;118:6200-6208.

98. Ripoche MA, Kress C, Poirier F et al. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element, Genes Dev 1997;11:1596-1604.

99. Gabory A, Ripoche MA, Le Digarcher A et al. H19 acts as a trans regulator of the imprinted gene network controlling growth in mice, Development 2009;136:3413-3421.

100. Esteves LI, Javaroni AC, Nishimoto IN et al. DNA methylation in the CTCF-binding site I and the expression pattern of the H19 gene: does positive expression predict poor prognosis in early stage head and neck carcinomas?, Mol Carcinog 2005;44:102-110.

101. Guo G, Kang Q, Chen Q et al. High expression of long non-coding RNA H19 is required for efficient tumorigenesis induced by Bcr-Abl oncogene, FEBS Lett 2014;588:1780-1786.

102. Leith CP, Kopecky KJ, Godwin J et al. Acute myeloid leukemia in the elderly: assessment of multidrug resistance (MDR1) and cytogenetics distinguishes biologic subgroups with remarkably distinct responses to standard chemotherapy. A Southwest Oncology Group study, Blood 1997;89:3323-3329.

103. Cancer, Leukemia Group B, Farag SS et al. Pretreatment cytogenetics add to other prognostic factors predicting complete remission and long-term outcome in patients 60 years of age or older with acute myeloid leukemia: results from Cancer and Leukemia Group B 8461, Blood 2006;108:63-73.

104. Frohling S, Schlenk RF, Kayser S et al. Cytogenetics and age are major determinants of outcome in intensively treated acute myeloid leukemia patients older than 60 years: results from AMLSG trial AML HD98-B, Blood 2006;108:3280-3288.

105. Schlenk RF, Dohner K, Krauter J et al. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia, N Engl J Med 2008;358:1909-1918.

106. Becker H, Marcucci G, Maharry K et al. Favorable prognostic impact of NPM1 mutations in older patients with cytogenetically normal de novo acute myeloid leukemia and associated gene- and microRNA-expression signatures: a Cancer and Leukemia Group B study, J Clin Oncol 2010;28:596-604.

107. Whitman SP, Maharry K, Radmacher MD et al. FLT3 internal tandem duplication associates with adverse outcome and gene- and microRNA-expression signatures in patients 60 years of age or older with primary cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study, Blood 2010;116:3622-3626.

108. Tseng YY, Moriarity BS, Gong W et al. PVT1 dependence in cancer with MYC copy-number increase, Nature 2014;512:82-86.

109. Wang XS, Zhang Z, Wang HC et al. Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma, Clin Cancer Res 2006;12:4851-4858.

110. Fan Y, Shen B, Tan M et al. Long non-coding RNA UCA1 increases chemoresistance of bladder cancer cells by regulating Wnt signaling, FEBS J 2014;281:1750-1758.

111. Annunziata CM, Davis RE, Demchenko Y et al. Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma, Cancer Cell 2007;12:115-130.

112. Kirstetter P, Schuster MB, Bereshchenko O et al. Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells, Cancer Cell 2008;13:299-310.

113. Pasmant E, Laurendeau I, Heron D et al. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF, Cancer Res 2007;67:3963-3969.

114. Khoury H, Suarez-Saiz F, Wu S et al. An upstream insulator regulates DLK1 imprinting in AML, Blood 2010;115:2260-2263.

115. Li L, Liu B, Wapinski OL et al. Targeted disruption of Hotair leads to homeotic transformation and gene derepression, Cell Rep 2013;5:3-12.

116. Gupta RA, Shah N, Wang KC et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis, Nature 2010;464:1071-1076.

117. Nie Y, Liu X, Qu S et al. Long non-coding RNA HOTAIR is an independent prognostic marker for nasopharyngeal carcinoma progression and survival, Cancer Sci 2013;104:458-464.

118. Heubach J, Monsior J, Deenen R et al. The long noncoding RNA HOTAIR has tissue and cell type-dependent effects on HOX gene expression and phenotype of urothelial cancer cells, Mol Cancer 2015;14:108.

119. Zhang H, Diab A, Fan H et al. PLK1 and HOTAIR Accelerate Proteasomal Degradation of SUZ12 and ZNF198 during Hepatitis B Virus-Induced Liver Carcinogenesis, Cancer Res 2015;75:2363-2374.

120. Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with Down's syndrome, Lancet 2000;355:165-169.

121. Athale UH, Razzouk BI, Raimondi SC et al. Biology and outcome of childhood acute megakaryoblastic leukemia: a single institution's experience, Blood 2001;97:3727-3732.

122. Creutzig U, Reinhardt D, Diekamp S et al. AML patients with Down syndrome have a high cure rate with AML-BFM therapy with reduced dose intensity, Leukemia 2005;19:1355-1360.

123. Klusmann JH, Li Z, Bohmer K et al. miR-125b-2 is a potential oncomiR on human chromosome 21 in megakaryoblastic leukemia, Genes Dev 2010;24:478-490.

124. Mullighan CG. Molecular genetics of B-precursor acute lymphoblastic leukemia, J Clin Invest 2012;122:3407-3415.

125. Bousquet M, Broccardo C, Quelen C et al. A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5, Blood 2007;109:3417-3423.

126. Mullighan CG, Miller CB, Radtke I et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros, Nature 2008;453:110-114.

127. Mullighan CG, Su X, Zhang J et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia, N Engl J Med 2009;360:470-480.

128. Coyaud E, Struski S, Prade N et al. Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study, Blood 2010;115:3089-3097.

129. Medina PP, Nolde M, Slack FJ. OncomiR addiction in an *in vivo* model of microRNA-21-induced pre-B-cell lymphoma, Nature 2010;467:86-90.

130. Smith CM, Steitz JA. Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5 '-terminal oligopyrimidine gene family reveals common features of snoRNA host genes, Mol Cell Bio 1998;18:6897-6909.

131. Raho G, Barone V, Rossi D et al. The gas 5 gene shows four alternative splicing patterns without coding for a protein, Gene 2000;256:13-17.

132. Tanaka R, Satoh H, Moriyama M et al. Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma, Genes to Cells 2000;5:277-287.

133. Nakamura Y, Takahashi N, Kakegawa E et al. The GAS5 (growth arrest-specific transcript 5) gene fuses to BCL6 as a result of t(1;3)(q25;q27) in a patient with B-cell lymphoma, Cancer Genet Cytogenet 2008;182:144-149.

134. Kino T, Hurt DE, Ichijo T et al. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor, Sci Signal 2010;3:ra8.

135. Hu G, Lou Z, Gupta M. The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation, PLoS ONE 2014;9:e107016.

136. Lee CS, Ungewickell A, Bhaduri A et al. Transcriptome sequencing in Sezary syndrome identifies Sezary cell and mycosis fungoides-associated lncRNAs and novel transcripts, Blood 2012;120:3288-3297.

137. Garding A, Bhattacharya N, Claus R et al. Epigenetic upregulation of lncRNAs at 13q14.3 in leukemia is linked to the In Cis downregulation of a gene cluster that targets NF-kB, PLoS Genet 2013;9:e1003373.

138. Blume CJ, Hotz-Wagenblatt A, Hullein J et al. p53-dependent non-coding RNA networks in chronic lymphocytic leukemia, Leukemia 2015;29:2015-2023.

CHAPTER II:

"LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia"

(reprint)

Molecular Cancer Research

LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia

Thilini R. Fernando[1], Norma I. Rodriguez-Malave[1,2], Ella V. Waters[1], Weihong Yan[3], David Casero[1], Giuseppe Basso[4], Martina Pigazzi[4], and Dinesh S. Rao[1,5,6]

Abstract

Long noncoding RNAs (lncRNA) have been found to play a role in gene regulation with dysregulated expression in various cancers. The precise role that lncRNA expression plays in the pathogenesis of B-acute lymphoblastic leukemia (B-ALL) is unknown. Therefore, unbiased microarray profiling was performed on human B-ALL specimens, and it was determined that lncRNA expression correlates with cytogenetic abnormalities, which was confirmed by qRT-PCR in a large set of B-ALL cases. Importantly, high expression of BALR-2 correlated with poor overall survival and diminished response to prednisone treatment. In line with a function for this lncRNA in regulating cell survival, BALR-2 knockdown led to reduced proliferation, increased apoptosis, and increased sensitivity to prednisolone treatment. Conversely, overexpression of BALR-2 led to increased cell growth and resistance to prednisone treatment. Interestingly, BALR-2 expression was repressed by prednisolone treatment and its knockdown led to upregulation of the glucocorticoid response pathway in both human and mouse B cells. Together, these findings indicate that BALR-2 plays a functional role in the pathogenesis and/or clinical responsiveness of B-ALL, and that altering the levels of particular lncRNAs may provide a future direction for therapeutic development.

Implications: lncRNA expression has the potential to segregate the common subtypes of B-ALL, predict the cytogenetic subtype, and indicate prognosis. *Mol Cancer Res; 13(5); 839–51. ©2015 AACR.*

Introduction

The advent of high-throughput techniques to study gene expression has led to the recognition that nearly 30% to 50% of the human genome is transcribed, and noncoding RNA represents a significant portion of this transcriptome (1, 2). Perhaps the clearest example of functional noncoding RNA is miRNA, which are dysregulated in cancer (reviewed in ref. 3). In oncogenesis, individual miRNAs have been found to act as either tumor suppressor genes or oncogenes, based on our work and that of others. A new addition to the repertoire of noncoding RNAs is long noncoding RNAs (lncRNA; ref. 4). These RNAs are frequently found in intergenic regions, lack open reading frames, and have been detected in the transcriptome by high-throughput technologies (4, 5). Although several classes of noncoding RNA species are being described, lncRNAs are unique in that there are epigenetic marks in their promoter regions (H3K4me3) and along transcribed regions (H3K36me3), identifying them as unique gene structures (6). At the molecular level, lncRNAs regulate gene expression via transcription, repress miRNA activity by competitive binding, regulate splicing, and activate transcription (7–12). At the cellular and organismal level, lncRNAs have been implicated in physiologic processes, such as maintenance of embryonic stem cells and erythroid development during hematopoiesis (13–18). Extending their putative roles to pathologic conditions, prior studies have examined lncRNA expression in cell lines as well as in epithelial malignancies, finding dysregulated expression (19).

Given that many hematopoietic malignancies result from mutations that cause dysregulation of gene expression, we reasoned that lncRNAs may play a role in the pathogenesis of these malignancies. Recent studies have demonstrated dysregulation of certain lncRNAs in hematopoietic malignancies, but the majority of these studies were relatively limited in their scope (reviewed in ref. 20). Among malignancies arising from hematopoietic precursors, B-acute lymphoblastic leukemia (B-ALL), a malignancy of precursor B cells, has previously been shown to harbor mutations and translocations resulting in dysregulated gene expression (21). Much progress has been made in understanding the molecular pathogenesis of B-ALL, including the role of chromosomal translocations. It is now well recognized that commonly observed recurrent chromosomal abnormalities have distinct pathogenetic and prognostic implications (21, 22). However, to date, there has been no understanding of how lncRNAs may participate in the pathogenesis of

[1]Department of Pathology and Laboratory Medicine, University of California, Los Angeles, Los Angeles, California. [2]Cellular and Molecular Pathology Ph.D. Program, University of California, Los Angeles, Los Angeles, California. [3]Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California. [4]Women and Child Health Department-Hematology-Oncology Laboratory, University of Padova, Padova, Italy. [5]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Los Angeles, California. [6]Broad Stem Cell Research Center, University of California, Los Angeles, Los Angeles, California.

Note: Supplementary data for this article are available at Molecular Cancer Research Online (http://mcr.aacrjournals.org/).

T.R. Fernando and N.I. Rodriguez-Malave contributed equally to this article.

Corresponding Author: Dinesh S. Rao, Department of Pathology and Laboratory Medicine, University of California, Los Angeles, 650 Charles E Young Drive, 12-272 Factor, Los Angeles, CA 90095-1752. Phone: 310-825-1675; Fax: 310-825-0814; E-mail: drao@mednet.ucla.edu

doi: 10.1158/1541-7786.MCR-15-0006-T

this disease. Hence, we undertook a comprehensive study examining lncRNA expression in B-ALL, examining correlations with clinicopathologic parameters, and querying the functional consequences of lncRNA expression.

Here, we found that overall lncRNA expression corresponds with specific cytogenetic abnormalities in B-ALL and that a subset of lncRNAs could correctly predict the cytogenetic subtype of B-ALL among the three most common abnormalities. We termed these B-ALL–associated long RNAs (BALR). The four most differentially regulated lncRNAs were mapped to their chromosomal locations and transcripts originating from these genomic loci were sequenced. Interestingly, high expression of one lncRNA, BALR-2, was correlated with a poor patient response to prednisone and worse overall survival. Knockdown of BALR-2 caused an increase in apoptosis of B-ALL cell lines alone and in combination with prednisolone. Interestingly, BALR-2 was repressed when human B-ALL cell lines were treated with prednisolone. Gene expression analyses of cells with knockdown of BALR-2 revealed activation of several genes involved in the response of B cells to glucocorticoid receptor engagement. Finally, the expression of BALR-2 in B-ALL cell lines with low basal expression of this lncRNA led to resistance to prednisolone treatment and concordant changes in gene expression. Together with our observation that BALR-2 is repressed by prednisolone treatment, our findings indicate a novel role for a noncoding RNA as a modulator of a treatment response. These data represent some of the first insights into long noncoding RNA expression in B-ALL and reveal that they play a role in pathogenesis, disease severity, and measurement/alteration of their levels may be useful in prognosis and/or treatment of this disease, respectively.

Materials and Methods

Patients and samples

The patient cohort consisted of 160 children consecutively admitted to the Pediatric Oncologic Department at the University of Padua (Padova, Italy) from 2000 to 2008 with the diagnosis of B-ALL. Following analysis for RNA quality, we had a total of 44 patients for the microarray studies (cytogenetic information only) and 93 patients for qRT-PCR studies (comprehensive clinicopathologic information; See Supplementary Methods and Supplementary Table S1). For the initial microarray studies, we utilized 20 patients [7 patients t(4;11) *MLL (KMT2A)* rearranged, 6 t(12;21) *TEL-AML1 (ETV6-RUNX1)* translocated and 7 t(1;19) *E2A-PBX1 (TCF3-PBX1)* translocated]. For validation microarrays, an additional 24 samples were hybridized ($n = 7$ patients for each translocation). For independent validation by qRT-PCR, we utilized 93 samples of *de novo* B-ALL without selection criteria. For each case, we had data on the following clinicopathologic parameters: cytogenetic markers, immunophenotype, age, risk stratification by minimal residual disease (23), response to prednisone, occurrence of recurrence/relapse, time to recurrence/relapse, overall survival, and time to death. Peripheral blood mononuclear cells derived from anonymized donors were obtained from the Center for AIDS Research Virology Core Lab at the University of California, Los Angeles (UCLA, Los Angeles, CA), and the Flow Cytometry and Bone Marrow Laboratory at the Department Pathology and Laboratory Medicine (UCLA, Los Angeles, CA). All procedures were approved by the local institutional review boards, and the study was considered exempt from review at UCLA (Los Angeles, CA).

Microarray data analysis

Agilent SurePrintG3 Human GE 8 × 60K microarrays (product # G4851A; http://www.genomics.agilent.com/article.jsp?crumbAction=push&pageId=1516) were hybridized at the Caltech microarray core facility. These arrays target 27,958 Entrez Gene RNAs, based on RefSeq Build 36.3, Ensemble Release 52, Unigene Build 216, GenBank (April 2009), as well as 7,419 lncRNAs, based on the initial discovery set from the Broad Institute (Cambridge, MA; ref. 4). Data analysis was implemented in the R statistics package (24). The data from two microarray experiments (20 in the initial set and 24 in the validation set) were analyzed independently but following the same protocol. The Agilent feature extraction raw data files were loaded into the R environment and analyzed using the R library of Linear Models for Microarray Data (LIMMA; ref. 25). The raw data were preprocessed for background correction and normalized between arrays using the quantile method, summarized by taking the average of replicates for each gene, and subsequently $\log_2$-transformed. Additional microarray data were generated from 4 different cultures of RS4;11 cells with or prednisolone treatment and with a siRNA against BALR-2 (or control vector). Samples were hybridized at the UCLA Microarray Core Facility using Affymetrix HG-U133_Plus_2 microarray. Raw hybridization intensities were summarized and normalized using the RMA method (26) as implemented in Matlab (The Mathworks Inc.). Differential expression analysis was performed as previously described (27). Functional annotation of differentially expressed genes was retrieved from www.ingenuity.com. All data from the microarray studies have been deposited in the Gene Expression Omnibus (GEO) resource (GEO Series ID GSE65647).

Clinicopathologic data analysis

All data analysis was completed using SPSS software. Clinicopathologic parameters (available for 93 patients) were correlated with the continuous quantitative PCR (qPCR) data obtained for each of the lncRNAs using a Pearson χ^2, and two-tailed t tests (for dichotomous variables). Next, Kaplan–Meier analyses were performed using dichotomized variables as described in Supplementary Materials and Methods.

Vectors, cell culture, and flow cytometric analysis

mmu-miR-155 formatted siRNAs were cloned into BamHI and ApaI or XhoI sites in the pHAGE2-CMV-ZsGreen-W vector (28), using the strategy that we have previously described to generate knockdown vectors against protein coding genes (29, 30). The BALR-2 sequence was cloned into a bicistronic MSCV vector with PGK driving GFP expression. RS4;11 and MV4;11 (*MLL-AF4*–translocated; ATCC CRL-1873 and CRL-9591), Reh (*TEL-AML1*–translocated; CRL-8286), 697 (*E2A-PBX1*–translocated), Nalm-6, and 70Z/3 (ATCC TIB-158) murine pre–B-cell leukemic cell line, and the HEK 293T cell line (ATCC CRL-11268) were grown as previously described (30). Lentiviruses and MSCV-based retroviral vectors were produced to generate knockdown and overexpression constructs as previously described (28, 30). A total of 5.0×10^5 cells were spin-infected at 30°C for 90 minutes in the presence polybrene (4 μg/mL). Transduced cell lines were sorted using a BD FACSAriaII cell sorter.

Apoptosis, proliferation, cell cycle, and drug response assays

To measure proliferation, cells were plated in 96-well plates with MTS reagent (Promega CellTiter 96) and cells were

incubated at 37°, 5% CO_2 for 4 hours before absorbance was measured at 490 nm. For drug response assays, cells were plated and treated with corresponding chemotherapeutic agent (250 μg/mL prednisolone dissolved in DMSO, 75 μg/mL of dexamethasone dissolved in DMSO, and 100 μg/mL of doxorubicin dissolved in water) for 24 hours, then harvested for RNA isolation or cell proliferation measurement. To measure apoptosis, chemotherapy-treated cells were stained with APC-tagged Annexin V, propidium iodide to assess the sub-G_0 fraction, or lysed for protein isolation, and analyzed by FACS. For caspase-3 assays, colorimetric reagents were added to 50 μg of protein lysates (Biovision), and absorbance was measured at 405 nm. Stably transduced cells were synchronized by serum starvation for 12 hours, replated into fresh media with FBS, collected after 24 hours, fixed with 70% EtOH, and stained with 1× propidium iodide (PI) solution in PBS. The DNA content of the cells was analyzed by flow cytometry using BD FACSDiva software and FlowJo software.

qRT-PCR and Western blot analysis

RNA collected from human samples was reverse transcribed using iScript Reagent (Quanta Biosciences). RNA from cell lines was reverse transcribed using qScript (Quanta Biosciences). Real-time quantitative PCR was performed with the StepOnePlus Real-Time PCR System (Applied Biosystems) using PerfeCTa SYBR Green FastMix reagent (Quanta Biosciences). Primer sequences used are listed in Supplementary Table S4. For Western blotting, cells were lysed in radioimmunoprecipitation assay buffer (Boston BioProducts) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific), and Western blot analysis was performed according to standard procedures. Antibodies were purchased from Cell Signaling Technology: c-Jun (60A8) Rabbit monoclonal antibody (mAb); Bim (C34C5) Rabbit mAb; c-Fos (9F6) Rabbit mAb. Monoclonal anti–β-actin (AC-15) antibody was purchased from Sigma Aldrich. Fold-change values were quantitated using ImageJ software and normalizing to actin levels.

Results

LncRNA expression segregates three common cytogenetic subtypes of B-ALL and can predict the cytogenetic subtype

LncRNAs have been ascribed functions in both cancer causation and development of the hematopoietic system. Here, we undertook a microarray study to define patterns of lncRNA expression in different subsets of B-ALL using the Agilent SurePrintG3 Human GE 8 × 60K platform (refs. 4, 7, 31; see Supplementary Fig. S1 for experimental pipeline). In our initial sample set, we hybridized 20 cases of B-ALL with three different common translocations. Following correction for multiple hypothesis testing, we performed unsupervised hierarchical clustering with significantly differentially expressed protein-coding genes ($P_{adj} \leq 0.01$), finding that the three cytogenetic subtypes separated into three subsets (Fig. 1A). Interestingly, when we focused exclusively on lncRNAs ($P_{adj} \leq 0.01$), lncRNAs were differentially expressed across the 3 different cytogenetic subtypes, although *MLL*-translocated cases appeared to have the most distinctive lncRNA expression pattern (Fig. 1B). Interestingly, lncRNA expression seemed to be highly predictive of the cytogenetic abnormality in B-ALL. Using nearest shrunken centroid analysis, we determined that lncRNA expression is at least equivalent to protein coding gene expression at differentiating these three common cytogenetic abnormalities in an independent set of 24 cases of B-ALL (Supplementary Fig. S2). For instance, one case of ALL became misclassified when 113 protein-coding genes were used to classify the samples, whereas one could use 27 lncRNAs before the same case was misclassified (Supplementary Fig. S2G and S2H; see Supplementary Figure legend for complete description of results in Supplementary Fig. S2). We then compared expression of these lncRNAs between B-ALL and normal $CD10^+CD19^+$ cells (data not shown), narrowing down the list to 10 putative oncogenic and tumor suppressor lncRNAs. These were designated as B-ALL–associated long RNAs (BALR), unless they had already been assigned LINC designations. Four lncRNAs were subsequently selected on the basis of higher expression in leukemic versus normal B-cell precursors in the microarray, and the fact that these lncRNAs were located in well-annotated loci those that showed multiple ESTs or mRNA transcripts and were conserved in human and mouse genomes (Supplementary Fig. S1). These lncRNAs were BALR-1, BALR-2, BALR-6, and LINC00958, which were also the most differentially expressed in the different cytogenetic subtypes (Fig. 1C–F). Hence, lncRNAs are differentially regulated in B-ALL with specific translocations.

LncRNA expression is correlated with clinicopathologic parameters in a large set of B-ALL cases

qRT-PCR was used to confirm the expression pattern of four individual lncRNAs in a subset of cases from the microarray cohort as well as in an independent set of B-ALL samples (Fig. 2A–D). qPCR primer sets with the most consistent behavior across technical replicates and serial dilutions were chosen (Supplementary Table S4, qPCR primers). LncRNA expression was most consistently different in the subset of cases with *MLL* translocations, which is associated with a bad prognosis. For all four lncRNAs, we confirmed that *MLL*-translocated cases showed significantly different expression levels, compared with cases showing *TEL-AML1*, *E2A-PBX1*, or *BCR-ABL1* translocations. In addition, BALR-2 and BALR-6 expression was significantly upregulated in all subsets of patient samples when compared with normal $CD19^+$ cells (Fig. 2B and C).

To examine the possibility that lncRNA expression is predictive of patient outcomes, we performed statistical analyses on available clinicopathologic parameters associated with these cases. Interestingly, BALR-2 expression was significantly higher in B-ALL patients that were unresponsive to prednisone as opposed to those who had a response to prednisone (Fig. 2E, *t* test, two-tailed $P = 0.0003$). We dichotomized BALR-2 expression using a cut point based on cluster analysis (Supplementary Fig. S3B), and found that high BALR-2 expression was associated with inferior overall survival (Fig. 2F; Kaplan–Meier survival analysis; log-rank test, $P = 0.005$). Although BALR-2 expression did correlate with translocations, it should be noted that four of the poor-prognosis cases that had high BALR-2 expression did not contain a detectable translocation. When we attempted multivariate logistic regression, BALR-2 was not established as an independent prognostic variable (Supplementary Table S3). This may be due to insufficient numbers of cases with high BALR-2 expression in our cohort and requires further study.

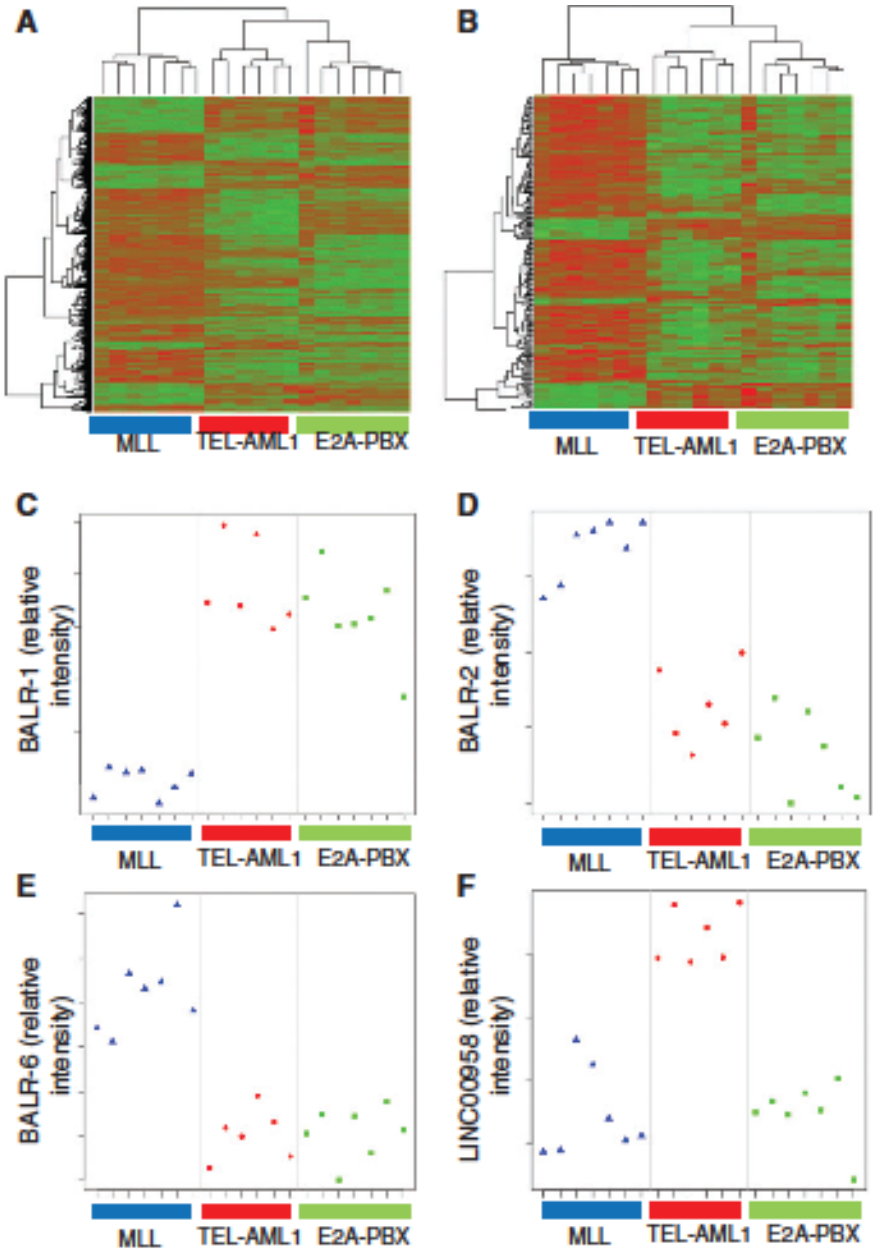


Figure 1.
LncRNA expression segregates with ALL cytogenetic subtypes. A, hierarchical clustering of differentially regulated protein-coding gene expression data (adjusted *P* value after correction for multiple hypothesis testing, $P_{adj} < 0.01$) in 20 B-ALL samples with three common translocations, namely, t(12; 21), *TEL-AML1* ($n = 6$); t(1;19) *E2A-PBX1* ($n = 7$); and t(4;11) *MLL-AF4* ($n = 7$). Genes that are relatively upregulated appear in green, and those that are relatively downregulated appear in red. B, hierarchical clustering of lncRNAs that were differentially expressed ($P_{adj} < 0.01$) showed distinct separation into three subsets of B-ALL, corresponding to the cytogenetic abnormalities. C–F, plots of normalized intensity ratios of BALR-1, BALR-2, BALR-6, and LINC00958 in individual cases of B-ALL, respectively.

BALR-2 is a nucleocytoplasmic lncRNA whose expression is regulated by glucocorticoid treatment

Given the interesting clinicopathologic findings with BALR-2, we began to study whether the expression of lncRNAs was regulated during treatment of B-ALL cells with chemotherapeutic agents. Interestingly, while the expression of these lncRNAs showed some variation with treatment, only BALR-2 was specifically regulated by glucocorticoid treatment (Fig. 3A–D). Specifically, BALR-2 was downregulated in two different B-ALL cell lines, RS4;11 and NALM6, following treatment with either prednisolone or dexamethasone, but not with doxorubicin (Fig. 3B). This finding is in line with the idea that prednisolone treatment downregulates BALR-2, but in B-ALL with constitutively high BALR-2, the cells become resistant. Next, we characterized the subcellular localization of these four lncRNAs, finding that in contrast to the other three lncRNAs, BALR-2 expression was on the same order of magnitude in nuclear and cytoplasmic fractions, and was in fact higher in the cytoplasm in one set of experiments (Fig. 3E–H). This was confirmed in multiple B-ALL cell lines, including RS4;11, REH, and 697, which show differing levels of expression of BALR-2 (Fig. 3F). As expected, GAPDH mRNA is primarily localized to the cytoplasm (Fig. 3I) and the nuclear RNA CELF4 is primarily in the nuclear fraction (Fig. 3J). Hence, BALR-2 expression appears to be regulated by glucocorticoid treatment and significant amounts of BALR-2 are found in the cytoplasm, making this lncRNA amenable to siRNA-mediated knockdown.

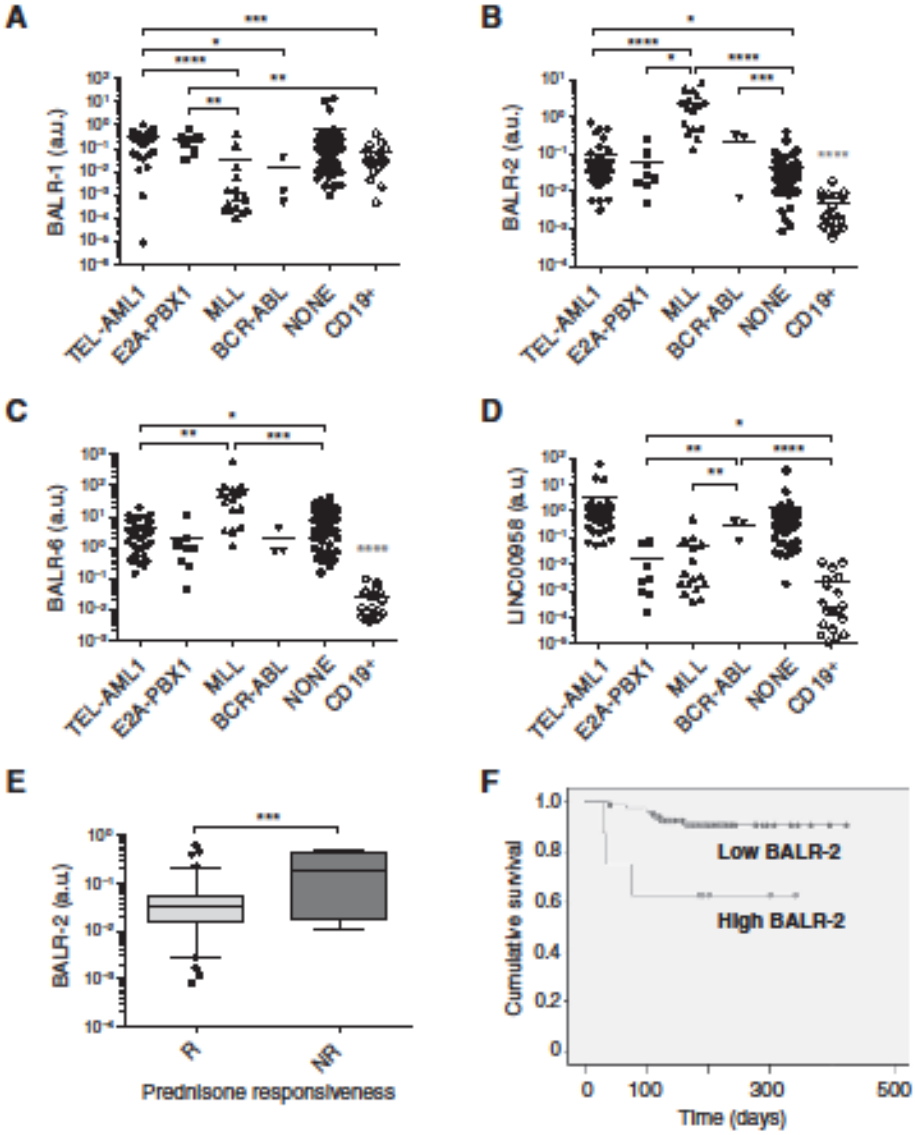


Figure 2.
lncRNAs exhibit differential expression in human B-ALL samples and BALR-2 expression correlates with prednisone response. A–D, expression levels of BALR-1, BALR-2, BALR-6, and LINC00958, respectively, normalized to actin. For these analyses, 118 samples were analyzed on the basis of the presence of good quality RNA, including both the discovery and validation cohorts. qRT-PCR was performed with specific primers for these lncRNAs, showing differential expression between the three cytogenetic subtypes of B-ALL used for the initial microarray experiments and an independent cohort of clinical samples. Number of cases used in this analysis: *TEL-AML1*-translocated, $n = 38$; *E2A-PBX1*-translocated, $n = 8$; *MLL*-translocated, $n = 16$; *BCR-ABL1*-translocated, $n = 3$; karyotypically normal cases, $n = 53$; and $CD19^+$ cells isolated from normal donors, $n = 19$. $CD19^+$ cells showed significantly differential expression of BALR-2 and BALR-6 from all the subsets of B-ALL cases depicted here. E, analysis of BALR-2 expression data and response to prednisone treatment shows that BALR-2 expression is significantly higher in B-ALL patients that are not responsive (NR) to prednisone compared with those who do respond to prednisone (R). Number of cases used in this analysis: responsive, $n = 81$; nonresponsive, $n = 8$. qRT-PCR assays were normalized to actin. Comparisons made using a two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ****, $P < 0.0001$. F, Kaplan–Meier survival analysis for two patient groups shows that high BALR-2 expression is associated with poor overall survival [overall survival (OS) high = 62.5% ($n = 8$), OS low = 91.5% ($n = 82$), log-rank test, $P = 0.005$]. The two groups were dichotomized on the basis of two-step cluster analysis using SPSS software.

Knockdown of BALR-2 results in growth inhibition and increased apoptosis

Given the correlation of BALR-2 with overall survival, we first characterized its chromosomal location, 7q21.2. The area surrounding BALR-2 is conserved in mammals and mRNA NR_110088 is localized to this region (Fig. 4A). Similar gene diagrams are provided for the other three lncRNAs in Supplementary Fig. S4A–S4C. To map the transcript, we carried out 5′RACE and 3′RACE using mRNA extracted from B-ALL cell line RS4;11. We confirmed the annotated mRNA and an alternate splice form with a previously unannotated exon (Supplementary Fig. S4D). 3′ RACE confirmed the transcript as being

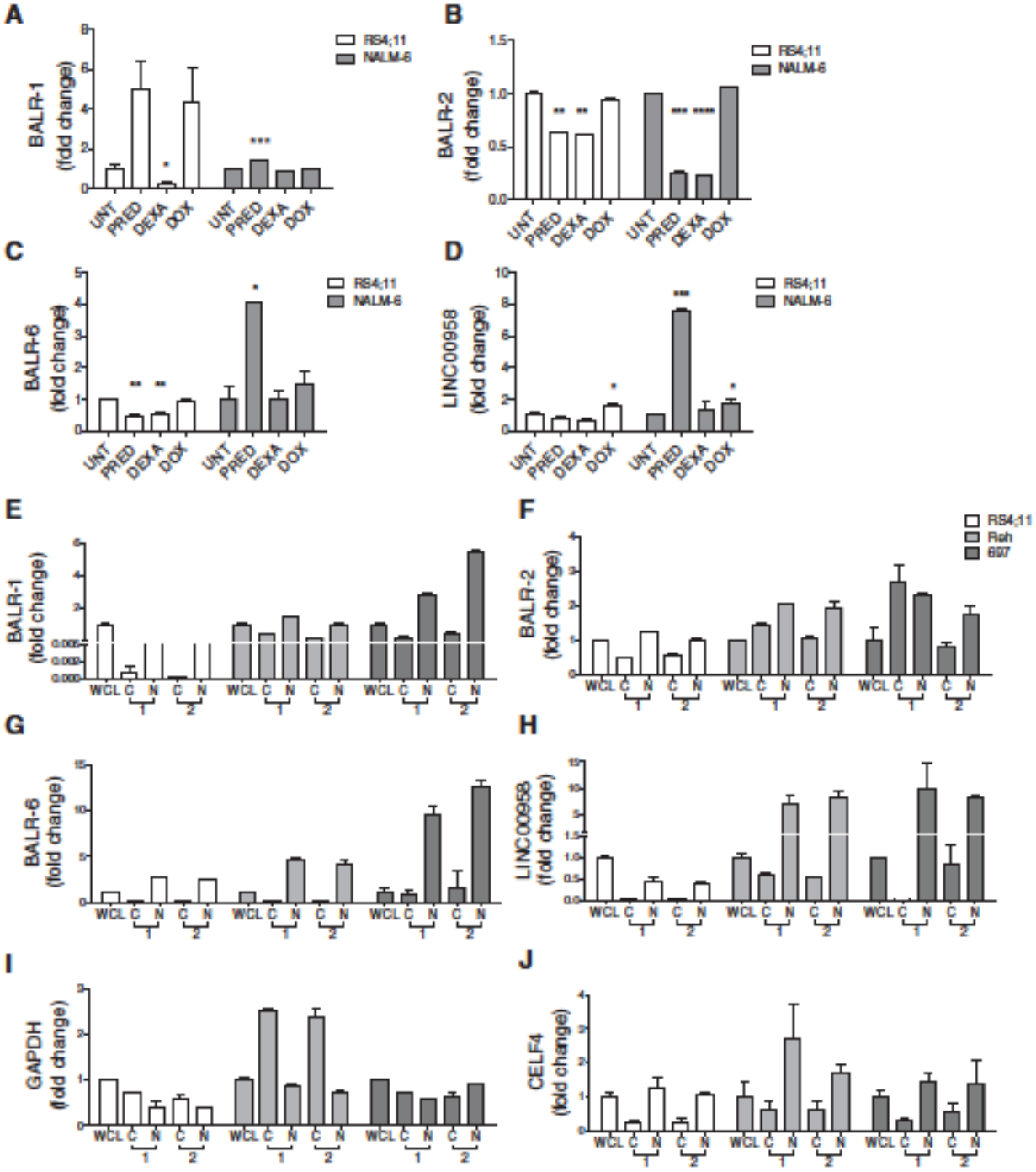


Figure 1

LncRNA expression following glucocorticoid treatment and subcellular localization of lncRNA. A-D, expression of BALR-1, BALR-2, BALR-6, and LINC00958 following treatment of RS4;11 and NALM6 cells with prednisolone (PRED), dexamethasone (DEXA), and doxorubicin (DOX). Control cells are denoted as untreated (UNT). E-H, subcellular localization of lncRNA transcripts. qRT-PCR data showing expression of BALR-2 (E), BALR-6 (F), BALR-1 (G), and LINC00958 (H) in cytoplasm and nuclear fractionations of from RS4;11 (left, white bars), Reh (middle, light gray bars) and 697 (right, dark gray bars). All 4 lncRNAs are expressed in the nucleus while BALR-2 is also expressed in the cytosolic fraction. GAPDH (cytoplasmic; I) and CELF4 (nuclear; J) mRNAs were used to demonstrate successful fractionation. The data are represented as fold-change values over the relative expression of each gene in the whole-cell lysate. Each panel represents two experiments, with overall similar results. Comparisons made using a two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ****, $P < 0.0001$. Abbreviations: WCL, whole cell lysate; C, cytoplasmic; N, nuclear.

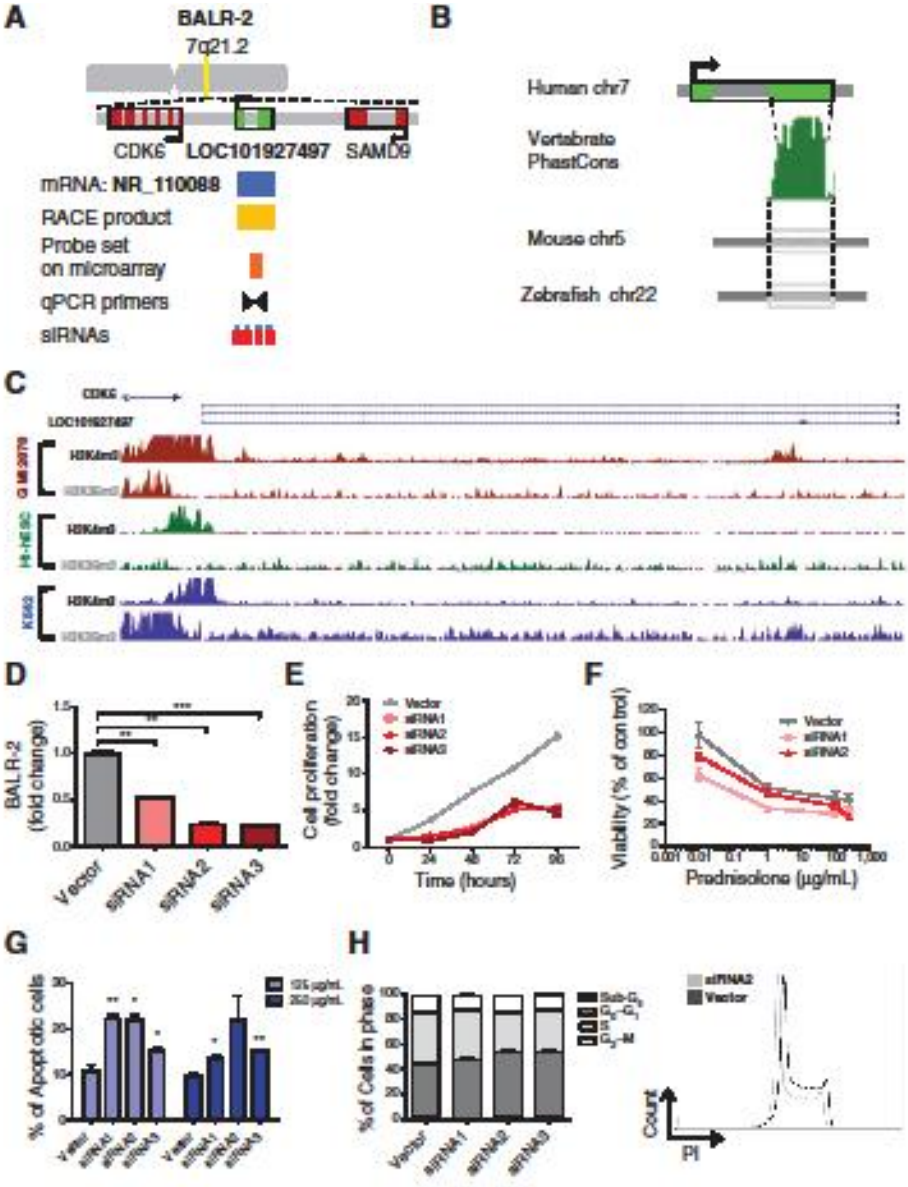
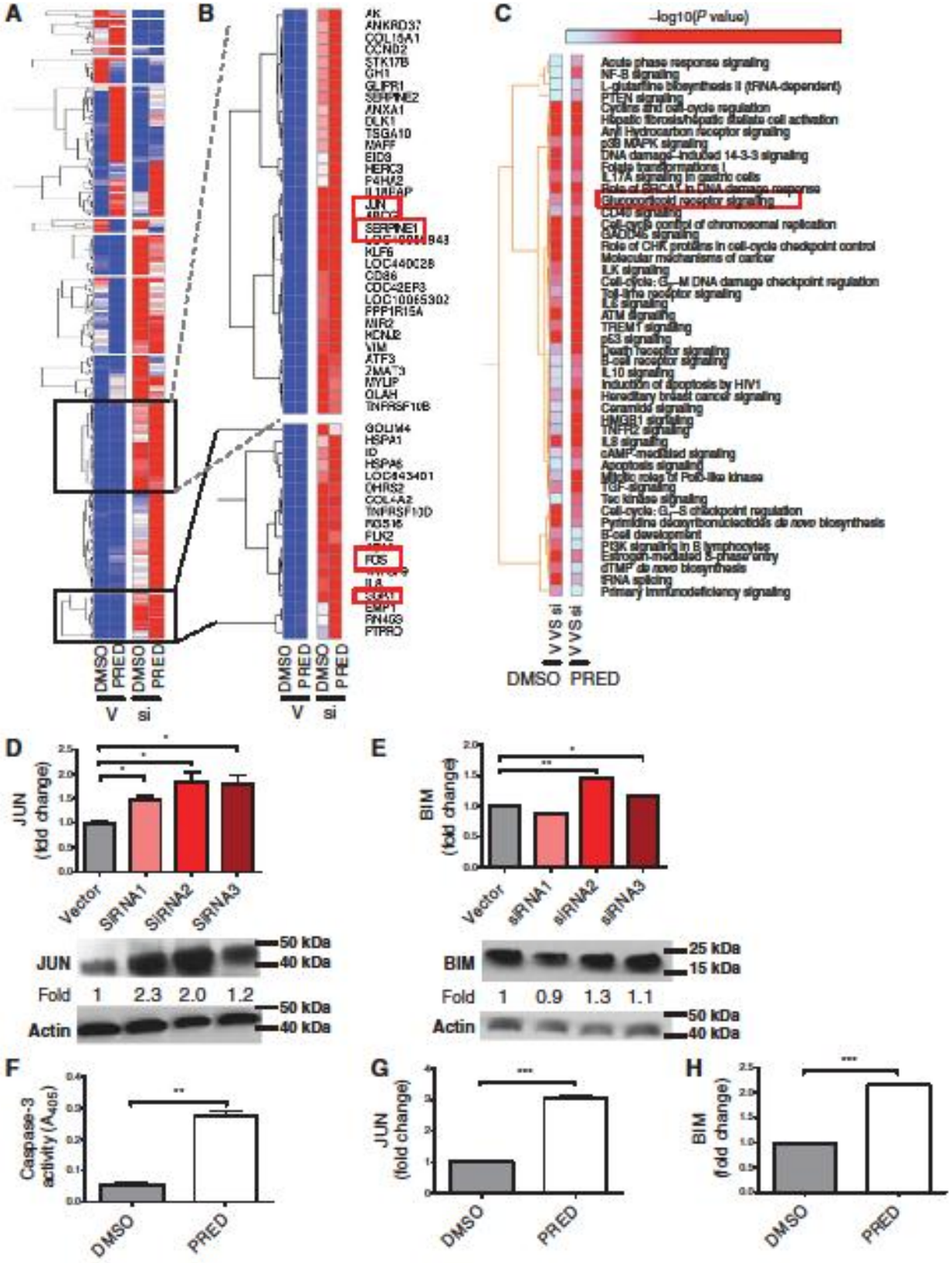


Figure 4.

BALR-2 shows a functional role in human B-ALL cell lines. A, map showing the position of BALR-2 in the genome, including the locations of neighboring genes, corresponding annotated mRNA, RACE product confirmation, probe set on microarray, qPCR primers, and siRNAs targeting the lncRNA. B, the Vertebrate PhastCons plot from the UCSC whole-genome alignments to mouse and zebrafish shows conserved regions within the terminal exon, including a region highly conserved among 91 vertebrates. C, Chip-seq histone modification map from the ENCODE/Broad Institute, taken from UCSC genome browser, shows H3K4m3 and H3K36m3 pattern at the BALR-2 locus in three different cell types indicating active promoter and transcription of the lincRNA. D, siRNA-mediated knockdown of BALR-2 in RS4;11 cell line, shown by qRT-PCR (normalized to actin). E, reduction of cell proliferation in RS;411 cells stably transduced with siRNAs against BALR-2, measured by MTS assay. Absorbance was normalized to the 0 hour timepoint ($P < 0.05$ for all non-zero hour timepoints). F, *in vitro* prednisolone response curves for RS4;11 cell lines transduced with vector control, siRNA1 and siRNA2. Prednisolone response was assessed by MTS assay after 24 hours of plating. Untreated samples were set to 100%. Absorbance readings were normalized to untreated samples. G, Increased apoptosis measured by AnnexinV staining. RS4;11 cells stably transduced with siRNAs against BALR-2 were treated with 125 μg/mL, and 250 μg/mL prednisolone for 24 hours. H, PI staining of RS4;11 cells treated with siRNAs, showed an increase in G_0–G_1 cells ($P < 0.05$ for all siRNAs), as well as a decrease in S cells ($P < 0.05$ for all siRNAs), and G_2–M cells ($P < 0.05$ for siRNAs1-3) The right panel shows a representative histogram of PI staining. Comparisons made using a two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$.

approximately 500 base pairs in length (Supplementary Table S4). This transcript lacked an open reading frame and translation initiation site, and was predicted to be noncoding (32). Vertebrate PhastCons analysis demonstrated significant conservation within the terminal exon in 91 vertebrate species, suggesting a functional transcript (Fig. 4B). Moreover, analysis


A
B
C
DMSO
PRED
V
si
AK
ANKRD37
COL15A1
CCND2
STK17B
GH1
GLIPR1
SERPINE2
ANXA1
DLK1
TSGA10
MAFF
EID3
HERC3
P4HA2
IL1RAP
JUN
SERPINE1
KLF6
LOC440028
CD86
CDC42EP3
LOC100653302
PPP1R15A
MIR2
KCNJ2
VIM
ATF3
2MA"3
MYLIP
OLAH
TNFRSF10B
GOLIM4
HSPA1
ID
HSPA6
LOC643401
DHRS2
COL4A2
TNFRSF10D
FOS
FLN2
IL8
EMP1
RN4G9
PTPRO
–log10(P value)
Acute phase response signaling
NF-B signaling
L-glutamine biosynthesis II (tRNA-dependent)
PTEN signaling
Cyclins and cell-cycle regulation
Hepatic fibrosis/hepatic stellate cell activation
Aryl Hydrocarbon receptor signaling
p38 MAPK signaling
DNA damage–induced 14-3-3 signaling
Folate transformations I
IL17A signaling in gastric cells
Role of BRCA1 in DNA damage response
Glucocorticoid receptor signaling
CD40 signaling
Cell cycle control of chromosomal replication
GADD45 signaling
Role of CHK proteins in cell-cycle checkpoint control
Molecular mechanisms of cancer
ILK signaling
Cell cycle: G2–M DNA damage checkpoint regulation
Toll-like receptor signaling
IL6 signaling
ATM signaling
TREM1 signaling
p53 signaling
Death receptor signaling
B cell receptor signaling
IL10 signaling
Induction of apoptosis by HIV1
Hereditary breast cancer signaling
Ceramide signaling
HMGB1 signaling
TNFR2 signaling
IL8 signaling
cAMP-mediated signaling
Apoptosis signaling
Mitotic roles of Polo-like kinase
TGF-signaling
Tec kinase signaling
Cell cycle: G1–S checkpoint regulation
Pyrimidine deoxyribonucleotides de novo biosynthesis
B cell development
PI3K signaling in B lymphocytes
Estrogen-mediated S-phase entry
dTMP de novo biosynthesis
tRNA splicing
Primary Immunodeficiency signaling
V
si
DMSO
PRED
D
JUN
(fold change)
Vector
SiRNA1
SiRNA2
SiRNA3
50 kDa
40 kDa
JUN
Fold 1 2.3 2.0 1.2
Actin
E
BIM
(fold change)
siRNA1
siRNA2
siRNA3
25 kDa
15 kDa
BIM
Fold 1 0.9 1.3 1.1
F
Caspase-3
activity (A405)
G
H

of H3K4m3 and H3K36m3 revealed that signals along the promoter and gene body are consistent with a transcriptional element (Fig. 4C) (33). To define a functional role for this transcript in B-ALL, siRNAs against BALR-2 were cloned into a modified lentiviral vector using miRNA-formatted flanking and loop sequences (28, 30). These lentiviral constructs were used to transduce the RS4;11 cell line (Fig. 4D). RS4;11 cells stably expressing siRNA against BALR-2 showed significantly decreased proliferation, measured by MTS (Fig. 4E). Decreased cell viability and increased apoptosis was observed both at baseline (data not shown) and after treatment with prednisolone (Fig. 4F and G; refs. 34, 35). In line with cell proliferation data, RS4;11 cells treated with siRNAs1–3, revealed an increase in G_0–G_1 phase, as well as a decrease in S- and G_2–M phase, as measured by PI staining. A significant increase in apoptotic cells (cells in sub-G_0 phase) was observed in cells treated with siRNAs2–3 (Fig. 4H).

BALR-2 regulates apoptosis by modulating the glucocorticoid receptor signaling pathway

To examine the mechanism by which BALR-2 affects proliferation and apoptosis in B-ALL cell lines, we examined gene expression in RS4;11 cell lines stably transduced with siRNA2 (which showed the greatest degree of knockdown against BALR-2), with and without prednisolone treatment. Hierarchical clustering analysis identified clusters (Fig. 5A and B) of genes that were upregulated in the siRNA group, both with and without prednisolone treatment. Several of these clusters consist of genes involved in the glucocorticoid receptor signaling pathway, confirmed by functional annotation results (2–4-fold increases in *FOS*, *HSPA6*, *SGK1*, *IL8*, *JUN*, *SERPINE1*, *CDKN1A*, and *ICAM1* in siRNA group, both with and without prednisone treatment; Fig. 5B and C). We confirmed knockdown of BALR-2 and upregulation of FOS, JUN, SGK1, and SERPINE1 by qRT-PCR (Supplementary Fig. S5A–S5E). When we tried to validate the findings with multiple siRNAs, expression of JUN showed the most consistent changes (Fig. 5D). We also observed upregulation of the proapoptotic protein BIM, which is a downstream target of JUN and an important mediator of glucocorticoid-induced apoptosis of lymphocytic cells (ref. 36; Fig. 5E). Hence, BALR2 knockdown mirrored the effects of prednisolone treatment, which also resulted in induction of apoptosis and increased expression of JUN and BIM (Fig. 5F–H).

To confirm our findings, we tested additional cell lines that show high-level expression of BALR-2, including Nalm-6, Reh, and MV4;11. The glucocorticoid signaling pathway is operant in Nalm-6 cells, despite this cell line not carrying an MLL translocation (Supplementary Fig. S5F–S5H). Consistent with our data in RS4;11 cells, knockdown of BALR-2 in Nalm-6 cells resulted in decreased cell proliferation, increased apoptosis, and increased expression of FOS (siRNA1 and 2), JUN (siRNA1), and BIM (siRNA1 and 2; Supplementary Fig. S6A–S6D). Knockdown of BALR-2 in Reh cells by targeting of the BALR-2 splice junction by siRNA (sisplice1), as well as with siRNA3, resulted in increased apoptosis, decreased cell proliferation, and significant upregulation of FOS, JUN, and BIM (Supplementary Fig. S6E–S6J; ref. 37). Moreover, MV-4-11 cells, which carry the t(4;11) translocation, showed an increased number of apoptotic cells and increased expression of glucocorticoid response genes, *FOS*, and *JUN*, following knockdown of BALR-2 with siRNA3 (Supplementary Fig. S6K–S6M). Overall, these findings demonstrated parallel effects of prednisolone treatment and BALR-2 knockdown in multiple cell lines, suggesting that BALR-2 is an important regulator of this pathway.

The function of BALR-2 is conserved in human and mouse cells

To assess functional conservation of BALR-2, we mapped and characterized murine Balr-2 to 5qA1, and the murine transcript demonstrates 90% homology to the human sequence (Fig. 6A). RACE was performed and two products were identified at this locus (Supplementary Fig. S4E). We generated miRNA-formatted siRNAs (30) against the murine transcript and confirmed decreased expression in the murine pre-B-cell line 70Z/3 (Fig. 6B). Similar to what we observed in human cell lines, Balr-2 knockdown led to an upregulation of Fos and Jun in all three of the cell lines with upregulation of Bim in two out of three cell lines by qRT-PCR (Fig. 6C–E). We further confirmed the upregulation of these proteins by Western blot analysis, finding increased expression Fos, Jun, and Bim with all three knockdown constructs (Fig. 6F). Prednisolone treatment in the murine cells led to downregulation of Balr-2 concomitant with increased apoptosis, and upregulation of Fos, Jun, and Bim (Fig. 6G–K), mirroring our observations in the human cells. Our data indicate that BALR-2 plays a key role in regulating the glucocorticoid receptor signaling pathway, thereby regulating the cellular response to prednisolone treatment (a putative schematic mechanism is presented in Supplementary Fig. S7).

Enforced expression of BALR-2 promotes cell growth and inhibits apoptosis

Given that normal and malignant B-lymphoid cells show markedly differential expression of BALR-2, we wanted to investigate the effects of gain-of-function of BALR-2. We designed and validated a MSCV-based two-promoter vector to constitutively express BALR-2 in E2A-PBX–translocated 697 cells, which show very low basal levels of BALR-2 (Fig. 7A). In juxtaposition to knockdown, overexpression of BALR-2 resulted in increased growth of the cells, as measured by MTS assay (Fig. 7B). Remarkably, overexpression of BALR-2 led to a partial rescue of prednisolone-induced apoptosis in 697 cells (Fig. 7C). Baseline levels of apoptosis were also somewhat lower with overexpression, but this finding was not statistically significant. Also in line with its

Figure 5.

BALR-2 plays a role in the glucocorticoid response pathway. A, hierarchical gene clustering of microarray data from RS4;11 cells treated with or without siRNA2 against BALR-2 and with or without prednisolone. Abbreviations, V, Vector; si, siRNA 2 against BALR-2; DMSO, dimethylsulfoxide (used to solubilize prednisolone); PRED, prednisolone. B, two clusters of genes significantly overexpressed in siRNA2-treated cells include genes involved in glucocorticoid response (*FOS*, *JUN*, *SGK1*, and *SERPINE1*). C, functional analysis of genes differentially expressed in siRNA2-treated cells shows significant enrichment of various canonical pathways, including glucocorticoid receptor signaling. D and E, qRT-PCR (top), showing upregulation of JUN (D), and BIM (E) in multiple siRNA knockdown lines, normalized to actin, with corresponding Western blot analysis (bottom). Fold-change values are as quantitated by ImageJ and normalized to actin. F–H, prednisolone treatment of RS4;11 cells resulted in induction of apoptosis, as measured by caspase-3 activity (F) and upregulation of JUN (G) of BIM (H), as measured by qRT-PCR, normalized to actin. Overall, the effects of the siRNA are similar to those induced by prednisolone treatment. Comparisons made using a two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$.

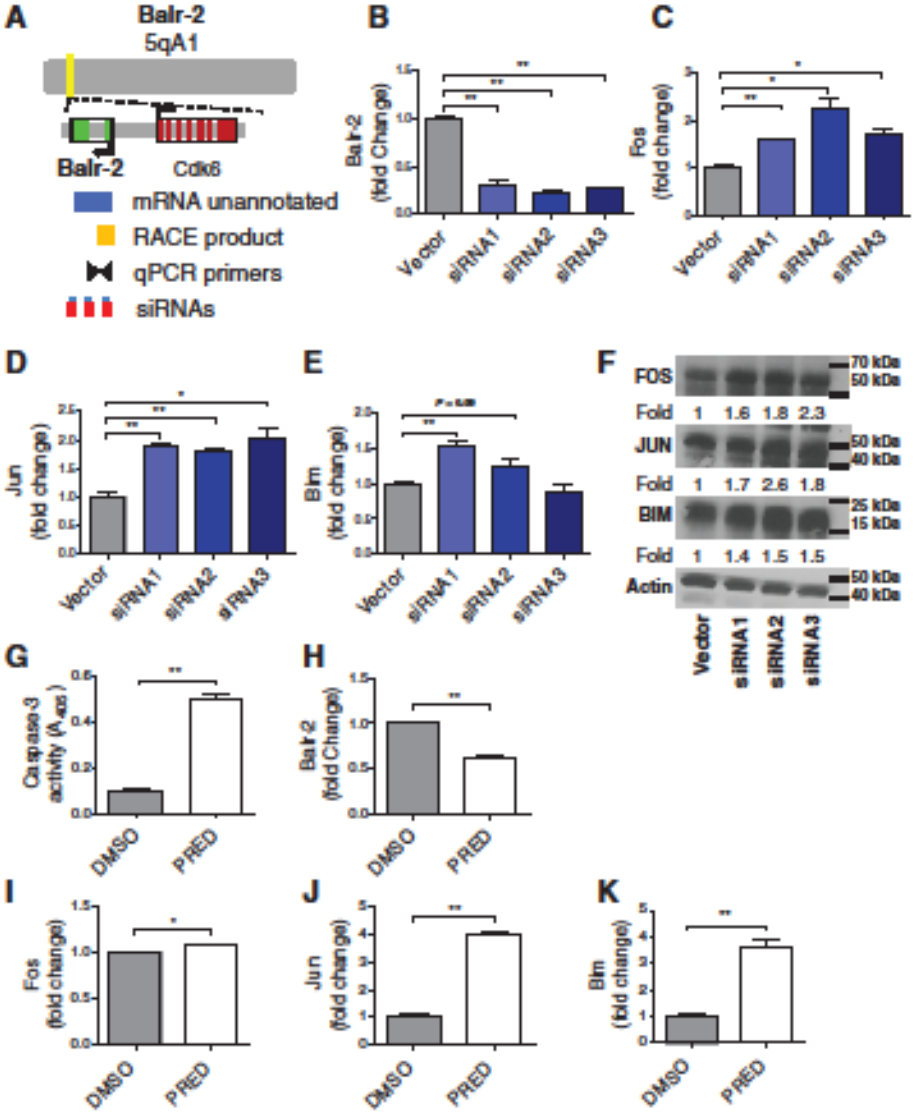


Figure 6.
The mouse homolog of BALR-2, Balr-2, shows a functional role in mice B-ALL cell lines. A, map showing the position of Balr-2 in the genome, including the locations of neighboring genes, unannotated mRNA, RACE product confirmation, qPCR primers, and siRNAs targeting the mouse lncRNA. B, siRNA-mediated knockdown of Balr-2 in the 70Z/3 mouse cell line, shown by qRT-PCR (normalized to L32). C–F, expression of glucocorticoid response genes *Fos* (C and F), *Jun* (D and F) and its target *Bim* (E and F), upon siRNA-mediated knockdown of the mouse lncRNA. Expression was analyzed by qRT-PCR (normalized to L32; C–E), and Western blot analysis (F). Fold-change values as quantitated by ImageJ and normalized to actin. G–K, caspase-3 activity, expression of mouse Balr-2 (H), and glucocorticoid response genes Fos, Jun, and Bim, respectively, are altered in 70Z/3 cells upon prednisolone treatment for 6 hours. Comparisons made using a two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$.

anticipated function, the expression of FOS, JUN, and BIM was reduced (Fig. 7D–F). These findings suggest that BALR-2 plays a similar function in primary human B-ALL, leading to prednisone resistance and a poor prognosis. Next, we used the retroviral vector system to overexpress Balr-2 in murine 70Z/3 cells (Fig. 7G). This led to similar changes in gene expression as observed in 697 cells. Moreover, the expression changes of Fos, Jun, and Bim were the opposite of those seen with knockdown of Balr-2 (Fig. 7H–J). Hence, knockdown and overexpression of BALR-2 led to opposing phenotypes and gene expression changes in human and mouse cells, consistent with the role we propose here, and once again suggesting a conserved function for this lncRNA.

Discussion

In this report, we describe how long noncoding RNAs may play a role in B-ALL pathogenesis. We find that patterns of lncRNA expression are correlated with and can predict certain common chromosomal translocations in B-ALL. We found that expression of one lncRNA, BALR-2, correlated with overall survival and with response to prednisone. To characterize this functionally, we successfully knocked down BALR-2, finding increased apoptosis with chemotherapy. Most interestingly, we demonstrate a putative mechanism for BALR-2 in regulating cell survival in B-ALL, namely, that it is downregulated by glucocorticoid receptor engagement, and that its downregulation results in activation of the glucocorticoid receptor signaling pathway. These findings were conserved in human and murine B-ALL–derived cells and were achieved with multiple siRNA sequences targeting either the exonic mRNA sequence that we have defined here, or via targeting of exon–intron boundaries (37).

Our results present some of the first clinicopathologic correlations with lncRNA expression data in B-ALL, along with a recent publication that reports dysregulated lncRNA expression in B-ALL (38). Because of the limitations of the microarray

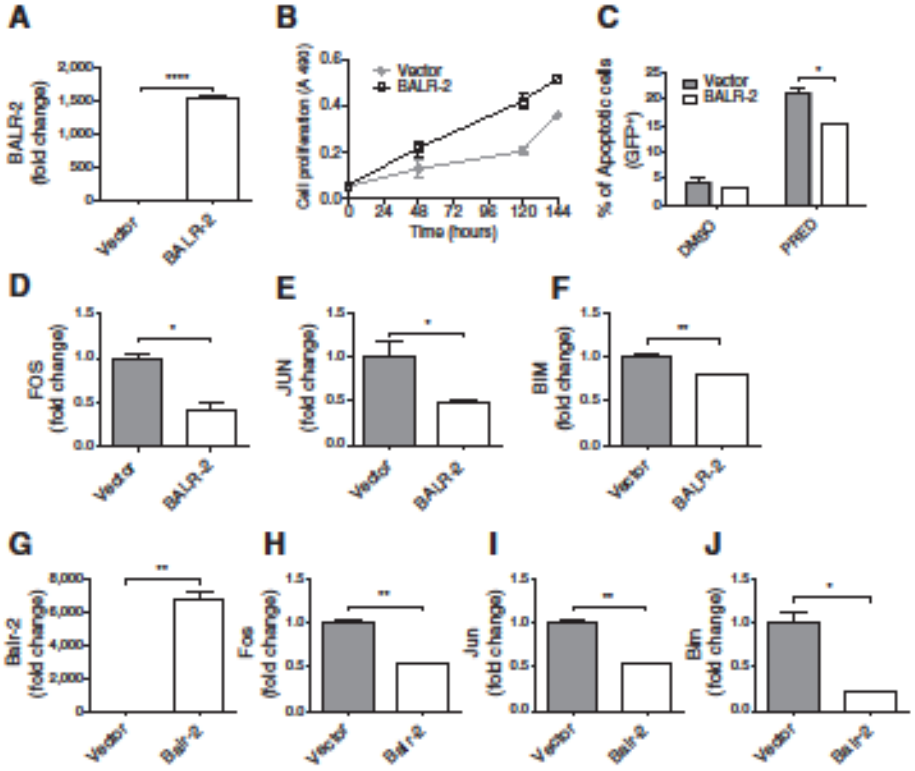


Figure 7.
Enforced expression of BALR-2 led to increased cell growth and resistance to apoptosis. A, overexpression of BALR-2 in 697 cells shown by qRT-PCR (normalized to actin). B, increased cell proliferation in 697 cells stably transduced with BALR-2, measured by MTS assay. C, annexinV staining showed that 697 cells stably transduced with BALR-2 resulted in reduction of apoptosis upon treatment with 250 μg/mL prednisolone for 24 hours. D–F, consistent with the function, the expression of FOS (D), JUN (E), and BIM (F) are downregulated with the overexpression of BALR-2, as measured by qRT-PCR, normalized to actin. G–J, overexpression of Balr-2 (G) in mouse 70Z/3 cells rescued the gene expression of glucocorticoid responsive genes *Fos* (H), *Jun* (I), and *Bim* (J), respectively, as measured by qRT-PCR, normalized to L32. Comparisons made using a two-tailed *t* test; *, $P < 0.05$; **, $P < 0.005$; ****, $P < 0.0001$.

platform, we may not have completely profiled all lncRNAs, and future efforts via high throughput RNA sequencing as well as alternate array platforms may yield additional relevant lncRNAs. Nonetheless, we found several correlations between the expression level of a few specific lncRNAs and clinicopathologic parameters. These studies validate the idea that lncRNAs are important players in biologic processes and are tied to oncogenesis. Although BALR-2 expression segregated mainly with MLL-translocated cases in the curated dataset, there were several cases of B-ALL without translocations that also showed high level expression of BALR-2. We suggest that the expression level of BALR-2 may represent another variable for prognostication in this disease. Although high-risk cytogenetic abnormalities characterize the majority of patients who do not respond well to therapy, we suggest further study of BALR-2 expression, with a view toward prospectively identifying patients who will not respond well to standard therapeutic interventions. This is particularly relevant in the context of responsiveness to prednisone. Although not addressed in this study, lncRNA expression in other types of leukemia may also represent a prognostic and therapeutically relevant variable, and warrants further study.

In addition to the clinical relevance of our findings, the functional studies present a compelling argument for the further study of lncRNA in cancer causation and therapy. Together with the clinicopathologic correlation between high BALR-2 and poor responsiveness to prednisone, our functional studies suggest that BALR-2 expression plays a role in causing resistance to apoptosis. This is supported by both loss-of-function and gain-of-function approaches in human and murine B-cells, presented herein. Although the mechanism is incompletely understood, prednisone is thought to cause apoptosis in lymphoid cells by binding to the glucocorticoid receptor with subsequent activation of the intrinsic apoptosis pathway by repression of antiapoptotic proteins such as BCL2 or activation of proapoptotic BCL2 protein family members, such as BIM (39–41). Here, we find that knockdown of BALR-2 by siRNA caused an increase in expression of the BCL2 family member BIM and increased cell death. These findings are particularly interesting in light of the proposed critical role for BIM in the regulation of the glucocorticoid response in B-ALL. Previous work has demonstrated that BIM is upregulated in leukemic blasts isolated from pediatric B-ALL patients undergoing treatment and that it is downregulated in patients who have poor responses to prednisone (42, 43). Hence, BALR-2 may be a determinant of glucocorticoid response in patients, and our study points to therapeutic inhibition of BALR-2 as a possible strategy to overcome resistance to glucocorticoid treatment in B-ALL.

While our data suggest that BALR-2 acts in promoting cell survival via inhibition of genes downstream of the glucocorticoid receptor, such as Fos, Jun, and Bim, the molecular mechanism of its action remains to be defined. Recent work has demonstrated that noncoding RNA can lead to recruitment of transcriptional regulatory complexes to chromosomal loci (7, 8, 44). In addition, the expression of certain lncRNAs affects the expression of nearby protein coding genes (37, 45). In humans and mice, BALR-2 is located in a chromosomally adjacent region to CDK6. This presents an attractive target as CDK proteins have important roles in cell-cycle progression and cancer (46). However, alternate modes of action, including interactions with proteins or other RNAs, have been ascribed to lncRNAs, and BALR-2 may act via any of these mechanisms to

regulate the glucocorticoid response pathway (8, 10, 11, 47, 48). To comprehensively address the molecular mechanism of BALR-2's function, future studies are required to study its function as a *cis*-regulatory element of cell-cycle regulation in human cells and in murine model systems, given conservation of its function in human and mouse B cells.

Future directions will include examination of the prognostic relevance of BALR-2 in B-ALL, assessment of BALR-2–mediated oncogenesis and delineating the therapeutic utility of BALR-2 knockdown in mouse models of B-ALL. Moreover, a detailed exploration of the molecular mechanisms of action of BALR-2 is warranted. Given some initial studies reporting the involvement of lncRNA in hematopoiesis, another important question is whether BALR-2 and other lncRNAs that are differentially regulated in B-ALL also play a role in normal development of B lymphocytes (17). Hence, these findings open up a new area of research into the role of lncRNAs in B-cell development and malignancy, and promise to further refine prognostication and therapeutic intervention in B-ALL.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: T.R. Fernando, N.I. Rodriguez-Malave, W. Yan, D.S. Rao

Development of methodology: T.R. Fernando, N.I. Rodriguez-Malave, W. Yan, D.S. Rao

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): N.I. Rodriguez-Malave, E.V. Waters, G. Basso, M. Pigazzi, D.S. Rao

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): T.R. Fernando, N.I. Rodriguez-Malave, E.V. Waters, W. Yan, D. Casero, D.S. Rao

Writing, review, and/or revision of the manuscript: T.R. Fernando, N.I. Rodriguez-Malave, E.V. Waters, W. Yan, G. Basso, M. Pigazzi, D.S. Rao

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): T.R. Fernando, N.I. Rodriguez-Malave, D.S. Rao

Study supervision: T.R. Fernando, D.S. Rao

Acknowledgments

The authors thank Vijaya Rao and Igor Antoshechkin at the California Institute of Technology Millard and Muriel Jacobs Genetics and Genomics Laboratory for performing the microarray hybridization experiments, Alejandro Balazs at Caltech for lentiviral vector backbones, Gay Crooks and Kathy Sakamoto for cell lines, Subba S. Rao for statistical consultation, Ken Dorshkind, Jayanth Kumar, and Jorge Contreras for helpful discussions. The authors also thank Tiffany Tran, Parth Patel, and Jasmine Gajeton for their technical support.

Grant Support

This study was supported by the grant K08CA133521 from the NIH, Sidney Kimmel Foundation (Translational Award SKF-11-013; to D.S. Rao), the Irving Reintech Family Foundation/Tower Cancer Research Foundation Research Grant, the University of California Cancer Research Coordinating Committee (to D.S. Rao) and the Stein-Oppenheimer Endowment Award, Tumor Biology Training grant NIH T32CA009056 from the NIH (to D.S. Rao), and Eugene V. Cota-Robles Fellowship from UCLA and the Graduate Research Fellowship Program from the National Science Foundation (to N.I. Rodriguez-Malave).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 6, 2015; accepted January 27, 2015; published OnlineFirst February 11, 2015.

References

1. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP, et al. Large-scale transcriptional activity in chromosomes 21 and 22. Science 2002;296:916–9.
2. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 2007;316:1484–8.
3. Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD. MicroRNAs: new regulators of immune cell development and function. Nat Immunol 2008;9:839–45.
4. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223–7.
5. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011;25:1915–27.
6. Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project. Post-transcriptional processing generates a diversity of 5[prime]-modified long and short RNAs. Nature 2009;457:1028–32.
7. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129:1311–23.
8. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 2010;142:409–19.
9. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 2011;147:358–69.
10. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell 2010;39:925–38.
11. Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature 2012;491:454–7.
12. Gong C, Maquat LE. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3[prime] UTRs via Alu elements. Nature 2011;470:284–8.
13. Brenyo A, Rao M, Koneru S, Hallinan W, Shah S, Massey HT, et al. Risk of mortality for ventricular arrhythmia in ambulatory LVAD patients. J Cardiovasc Electrophysiol 2012;23:515–20.
14. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res 2008;18:1433–45.
15. Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA 2010;16:324–37.
16. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 2011;477:295–300.
17. Hu W, Yuan B, Flygare J, Lodish HF. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. Genes Dev 2011;25:2573–8.
18. Panikar VR, Weiss MJ. A new 'Linc' between noncoding RNAs and blood development. Genes Dev 2011;25:2555–8.
19. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotech 2011;29:742–9.
20. Garitano-Trojaola A, Agirre X, Prosper F, Fortes P. Long non-coding RNAs in haematological malignancies. Int J Mol Sci 2013;14:15386–422.
21. Mullighan CG. Molecular genetics of B-precursor acute lymphoblastic leukemia. J Clin Invest 2012;122:3407–15.

22. Borowitz MJ, Chan JKC. B lymphoblastic leukaemia/lymphoma. In: Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW, editor. WHO Classification of tumours of haematopoietic and lymphoid Tissues. 4 ed. Volume 2: International Agency for Research on Cancer; 2008. p 168–75.
23. Conter V, Bartram CR, Valsecchi MG, Schrauder A, Panzer-Grumayer R, Moricke A, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood 2010;115:3206–14.
24. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2008.
25. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:Article3. Epub 2004 Feb 12.
26. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.
27. Baldi P, Long AD. A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. Bioinformatics 2001;17:509–19.
28. O'Connell RM, Balazs AB, Rao DS, Kivork C, Yang L, Baltimore D. Lentiviral vector delivery of human interleukin-7 (hIL-7) to human immune system (HIS) mice expands T lymphocyte populations. PLoS ONE 2010;5:e12009.
29. O'Connell RM, Chaudhuri AA, Rao DS, Baltimore D. Inositol phosphatase SHIP1 is a primary target of miR-155. Proc Natl Acad Sci U S A 2009; 106:7113–8.
30. Rao DS, O'Connell RM, Chaudhuri AA, Garcia-Flores Y, Geiger TL, Baltimore D. MicroRNA-34a perturbs B lymphocyte development by repressing the forkhead box transcription factor Foxp1. Immunity 2010;33:48–59.
31. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci USA 2009;106:11667–72.
32. Niazi F, Valadkhan S. Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3' UTRs. RNA 2012;18:825–43.
33. Ram O, Goren A, Amit I, Shoresh N, Yosef N, Ernst J, et al. Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. Cell 2011;147:1628–39.
34. Vangipuram SD, Buck SA, Lyman WD. Wnt pathway activity confers chemoresistance to cancer stem-like cells in a neuroblastoma cell line. Tumour Biol 2012;33:2173–83.
35. Tissing WJE, den Boer ML, Meijerink JPP, Menezes RX, Swagemakers S, van der Spek PJ, et al. Genomewide identification of prednisolone-responsive genes in acute lymphoblastic leukemia cells. Blood 2007;109:3929–35.
36. Heidari N, Miller AV, Hicks MA, Marking CB, Harada H. Glucocorticoid-mediated BIM induction and apoptosis are regulated by Runx2 and c-Jun in leukemia cells. Cell Death Dis 2012;3:e349.
37. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 2011;147:1537–50.
38. Fang K, Han BW, Chen ZH, Lin KY, Zeng CW, Li XJ, et al. A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: biology and epigenetic target. Hum Mol Genet 2014; 23:3278–88.
39. Alnemri ES, Fernandes TF, Haldar S, Croce CM, Litwack G. Involvement of BCL-2 in glucocorticoid-induced apoptosis of human pre-B-leukemias. Cancer Res 1992;52:491–5.
40. Casale F, Addeo R, D'Angelo V, Indolfi P, Poggi V, Morgera C, et al. Determination of the in vivo effects of prednisone on Bcl-2 family protein expression in childhood acute lymphoblastic leukemia. Int J Oncol 2003; 22:123–8.
41. Wang Z, Malone MH, He H, McColl KS, Distelhorst CW. Microarray analysis uncovers the induction of the proapoptotic BH3-only protein Bim in multiple models of glucocorticoid-induced apoptosis. J Biol Chem 2003;278:23861–7.
42. Schmidt S, Rainer J, Riml S, Ploner C, Jesacher S, Achmuller C, et al. Identification of glucocorticoid-response genes in children with acute lymphoblastic leukemia. Blood 2006;107:2061–9.
43. Bachmann PS, Gorman R, Papa RA, Bardell JE, Ford J, Kees UR, et al. Divergent mechanisms of glucocorticoid resistance in experimental models of pediatric acute lymphoblastic leukemia. Cancer Res 2007;67: 4482–90.
44. Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, et al. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature 2008;451:202–6.
45. Zhang B, Arun G, Mao Yuntao S, Lazar Z, Hung G, Bhattacharjee G, et al. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep 2012;2:111–23.
46. Placke T, Faber K, Nonami A, Putwain SL, Salih HR, Heidel FH, et al. Requirement for CDK6 in MLL-rearranged acute myeloid leukemia. Blood 2014;124:13–23.
47. Kaneth FA, Tay Y, Perna D, Ala U, Tan SM, Rust AG, et al. In vivo identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. Cell 2011;147:382–95.
48. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 2011;146:353–8.

SUPPLEMENTARY METHODS

Further information on patients and samples

The enrollment criteria for the study were: newly diagnosed B-ALL, age range from 0 to 18 years, and written informed consent of the parents following the AIEOP (Italian Association of Pediatric Hematology and Oncology) and the BFM (Berlin-Frankfurt-Muenster) ALL-2000 trial. The diagnosis of B-ALL was established by morphology, immunophenotyping and molecular genetics. The Philadelphia chromosome t(9;22)BCR-ABL, t(12;21)TEL-AML1, t(1;19)E2A-PBX, and t(4;11)MLL-AF4 were detected by karyotype, FISH, or RT-qPCR. Of the 160 patients, we had cytogenetic data only for 60 patients, and more comprehensive clinicopathologic information for 100 patients.

Clinicopathologic data analysis

Data analysis was completed using SPSS software. Initially, existing clinicopathologic parameters (available for approximately 93 patients) were correlated with the continuous qPCR data obtained for each of the lncRNAs using a Pearson's Chi-square, and two-tailed T-tests (for dichotomous variables). We also performed correlational and Kaplan-Meier survival analyses between the various clinicopathologic parameters and internally validated the data set. This internal validation consisted of Pearson's correlational analyses (Supp Table 2) which showed expected correlations between translocation and minimal residual disease (MRD), MRD and prednisone response, MRD and recurrence, MRD and death, prednisone response and recurrence, prednisone response and death. Moreover, Kaplan-Meier survival analyses showed that all four parameters (translocation, MRD, recurrence, and prednisone response) were significantly correlated

with overall survival (Log-Rank test, $p<0.01$ for all comparisons; data not shown). Next, for survival analyses, we dichotomized BALR-2 expression into low and high expressors, based on the finding of two clusters of data within the distribution. To confirm that there were two clusters in the distribution, a two-sample, non-parametric Kolmogorov-Smirnov test was applied to these two clusters which showed that the clusters were statistically significant (two-tailed p-value approaching 0). This allowed us to compare overall survival in these two subsets of patients. Expression levels for BALR-1, BALR-6, and LINC00958 showed a less certain cut-point based on statistical considerations; we used a visually based cut-point to dichotomize these data.

Microarray data analysis for the class prediction

A linear model was fitted to the expression data using lmFit function and the empirical Bayes (eBayes) method was employed to rank genes, and adjusted p-values were obtained after the Benjamini and Hochberg method was applied to the results from the eBayes method. Supervised class prediction using lncRNA expression profile or coding gene expression profile was carried using the R library of prediction analysis for microarrays (PAM) [1]. The expression data from 20 microarrays were used as the training data and the data from 24 other microarrays were used for the class prediction. In brief, the training datasets were trained using the pamr.train program and the training result was evaluated by 10-fold cross-validation. Misclassification error from the cross-validation of training data was examined and the shrinkage threshold values that gave the least classification error were chosen and used to classify the data from 24 microarrays.

Rapid Amplification of cDNA Ends (RACE)

To determine the 5' and 3' transcript ends of the lncRNAs, we performed RACE using FirstChoice RLM-RACE kit (Ambion). Using the sequence information from 5' and 3' RACE products, we cloned full length transcripts into P6CZUL, P6UZCL and P6CZML vectors (variants of the third generation lentiviral vector system (28)) between the NotI and BamHI sites. Primer sequences used in RACE and cloning are listed in Supp Table 4.

Cell fractionation for RNA isolation

10 X 10^6 cells were spun down and wash with PBS. Cell pellet was resuspended in NP-40 lysis buffer (0.5% NP-40, 10 mM Tris (pH 7.4), 10 mM NaCl, 3 mM MgCl2, and 1 mM DTT) and incubated on ice for 5 min. Suspension was spun at 1100 rpm at 4 °C for 5 minutes. The supernatant (containing the cytoplasm) was transferred to a fresh tube without disturbing the nuclear pellet. Each fraction was resuspended in Trizol and RNA extractions were carried out.

$CD19^+$ cell isolation

Peripheral blood and peripheral blood mononuclear cells (PBMCs) were obtained from healthy donors. Leukocytes from peripheral blood were isolated from ficoll gradient. $CD19^+$ cells were separated using human CD19 Microbeads, LS columns and MACS separator (MACS Miltenyi Biotec). Purity of the CD19+ cells was assessed by FACS. RNA was isolated as per protocols described in the main methods section.

Data Sources

Human genome assembly GRCh37/hg19, and mouse genome assembly GRCm38/mm10 was used throughout the study. Chip-seq data from three cell lines was obtained from UCSC genome browser generated by the Broad/MGH ENCODE group. Peak viewing range set from 1 to 50 for H3K4m3 modification, and 1 to 15 for H3K36m3

modifications. Genome alignments of RefSeq transcripts from human, mouse, and other organisms, GenBank mRNAs and ESTs, as well as PhastCons scores were obtained from the UCSC genome browser.

SUPPLEMENTARY FIGURE LEGENDS

Supplemental Figure 1: Schematic depiction of the pipeline to determine pathophysiologically important lncRNAs. Abbreviations used: FC, fold change; adj. p, adjusted p-value; TS, putative tumor suppressor lncRNA; Onc, putative oncogenic lncRNA. For this study we focused on BALR-2 because of its unique properties, including a cytosolic localization and its downregulation with prednisone treatment.

Supplemental Figure 2: LncRNA expression can predict the cytogenetic subtype of B-ALL. Class prediction of the subtypes of B-ALL using the nearest shrunken centroid method. **(A-D)** Using the initial 20 cases as training data, subsets of protein coding genes **(A, C)** or lncRNAs **(B, D)** can distinguish B-ALL subtypes. The misclassification error and the number of genes for each threshold were computed using the R library of prediction analysis for microarrays (PAM). Individual **(A-B)** and cumulative **(C-D)** cross-validation error of PAM model are shown as a function of the threshold. Error bars show the standard error. **(E-F)** Scatter plot showing the number of genes as a function of the threshold. Number of protein-coding genes ranges from 1112 to 4 for the thresholds of 2.125 to 9.351 **(E)** while number of lncRNAs was only 27 for the thresholds of 3.939 to 6.401**(F)**. **(G-H)** Prediction results of the 24 independent samples of B-ALL. One of the 8 MLL samples was misclassified as TEL-AML1 when the threshold was set at 4.676 for protein coding genes or 3.969 for lncRNAs. This analysis showed that the

misclassification errors reached a minimum between the thresholds 2.125 to 9.351 (corresponding to gene numbers of 1112 to 4) for protein-coding genes and 3.939 to 6.401 (gene numbers of 27 to 4) for lncRNAs, respectively. We then proceeded to examine the classification of 24 independent samples of B-ALL using the thresholds that produced the minimum error rate. When the threshold was set at 4.676 (number of protein coding genes=113), one case of ALL became misclassified; whereas one could use 27 lncRNAs at the threshold 3.969 before the same case was misclassified.

Supplemental Figure 3: Two step cluster analysis identified two clusters of expressors. **(A-D)** Histogram showing the distribution of BALR-1 **(A)** BALR-2 **(B)**, BALR-6 **(C)** and LINC00958 **(D)** expression. Two step clustering analysis identified two clusters (high and low expression) of data within the distribution for BALR-2. For BALR-1, BALR-6 and LINC00958, low and high expressors were defined by visual inspection of the data, which may be somewhat subjective. Bars indicate the frequency of cases within each bin (n = 90) **(E-G)** Kaplan Meier survival analysis for the high and low expression groups of BALR-1 **(E**, overall survival (OS) high = 100%, OS low = 88.5%), BALR-6 **(F**, OS high = 66.7%, OS low = 89.7%) and LINC00958 (G, OS high = 100%, OS low = 88.4%).

Supplemental Figure 4: LncRNA positional information and molecular characterization of BALR-2. **(A-C)** Maps showing the positions of BALR-1 **(A)**, BALR-6 **(B)** and LINC00958 **(C)** in the human genome, including the locations of neighboring genes (exons shown in green or red), corresponding annotated mRNA, RACE product confirmation (BALR-6 and LINC00958 only), probe set on microarray and qPCR primers. **(D-E)** Diagrams showing the BALR-2 loci with annotated exons (in

green) and the RACE sequence products obtained from the human (**D**) and mouse homolog of BALR-2 (**E**). 5' and 3' RACE primers are shown in blue and yellow, respectively, with the newly discovered exon shown in red. RACE gel confirmation is shown on the bottom of the each diagram.

Supplemental Figure 5: BALR-2 has an expression pattern similar to known glucocorticoid genes. The expression of BALR-2 **(A),** and glucocorticoid response genes SGK1 **(B)**, SERPINE1 **(C),** FOS **(D),** and JUN **(E),** are shown. This data is a technical confirmation of the microarray results in Figure 5 by RT-qPCR. Comparisons made using a two-tailed T-test, $p<0.05$ (*); $p<0.005$ (**). **(F-H)** Expression of glucocorticoid response genes FOS (F), JUN (G), and BIM (H), as measured by RT-qPCR, shows a similar pattern to BALR-2, but not the other 3 lncRNAs. RS4;11 and Nalm-6 cells were treated for 24 hours with 250 μg/mL of prednisolone, 100 μg/mL of doxorubicin, 75 μg/mL of dexamethasone, or DMSO as the vehicle.

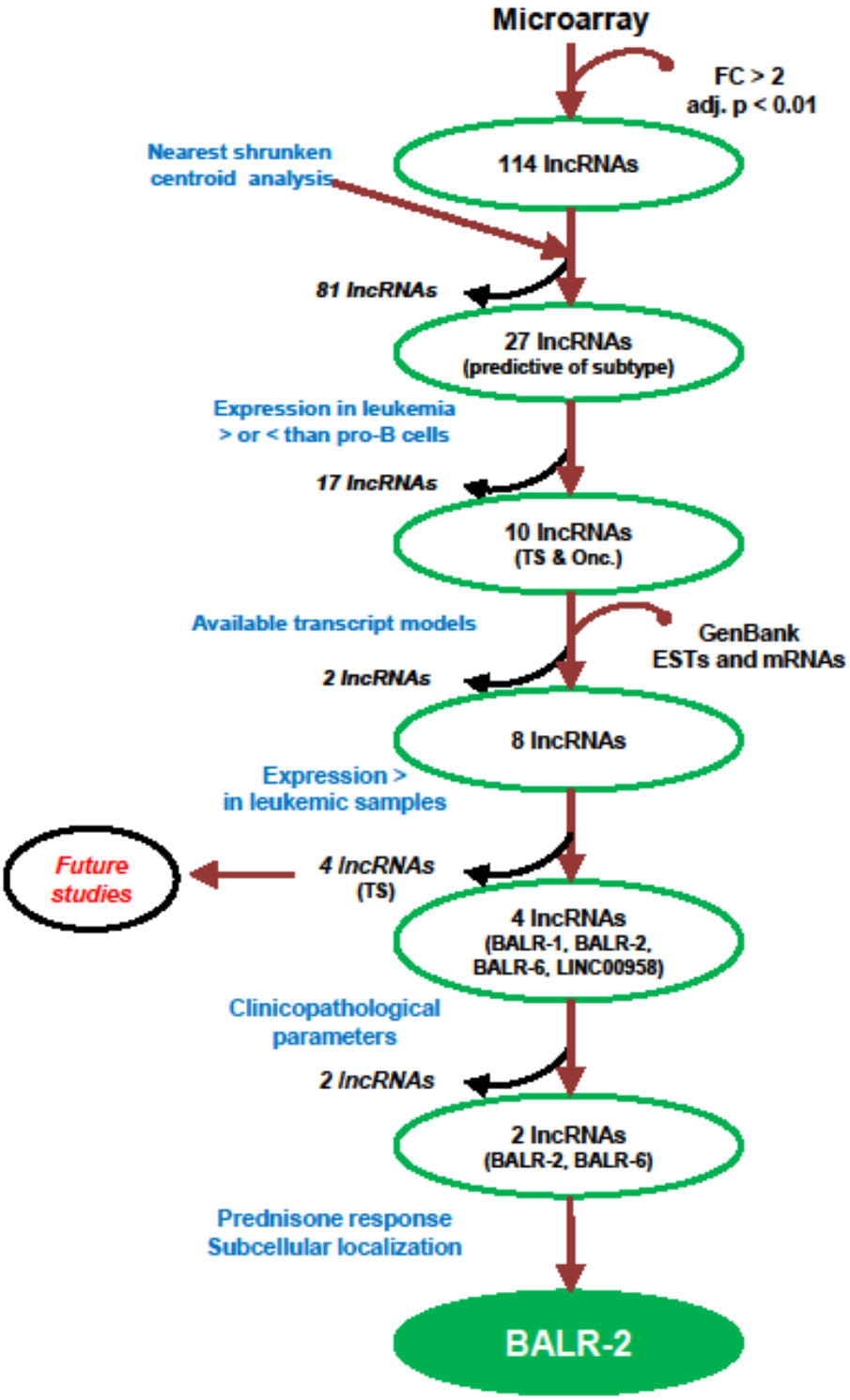
Supplemental Figure 6: Critical glucocorticoid responsive genes show altered expression after knockdown of BALR-2 in leukemic cell lines, mirroring the effects of glucocorticoid receptor engagement. (A-D) Knockdown of BALR-2 using all three siRNAs in Nalm-6 cells (A) resulted in decreased cell proliferation measured by MTS assay (B), increased apoptosis when the transduced cell lines were treated with 125 μg/mL, and 250 μg/mL prednisolone for 24 hrs, as measured by AnnexinV staining (C) and increased expression of FOS (siRNA1 and 2) , JUN (siRNA1), and BIM (siRNA1 and 2) (D). **(E-J)** Additionally, Reh cells stably transduced with siRNA against the splice junction of BALR-2 showed knockdown of the lncRNA (E), and increased apoptosis measured by caspase-3 activity (F). Expression of glucocorticoid response genes, FOS,

JUN, and JUN's target gene BIM (G), increased upon siRNA mediated knockdown of BALR-2. Reh cells stably transduced with siRNA3 showed knockdown of BALR-2 (H), and increased apoptosis measured by caspase-3 activity (I). Expression of glucocorticoid response genes, FOS, and JUN (J), increased upon siRNA mediated knockdown of BALR-2. Knockdown with siRNA3 and splice-siRNA are shown separately, as these represent independent experiments. **(K-M)** MV-4-11 cells, which contain t(4;11), stably transduced with siRNA3 showed knockdown of the lncRNA (K), and increased number of apoptotic cells and measure by Annexin V staining (L). Glucocorticoid response genes, FOS, and JUN (M) show increased expression upon knockdown of BALR-2. Comparisons made using a two-tailed T-test, $p \leq 0.05$ (*); $p \leq 0.005$ (**); $p \leq 0.0005$ (***).

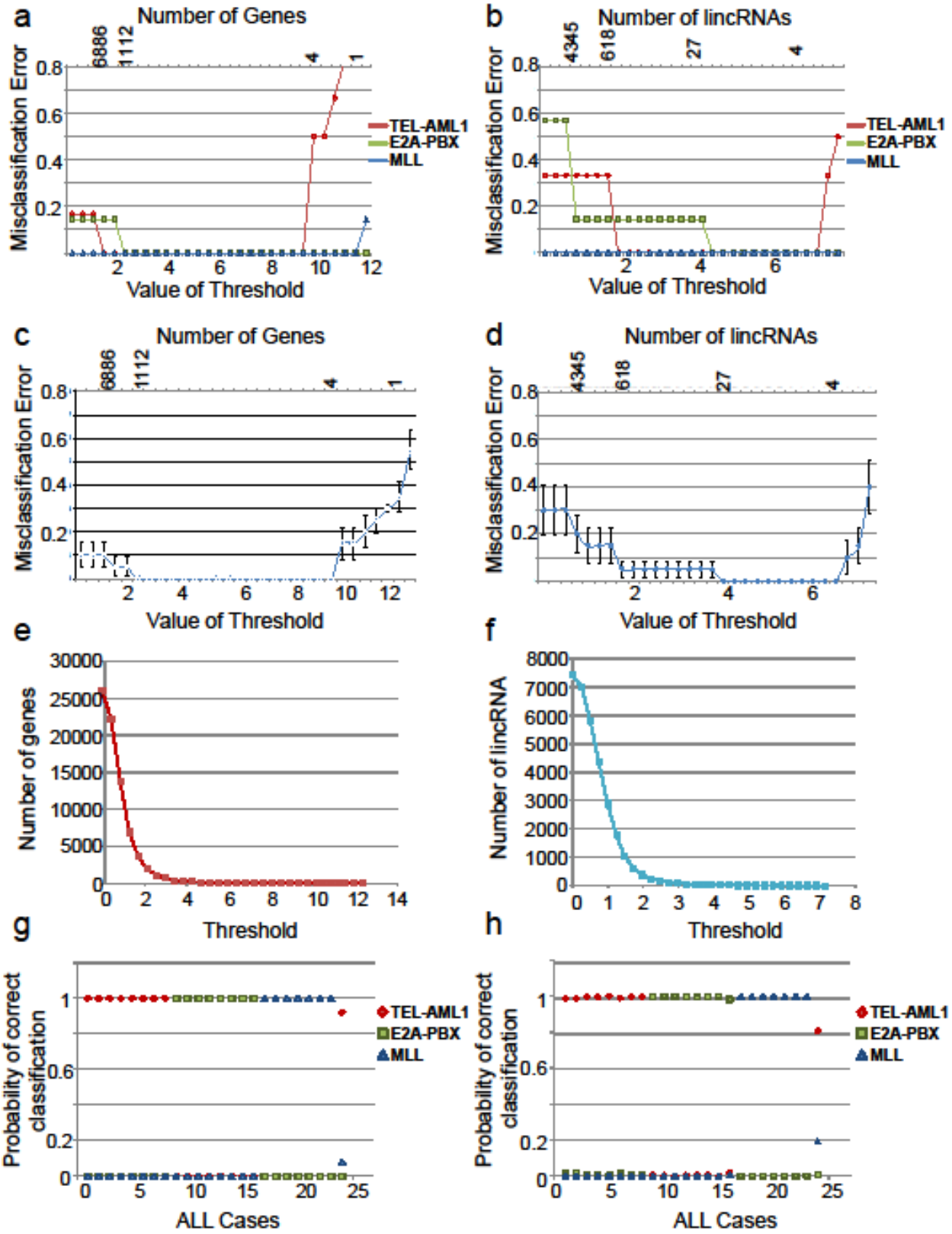
Supplemental Figure 7: Proposed mechanism of action of BALR-2 in the glucocorticoid response pathway. BALR-2 inhibits expression of FOS and JUN genes. Upon prednisone treatment BALR-2 is inhibited, releasing the block on FOS and JUN. JUN is expressed and in turn activates expression of BIM which is a well-known pro-apoptotic gene.

Reference:

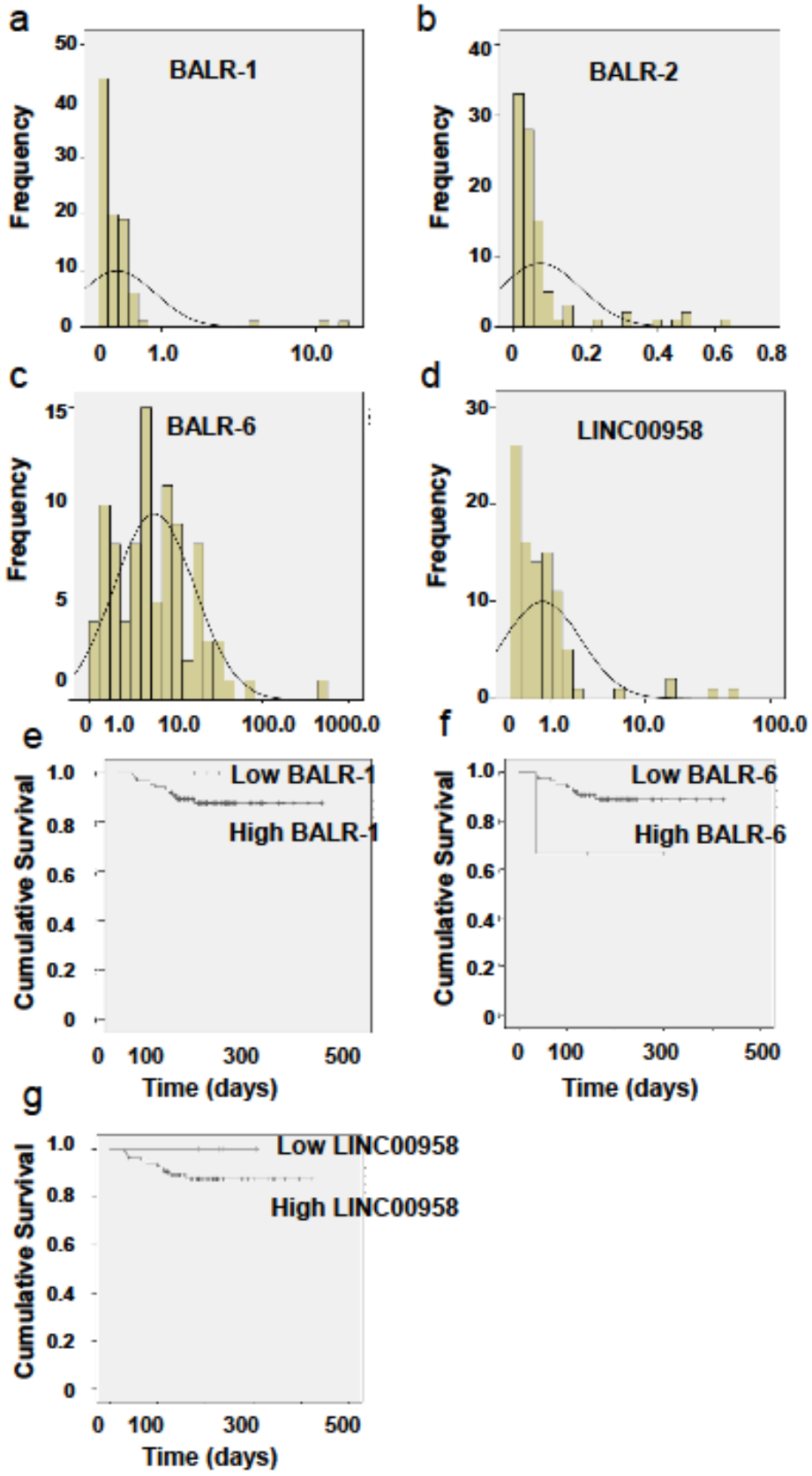
1. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A*. 2002;99(10):6567-6572.



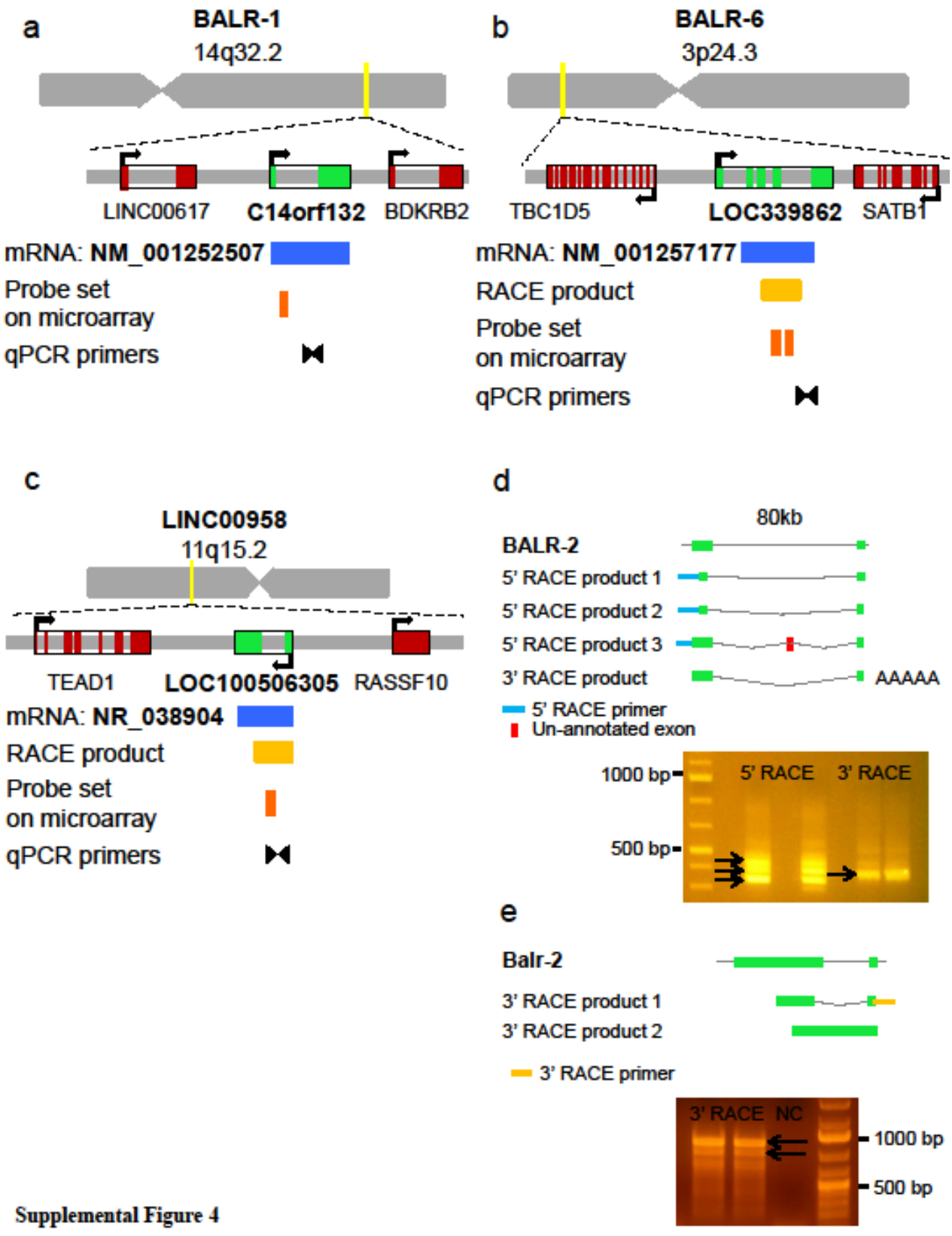
Supplemental Figure 1



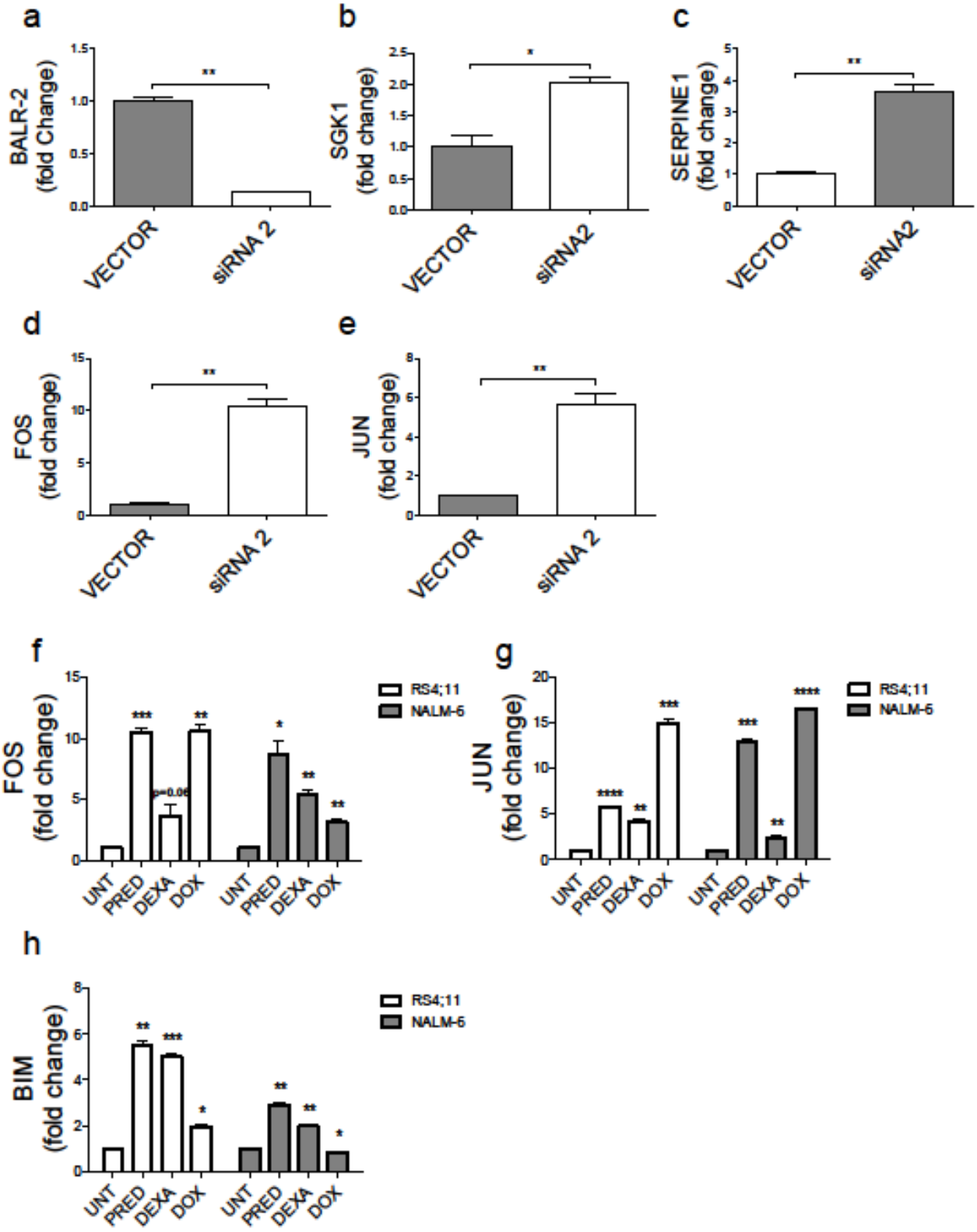
Supplemental Figure 2



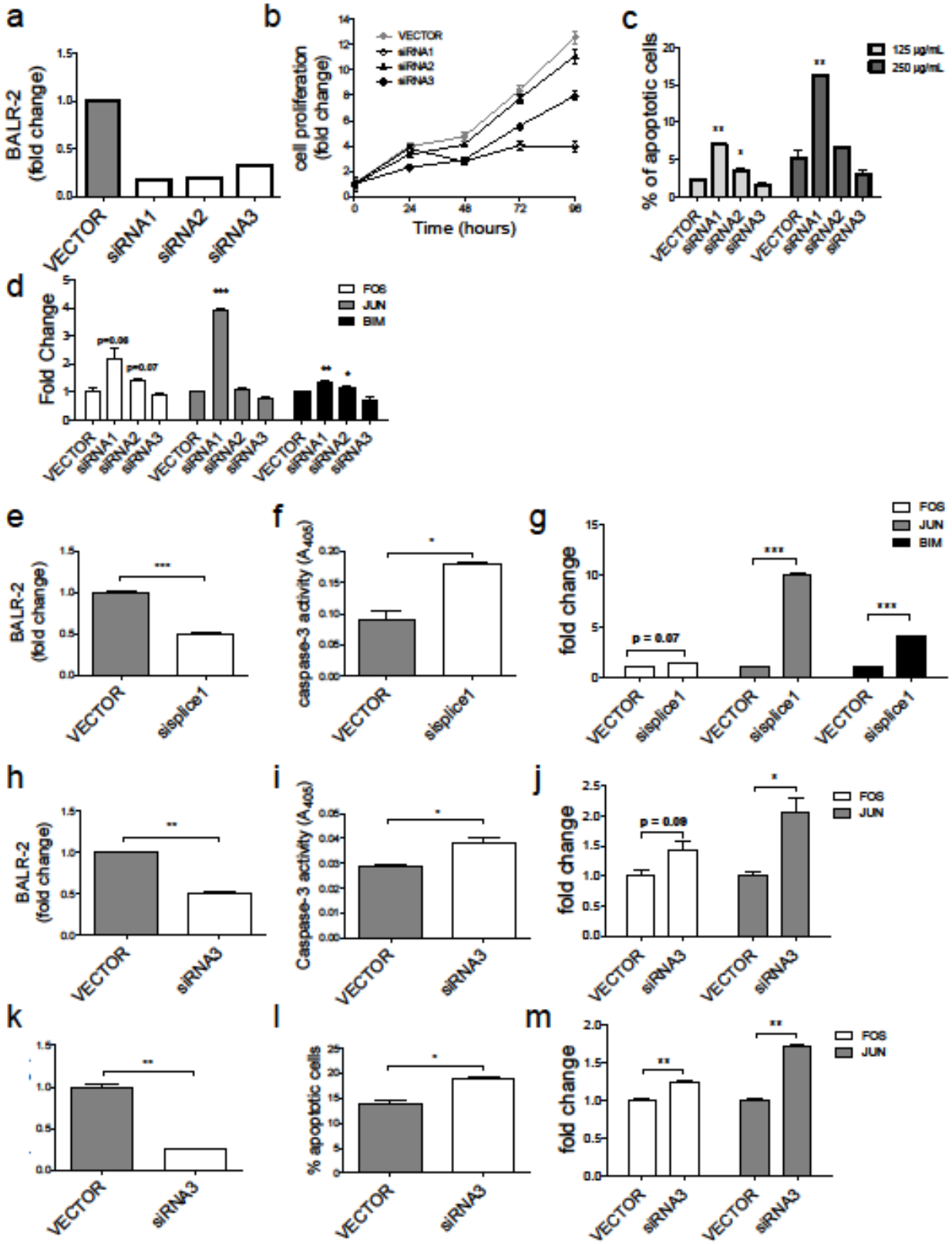
Supplemental Figure 3



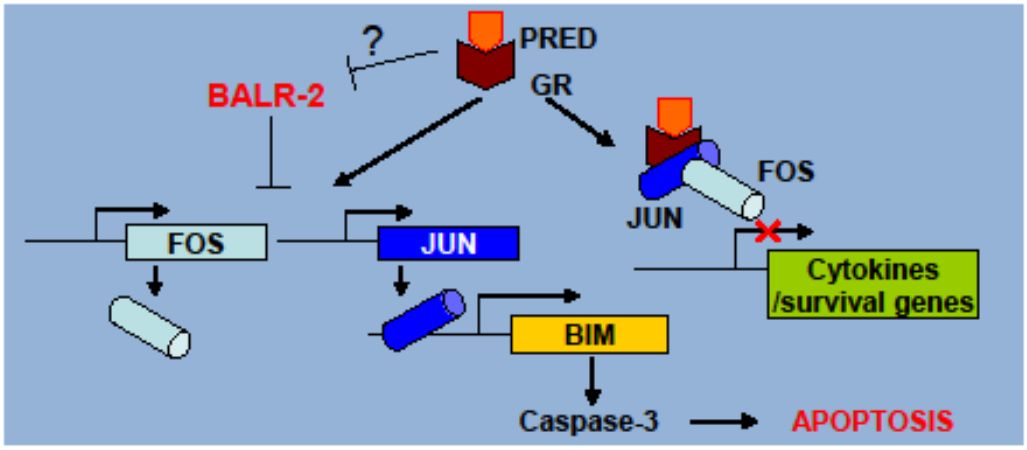
Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure 7

Pt. ID	Age (yrs)	Karyotype	Phenotype	MRD	Recurrence	Status (8 yr. F/U)	Prednisone response
Cohort 1: Discovery							
A1	ND	12;21	ND	ND	ND	ND	ND
A2	ND	12;21	ND	ND	ND	ND	ND
A3	ND	12;21	ND	ND	ND	ND	ND
A4	ND	12;21	ND	ND	ND	ND	ND
A5	ND	12;21	ND	ND	ND	ND	ND
A6	ND	12;21	ND	ND	ND	ND	ND
A7	ND	1;19	ND	ND	ND	ND	ND
A8	ND	1;19	ND	ND	ND	ND	ND
A9	ND	1;19	ND	ND	ND	ND	ND
A10	ND	1;19	ND	ND	ND	ND	ND
A11	ND	1;19	ND	ND	ND	ND	ND
A12	ND	1;19	ND	ND	ND	ND	ND
A13	ND	1;19	ND	ND	ND	ND	ND
A14	ND	V;11	ND	ND	ND	ND	ND
A15	ND	V;11	ND	ND	ND	ND	ND
A16	ND	V;11	ND	ND	ND	ND	ND
A17	ND	V;11	ND	ND	ND	ND	ND
A18	ND	V;11	ND	ND	ND	ND	ND
A19	ND	V;11	ND	ND	ND	ND	ND
A20	ND	V;11	ND	ND	ND	ND	ND
Cohort 2: Replication							
B1	ND	12;21	ND	ND	ND	ND	ND
B2	ND	12;21	ND	ND	ND	ND	ND
B3	ND	12;21	ND	ND	ND	ND	ND
B4	ND	12;21	ND	ND	ND	ND	ND
B5	ND	12;21	ND	ND	ND	ND	ND
B6	ND	12;21	ND	ND	ND	ND	ND
B7	ND	12;21	ND	ND	ND	ND	ND
B8	ND	12;21	ND	ND	ND	ND	ND
B9	ND	1;19	ND	ND	ND	ND	ND
B10	ND	1;19	ND	ND	ND	ND	ND
B11	ND	1;19	ND	ND	ND	ND	ND
B12	ND	1;19	ND	ND	ND	ND	ND
B13	ND	1;19	ND	ND	ND	ND	ND
B14	ND	1;19	ND	ND	ND	ND	ND
B15	ND	1;19	ND	ND	ND	ND	ND
B16	ND	1;19	ND	ND	ND	ND	ND
B17	ND	V;11	ND	ND	ND	ND	ND
B18	ND	V;11	ND	ND	ND	ND	ND
B19	ND	V;11	ND	ND	ND	ND	ND
B20	ND	V;11	ND	ND	ND	ND	ND
B21	ND	V;11	ND	ND	ND	ND	ND
B22	ND	V;11	ND	ND	ND	ND	ND
B23	ND	V;11	ND	ND	ND	ND	ND
B24	ND	V;11	ND	ND	ND	ND	ND

Cohort 3: Validation Cohort							
Pt ID	Age (yrs)	Karyotype	Phenotype	MRD	Recurrence	Status (8 yr. F/U)	Prednisone response
1	2.9	12;21	ND	ND	NO	ALIVE	GPR
2	4.6	1;19	ND	ND	NO	ALIVE	GPR
3	2.8	12;21	B-II	SR	NO	ALIVE	GPR
4	9.7	12;21	B-II	MR	Yes	ALIVE	GPR
5	4.2	12.21	B-II	MR	Yes	ALIVE	GPR
6	2.4	12;21	B-II	SR	NO	ALIVE	GPR
7	3.9	12;21	B-II	MR	Yes	DEAD	GPR
8	6.4	4;11	B-I	MR	NO	ALIVE	GPR
9	3.9	4;11	B-I	HR	NO	ALIVE	PPR
10	4.0	12;21	B-II	MR	ND	ND	GPR
11	7.1	12;21	B-II	SR	NO	ALIVE	GPR
12	1.1	4;11	B-I	HR	Yes	DEAD	PPR
13	3.4	12;21	B-II	MR	NO	ALIVE	GPR
14	5.4	12;21	B-II	SR	NO	ALIVE	GPR
15	7.9	12;21	B-II	SR	NO	ALIVE	GPR
16	2.4	9;11	B-I	SR	NO	ALIVE	GPR
17	11.7	NEG	B-II	MR	NO	ALIVE	GPR
18	5.7	12;21	B-II	MR	NO	ALIVE	GPR
19	6.6	12;21	B-II	MR	NO	ALIVE	GPR
20	4.1	12;21	B-II	MR	NO	ALIVE	GPR
21	10.6	NEG	B-II	MR	NO	ALIVE	GPR
22	5.3	NEG	B-II	MR	NO	ALIVE	GPR
23	11.6	NEG	B-II	MR	NO	ALIVE	GPR
24	4.4	NEG	B-II	MR	Yes	ALIVE	GPR
25	12.2	NEG	B-II	HR	ND	ND	GPR
26	6.6	12;21	B-II	MR	Yes	DEAD	GPR
27	2.6	NEG	B-II	MR	NO	ALIVE	GPR
28	5.1	NEG	B-II	SR	NO	ALIVE	GPR
29	7.9	NEG	B-I/B-II	HR	NO	ALIVE	GPR
30	8.4	NEG	B-II	MR	NO	ALIVE	GPR
31	5.8	NEG	B-II	MR	NO	ALIVE	GPR
32	2.5	12;21	B-II	MR	NO	ALIVE	GPR
33	2.2	9;22	B-II	HR	Yes	DEAD	PPR
34	3.1	12;21	ND	MR	NO	ALIVE	GPR
35	5.1	NEG	B-II	MR	NO	ALIVE	GPR
36	4.9	NEG	B-II	MR	NO	ALIVE	GPR
37	4.3	NEG	B-II	MR	NO	ALIVE	GPR
38	9.1	NEG	B-I	MR	NO	ALIVE	GPR
39	3.8	NEG	B-II	MR	NO	ALIVE	GPR
40	4.5	12;21	B-II	MR	NO	ALIVE	GPR
41	6.3	NEG	PreB	MR	NO	ALIVE	GPR
42	2.1	NEG	B-II	MR	NO	ALIVE	GPR
43	3.8	NEG	B-II	MR	NO	ALIVE	GPR
44	4.6	12;21	B-II	MR	Yes	DEAD	GPR
45	-3.1	9;22	B-II	ND	Yes	DEAD	GPR
46	5.1	NEG	B-II	MR	Yes	DEAD	GPR
47	2.4	NEG	B-II	MR	NO	ALIVE	GPR
48	8.0	9;22	B-II	HR	NO	DEAD	PPR
49	3.3	NEG	B-II	HR	NO	ALIVE	PPR
50	4.2	12;21	B-II	MR	NO	ALIVE	GPR
51	3.7	NEG	B-II	MR	NO	ALIVE	GPR
52	8.6	NEG	BD	SR	NO	ALIVE	GPR
53	1.5	NEG	B-II	MR	NO	ALIVE	GPR
54	16.4	NEG	B-II	MR	Yes	DEAD	GPR
55	2.7	NEG	B-III	MR	NO	ALIVE	GPR
56	3.5	NEG	B-I/B-II	MR	NO	ALIVE	GPR

57	13.5	NEG	B-I/B-II	SR	NO	ALIVE	GPR
58	6.2	NEG	B-II	HR	NO	ALIVE	GPR
59	15.9	NEG	B-II	MR	NO	ALIVE	GPR
60	4.4	NEG	ND	MR	NO	ALIVE	GPR
61	1.6	NEG	B-II	SR	NO	ALIVE	GPR
62	3.5	NEG	B-II	MR	NO	ALIVE	GPR
63	3.3	NEG	B-II	SR	NO	ALIVE	GPR
64	8.4	12;21	B-II	MR	NO	ALIVE	GPR
65	6.0	NEG	B-II	MR	NO	ALIVE	GPR
66	4.1	12.21	B-II	MR	NO	ALIVE	GPR
67	5.4	NEG	B-I/B-II	HR	NO	ALIVE	PPR
68	9.6	NEG	B-II	MR	NO	ALIVE	GPR
69	8.2	NEG	Pre-B	MR	NO	ALIVE	GPR
70	3.9	12;21	B-II	SR	NO	ALIVE	GPR
71	2.5	NEG	B-III	MR	NO	ALIVE	GPR
72	3.0	NEG	B-II	ND	Yes	DEAD	GPR
73	3.5	12;21	B-III	MR	NO	ALIVE	GPR
74	1.0	NEG	B-II	HR	NO	ALIVE	PPR
75	6.9	NEG	ND	ND	NO	ALIVE	GPR
76	3.4	NEG	B-I	ND	NO	ALIVE	GPR
77	10.4	NEG	B-II	MR	NO	ALIVE	GPR
78	14.5	1;19	B-II	MR	NO	ALIVE	GPR
79	5.4	NEG	B-II	MR	NO	ALIVE	GPR
80	11.5	NEG	B-II	MR	NO	ALIVE	GPR
81	3.9	12;21	B-II	MR	NO	ALIVE	GPR
82	3.9	12;21	B-II	SR	NO	ALIVE	GPR
83	13.1	NEG	B-II	MR	NO	ALIVE	GPR
84	1.8	NEG	B-II	MR	NO	ALIVE	GPR
85	5.6	NEG	B-II	SR	NO	ALIVE	GPR
86	16.5	NEG	B-I/B-II	MR	NO	ALIVE	GPR
87	1.8	NEG	B-I/B-II	MR	NO	ALIVE	GPR
88	10.8	12;21	B-I/B-II	MR	NO	ALIVE	GPR
89	14.6	NEG	B-II	MR	NO	ALIVE	GPR
90	2.7	NEG	B-II	HR	Yes	ALIVE	PPR
91	3.1	12;21	B-II	SR	NO	ALIVE	GPR
92	6.4	NEG	B-I	MR	NO	ALIVE	GPR
93	7.5	NEG	B-II	SR	ND	ND	GPR

Supplemental Table 1: Patient characteristics for use in this study. Cohorts 1 and 2 only had cytogenetic results available; these were used for the discovery and replication by microarray. All three cohorts were used to confirm the expression patterns by qPCR. Cohort 3 was used to assess clinicopathologic correlations. Abbreviations: MRD, Minimal residual disease risk stratification (MR- minimal risk, SR-standard risk, HR- high risk);Phenotype: B-I (CD10-), B-II (CD10+), B-III or B-IV (cytoplasmic Ig+); NEG, Karyotype normal (does not exclude sub-karyotypic rearrangement; GPR: Good Prednisone Response; PPR: Poor Prednisone Response.

		Translocation	MRD	Prednisone Response	Recurrence	Death
Translocation	Pearson's Correlation	N.A.	.239	.127	-.104	-.061
	Two-tailed p-value		.026	.226 (NS)	.331(NS)	.570 (NS)
MRD	Pearson's Correlation		N.A	.315	.352	.401
	Two-tailed p-value			.003	.001	<.001
Prednisone Response	Pearson's Correlation			N.A	.234	.293
	Two-tailed p-value				.028	.005
Recurrence	Pearson's Correlation				N.A	.760
	Two-tailed p-value					<.001
Death	Pearson's Correlation					N.A.
	Two-tailed p-value					

Supplemental Table 2: Correlational bivariate analyses between ordinal variable clinicopathologic characteristics. Parametric and non-parametric tests yielded similar results. Note that the translocation type did not yield significant correlations here, but was a statistically correlated with death in Kaplan Meier survival analyses.

Variable	Regression Coefficient (b)	Standard error SE(b)	p-value	Hazard Ratio (e^b)	95% Confidence Interval	
					Lower	Upper
BALR-2	-1.111	2.632	0.673	0.329	0.002	57.280
Translocation	1.361	0.406	0.001	3.901	1.759	8.650
Prednisone Response	1.396	0.888	0.116	4.04	0.708	23.044

Supplemental Table 3: Results of Cox Regression analysis using BALR-2, Translocation, and Prednisone Response as co-variates. In this sample set, the translocation was the only significant predictor of death. MRD was not included in this regression analysis because it is highly correlated with prednisone response.

RT-qPCR primers		
BALR-1	FOW	5' GGGACCTGGCCCCTCACCAA 3'
	REV	5' AGGACTGGGCACATGGAAAAAGGT 3'
BALR-2	FOW	5' AGCAGCAAAGCAAAGCCTGGGA 3'
	REV	5' CACGGCGTGGCAGCTTTCAG 3'
BALR-6	FOW	5' CGTGTGCTGGGGAAGGCACTG 3'
	REV	5' CCAGGCTCAGAGCAACACAGGGA 3'
LINC00958	FOW	5' GCTGGAGTGTGTGTGAGTGAACCA 3'
	REV	5' GCTGAGTCTCTCCACTCAGGGGG 3'
ACTIN	FOW	5' CATGTACGTTGCTATCCAGGC 3'
	REV	5' CTCCTTAATGTCACGCACGAT 3'
FOS	FOW	5' GGGGCAAGGTGGAACAGTTAT 3'
	REV	5' AGGTCATCAGGGATCTTGCAG 3'
JUN	FOW	5' TCCAAGTGCCGAAAAAGGAAG 3'
	REV	5' CGAGTTCTGAGCTTTCAAGGT 3'
BIM	FOW	5' TAAGTTCTGAGTGTGACCGAGA 3'
	REV	5' GCTCTGTCTGTAGGGAGGTAGG 3'
CELF4 (unspliced)	FOW	5' AACTGCTCTCTGGGACTCCA 3'
	REV	5' CAGCACATTAGGTGCAGAGC 3'
Balr-2	FOW	5' CGCTGGTGATGTCTGTTGTC 3'
	REV	5' GAGGCCTTGCTTTCACTGAG 3'
L32 (mouse)	FOW	5' AAGCGAAACTGGCGGAAAC 3'
	REV	5' TAACCGATGTTGGGCATCAG 3'
Fos	FOW	5' CGGGTTTCAACGCCGACTA 3'
	REV	5' TGGCACTAGAGACGGACAGAT 3'
Jun	FOW	5' TTCCTCCAGTCCGAGAGCG 3'
	REV	5' TGAGAAGGTCCGAGTTCTTGG 3'
Bim	FOW	5' CGACAGTCTCAGGAGGAACC 3'
	REV	5' CCTTCTCCATACCAGACGGA 3'
Cloning primers for P2CZL		
BamHI site	FOW	5' ATCGGCTGAGTCGACGGATCCCTGGAGGCTTGCTGAAGGCTGTATGCTG 3'
XhoI site	FOW	5' ATCGGCTGAGTCGACCTCGAGCTGGAGGCTTGCTGAAGGCTGTATGCTG 3'
	REV	5' ATCGCAATTGCTCGAGTGGGCCATTTGTTCCATGTGAGTGCTAGTAACAGGCCTTGTGTC 3'
XbaI site	FOW	5' ATCGGCTGAGTCGACTCTAGACTGGAGGCTTGCTGAAGGCTGTATGCTG 3'
	REV	5' ATCGCAATTGTCTAGATGGGCCATTTGTTCCATGTGAGTGCTAGTAACAGGCCTTGTGTC 3'
NheI site	REV	5' ATCGCAATTGGCTAGCTGGGCCATTTGTTCCATGTGAGTGCTAGTAACAGGCCTTGTGTC 3'
ApaI site	REV	5' ATCGCAATTGGGGCCCTGGGCCATTTGTTCCATGTGAGTGCTAGTAACAGGCCTTGTGTC 3'
linker sequence	FOW	5' GATCCACCTCGAGTATCTAGAATGCTAGCTTGGGCCCACT 3'
	REV	5' GATCAGTGGGCCCAAGCTAGCATTCTAGATACTCGAGGTG 3'
Primers for sequencing P2CZL		
P2CZL seq1	FOW	5' TGGCACCTGACCGAGCACG 3'
P2CZL seq2	FOW	5' TGACCCGCGAGGACCGC 3'

Cloning primers for P6UZCL		
NotI site1	FOW	5' ACAAGCGGCCGCAGTCGCGTCGGGCCTCCCGAG 3'
NotI site2	FOW	5' ACAAGCGGCCGCAGTCGCGTCGGGCCTCCCGA 3'
BamHI site1	REV	5' AAAGGATCCGTGGTAACAACGCAGAGTACTT 3'
BamHI site2	REV	5' AAAGGATCCCAAAGTCGCAATTTGCCTTTAATGTG 3'
Primers for sequencing P6UZCL		
phage6 seq1	FOW	5' CAGTGCAGGGGAAAGAATAGTAGA 3'
phage6 seq2	REV	5' CAAAGGCATTAAAGCAGCGTATCC 3'
phage6 seq3	REV	5' GGCGTAGCGCTTCATGGCTTTGT 3'
Cloning primers for MSCV vector		
BglII site	FOW	5' ATGGGCTTAGCTAGATCTGGGCCCAGTCGCGTCGG 3'
XhoI site	REV	5' ATGGCAATTATCCTCGAGTTTTTTTTTGTGGTAACAACGCAGAGTACTTT 3'
NotI site-si	FOW	5' ATCGGCTGAGTCGACGCGGCCGCCTGGAGGCTTGCTGAAGGCTGTATGCTG 3'
XhoI site-si	REV	5' ATCGCAATTGCTCGAGTGGGCCATTTGTTCCATGTGAGTGCTAGTAACAGGCCTTGTGTC 3'
Primers for sequencing MSCV vector		
MSCV seq1	FOW	5' CCGACAACCACTACCTGAGC 3'
MSCV seq2	FOW	5' CTTTATCCAGCCCTCACTCCTTCTCT 3'
RACE primers		
BALR-2 3'RACE-1	FOW	5' AGCAGCAAAGCAAAGCCTGGGA 3'
BALR-2 3'RACE-2	FOW	5' GCAGCAAAGCAAAGCCTGGGA 3'
BALR-2 3'RACE-3	FOW	5' CATGCCAACCTAATCTGTGTTAAAATGC 3'
BALR-2 3'RACE-4	FOW	5' GCATATGAAGGTCTTGACCTGAGAAAACC 3'
BALR-2 5'RACE-1	REV	5' CACGGCGTGGCAGCTTTCAG 3'
BALR-2 5'RACE-2	REV	5' TCACGGCGTGGCAGCTTTCAG 3'
BALR-2 5'RACE-3	REV	5' GATCAATTTAAGGTAAGTGGCAGGC 3'
BALR-2 3'RACE-1	FOW	5' GGTAACCAGGGCAAGGAAATGCAA 3'
BALR-2 3'RACE-2	FOW	5' GCAAACAGTAGAATCATGCCAACGT 3'
mmu-miR-155 formatted siRNA oligos		
BALR-2-siRNA1		5'GAAGGCTGTATGCTGAGATTAGGTTGGCATGATTCTGTTTTGGCCACTGACTGACAGAATC ATCAACCTAATCTCAGGACACAAGGCCTG3'
BALR-2-siRNA2		5'GAAGGCTGTATGCTGTTTACTGAAATCTCCTAGGTGGTTTTGGCCACTGACTGACCACCTA GGATTTCAGTAAACAGGACACAAGGCCTG3'
BALR-2-siRNA3		5'GAAGGCTGTATGCTGTTAATCTCAAAGTGGCTGATCGTTTTGGCCACTGACTGACGATCAG CCTTTGAGATTAACAGGACACAAGGCCTG3'
BALR-2-siRNA-splice1		5'GAAGGCTGTATGCTGCAGAGTCTGATTACCTGCTCCGTTTTGGCCACTGACTGACGGAGCA GGATCAGACTCTGCAGGACACAAGGCCTG3'
Balr-2-siRNA1		5'GAAGGCTGTATGCTGAAATGGTTTCCTCAGGTCAAGGTTTTGGCCACTGACTGACCTTGAC CTGGAAACCATTTCAGGACACAAGGCCTG3'
Balr-2-siRNA2		5'GAAGGCTGTATGCTGTAAGGTAAGTGGCAGGCGCTTGTTTTGGCCACTGACTGACAAGCGC CTCACTTACCTTACAGGACACAAGGCCTG3'
Balr-2-siRNA3		5'GAAGGCTGTATGCTGATTTAAGGTAAGTGGCAGGCGGTTTTGGCCACTGACTGACCGCCTG CCTTACCTTAAATCAGGACACAAGGCCTG3'

RACE Sequence BALR-2	5'AGTCGCGTCGGGCCTCCCGAGGGGGCTGCGAGTGTCAGTCGGCTCTCCGCACGTGTCCGCG GCCTCGCGGAGCAGCACTGCAAGCTATTTGGCATATGAAGGTCTTGACCTGAGAAAACCATCT TGGATAACTGCAGCAAGGAAAAGGAAAAATGCAACACCTAGGAGATTTCAGTAAACAGTAGAA TCATGCCAACCTAATCTGTGTTAAAATGCTTGGAATGTGGGAGCCGCTGATGATGCCTCTTGT CTGTGTGTCTGACTGAATCCTTTCTTTTCTCAGAGCAGCAAAGCAAAGCCTGGGAACCAGGCC AAATGCCTGCCACTTACCTTAAATTGATCAGCCACTTTGAGATTAAAACCCCTGAAAGCTGCC ACGCCGTGAAAACAAGGCCTCCTTCACATTAAAGGCAAATTGCGACTTTGAAAAAAAAAAAAA AGAAAAAAAAAAAAAAAGTACTCTGCGTTGTTACCAC3'
RACE Sequence Balr-2	5'CGCAGCAAAGCCAAGCCTGGGAACCAGGCCAAGCGCCTGCCACTTACCTTAAATTGATCAGCCACTT TGAGATTAAAACCCCTGAAAGCTGCCGCTCAGTGAAAGCAAGGCCTCTTTCACATTAAAGGCAAATTGC GACTGTGGGTTTGCTGTGATCGCCCTCTCTCTCTCTCTCTCCTTTTTTCCCCCTATTTTACCCGTTTTT TTCAGTGTGGACTTTTTTCTCTCTCCTCTTTCCATTTATGCTTCCATGTAGAAGCCAATAGTTATAGCA TCTTAGCCAAGCATTCATTAAGGTTTATAAGAACAAAATTTATGGTTATTTAGAAGTTGGCACAGAGAT CAATTGTTGCTAGAAAAAAAAAAGAAAGAAA3'

Supplemental Table 4: Primers and RACE sequence for human and mouse BALR-2.

CHAPTER III:

“BALR-6 Regulates Cell Growth and Cell Survival in B-Lymphoblastic Leukemia”

Abstract

Background: A new class of non-coding RNAs, known as long non-coding RNAs (lncRNAs), has been recently described. These lncRNAs are implicated to play pivotal roles in various molecular processes, including development and oncogenesis. Gene expression profiling of human B-ALL samples showed differential lncRNA expression in samples with particular cytogenetic abnormalities. One of the most promising lncRNAs identified, designated B-ALL associated long RNA-6 (BALR-6), had the highest expression in patient samples carrying the MLL rearrangement, and is the focus of this study.

Results: Here, we performed a series of experiments to define the function of BALR-6, including several novel splice forms that we identified. Functionally, siRNA-mediated knockdown of BALR-6 in human B-ALL cell lines caused reduced cell proliferation and increased cell death. Conversely, overexpression of BALR-6 isoforms in both human and mouse cell lines caused increased proliferation and decreased apoptosis. Overexpression of BALR-6 in murine bone marrow transplantation experiments caused a significant increase in early hematopoietic progenitor populations, suggesting that its dysregulation may cause developmental changes. Notably, the knockdown of BALR-6 resulted in the dysregulated expression of a set of genes enriched for leukemia-associated genes, as well as the transcriptome regulated by Specificity Protein 1 (SP1). We confirmed changes in the expression of SP1, as well as its known interactor and downstream target CREB1. Luciferase reporter assays demonstrated an enhancement of SP1-mediated transcription in the presence of BALR-6. These data provide a putative mechanism for regulation by BALR-6 in B-ALL.

Conclusions: Our findings support a role for the novel lncRNA BALR-6 in promoting cell survival in B-ALL. Furthermore, this lncRNA influences gene expression in B-ALL in a manner consistent with a function in transcriptional regulation. Specifically, our findings suggest that

BALR-6 expression regulates the transcriptome downstream of SP1, and that this may underlie the function of BALR-6 in B-ALL.

Background

The human genome produces thousands of non-coding transcripts [1]. These include the recently described class of long non-coding RNAs (lncRNAs), which have distinct chromatin signatures and epigenetic marks, designating them as unique structures that are conserved in mammals [2, 3]. More recently, comparison of lncRNA expression in zebrafish to that of mammals has suggested that although these structures retain limited overall sequence conservation among vertebrates, they show strong conservation of short stretches of sequence, chromosomal synteny, and functional conservation [4]. Prior studies have shown that lncRNAs play a variety of roles in the regulation of transcription, splicing, and miRNA function [5-7]. This may not be an exhaustive description of the functions of lncRNAs, as new functions are being discovered in other cellular processes [8, 9]. As might be expected considering their roles in critical cellular functions, lncRNAs have been found to be dysregulated in cancer, with functional roles in oncogenesis described for a handful of lncRNAs so far [10-13].

B-lymphoblastic leukemia (B-acute lymphoblastic leukemia, B-ALL) is a malignancy of precursor B-cells harboring mutations and translocations that result in dysregulated gene expression [14, 15]. We have recently completed a comprehensive description of lncRNAs in B-ALL and analyzed the association of lncRNA expression with clinicopathologic parameters. Our study showed differential lncRNA expression in samples with particular cytogenetic abnormalities [16]. One of the lncRNAs from our study, designated B-ALL associated long RNA-6 (BALR-6), was significantly upregulated in all subsets of patient samples when compared to normal CD19+ cells. Interestingly, the highest expression of BALR-6 was seen in patient

samples carrying the MLL rearrangement [16]. MLL rearranged B-ALL cases have a very poor prognosis and occur in infants, making them particularly hard to treat [17].

Located on chromosome 3p24.3 in humans, *BALR-6* exists in a syntenic gene block with neighboring genes *SATB1* and *TBC1D5* that is conserved in several vertebrate species (Figure 1A-B, 1D) [16]. Analysis of publically available data from the Broad Institute/ENCODE shows H3K4m3 and H3K36m3 modifications along the promoter and gene body at *LOC339862*, where BALR-6 resides, indicating that it is a transcriptional element (Figure 1A) [4, 16, 18-20]. Alternative splicing analysis by the Swiss Institute of Bioinformatics predicted multiple transcripts expressed at this gene locus (Supp Fig 1A) [21]. Moreover, 100 Vertebrate PhastCons analyses of the *BALR-6* locus demonstrated significant conservation of the gene body, suggesting a functional transcript (Figure 1C) [22].

To further study this lncRNA we undertook loss-of-function analyses in B-ALL cell lines and gain-of-function analyses *in vivo*. We found that BALR-6 is a pro-survival factor for B-ALL cell lines, and that its knockdown led to decreased growth and increased apoptosis of these cells. *In vivo*, overexpression of BALR-6 led to an alteration of hematopoiesis with a shift to more immature progenitor populations. Gene expression analyses of knockdown cell lines showed a differentially expressed gene set in BALR-6 knockdown cells, with enrichment for SP1 transcriptional targets and leukemogenic genes. Finally, luciferase assays demonstrated and increase in transcriptional activity when SP1 and BALR-6 were co-expressed. Together, these findings point to a role for BALR-6 in cellular survival, leukemogenesis, and highlight the role of novel elements of gene regulation in B-ALL.

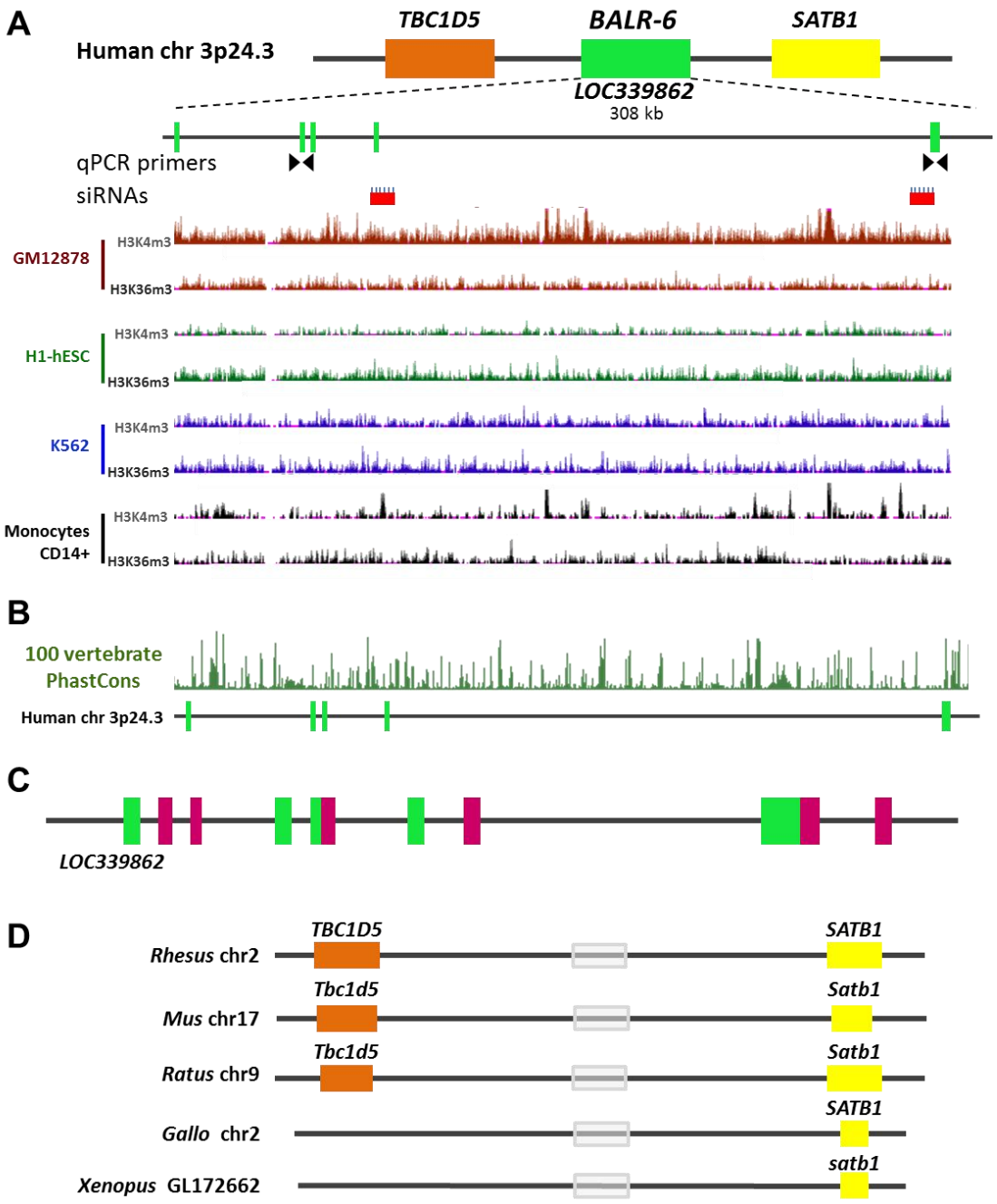


Figure 1: Molecular characterization of BALR-6.

(**A**) Top: Chromosomal location of *BALR-6* in the human genome, surrounding genes, qPCR primers, siRNAs, known annotated exons (green boxes), known introns (black lines) are shown. Bottom: Chip-Seq histone modification map from the ENCODE/Broad institute, taken from UCSC genome browser, shows H3K4m3 and H3K36m3 patterns at *LOC339862* in four different cell types indicating active transcription of the lncRNA. (**B**) The 100 Vertebrate PhastCons plot from the UCSC whole-genome shows conserved regions among 98 vertebrates including mice and zebrafish throughout the locus. (**C**) RACE discovered unannotated exons (magenta) depicted with known annotated exons (green) at *LOC339862*. (**D**) Schematic depicting genomic conservation of the syntenic block among multiple vertebrates, as analyzed by BLAT. Grey box indicates location of homology to *BALR-6*.

Results

BALR-6 knockdown inhibits proliferation of human B-ALL cell lines

To comprehensively study the function for this novel lncRNA, we first characterized the transcripts originating at the genomic locus corresponding to *BALR-6*. Using RS4;11 cell line mRNA, Rapid Amplification of cDNA Ends (RACE) uncovered multiple isoforms; from these, three were cloned and sequenced corresponding to the genomic locus as shown (Supp Fig 1A-B). Northern Blot analysis of RS4;11 DNAse treated RNA revealed the expression of two isoforms containing exon 3 and exon 5 sequences, one sized at ~3.8 Kb and the other at ~1.2 Kb (Supp Fig 1C). The annotated mRNA and new alternative splice forms, including unannotated exons, were confirmed as depicted in Figure 1C. Isoform 1 contains several small open reading frames (ORFs), however no Kozak sequences are found in their initial ATG region, and the predicted ORFs do not resemble any known functional proteins or peptide [23]. Isoforms 2 and 3 lacked open reading frames and translation initiation sites as evaluated by EMBOSS Transeq, predicting them to be non-coding transcripts (Supp Fig 1D).

To map the murine homologous transcript, we carried out 5'RACE and 3'RACE using mRNA extracted from murine pre B-ALL cell line 70Z/3. The sequences uncovered match the human BALR-6 sequence, confirming that there is a murine transcript originating from this same locus (Supp Fig 1E). Further analysis by BLAT showed genomic conservation of syntenic blocks in a variety of vertebrates, including *Xenopus tropicalis* (Figure 1D). Together, these data demonstrate a highly conserved, functional, and complex gene locus that expresses multiple non-coding transcripts, some yet to be discovered. During normal B cell development, BALR-6 is dynamically expressed, with high expression in pre-B cells and subsequent downregulation (Figure 2A). This suggests that the high expression of BALR-6 in B-ALL could represent a stage-specific expression pattern in leukemia derived from early stages of B-cell development. To elucidate a cellular function for BALR-6, we first evaluated the expression levels of the

transcripts in human B-ALL cell lines. BALR-6 expression was highest in RS4;11 cells and MV(411) cells, which carry the MLL-AF4 rearrangement, when compared to other lines (Figure 2B). Additionally, RS4;11 cells treated with bromodomain and extra-terminal (BET) motif binding protein inhibitor I-BET151 [24] showed decreased levels of BALR-6 in a dose-dependent manner (Figure 2C). Given that I-BET151 has previously been shown to inhibit transcription downstream of MLL, we propose that BALR-6 expression is induced by MLL, although this effect may not be entirely specific to MLL-AF4.

Using the approach described previously, siRNAs against the splice junctions between exons of BALR-6 were cloned into a mmu-miR-155 expression cassette (Supp Fig 2A) [4, 16, 25, 26]. We observed knockdown of all the identified transcripts in multiple B-ALL cell lines (Figure 2D and Supp Fig 2B). Transduced B-ALL cells showed a reduction in proliferation as early as 48 hours after plating, with consistent reduction in proliferation observed over the full duration of the assay (up to 144 hours) (Figure 2E-F and Supp Fig 2C). siRNA-transduced B-ALL cells had significantly higher levels of apoptosis, as measured by AnnexinV, when compared with vector-transduced lines (Figure 2G-H and Supp Fig 2D). Flow cytometry demonstrated that the siRNA2-transduced RS4;11 cell lines had an increase in Sub-G0 cells and a decrease in all other cell stages, consistent with increased apoptosis and decreased flux through the cell cycle (Figure 2I). Together, these findings suggest a modest yet conserved role for BALR-6 in the regulation of B-ALL cell survival and proliferation.

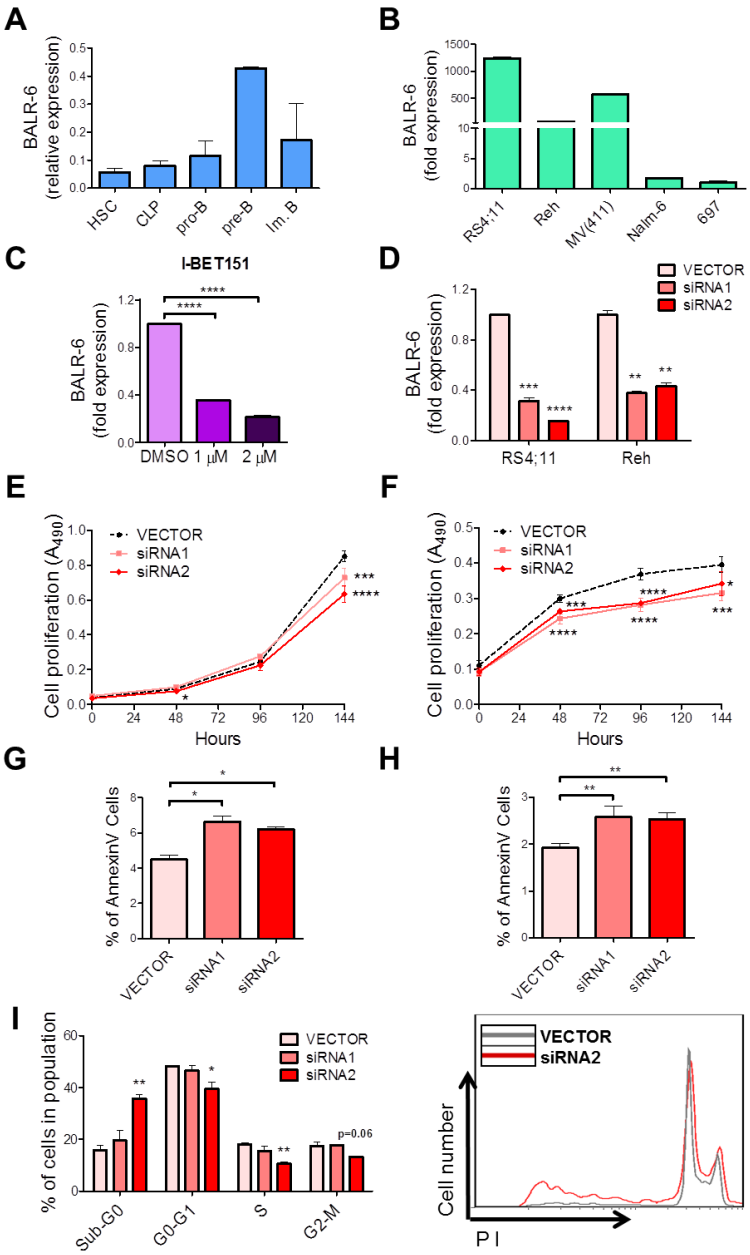


Figure 2: BALR-6 knockdown reduces cell proliferation and increases apoptosis in human B-ALL cells. (**A**) BALR-6 expression in human bone marrow B cell subsets by qRT-PCR. Normalized to ACTIN. (**B**) Quantitation of BALR-6 expression in human B-ALL cell lines by qRT-PCR confirming elevated levels in MLL translocated cell lines RS4;11, and MV(411). Normalized to ACTIN. (**C**) RS4;11 cell lines treated with 1µM, and 2µM of I-BET151 inhibitor for 36 hours, presented a decrease in BALR-6 expression levels. Normalized to ACTIN. (**D**) qRT-PCR quantification of BALR-6 in RS4;11 and Reh cell lines transduced with vector control, siRNA1, or siRNA2. Normalized to ACTIN. (**E-F**) Decreased cell proliferation, upon siRNA mediated knockdown of BALR-6 in RS4;11 cells (**E**), and Reh cells (**F**) as measured by MTS. (**G-H**) AnnexinV staining showed that siRNA mediated knockdown of BALR-6 in RS4;11 cells (**G**), and Reh cells (**H**) resulted in an increase of apoptosis. (**I**) Propidium iodide staining of RS4;11 knockdown cell lines showed an increase in Sub-G0 and a decrease in G0-G1, S, and G2-M cells. Representative histogram of **I** confirms cell cycle changes, shown to the right. HSC, hematopoietic stem cell; CLP, common lymphoid progenitor; pro-B, progenitor B; pre-B, precursor B; DMSO, dimethyl sulfoxide. Evaluations were made using a two-tailed T-test, $p<0.05$ (*); $p<0.005$ (**); $p<0.0005$ (***); $p<0.0001$ (****).

Constitutively expressed BALR-6 supports cell survival and proliferation

To examine the effects of BALR-6 gain of function, we overexpressed the previously identified isoforms in the human B-ALL cell line Nalm-6, which has relatively low endogenous levels of the transcript (Figure 3A, Figure 2B). Gene transfer was conducted via a lentiviral expression system that has proven successful in our previous studies (Supp Fig 2E) [16]. Constitutive overexpression of BALR-6 Isoforms 2 and 3 led to a significant increase in proliferation as measured by MTS (Figure 3C). In addition to an observed increase in overall growth rate, BALR-6 Isoforms 2 and 3 caused an increase in S phase cells and G2-M cells (Figure 3D). Furthermore, AnnexinV staining showed significantly lower numbers of apoptotic cells under basal growth conditions in cell lines overexpressing any of the BALR-6 isoforms (Supp Figure 2G).

To overexpress BALR-6 in mouse cells, we constructed a set of MSCV-based bicistronic vectors (Figure 3B, Supp Fig 2F). Successful overexpression of these constructs in murine pre B-ALL 70Z/3 cells led to a modest increase in proliferation (Figure 3E-F). Cell cycle analysis of these lines showed an increase of S phase cells, G2-M cells (in Isoform 3 overexpressing lines) and a reduction in Sub-G0 cells, similar to the effects in Nalm-6 cells (Figure 3G-H). Analysis by AnnexinV staining confirmed the lower number of apoptotic cells in Isoform 3 expressing cell lines (Supp Fig 2H). Moreover, these 70Z/3 Isoform 3 overexpression lines were less vulnerable to prednisolone-induced apoptosis (Supp Fig 2I). Conversely, siRNA-transduced RS4;11 cells were more prone to prednisolone-induced apoptosis (Supp Fig 2I). Therefore, knockdown and overexpression of BALR-6 had opposing phenotypes in B-ALL cell lines, and gain of function phenotypes were conserved in both human and mouse cells.

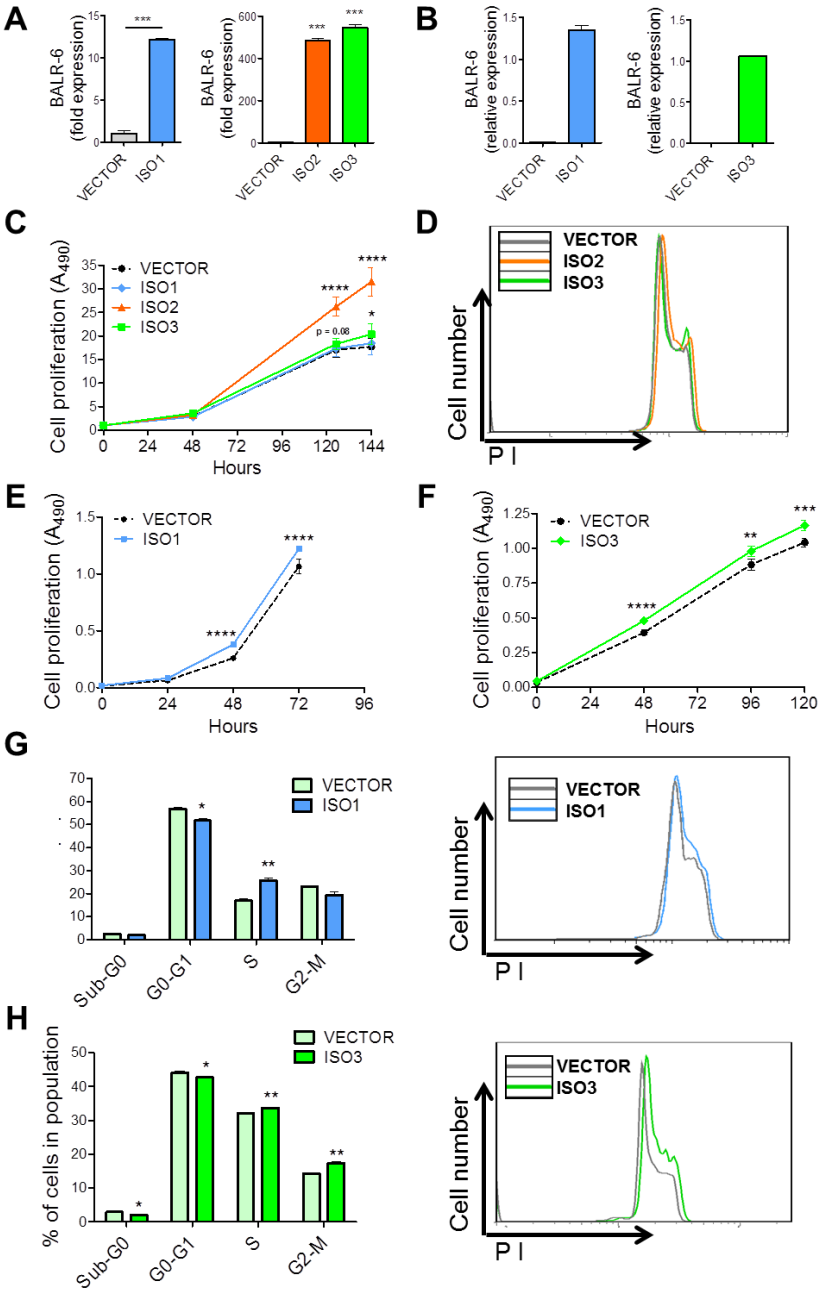


Figure 3: BALR-6 overexpression increases proliferation in human Nalm-6 and murine 70Z/3 cells. (**A**) qRT-PCR quantitation of BALR-6 isoform expression in Nalm-6 human pre B-ALL cell line. Normalized to ACTIN. (**B**) qRT-PCR quantification of BALR-6 isoforms in 70Z/3 mouse pre B-ALL cell line. Normalized to Actin (ISO1) or L32 (ISO3). (**C**) Increased cell proliferation in BALR-6 overexpressing Nalm-6 cell lines, as measured by MTS. (**D**) Representative histogram of Nalm-6 overexpression lines, stained with propidium iodide, shows an increase in S phase cells and G2-M cells. (**E-F**) Increased cell proliferation in BALR-6 Isoform 1 (**E**) and Isoform 3 (**F**) overexpressing 70Z/3 cell lines, as measured by MTS. (**G-H**) Propidium iodide staining of 70Z/3 cells overexpressing BALR-6 Isoform 1 (**G**) and Isoform 3 (**H**), shows a consistent increase in G2-M cells, and a decrease in Sub-G0 cells. Representative histogram of figures **G-H** confirmed the increase in cells in the G2-M when compared to the empty vector, shown to the right. ISO1, Isoform 1; ISO2, Isoform 2; ISO3, Isoform 3. Evaluations were made using a two-tailed T-test, $p<0.05$ (*); $p<0.005$ (**); $p<0.0005$(***); $p<0.0001$(****).

Enforced BALR-6 expression promotes expansion of hematopoietic progenitor populations *in vivo*

Since BALR-6 is highly expressed in B-ALL, we tested the effects of constitutive expression in an in vivo model [16]. 5-FU enriched bone marrow was transduced with retrovirus expressing the BALR-6 Isoform 3 and transplanted into lethally irradiated hosts (Figure 3 and Supp Fig 2F, 2H). Mice were followed with peripheral bleeds for 16 weeks and then sacrificed for analysis. Peripheral white blood cell counts were not statistically different between the control and experimental groups. However, mice with enforced expression of BALR-6 showed a trend towards lower red blood cell counts, hematocrit and platelet counts (Supp Fig 3A). Flow cytometry revealed a lower percentage of CD11b+ myeloid cells and a higher percentage of B220+ B cells, but no difference in CD3ε+ T cell percentage in the eGFP+ population of experimental mice (Supp Fig. 3B, 3C).

Mice were sacrificed following 4 months of reconstitution. Gross analysis showed no changes in the thymus, spleen, livers, or kidneys. Microscopic inspection of hematoxylin and eosin – stained tissues did not reveal any differences (Supp Fig. 3D). In the bone marrow, qRT-PCR confirmed successful overexpression of BALR-6 (Supp Fig. 4A-B). Analysis by flow cytometry revealed an increase in precursor cell populations in the GFP+ population of the experimental mice, when compared to the control group (Figure 4A-B, Supp Fig 5C). After exclusion of differentiated cells in the bone marrow, we observed increased relative proportion of Lin-Sca1+c-Kit+ (LSK) cells, hematopoietic stem cells (HSCs), and lymphoid-primed multipotent progenitors (LMPPs) in mice overexpressing BALR-6 (Figure 4B). An increase in the relative population of Lin-Sca1loc-Kitlo cells and a trend towards increased relative population of common lymphoid progenitors (CLPs) was also observed (Supp Fig 4C). The developmental pathway of B-cells in the bone marrow was investigated by the method of Hardy et al [27]. Once again, trends towards higher relative proportions of these B-cell developmental stages were

observed (fractions A-F, Supp Fig 4D). Taken together, these results suggest that BALR-6 overexpression leads to an enrichment of early developmental stage cells in murine bone marrow, indicating that its expression confers a survival advantage or increased proliferation for cells in these earlier stages.

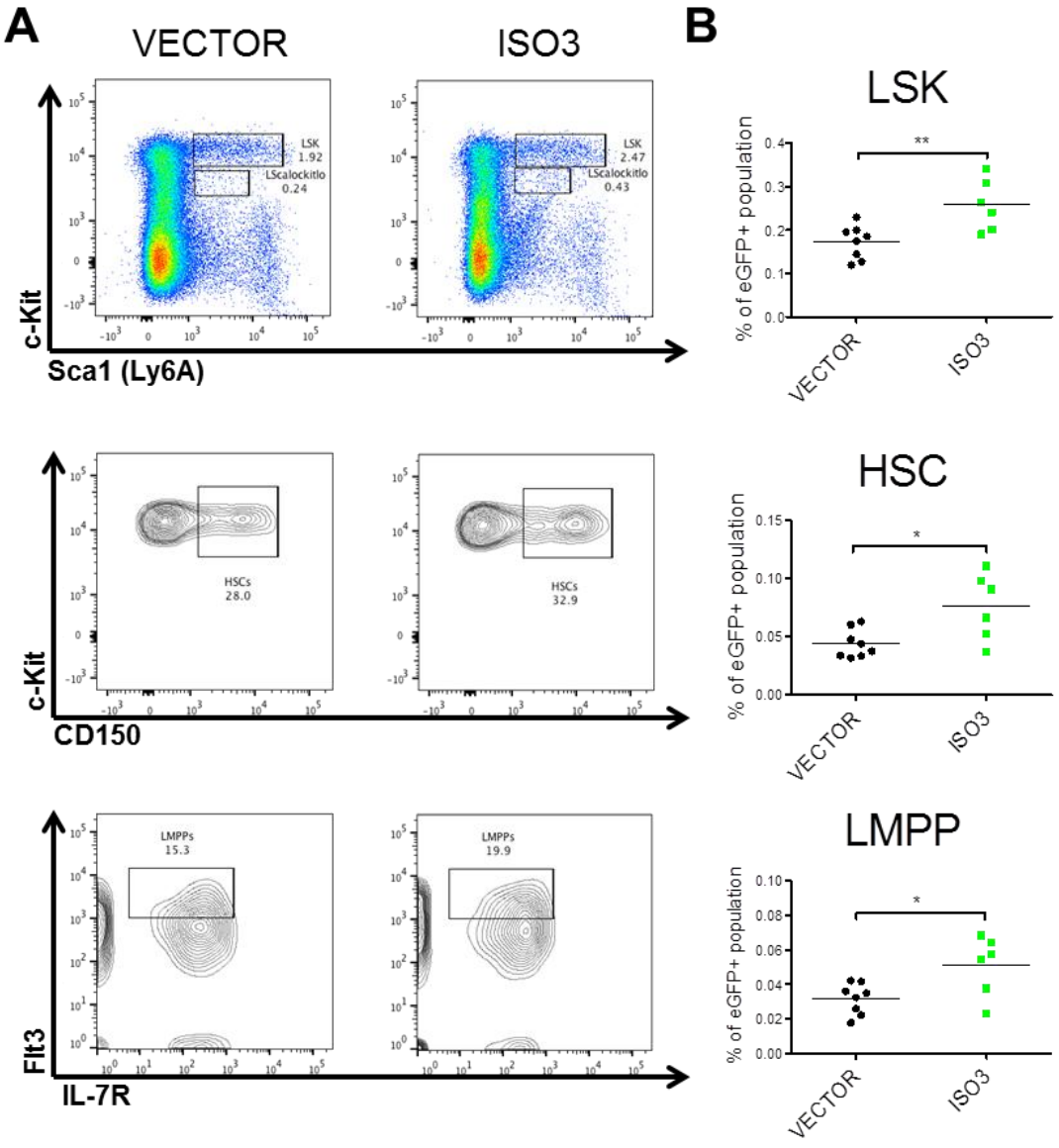


Figure 4. BALR-6 overexpression causes an increase in hematopoietic precursor cells *in vivo*.
(**A**) Representative FACS plots of hematopoietic progenitor populations LSK, HSC, and LMPP in bone marrow transfer mice. (**B**) Quantitation of progenitor populations showing a significant increase in experimental mice when compared to control. Number of mice used in this analysis: VECTOR, n=8; ISO3, n=6. ISO3, Isoform 3; HSC, hematopoietic stem cell; LMPP, lymphoid primed multipotent progenitor; LSK, lineage- Sca1+ c-Kit+. Evaluations made using a two-tailed T-test, $p<0.05$ (*); $p<0.005$ (**).

BALR-6 regulates expression of genes involved in multiple biological processes

At the molecular level, several studies have demonstrated that many lncRNAs act as transcriptional regulators [5, 11, 23, 28, 29]. To explore whether or not BALR-6 regulates gene expression, RNA isolated from knockdown cell lines was analyzed by microarray [30, 31]. Upon siRNA mediated knockdown of BALR-6, 2499 probes showed differential expression. Of these, 1862 unambiguously mapped to 1608 unique Entrez Gene IDs. Unsupervised hierarchical clustering analysis identified differentially expressed genes in the siRNA-expressing cell lines (Figure 5A).

Further data analysis was carried out using WebGESTALT [32, 33]. Gene Ontology (GO) slim classification of differentially expressed genes by molecular function was utilized to provide insight into the pathways in which BALR-6 is involved, with protein binding function category having the most dysregulated genes (Figure 5B). A number of biological processes, as annotated in the GO database, were significantly dysregulated in BALR-6 knockdown cell lines, including cell death and cell proliferation (Figure 5C). Disease associated enrichment analysis, which was inferred using GLAD4U, showed an enrichment of genes known to be dysregulated in various disease states (Figure 5E). Of the 38 significantly associated disease states, 14 were of leukemic origin. Transcription factor enrichment analysis showed a significant enrichment of genes that are predicted to be targeted by SP1, among other transcription factors (Figure 5D). Taken together, these data revealed the biological importance of BALR-6. A detailed description of the microarray analyses can be found in the methods.

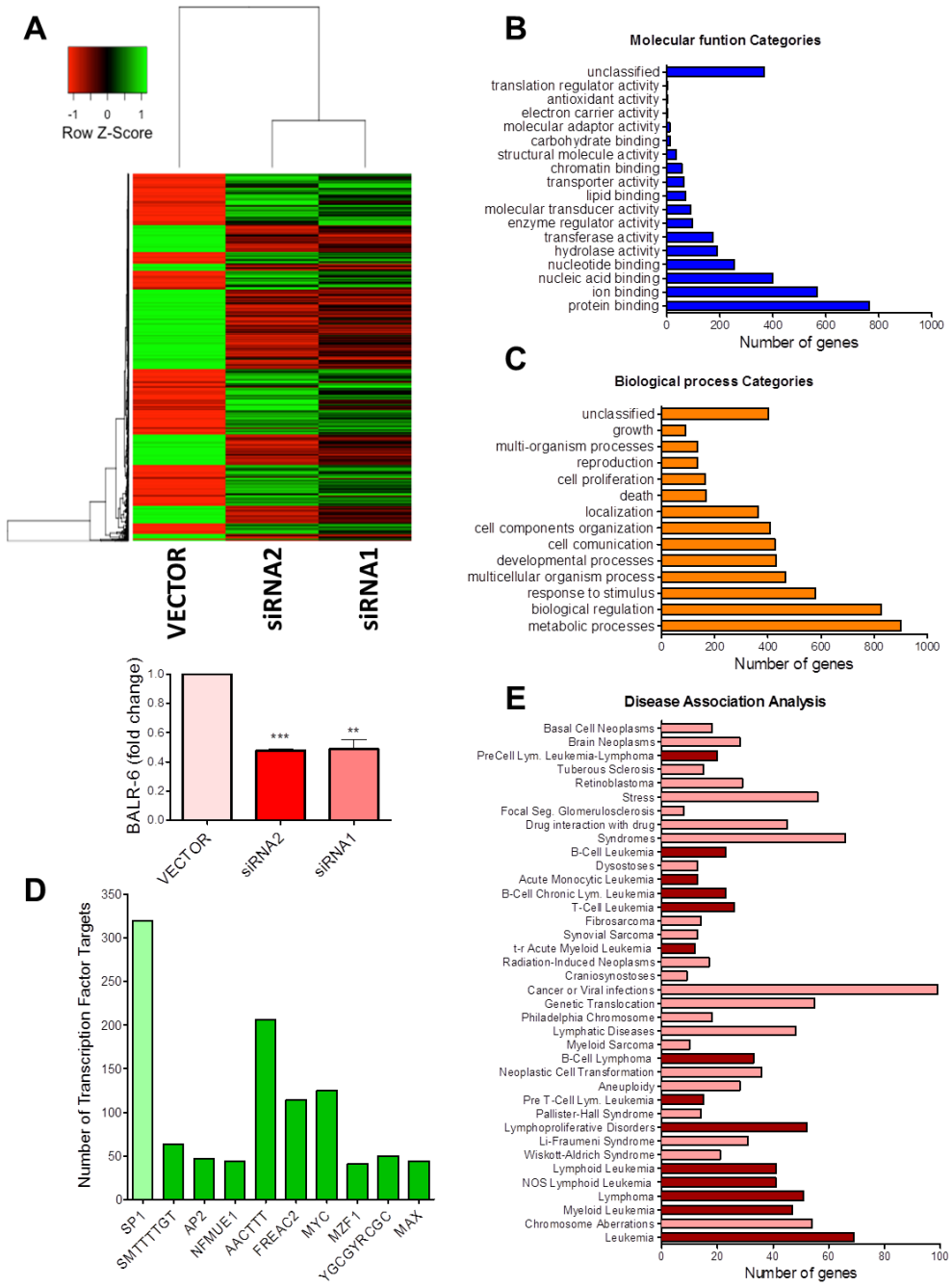


Figure 5. BALR-6 knockdown leads to global differential expression of genes.
(**A**) Unsupervised hierarchical gene clustering of differentially expressed genes upon BALR-6 siRNA mediated knockdown in RS4;11 cells (PPDE > 95%, fold change > 1.5). qRT-PCR confirmation of BALR-6 knockdown shown below. (**B-C**) Bar graphs of GO Slim classification enrichment analysis of differentially expressed genes by molecular function (**B**), and biological processes (**C**). (**D**) Enrichment analysis of transcription factor targets. Top ten transcription factors with a p value < 0.0001 are shown. For unknown transcription factors, transcription site sequence is shown. (**E**) Disease association analysis by GLAD4U, revealing enrichment of genes associated to various malignancies, in particular, hematological malignancies (dark red). Diseases with a p value $\leq$0.05 are shown. ; PPDE, posterior probability of differential expression. Evaluations were made using a two-tailed T-test, p<0.005 (**); p<0.0005(***).

SP1 transcriptome is modulated by BALR-6

As indicated by the transcription factor enrichment analysis, we confirmed that the expression SP1 and CREB1, a target and interactor of SP1, were dysregulated upon BALR-6 knockdown (Figure 6A). The strongest phenotype was seen in the siRNA2 mediated knockdown, which also showed the strongest cellular phenotypes in the majority of pre B-ALL cell lines (Figure 6A, Supp Fig 5A-B). Conversely, increased levels of SP1 and CREB1 correlated with overexpression of BALR-6 isoforms in both human and murine cell lines (Nalm-6 and 70Z/3) (Figure 6B).

To confirm our findings, a second microarray analysis was carried out with technical duplicates of RS4;11 cell lines transduced with empty vector or siRNA2. 2756 probes showed differential expression. Of these, 2280 unambiguously mapped to 2128 Entrez Gene IDs and were analyzed by hierarchical clustering (Supp Fig 6A). Enrichment analysis in WebGESTALT revealed similar GO slim classifications (Supp Fig 6B-C), and transcription factor target enrichment analysis confirmed significant enrichment of SP1 targets (Supp Fig 6D). Additionally, enrichment of CREB1 targets was significant (Supp Fig 6D). Notably, leukemic diseases were the only ones significantly enriched in the disease association analysis (Supp Fig 6E). Together, these findings indicated a consistent change in the transcriptome, particularly downstream of SP1, upon knockdown of BALR-6 in MLL rearranged B-ALL.

To further understand the relationship of BALR-6 and SP1, we examined promoter regions of known SP1 targets (CREB1 and p21) and cloned these sequences into the luciferase reporter vector, pGL4.11 (Figure 6C). The p21 promoter contained 6 putative SP1 binding sites, while the CREB1 promoter contained 7 such sites (Supp Fig 5C-D). Luciferase reporter assays in HEK 293T cells with constitutive expression of SP1, Isoform 1, Isoform 3 or a combination of these vectors, revealed increased luciferase activity in both promoters (Figure 6D). Notably,

when SP1 and BALR-6 were co-overexpressed, we noted a strong increase in transcriptional activity with both the p21 and CREB1 promoter.

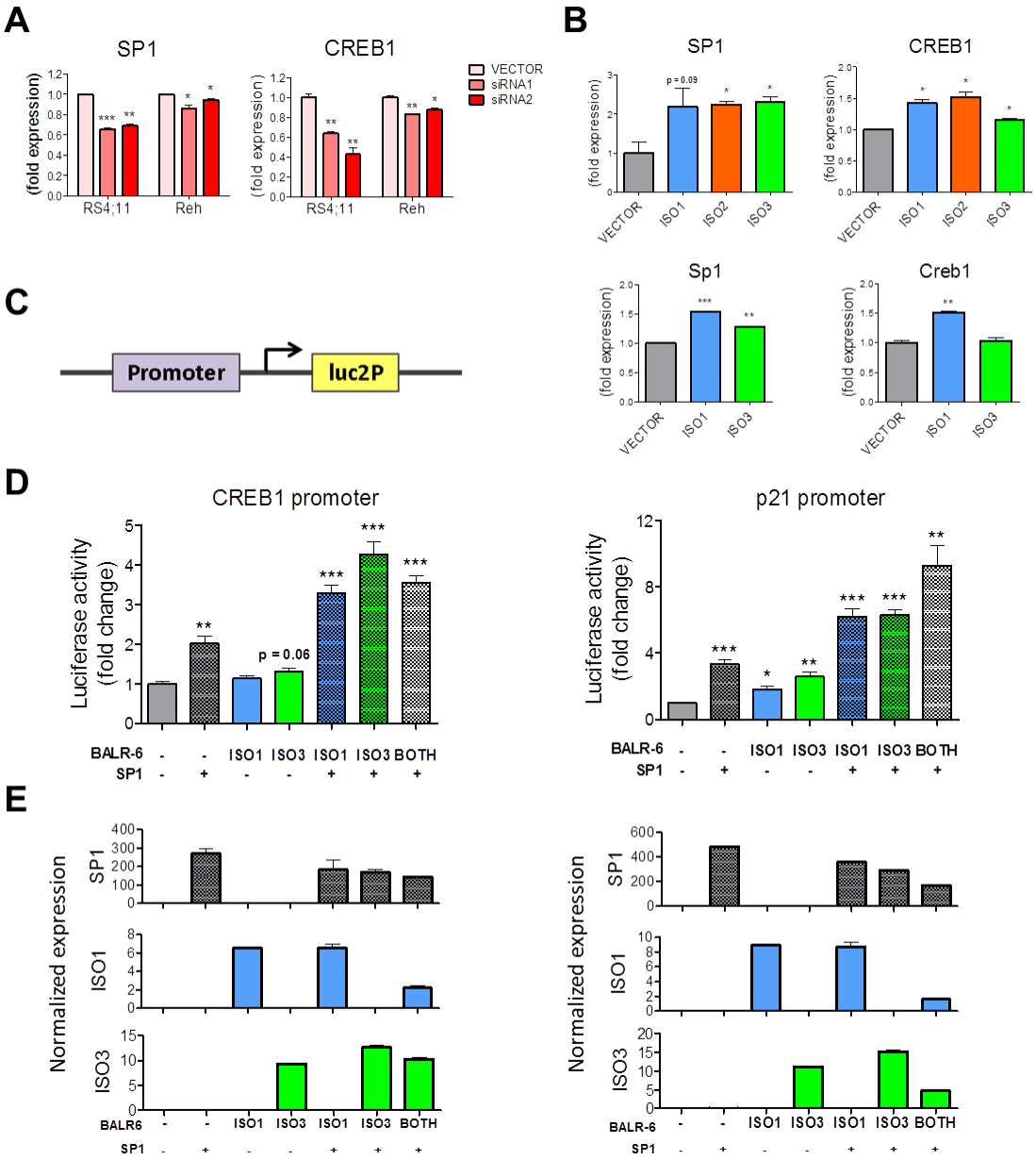


Figure 6: SP1 transcriptome is modulated by BALR-6.
(**A**) Confirmation of SP1 and CREB1 expression in RS4;11 microarray samples, as well as Reh knockdown cell lines. Normalized to ACTIN. (**B**) SP1 and CREB1 transcript level increase correlates with overexpression of BALR-6 in Nalm-6 cells (top) and 70Z/3 cells (bottom). Quantitation by qRT-PCR, normalized to ACTIN (Nalm-6 cells) or L32 (70Z/3 cells). (**C**) Schematic depicting location of cloned promoter sequences in the pGL4 vector system for luciferase assays. (**D**) Transcriptional activity at CREB1 (left) and p21 (right) promoter regions upon SP1 and/or BALR-6 overexpression, as measured by luciferase activity. (**E**) Quantitation of overexpression in luciferase assays (as seen in **D**) by qRT-PCR of respective transcripts, normalized to ACTIN. Evaluations were made using a two-tailed T-test, $p<0.05$ (*); $p<0.005$ (**); $p<0.0005$(***). luc2p, synthetic firefly luciferase.

Discussion

The discovery of lncRNAs has revolutionized how we think about gene expression. The genomic organization of many lncRNAs is indeed complex. Some are found in regions overlapping with protein coding genes, while others that are exclusively intergenic [2, 4]. Some lncRNAs contain microRNAs within either their exonic or intronic sequence [34, 35]. Here, we have characterized several isoforms of a lncRNA that is overexpressed in leukemia and shows dynamic expression in hematopoietic development [16]. Expressed from a locus adjacent to genes important in lymphocyte development, BALR-6 itself is dynamically regulated during human B-cell development [36-38]. Our work significantly adds to the known repertoire of RNA molecules that are expressed from this locus, and several of these appear to be functional within a cellular context.

In this manuscript, we describe the cellular function of a second lncRNA that was discovered as being overexpressed in MLL-translocated B-ALL. In some ways, BALR-6 shows some similarities with the other lncRNA we studied, BALR-2 [16]. Indeed, knockdown of both lncRNAs led to decreased cell growth and increased apoptosis, and overexpression led to increased growth and a partial resistance to prednisolone treatment. These findings are not altogether surprising given that these lncRNAs may be contributing to the poor clinical behavior of an aggressive cytogenetic subtype of B-ALL [17]. However, there are important differences between these lncRNAs— the genomic locus for BALR-6 is more complex, there are multiple isoforms, and no comparable murine transcript is described in publically available databases. Nonetheless, we have obtained fragments of a low-expression transcript from murine hematopoietic cell lines that encoded portions homologous to human BALR-6. Further characterization of the murine transcripts will be the goal of future studies.

Significantly, our study is amongst the few characterizations of lncRNA dysregulation in the hematopoietic system [16, 39-41]. LncRNAs have been ascribed functions in lymphopoiesis,

myelopoiesis, and erythropoiesis [42-45]. Additionally, their differential expression has been described in peripheral T-cell subsets [46]. Here, we discovered the effect of BALR-6 overexpression on early hematopoietic progenitors in the marrow, including LSK cells, HSCs, and LMPPs. Constitutive expression of BALR-6 isoforms led to increased survival or proliferation of normally transient bone marrow progenitor cells. Furthermore, Hardy fractions showed a trend towards being increased when compared to control, particularly those that developmentally precede the large pre B-cell stage (fraction C', early pre-B). The relative percentages of more mature B-lineage cells downstream of these developmental stages are largely normal. Despite increased proportions of early progenitor cells, passage through a checkpoint (such as the pre-BCR checkpoint) may reduce cell numbers back to baseline. This suggests that the function of BALR-6 *in vivo* may be in directing differentiation and adequate lymphoid cell development. The upregulation of this lncRNA causes a survival or proliferative advantage, a hallmark of leukemogenesis. Coupling BALR-6 overexpression with an appropriate oncogenic co-stimulus may lead to full-blown leukemogenesis or enhancement thereof, and this is currently an active area of investigation in the laboratory.

In line with a function in promoting the survival of early hematopoietic progenitors, BALR-6 clearly affects proliferation in cell line experiments. Upon siRNA mediated knockdown, we saw reduced cell proliferation and increased cell death. We observed the opposite effect when we constitutively expressed BALR-6 in human and murine B-ALL cell lines. Moreover, similar mechanisms may be operant in B-ALL with MLL translocations, and loss-of-function experiments in primary patient samples and mouse models of MLL-driven leukemia are areas for further investigation.

Given prior reports of lncRNAs serving to regulate transcriptional complexes, our finding that BALR-6 knockdown causes changes in the SP1 transcriptome is compelling. SP1 is a transcriptional regulator that is associated with dysregulated cell cycle arrest in multiple

myeloma [47-49]. CREB1 is a well-known proto-oncogene that promotes cellular proliferation in hematopoietic cells [50, 51]. Here we demonstrate that SP1-mediated transcription at the CREB1 and p21 promoters are positively regulated by BALR-6, providing a putative mechanism for our observations of BALR-6's role in B-ALL.

Conclusions

In this study, we demonstrate that the MLL-AF4-dysregulated lncRNA, BALR-6, plays a role in cell survival and regulates hematopoietic progenitors. At the molecular level, BALR-6 regulates the transcriptome of B-ALL cell lines, likely through regulating SP1-mediated transcription. In summary, our study has several novel and unique findings that help uncover a role for a poorly understood class of molecules in a pathogenetic process. This will undoubtedly have impacts on our understanding of molecular biology within cancer cells.

Methods

Cloning and Cell culture

mmu-miR-155 formatted siRNAs were cloned into BamHI and XhoI sites in the pHAGE2-CMV-ZsGreen-WPRE vector using the strategy that we have previously described to generate knockdown vectors [16, 25, 26, 52]. Using the sequence information from 5' and 3' RACE products we cloned full length transcripts into an MSCV viral vector between the BamHI and XhoI sites, as described previously, and into a pHAGE6-UBC-ZsGreen-CMV-LNC (P6UZCL) variant of the third generation lentiviral vector system, between the NotI and BamHI sites [16, 52]. Primer sequences used are listed in Supplemental Table 1 or mentioned previously [16]. RS4;11 and MV4;11, (MLL-AF4-translocated; ATCC CRL-1873 and CRL-9591), Reh (TEL-AML1-translocated; CRL-8286), 697 (E2A-PBX1-translocated), Nalm-6 , 70Z/3 (ATCC TIB-158)

murine pre B-cell leukemic cell line, and the HEK 293T cell line (ATCC CRL-11268) were grown in their corresponding media at 37oC in a 5% CO2 incubator as previously described [16, 53].

Rapid Amplification of cDNA Ends (RACE)

To determine the 5' and 3' transcript ends of the lncRNAs, we performed RACE using First Choice RLM-RACE kit (Ambion). Using the sequence information from 5' and 3' RACE products, we cloned full length transcripts into P6UZCL, and into the MSCV viral vector. Primer sequences used and isoform sequences obtained are listed in Supplemental Table 1.

Transduction and sorting of cell lines

Lentiviruses and MSCV-based retroviruses were produced to generate knockdown constructs as previously described [16, 25, 26, 52]. In brief, 5.0 x 10^5 cells were spin-infected at 30^{o}C for 90 minutes in the presence polybrene (4 µg/mL). Transduced cell lines were sorted for high green expression using a BD FACSAriaII cell sorter, and analysis was performed using BD FACSDiva software.

Biological assays

For pharmaco-induced assays, cells were cultured at a concentration of 1.0 X10^{6} cells per mL and treated for 36 hours. I-BET151 was dissolved in dimethyl sulfoxide to desired concentrations. After treatment, cells were harvested for RNA extraction. For MTS proliferation assays, cells were cultured for at least 5 days before plating. Cells were plated at a density of 2,500 cells per 100 µl of media in each well of a 96 well plate. Reagents were added according to the manufacturer's instructions (Promega CellTiter 96 Aqueous Non-Radioactive Cell Proliferation Assay kit) and cells were incubated at 37°C, 5% CO2 for 4 hours before absorbance was measured at 490 nm. For apoptosis assays, cells were plated at 5.0 X 10^5

cells/mL for 24 hours with or without prednisolone treatment. Prednisolone (TCI America) was dissolved in dimethyl sulfoxide to desired concentrations. Cells were harvested after 24 hours and stained with APC-tagged AnnexinV. For cell cycle analysis, cells were synchronized by serum starvation for 12 hours (human cell lines) or 4 hours (murine cell lines) then plated at 5.0 X 10^5 cells/mL and incubated at 37°C, 5% CO2 for 24 hours. Cells were harvested, fixed with EtOH and then stained with propidium iodide. AnnexinV stained and PI stained samples were analyzed using a BD FACS HTLSRII flow cytometer and further analysis was performed using FlowJo.

Luciferase Assays

Promoter sequences for CREB1 and p21 were cloned upstream of synthetic firefly luciferase (luc2p) in the pGL4.11 vector. Renilla luciferase is expressed in the pGL4.75 vector downstream of the PGK promoter. HEK 293T cells were transfected with the pGL4.75 and pGL4.11 containing reporter vectors at a 1:20 ratio (5ng: 100ng), along with a combination of MSCV vector (empty, Isoform-1, or Isoform-3) and pCMV3 (empty or SP1-HA, Sino Biological Inc.) vector at a 1:1 ratio (200ng:200ng). For the last condition SP1, Isoform1 and Isoform 3 were transfected together at a ratio of 2:1:1 (200ng:100ng:100ng). Co-transfections were performed with BioT (Bioland Scientific LLC) in 24 well plates as per the manufacturer's instructions. Cells were lysed after 32 hours and supernatant lysate was collected as per manufacturer's instructions (Promega). The dual luciferase assay kit (Promega) was used as substrates for Renilla and firefly luciferase activity. Luminescence was measured on a Glomax-Multi Jr (Promega). The ratio of firefly to Renilla luciferase activity was calculated for all samples. The luminescence for the MSCV empty vector with pCMV3 empty vector, was used as a normalization control.

qRT-PCR and PCR

RNA from cell lines was reverse transcribed using qScript (Quantas Biosciences). Real Time quantitative PCR was performed with the StepOnePlus Real-Time PCR System (Applied Biosystems) using PerfeCTa SYBR Green FastMix reagent (Quantas Biosciences). cDNA from mice samples was amplified using KOD Master Mix (EMD Millipore) and ran on a 1.2% agarose gel stained with ethidium bromide. Primer sequences used are listed in Supplemental Table 1.

Northern blot

Total RNA was separated on a 1.2 % (w/v) formaldehyde agarose gel and then blotted onto Hybond N+ nylon membranes (Amersham Biosciences) by semi-dry transfer (Bio-Rad, Trans-Blot SD Semi-Dry Transfer Cell). DNA probes were ordered from Integrated DNA Technologies (IDT, San Diego, CA) with digoxigenin incorporated at 3'end. For ACTIN we used the RNA probe provided in the DIG Northern Starter Kit (Roche). Membranes were hybridized overnight using ULTRAhyb-Oligo Buffer (Ambion) at 37^{o}C or 42^{o}C with probes. Visualization was done by X-Ray film using CDP-Star reagents (Roche). X-Ray film was scanned and saved as jpeg files. Brightness and contrast was increased by 20% for ease of visualization.

Data Sources

Human genome assembly GRCh37/hg19 and the mouse genome assembly GRCm38/mm10 were used. Methylation patterns for the four cell lines were obtained from Chip-Seq data available in the UCSC genome browser generated by the Broad/ENCODE group [18-20]. Peak viewing range set from 1 to 50 for H3K4m3 modifications, and 1 to 15 for H3K36m3 modifications. Alternative splice form information was obtained from the Swiss Institute of Bioinformatics, via UCSC Genome Browser [21]. Genome alignments of RefSeq transcripts

from human, mouse, and other vertebrates, GenBank mRNAs and ESTs, as well as PhastCons scores were obtained from the UCSC Genome Browser [22].

Microarray data analysis

Microarray data was generated from samples of 3 different transduced RS4;11 cell lines with siRNAs against BALR-6, or the control empty vector. Samples were hybridized at the UCLA Clinical Microarray Core facility using Affymetrix HG-U133_Plus_2 microarray. The Affymetrix raw data files (.cel files) were loaded into the R program for quality control analysis. Additionally, raw hybridization intensities were normalized using the MAS5 method with the affy package in R. Normalized values were sorted by detection p-value < 0.05. Differential expression analysis was performed using unpaired Bayesian comparison model (CyberT Website) [30, 31]. Data was then sorted for genes with a posterior probability of deferential expression (PPDE) > 95% and a fold change > 1.5. Analysis of differentially expressed genes was carried out using the WEB-based GEne SeT AnaLysis Toolkit (WebGESTALT, http://bioinfo.vanderbilt.edu/webgestalt/) [32, 33]. This online tool uses information from different public data sources for enrichment analysis, including the Gene Ontology data base, and GLAD4U. A second (validation) microarray was carried out, as described above, with technical duplicates for RS4;11 cell lines transduced with siRNA2 or the empty vector. For differential analysis the raw data files were loaded into the R environment and analyzed using the R library of Linear Models for Microarray Data (LIMMA). Pairwise comparison and eBayes fit was carried out. Data was then sorted for genes with a p value < 0.05. Further analysis was done as described above, using WebGESTALT.

Mice and bone marrow transplantation

Mice were housed under pathogen free conditions at the University of California, Los Angeles (UCLA). Donor mice were injected intraperitoneally with 200 mg/kg of 5-fluorouracil. After 5 days the mice were sacrificed and the bone marrow was collected under sterile conditions and plated in media enriched with IL-3, IL-6 and mSCF (Gibco). 24 hours after plating, the bone marrow was spin infected twice, at 30^{o}C for 90 minutes in the presence polybrene (4 μg/mL), with retroviruses expressing the empty MSCV vector or BALR-6 Isoform 3. Recipient mice were lethally irradiated and injected with donor bone marrow 6 hours after irradiation. 8 mice were used per group. One mouse in the ISO3 group died due to engraftment failure after 2 weeks post injection. These mice were bled at 8, 12, and 16 weeks post bone marrow injection. At 16 weeks the mice were sacrificed for full analysis. For statistical analysis, one mouse was excluded due to low eGFP expression. This experiment was repeated, and had similar results. All animal studies were approved by the UCLA Animal Research Committee (ARC).

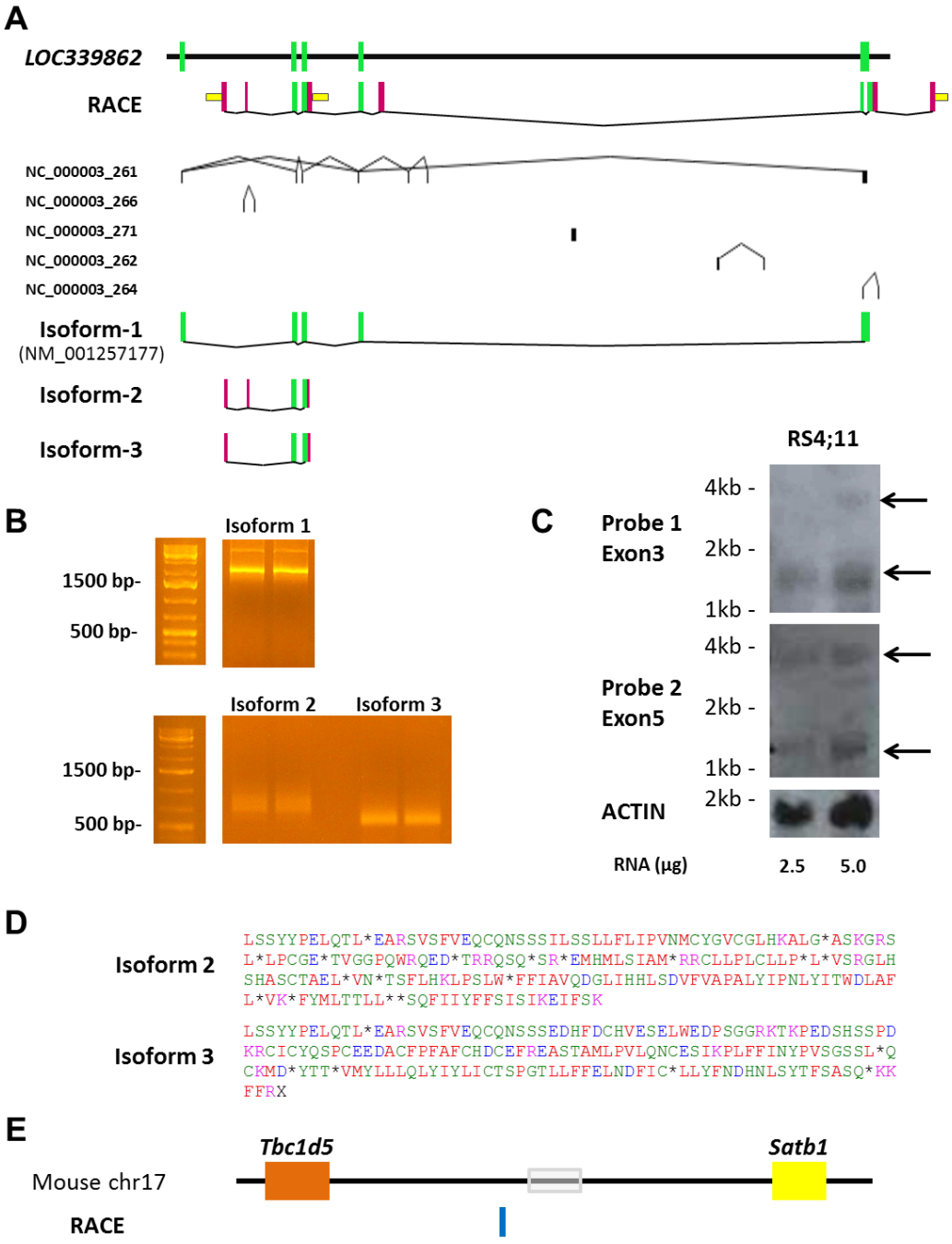
Flow cytometry of samples

At 16 weeks post bone marrow transplant, blood, bone marrow, thymus, and spleen were collected from the mice under sterile conditions [53]. Single cell suspensions were lysed in red blood cell lysis buffer. Fluorochrome conjugated antibodies were used for staining (antibodies obtained from eBiosciences, and Biolegend). Cells were stained with surface marker antibodies for 30 minutes at 4^{o}C, washed twice with PBS, and finally fixed with 1% PFA. Flow cytometry was performed at the UCLA Jonsson Comprehensive Cancer Center (JCCC) and at the BROAD Stem Cell Research Flow Core. Analysis was performed using FlowJo software. The lists of antibodies used and gating schematics are provided in Supplemental Table 2. Normal adult human bone marrow was obtained commercially from healthy adults (All Cells, Inc) as previously described [51]. CD34 enrichment from bone marrow was performed using the

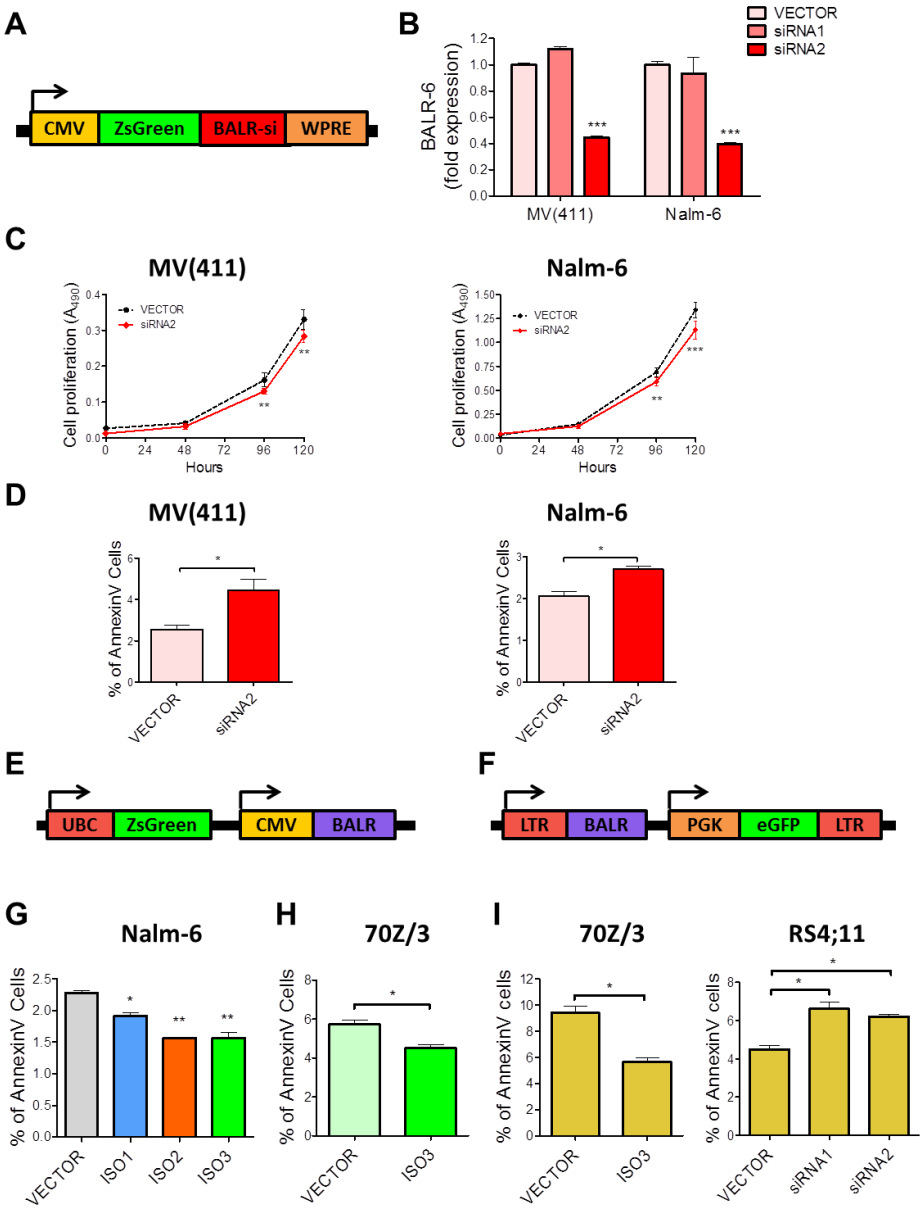
magnetic activated cell sorting (MACS) system (Miltenyi Biotec, San Diego, CA) prior to isolation of CD34+ subsets by flow cytometry. Bone marrow CD34 selected cells were incubated with the following antibodies: CD34-APC-Cy7 (581), CD38-APC (HIT2), CD10-PE-Cy7 (HI10a), CD20-FITC (2H7), CD45RA PerCPCy5.5 (HI-100), IgM Percpcy5.5 (NHM-88) (all from Biolegend, San Diego, CA), as well as the following FITC-labeled lineage depletion antibodies: CD3 (UCHT1), CD14 (M2E2), CD15 (HI-95), CD19 (SJ25C1), CD56 (B159), and CD235a (GA-R2)(Becton Dickinson, San Jose, CA). CD19 was not included in the lineage depletion cocktail used for sorting the progenitor B population. The following immunophenotypic definitions were used to isolate progenitors from bone marrow CD34 selected cells: CD34+CD38-lin- (HSC), CD34+CD10+CD45RA+lin- (CLP) and CD34+CD19+lin- (progenitor B); CD34-CD10+CD19+IgM-CD20- (precursor B), CD34-CD19+IgM+CD20+ (Immature B). All populations were purified using fluorescence-activated cell sorting on a FACSAria (355, 405, 488, 561 and 633 nm lasers) (BD Immunocytometry Systems).

Abbreviations

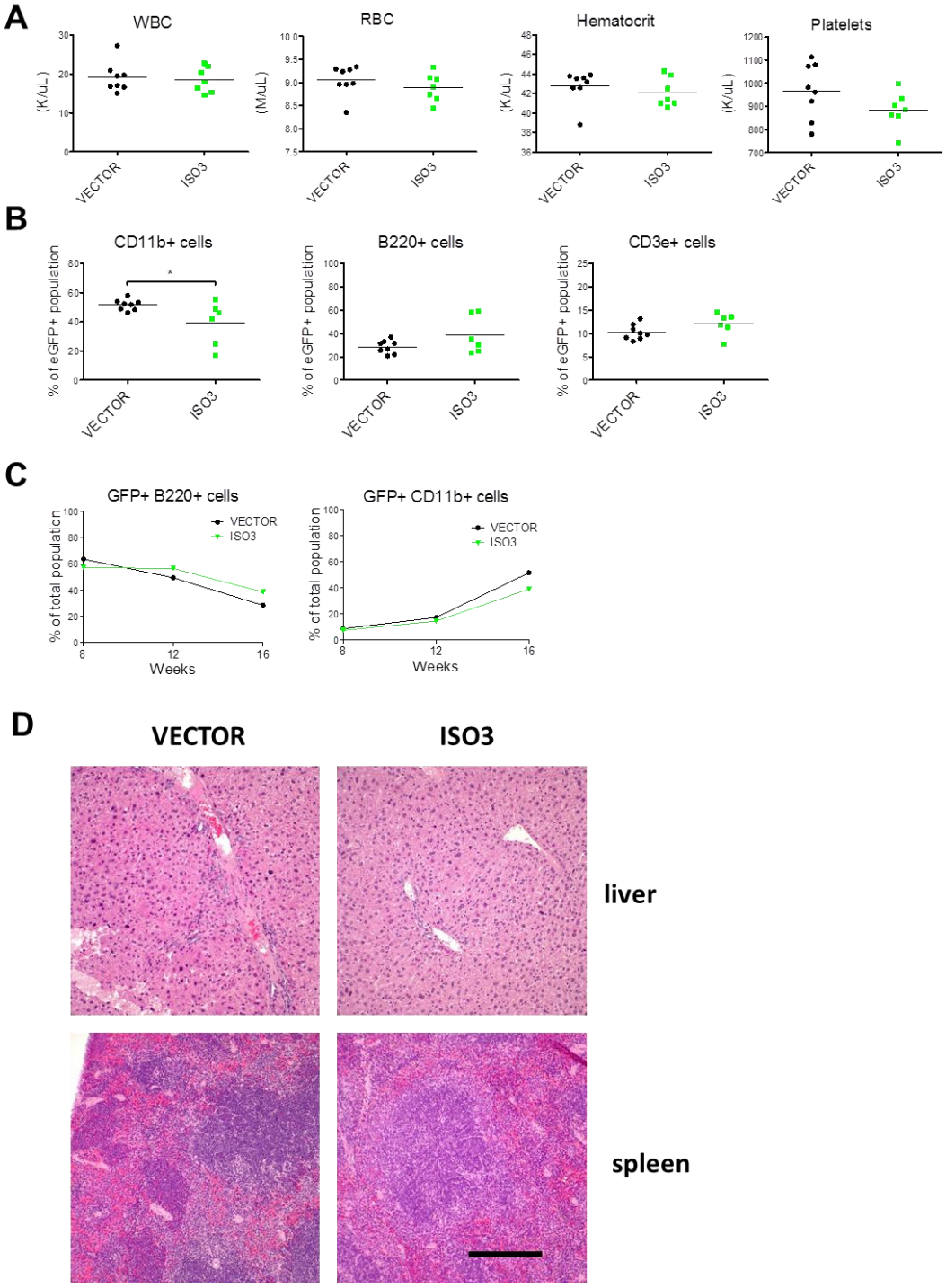
BALR, B-ALL associated long RNA; ORF, open reading frame; chr, chromosome; HSC, hematopoietic stem cell; CLP, common lymphoid progenitor; LMPP, lymphoid primed multipotent progenitor; LSK, lineage- $Sca1^{+}$ $c\text{-}Kit^{+}$; pro-B, progenitor B; pre-B, precursor B; DMSO, dimethyl sulfoxide; UBC, ubiquitin C promoter; ZsGreen, Zoanthus green fluorescent protein; CMV, cytomegalovirus promoter; LTR, long terminal repeats; PGK, phosphoglycerate kinase promoter; eGFP, enhanced green fluorescent protein; ISO1, Isoform 1; ISO2, Isoform 2; ISO3, Isoform 3; PFA, paraformaldehyde; PPDE, posterior probability of differential expression luc2p, synthetic firefly luciferase.



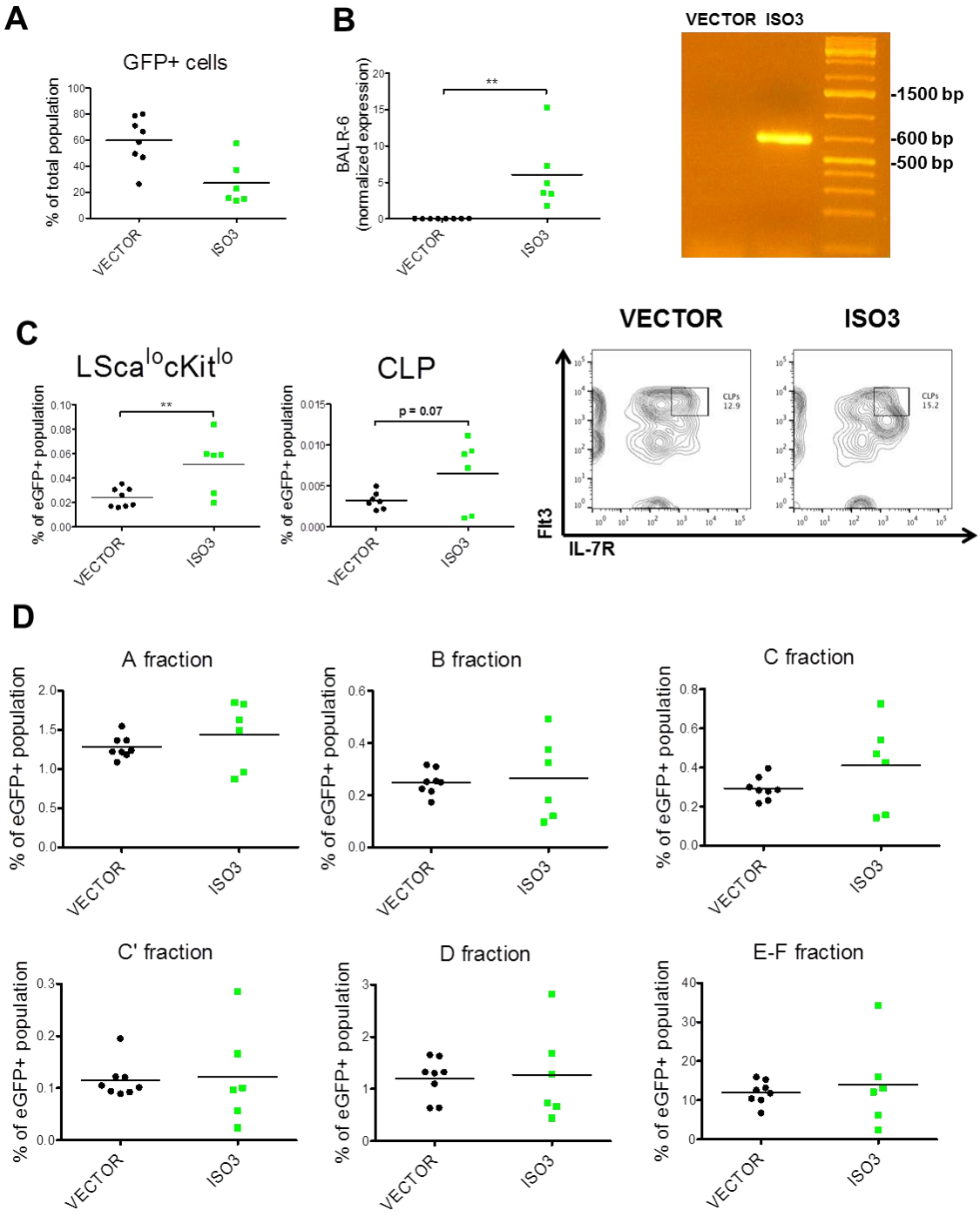
Supplemental Figure 1. BALR-6 locus encodes numerous alternative splice forms.
(**A**) Top: Diagram of RACE products obtained from *LOC339862*. 5' and 3' RACE primers are shown in yellow, with the newly discovered exons shown in magenta, as seen in Figure 1C. Known annotated exons are shown in green. Middle: Alternative splicing graph from the Swiss Institute of Bioinformatics of the predicted alternative splicing transcripts shown in the SIB Genes track. Blocks represent exons, lines indicate introns. Bottom: Schematic depiction of BALR-6 isoforms cloned from RACE sequences. Annotated exons in green, unannotated in magenta. (**B**) Gel confirmation of the isoforms cloned, including the annotated mRNA sequence (Isoform 1). (**C**) Northern blot of endogenous levels of two BALR-6 isoforms in RS4;11 cells. (**D**) EMBOSS analysis of the new isoforms confirmed lack of open reading frames, and lack of translation initiation sites. (**E**) Diagram of RACE products obtained from mouse cell lines with homology to BALR-6. chr, chromosome.



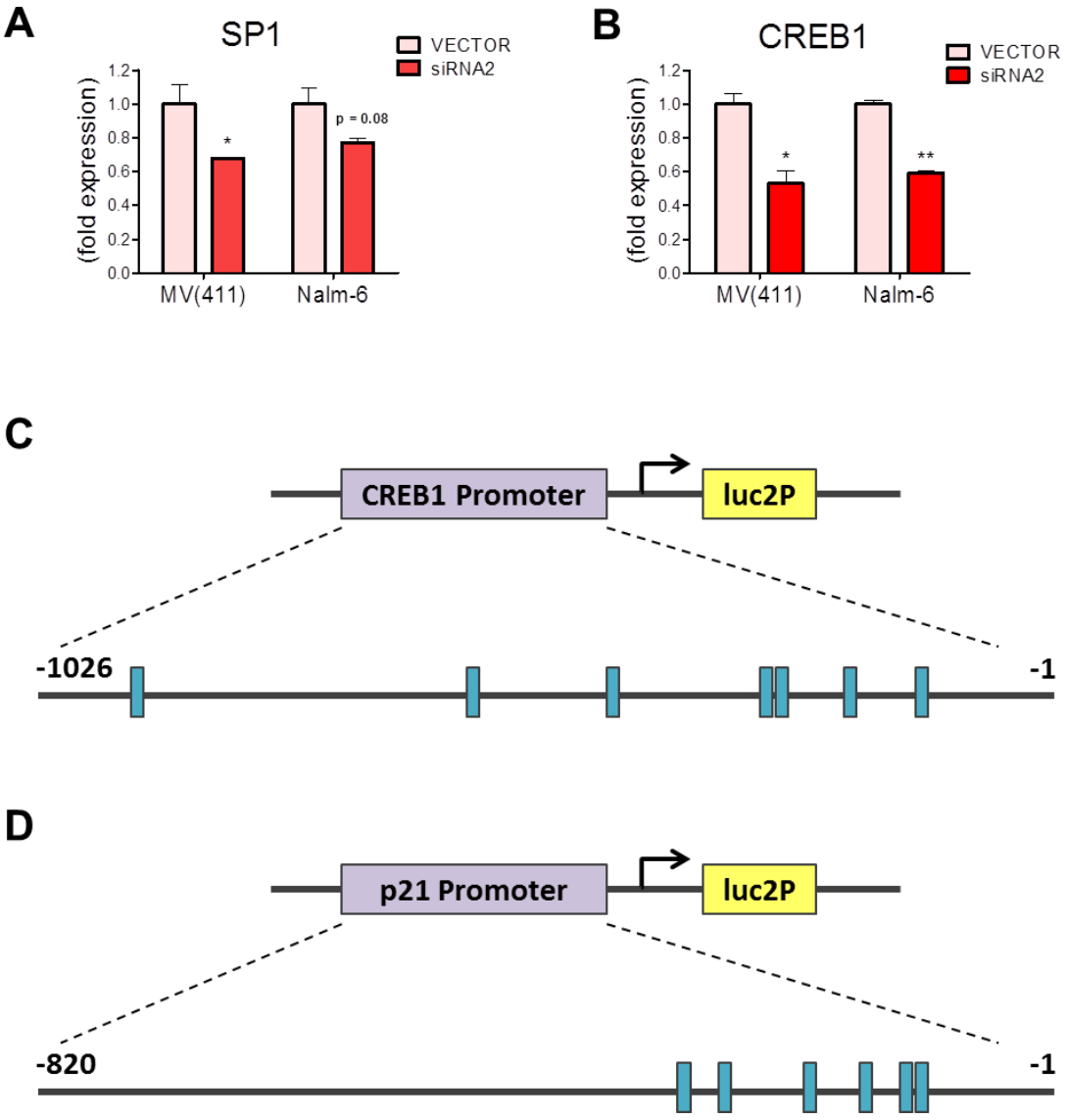
Supplemental Figure 2. Knockdown and overexpression of full length BALR-6 isoforms in mammalian cell lines. (**A**) Schematic representation of mmu-miR-115 knockdown expression cassette. (**B**) Successful knockdown of BALR-6 using siRNA2 in MV(411), and Nalm-6 cells. (**C**) Decreased cell proliferation in transduced MV(411), and Nalm-6 lines as measured by MTS assay. (**D**) Increased apoptosis at basal levels in MV(411), and Nalm-6 stable lines as measured by AnnexinV staining. (**E-F**) Schematic representation of dual promoter phage (**E**) and MSCV (**F**) expression cassettes. (**G**) AnnexinV staining showed that Nalm-6 stably transduced with BALR-6 isoforms, had lower number of apoptotic cells at basal level. (**H**) 70Z/3 cells overexpressing BALR-6 Isoform 3 had fewer apoptotic cells at basal level, as analyzed by AnnexinV staining. (**I**) 70Z/3 cells stably transduced with BALR-6 Isoform 3, resulted in reduction of apoptosis upon treatment with 250 µg/mL prednisolone for 6 hours. The opposite effect was seen with RS4;11 cells with siRNA mediated knockdown of BALR-6 and treated with 250 µg/mL prednisolone for 24 hours. Evaluations were made using a two-tailed T-test, $p<0.05$ (*); $p<0.005$ (**); $p<0.0005$(***).UBC, ubiquitin C promoter; ZsGreen, Zoanthus green fluorescent protein; CMV, cytomegalovirus promoter; LTR, long terminal repeats; PGK, phosphoglycerate kinase promoter; eGFP, enhanced green fluorescent protein.



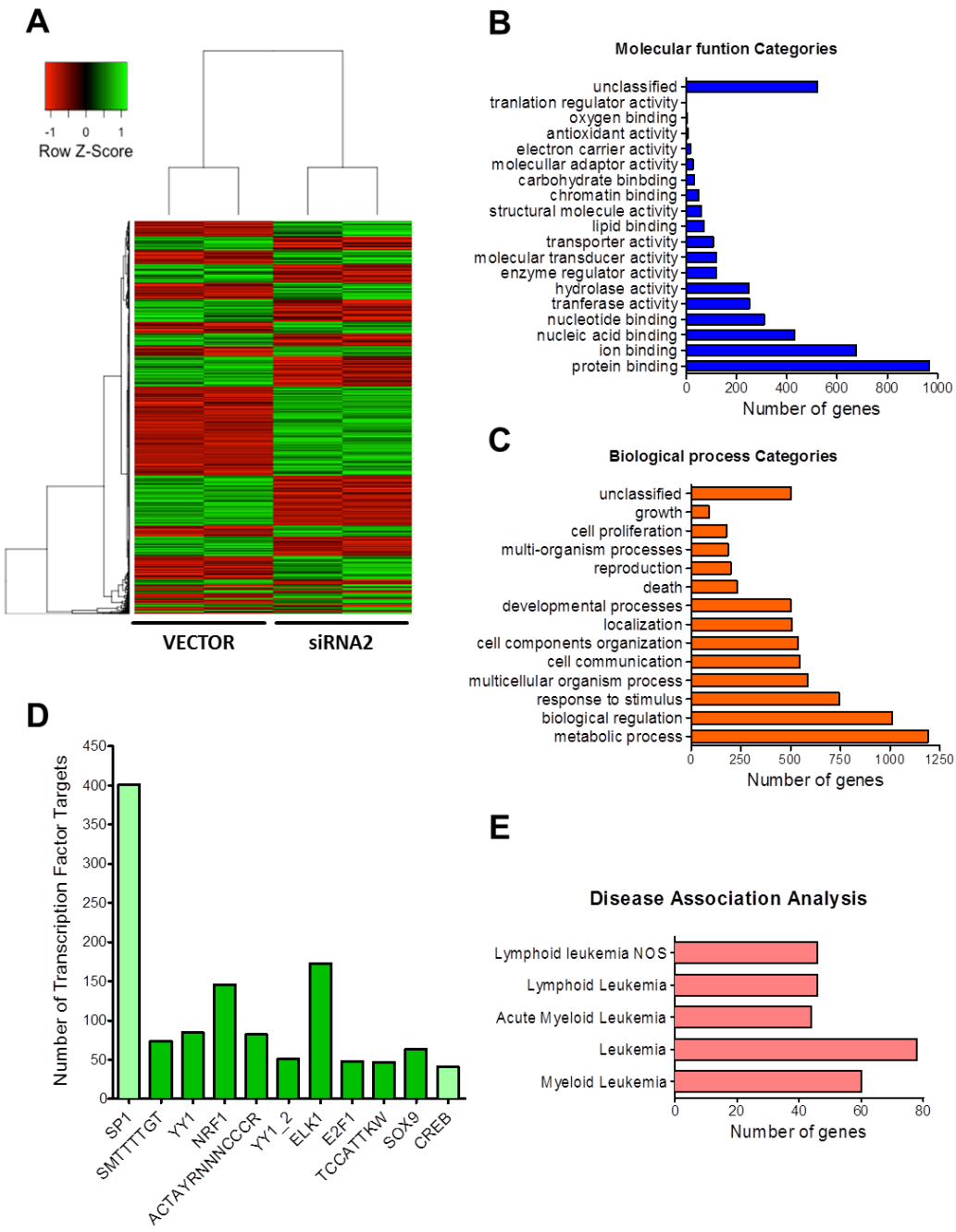
Supplemental Figure 3. Constitutive expression of BALR-6 in mice periphery.
(**A**) Peripheral white and red blood, hematocrit, and platelet cell counts. (**B**) Levels of B-cells (B220+), T-cells (CD3e+) and Myeloid cells (CD11b+) in the GFP+ compartment of the peripheral blood at 16 weeks. (**C**) Average levels of GFP+ B cells (B220+) and GFP+ Myeloid (CD11b+) cells in the peripheral blood throughout the experiment. Number of mice used in this analysis: VECTOR, n=8; ISO3, n=6. (**D**) Hematoxylin and eosin stained liver and spleen samples from bone marrow transfer mice. Scale bar, 200 mm. ISO3, Isoform 3. Evaluations made using a two-tailed T-test, $p<0.05$ (*).



Supplemental Figure 4. Elevated levels of immature B cell populations in mice with BALR-6 overexpression. (**A**) Levels of GFP+ cells in the bone marrow of experimental mice. (**B**) Expression levels of BALR-6 Isoform 3 in the experimental mice by qRT-PCR. Normalized to L32. Gel of cDNA PCR, obtained from bone marrow samples, confirming expression of full length transcript, shown to the right. (**C**) Quantitation of LSca^{lo}c-Kit^{lo} cells, and CLP cells in the GFP+ compartment of the experimental mice. Representative FACS plots of the population gating, shown to the right. (**D**) Percentage of cells in the Hardy fractions from the GFP+ compartment of experimental mice. Number of mice used in this analysis: VECTOR, n=8; ISO3, n=6. ISO3, Isoform 3. Evaluations made using a two-tailed T-test, p<0.005 (**).



Supplemental Figure 5. SP1 targets in siRNA mediated knockdown cell lines. (**A-B**) Transcript levels of SP1 (**A**) and CREB1 (**B**) in MV(411) and Nalm-6 knockdown cells. qRT-PCR quantitation of expression, normalized with ACTIN. Only expression levels upon siRNA2 mediated knockdown, which was successful, are shown. (**C-D**) Schematic representation of CREB1 (**C**) and p21 (**D**) promoter sequences cloned into the pGL4.11 luciferase expression vector. Promoter sequence distance shown in relation to the luc2p start codon. SP1 binding sites shown as blue boxes. luc2p, synthetic firefly luciferase.



Supplemental Figure 6. Confirmation of global differential expression findings seen in initial microarray. (**A**) Hierarchical gene clustering of differentially expressed genes in validation microarray upon siRNA2 mediated knockdown of BALR-6 in RS4;11 cells, p-value ≤0.05. Technical replicates of samples shown. (**B-C**) Bar graphs of GO Slim classification enrichment analysis of differentially expressed genes by molecular function (**B**) and biological processes (**C**) as analyzed by WebGESTALT. Proportions are highly similar to initial microarray. (**D**) Enrichment analysis of transcription factor targets. Top ten transcription factors with a p-value < 0.02 are shown. In addition, CREB1 was shown as a significantly enriched for its targets with p-value = 0.04. For unknown transcription factors, transcription site sequence is shown. SP1 (shown in light green) had the most dysregulated targets. (**E**) Disease association analysis by GLAD4U, revealed a significant enrichment of genes solely associated to leukemic diseases, p-value ≤ 0.05.

Supplemental Table 1: Primers and RACE sequences for BALR-6

RT-qPCR primers		
BALR-6 Set 1	FOW	5' CGTGTGCTGGGGAAGGCACTG 3'
	REV	5' CCAGGCTCAGAGCAACACAGGGA 3'
BALR-6 Set 2	FOW	5' GATCACTTTGATTGCCATGTGGA 3'
	REV	5' ATCTCTTATCTGGACTACGGTGAC 3'
ACTIN	FOW	5' CATGTACGTTGCTATCCAGGC 3'
	REV	5' CTCCTTAATGTCACGCACGAT 3'
SP1	FOW	5' TGGCAGCAGTACCAATGGC 3'
	REV	5' CCAGGTAGTCCTGTCAGAACTT 3'
CREB1	FOW	5' TTAACCATGACCAATGCAGCA 3'
	REV	5' TGGTATGTTTGTACGTCTCCAGA 3'
Sp1	FOW	5' AGGGTCCGAGTCAGTCAGG 3'
	REV	5' CTCGCTGCCATTGGTACTGTT 3'
Creb1	FOW	5' TGTAGTTTGACGCGGTGTGT 3'
	REV	5' GCTGGTTGTCTGCTCCAGAT 3'
L32 (mouse)	FOW	5' AAGCGAAACTGGCGGAAAC 3'
	REV	5' TAACCGATGTTGGGCATCAG 3'
Actin	FOW	5' GCTACAGCTTCACCACCACA 3'
	REV	5' GGGGTGTTGAAGGTCTCAAA 3'
Cloning primers for P6UZCL		
NotI site-1	FOW	5'ATGGGCTTAGCTGCGGCCGCTTTCTTCATACTATCCAGAGCTCCAAA3'
BamHI site-1	REV	5'ATGGCAATTATCGGATCCTTTTTTTTTTTTCGAAAAAATTTCTTTTATTGAGATGCT3'
NotI site-2	FOW	5'ATGGCA ATTGCGGCCGCCCACGCGTCCGGGACTGAGCA 3'
BamHI site-2	REV	5'ATGGGCTGATGATCATTTTTTTTTGGTTCATAGAAAGTATTTTCTTCTAGAGTCTC3'
Cloning primers for MSCV		
HindIII_eGFP	FOW	5'ATGGGCTTAGCTAAGCTTATGGTGAGCAAGGGCGAGGAGC3'
DraIII_eGFP	REV	5'ATGGCAATTATCCACCTGGTGTTTACTTGTACAGCTCGTCCATGCCGA3'
BclI_BglII site	FOW	5' ATGGGCTGAGGATCTCCACGCGTCCGGGACTGAGCA 3'
XhoI site-1	REV	5'ATGGCAATTCTCGAGTTTTTTTTTGGTTCATAGAAAGTATTTTCTTCTAGAGTCTC3'
BamHI_BglII site	FOW	5'ATGGGCTTAGCTGGATCCTTTCTTCATACTATCCAGAGCTCCAAA3'
XhoI site-2	REV	5'ATGGCAATTATCCTCGAGTTTTTTTTTTTTCGAAAAAATTTCTTTTATTGAGATGCT3'
Sequencing primers for MSCV		
Upstream_Puro	FOW	5' GCTGTTCTCCTCTTCCTCATCTCC 3'
Upstream_eGFP	FOW	5' CTTTATCCAGCCCTCACTCCTTCTCT 3'

RACE primers		
3-RACE_6	FOW	5' CCATGTGAAGAAGATGCTGGCTTC 3'
3-RACE_7	FOW	5' TAGGAAGCCAGAAGCGTCTCCTTT 3'
3-RACE_8	FOW	5' GGAGGCAGGAAGACTAAACCAGAA 3'
5-RACE_2	REV	5' CTCGCGAAACTCACAATCATGGCA 3'
5-RACE_4	REV	5' ATCTTCCATGTGCATGTGGCTGCA 3'
Northern probes		
BALR-6 probe 1		5'GGGCACAGAGTGTTGCATGCTCATTTCTGTTGATTTTTAATTAGCAGTAATTCATTT/3DiG_N/3'
BALR-6 probe 2		5'CTGGAAATCTAGGATCAGGACTAGCCTAAATTAGTAGATCTATGTGATAGTATATTGGTA/3DiG_N/3'
mmu-miR-155 formatted siRNA oligos		
siRNA1		5'GAAGGCTGTATGCTGGTGAACATACCACTTACCATTGTTTTGGCCACTGACTGACAATGGTAAGGTATGTTCACCAGGACACAAGGCCTG3'
siRNA2		5'GAAGGCTGTATGCTGGACTTCTGCACACCATGCCTGGTTTTGGCCACTGACTGACCAGGCATGGTGCAGAAGTCCAGGACACAAGGCCTG3'
BALR-6 sequences		
Isoform-2		5'TTTCTTCATACTATCCAGAGCTCCAAACTTTGTAGGAAGCCAGAAGCGTCTCCTTTGTTGAACAGTGCCAAAATAGCAGCTCTATCCTTTCCTCTCTCCTCTTTCTGATTCCAGTCAATATGTGTTATGGAGTCTGTGGTCTCCACAAGGCCTTGGGATAGGCATCCAAAGGAAGATCACTTTGATTGCCATGTGGAGAGTGAACTGTGGGAGGACCCCAGTGGAGGCAGGAAGACTAAACCAGAAGACAGTCACAGTAGTCCAGATAAGAGATGCATATGTTATCAATCGCCATGTGAAGAAGATGCTTGCTTCCCCTTTGCCTTCTGCCATGATTGTGAGTTTCGCGAGGCCTCCACAGCCATGCTTCCTGTACTGCAGAACTGTGAGTCAATTAAACCTCTTTTCTTCATAAATTACCCAGTCTCTGGTAGTTCTTTATAGCAGTGCAAGATGGACTAATACACCACCTAAGTGATGTATTTGTTGCTCCAGCTCTATATATACCTAATTTGTACATCACCTGGGACCTTGCTTTTCTTTGAGTTAAATGATTTTATATGTTAACTACTCTACTTTAATGATCACAATTTATCATATACTTTTTCAGCATCTCAATAAAAGAAATTTTTTCGAAA3'
Isoform-3		5'TTTCTTCATACTATCCAGAGCTCCAAACTTTGTAGGAAGCCAGAAGCGTCTCCTTTGTTGAACAGTGCCAAAATAGCAGCTCTGAAGATCACTTTGATTGCCATGTGGAGAGTGAACTGTGGGAGGACCCCAGTGGAGGCAGGAAGACTAAACCAGAAGACAGTCACAGTAGTCCAGATAAGAGATGCATATGTTATCAATCGCCATGTGAAGAAGATGCTTGCTTCCCCTTTGCCTTCTGCCATGATTGTGAGTTTCGCGAGGCCTCCACAGCCATGCTTCCTGTACTGCAGAACTGTGAGTCAATTAAACCTCTTTTCTTCATAAATTACCCAGTCTCTGGTAGTTCTTTATAGCAGTGCAAGATGGACTAATACACCACCTAAGTGATGTATTTGTTGCTCCAGCTCTATATATACCTAATTTGTACATCACCTGGGACCTTGCTTTTCTTTGAGTTAAATGATTTTATATGTTAACTACTCTACTTTAATGATCACAATTTATCATATACTTTTTCAGCATCTCAATAAAAGAAATTTTTTCGAAA3'

Supplemental Table 2: Antibodies used for flow cytometry analysis, and population gating schematics.

Marker	**Fluorochrome**
CD3e	PE
CD11b	PE-Cy7
B220	PerCP-Cy 5.5
CD117	APC-Cy7
Sca1	PerCP-Cy 5.5
CD135	APC
CD127	PE-Cy7
CD150	PE
IgM	PE
CD43	APC
CD24	PE-Cy7
Ly51	APC-Cy7
For lineage negative staining	
Biotin	CD3e, CD4, CD8, B220, NK1.1, Ter119, TCR beta, TCR gamma-delta
Streptavidin	eFluor 450 (pacific Blue)
Population	**Defined markers**
HSC	Lin- CD117 hi Sca1 hi CD150++
LMPP	Lin- CD117 hi Sca1 hi CD135+ CD127-
CLP	Lin- CD117 lo Sca1 lo CD135+ CD127+
A	B220+ IgM- CD43+ CD24- Ly51-
B	B220+ IgM- CD43+ CD24+ Ly51-
C	B220+ IgM- CD43+ CD24+ Ly51+
D	B220+ IgM- CD43-
E and F	B220+ IgM+

All antibodies were procured from eBiosciences (San Diego, CA) or Biolegend (San Diego, CA).

References

1. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, et al: RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 2007, 316:1484-1488.

2. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al: Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009, 458:223-227.

3. Niazi F, Valadkhan S: Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3' UTRs. *RNA* 2012, 18:825-843.

4. Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP: Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 2011, 147:1537-1550.

5. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY: Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 2007, 129:1311-1323.

6. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, et al: The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* 2010, 39:925-938.

7. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I: A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 2011, 147:358-369.

8. Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L, Santoro C, et al: Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* 2012, 491:454-457.

9. Gong C, Maquat LE: lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 2011, 470:284-288.

10. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, et al: Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 2010, 464:1071-1076.

11. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, Laxman B, Asangani IA, Grasso CS, Kominsky HD, et al: Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat Biotechnol* 2011, 29:742-749.

12. Gutschner T, Hammerle M, Eissmann M, Hsu J, Kim Y, Hung G, Revenko A, Arun G, Stentrup M, Gross M, et al: The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 2013, 73:1180-1189.

13. Isin M, Dalay N: LncRNAs and neoplasia. *Clin Chim Acta* 2015, 444:280-288.

14. Mullighan CG, Downing JR: Global genomic characterization of acute lymphoblastic leukemia. *Semin Hematol* 2009, 46:3-15.

15. Mullighan CG: Molecular genetics of B-precursor acute lymphoblastic leukemia. *J Clin Invest* 2012, 122:3407-3415.

16. Fernando TR, Rodriguez-Malave NI, Waters EV, Yan W, Casero D, Basso G, Pigazzi M, Rao DS: LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia. *Mol Cancer Res* 2015, 13:839-851.

17. Pui CH, Behm FG, Downing JR, Hancock ML, Shurtleff SA, Ribeiro RC, Head DR, Mahmoud HH, Sandlund JT, Furman WL, et al.: 11q23/MLL rearrangement confers a poor prognosis in infants with acute lymphoblastic leukemia. *J Clin Oncol* 1994, 12:909-915.

18. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR, et al: Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 2005, 120:169-181.

19. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, et al: Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011, 473:43-49.

20. Ram O, Goren A, Amit I, Shoresh N, Yosef N, Ernst J, Kellis M, Gymrek M, Issner R, Coyne M, et al: Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells. *Cell* 2011, 147:1628-1639.

21. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL: GenBank: update. *Nucleic Acids Res* 2004, 32:D23-26.

22. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al: Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005, 15:1034-1050.

23. Zhou Y, Zhong Y, Wang Y, Zhang X, Batista DL, Gejman R, Ansell PJ, Zhao J, Weng C, Klibanski A: Activation of p53 by MEG3 non-coding RNA. *J Biol Chem* 2007, 282:24731-24742.

24. Dawson MA, Prinjha RK, Dittmann A, Giotopoulos G, Bantscheff M, Chan WI, Robson SC, Chung CW, Hopf C, Savitski MM, et al: Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* 2011, 478:529-533.

25. O'Connell RM, Chaudhuri AA, Rao DS, Baltimore D: Inositol phosphatase SHIP1 is a primary target of miR-155. *Proc Natl Acad Sci U S A* 2009, 106:7113-7118.

26. Rao DS, O'Connell RM, Chaudhuri AA, Garcia-Flores Y, Geiger TL, Baltimore D: MicroRNA-34a perturbs B lymphocyte development by repressing the forkhead box transcription factor Foxp1. *Immunity* 2010, 33:48-59.

27. Hardy RR, Hayakawa K: B cell development pathways. *Annu Rev Immunol* 2001, 19:595-621.

28. Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM: Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* 2010, 38:662-674.

29. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, et al: Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* 2009, 106:11667-11672.

30. Baldi P, Long AD: A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. *Bioinformatics* 2001, 17:509-519.

31. Kayala MA, Baldi P: Cyber-T web server: differential analysis of high-throughput data. *Nucleic Acids Res* 2012, 40:W553-559.

32. Zhang B, Kirov S, Snoddy J: WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 2005, 33:W741-748.

33. Wang J, Duncan D, Shi Z, Zhang B: WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res* 2013, 41:W77-83.

34. Lerner M, Harada M, Loven J, Castro J, Davis Z, Oscier D, Henriksson M, Sangfelt O, Grander D, Corcoran MM: DLEU2, frequently deleted in malignancy, functions as a critical host gene of the cell cycle inhibitory microRNAs miR-15a and miR-16-1. *Exp Cell Res* 2009, 315:2941-2952.

35. Cai X, Cullen BR: The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* 2007, 13:313-316.

36. Dickinson LA, Joh T, Kohwi Y, Kohwi-Shigematsu T: A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* 1992, 70:631-645.

37. Will B, Vogler TO, Bartholdy B, Garrett-Bakelman F, Mayer J, Barreyro L, Pandolfi A, Todorova TI, Okoye-Okafor UC, Stanley RF, et al: Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. *Nat Immunol* 2013, 14:437-445.

38. Kohwi-Shigematsu T, Poterlowicz K, Ordinario E, Han HJ, Botchkarev VA, Kohwi Y: Genome organizing function of SATB1 in tumor progression. *Semin Cancer Biol* 2013, 23:72-79.

39. Fang K, Han BW, Chen ZH, Lin KY, Zeng CW, Li XJ, Li JH, Luo XQ, Chen YQ: A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: biology and epigenetic target. *Hum Mol Genet* 2014, 23:3278-3288.

40. Garitano-Trojaola A, Agirre X, Prosper F, Fortes P: Long non-coding RNAs in haematological malignancies. *Int J Mol Sci* 2013, 14:15386-15422.

41. Alvarez-Dominguez JR, Hu W, Gromatzky AA, Lodish HF: Long noncoding RNAs during normal and malignant hematopoiesis. *International Journal of Hematology* 2014, 99:531-541.

42. Venkatraman A, He XC, Thorvaldsen JL, Sugimura R, Perry JM, Tao F, Zhao M, Christenson MK, Sanchez R, Yu JY, et al: Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature* 2013, 500:345-349.

43. Aoki K, Harashima A, Sano M, Yokoi T, Nakamura S, Kibata M, Hirose T: A thymus-specific noncoding RNA, Thy-ncR1, is a cytoplasmic riboregulator of MFAP4 mRNA in immature T-cell lines. *BMC Mol Biol* 2010, 11:99.

44. Zhang X, Lian Z, Padden C, Gerstein MB, Rozowsky J, Snyder M, Gingeras TR, Kapranov P, Weissman SM, Newburger PE: A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. *Blood* 2009, 113:2526-2534.

45. Hu W, Yuan B, Flygare J, Lodish HF: Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev* 2011, 25:2573-2578.

46. Xia F, Dong F, Yang Y, Huang A, Chen S, Sun D, Xiong S, Zhang J: Dynamic transcription of long non-coding RNA genes during CD4+ T cell development and activation. *PLoS ONE* 2014, 9:e101588.

47. Fulciniti M, Amin S, Nanjappa P, Rodig S, Prabhala R, Li C, Minvielle S, Tai YT, Tassone P, Avet-Loiseau H, et al: Significant biological role of sp1 transactivation in multiple myeloma. *Clin Cancer Res* 2011, 17:6500-6509.

48. Amodio N, Di Martino MT, Foresta U, Leone E, Lionetti M, Leotta M, Gulla AM, Pitari MR, Conforti F, Rossi M, et al: miR-29b sensitizes multiple myeloma cells to bortezomib-induced apoptosis through the activation of a feedback loop with the transcription factor Sp1. *Cell Death Dis* 2012, 3:e436.

49. Wang X, Yan Z, Fulciniti M, Li Y, Gkotzamanidou M, Amin SB, Shah PK, Zhang Y, Munshi NC, Li C: Transcription factor-pathway coexpression analysis reveals cooperation between SP1 and ESR1 on dysregulating cell cycle arrest in non-hyperdiploid multiple myeloma. *Leukemia* 2014, 28:894-903.

50. Shankar DB, Cheng JC, Kinjo K, Federman N, Moore TB, Gill A, Rao NP, Landaw EM, Sakamoto KM: The role of CREB as a proto-oncogene in hematopoiesis and in acute myeloid leukemia. *Cancer Cell* 2005, 7:351-362.

51. Esparza SD, Chang J, Shankar DB, Zhang B, Nelson SF, Sakamoto KM: CREB regulates Meis1 expression in normal and malignant hematopoietic cells. *Leukemia* 2008, 22:665-667.

52. O'Connell RM, Balazs AB, Rao DS, Kivork C, Yang L, Baltimore D: Lentiviral vector delivery of human interleukin-7 (hIL-7) to human immune system (HIS) mice expands T lymphocyte populations. *PLoS ONE* 2010, 5:e12009.

53. Contreras JR, Palanichamy JK, Tran TM, Fernando TR, Rodriguez-Malave NI, Goswami N, Arboleda VA, Casero D, Rao DS: MicroRNA-146a modulates B-cell oncogenesis by regulating Egr1. *Oncotarget* 2015, 6:11023-11037.

CHAPTER IV:

Conclusions and Future Directions

Conclusions

The human genome produces thousands of non-coding transcripts in addition to those, which code for proteins[1]. Once thought as just transcriptional noise, many of these non-coding RNAs are indispensable for the regulation of cellular processes[2,3]. As our ability to probe the inner working of the cell has expanded, our understanding of these important regulatory noncoding roles has continuously grown. The advent of high-throughput techniques to study gene expression has uncovered that while 70% of the human genome is transcribed, just one percent codes for protein[4,5]. Significant progress has been made in the field of non-coding RNAs, elucidating the importance and function of these transcripts. From initial discoveries in *C. elegans* of the novel small RNA biogenesis pathway and the identification of RNA interference, the field has moved rapidly[3,6]. Research continues to uncover their far reaching effects, aiding in our understanding of the cell and how it skews from the traditional dogma of biology.

One class of these noncoding transcripts, termed microRNA (miRNA), has several important functional roles, acting as post-transcriptional regulators in diverse processes. miRNAs are about 19 – 23 nucleotides in length and act as post-transcriptional gene regulators by binding to partially complementary sequences in the 3' UTR on target messenger RNAs, thereby causing downregulation of the target. Victor Ambros first discovered them in 1993, during a study of lin-14 in *C. elegans*[6,7]. These miRNAs have emerged as significant modulators of gene expression, and regulate diverse physiologic processes. The involvement of miRNAs in hematopoiesis has now been demonstrated by numerous studies[8-10]. Dysregulation of these small RNAs can be seen in pathologic conditions of the hematopoietic and immune systems, including autoimmunity and cancer[11-13]. In addition, these miRNAs have been found to act as both tumor suppressor genes and oncogenes[14,15]. In the first appendix, we compiled a review detailing how they seem to regulate many aspects of hematopoietic development as well as their roles in hematological malignancies[16].

Long non-coding RNAs (lncRNAs) comprise a more recently characterized class. They are unique structures with distinct epigenetic marks in their promoter regions and gene body[17]. This means that they are transcribed in cells and subject to regulation; furthermore, they can be studied like any other genes. Guttman et al demonstrated their conservation in mammals by studying their distinctive epigenetic marks in mice. More recently, Ulitsky et al. showed conservation of these transcripts in vertebrates down to zebrafish. They have functional and genomic conservation despite rapid sequence evolution[18]. LncRNAs are involved in transcriptional regulation, chromatin remodeling, imprinting, splicing, and translation, among other critical functions in the cell[19-23]. One example is HOTAIR lncRNA, which regulates and suppresses expression of genes in the HOXD cluster by recruitment of the PRC2 complex[24]. Recent studies have elucidated the importance of lncRNAs in hematopoietic development[25-28]. Venkatrama et al has shown how H19 imprinting maintains stem cell quiescence[29]. Dysregulation of lncRNA expression is a feature of various diseases and cancers, and is also seen in hematopoietic malignancies[30-35]. In Chapter I we discussed work done to further understand the role of lncRNAs in hematopoietic malignancies, particularly those derived from the bone marrow.

Dysregulated expression of lncRNAs has been found in various cancers, but has not been comprehensively described in B lymphoblastic leukemia (B acute lymphoblastic leukemia; B-ALL)[31]. In our interest to uncover more about these lncRNAs, we completed a gene expression profiling study in human B-ALL samples, which showed differential lncRNA expression in samples with particular cytogenetic abnormalities[35]. In Chapter II, we found that lncRNA expression can discriminate B-ALL karyotypes, as well as, predict patient survival. We discovered four highly dysregulated lncRNAs, which we termed B-ALL associated long RNAs (BALRs), unless previously annotated. These transcripts were mostly localized to the nucleus, although BALR-2 did have comparable levels seen in the cytoplasm. Two lncRNAs from our

study had the highest expression in patient samples carrying the MLL rearrangement when compared to patients with other known B-ALL karyotypes and normal CD19+ cells. These were of particular interest since MLL rearranged B-ALL cases have a very poor prognosis and occur in infants, making them particularly hard to treat[36]. Our studies revealed that BALR-2 expression was significantly higher in patients unresponsive to prednisone treatment. Our assays revealed that knockdown of BALR-2 decreased survival and increased cell death, while overexpression had the opposing effect. Modulation of BALR-2 expression greatly affects the expression of glucocorticoid response pathway genes, in particular JUN and its pro-apoptotic partner BIM. Knockdown of BALR-2 increased the RNA and protein levels of these genes, while overexpression reduced them. These results demonstrate a potentially pivotal role for BALR-2 in the glucorticoid response pathway. This makes BALR-2 an interesting target in leukemia cases that are unresponsive to chemotherapeutic treatment, which demands further study.

After our studies on BALR-2 we decided to pursue analysis of BALR-6 to uncover its role in MLL rearranged B-ALL. In Chapter III, we found that siRNA mediated knockdown of BALR-6, in human B-ALL cell lines, caused decreased proliferation and increased apoptosis. Conversely, overexpression of BALR-6 isoforms in both human and mouse cell lines caused increased proliferation and decreased apoptosis. Additionally, overexpression of BALR-6 isoforms in mice showed a significant increase in hematopoietic precursors, specifically HSCs, CLPs, and LMPPs. We observed that BALR-6 expression was high in cell lines with t(4;11) translocation, similar to our findings in patients samples seen in Chapter II. Treatment of B-ALL cell lines with I-BET151, a bromodomain binding protein inhibitor, caused a reduction in BALR-6 levels, alluding to MLL-AF4 as a possible upstream regulator[37]. Differential expression analysis, from human knockdown samples, indicated an enrichment of genes involved in leukemia, as well as enrichment in genes targeted by SP1 transcription factor. Our studies on miR-146a uncovered how this transcript regulates the expression of Egr1, and in particular its known targets

(Appendix II). This demonstrated how a non-coding RNA can modulate a particular transcriptome[13]. We there for took a similar approach and studied how BALR-6 may regulate the SP1 transcriptome. Luciferase assays uncovered an enhancement in luciferase activity when both SP1 and BALR-6 are overexpressed. These data demonstrate that BALR-6 is functionally important in B-ALL cell survival and in transcriptional regulation of SP1 target genes.

Over the years, research has elucidated the importance of transcriptional regulation in B-cell development for successful progression through the developmental stages. It has also been demonstrated that non-coding RNAs have crucial roles in hematopoiesis including pluripotency, differentiation, and lineage commitment[38,39]. The developmental process of B-cells can be interrupted by mutations that effect essential transcriptional regulators, which can lead to B-ALL. Hence, a global pattern is emerging highlighting the importance of gene expression regulation by non-coding RNAs. This thesis identifies novel and interesting RNA transcripts with the potential to regulate gene expression and pathogenesis in B-ALL suggesting diagnostic, prognostic, and therapeutic implications.

Future Directions

In cis regulation of SATB1 and TBC1D5 by BALR-6

As mentioned previously, Ulitsky and colleagues demonstrated that lncRNAs and their surrounding genes are in synteny across vertebrates and that the surrounding genes can be regulated by the neighboring lncRNA[18]. BALR-6 is located on chromosome 3p24.3 in humans and exists in a syntenic gene block with neighboring genes SATB1 and TBC1D5. Although lncRNAs should theoretically have an intrinsic *cis*-regulatory capacity, only a few studies have described this type of regulation[40-42]. To explore whether BALR-6 regulates surrounding genes, we analyzed microarray data of MLL rearranged B-ALL samples (from Chapter II), finding that

expression of BALR-6 correlates with expression of surrounding genes SATB1 and TBC1D5 in MLL translocated cases (Figure 1A). Our preliminary studies show that SATB1 expression correlates with BALR-6 in human B cell developmental stages (Figure 1B, Chapter III). Knockdown and overexpression of BALR-6 caused an effect on the expression of surrounding genes SATB1 and TBC1D5 (Figure 1C-D). Previous findings have shown that dysregulated SATB1 has been seen in a variety of malignancies[43-46]. Future *in vivo* and *in vitro* experiments could delineate how BALR-6 regulates gene expression of surrounding genes SATB1 and TBC1D5 (Figure 1E). This information will be important to further understand lncRNA regulation in *cis.*

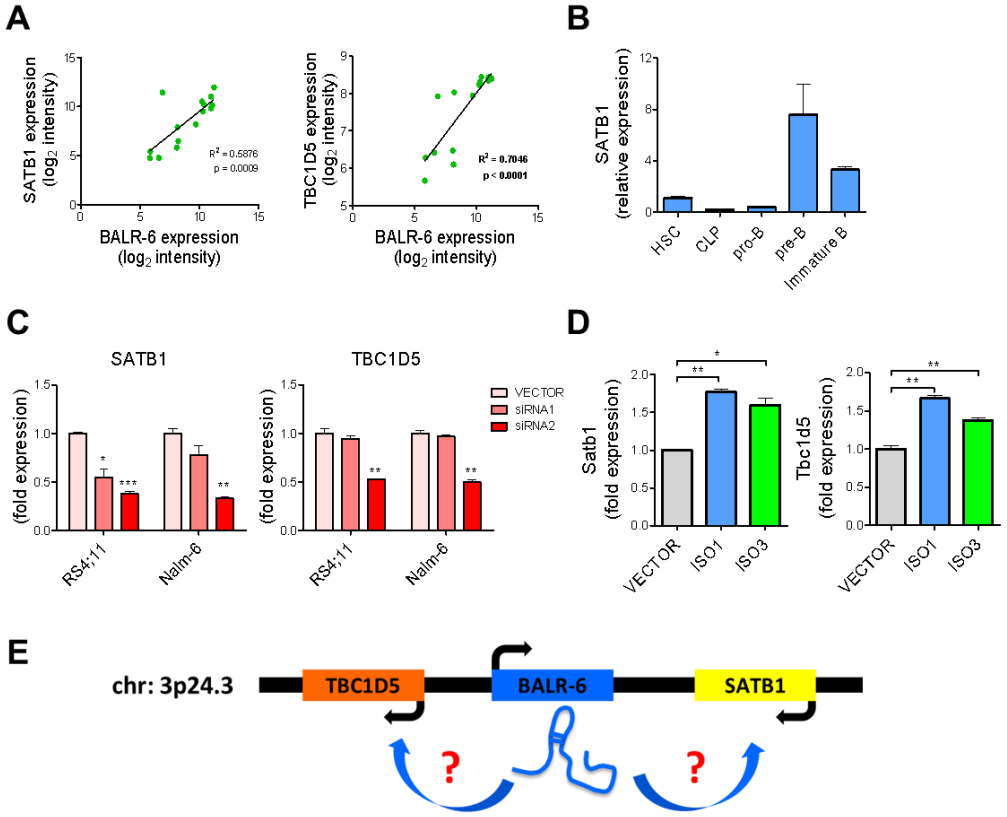


Figure 1: BALR-6 dysregulation causes changes in SATB1 and TBC1D5 expression. A) Correlation between BALR-6 expression, SATB1, and TBC1D5 in MLL translocated B-ALL cases, n=15. **B)** Quantitation of SATB1 expression in human B-cell subsets by qRT-PCR. **C)** Knockdown of BALR-6 by siRNA causes a decrease in SATB1 and TBC1D5 expression, by qRT-PCR, in B-ALL cell lines RS4;11 and Nalm-6. **D)** Expression of Satb1 and Tbc1d5 are elevated in stably transduced murine 70Z/3 cell lines overexpressing BALR-6 isoforms. **E)** Schematic of possible in *cis* mechanism by BALR-6.

Upstream regulators of BALR-6

BALR-6 is upregulated in patient samples and B-ALL cell lines that have MLL-AF4 translocations (Chapter II and III). Additionally, in Chapter III we showed that treatment with a bromodomain binding protein inhibitor (I-BET151) reduced the expression of this lncRNA. It is possible that BALR-6 is regulated by MLL-AF4, or other MLL genes, since I-BET151 was shown to inhibit transcription downstream of MLL fusion proteins[47]. To define whether BALR-6 expression is dependent upon the MLL-AF4 fusion protein, we will use stable cell lines transduced with MLL-AF4 translocation-specific shRNAs and evaluate them for expression of BALR-6. Furthermore, we will assess whether BALR-6 can rescue MLL-AF4 knockdown cells from apoptosis. To elucidate other possible regulators, we need to define the promoter region. A 2 kB fragment of the genome upstream of the most proximal 5' transcription start site, as described by RACE, will be analyzed via biocomputational methods for the presence of various promoter binding site consensus sequences. Then, we will clone this fragment into the pGL4 luciferase vector and transfect it into HEK293T cells along with a gene from the predicted set proposed to bind to this region of DNA. We anticipate that MLL-AF4 knockdown will lead to a decrease in BALR-6 expression. Moreover, we anticipate discovering numerous transcription factors that regulate BALR-6 expression, including MLL-AF4.

BALR-6 in hematopoiesis

The BALR-6 locus shows homology to the mouse genome as well as chromosomal synteny between the two species in this region of human chromosome 3/mouse chromosome 17. Using RACE, we have begun to characterize transcripts in murine pre-B-ALL 70Z/3 cells that originate from this locus. Following sequencing and cloning of 5' m7G-capped and polyadenylated transcripts that originate at this locus, we will design qRT-PCR primers to detect these transcripts in normal murine B-cells. Using RNA from murine B-cell developmental

subsets, we will be able to characterize the expression of BALR-6 at each stage. Subsequently, we can develop a knockout mouse model to look at changes in hematopoiesis upon BALR-6 deficiency. Alternatively, a siRNA-mediated knockdown can be carried out using an MSCV-based retrovirus system[10,48]. Mice will be observed carefully for morbidity and symptomatology. They will be bled retrorbitally repeatedly for 8 - 12 weeks and complete blood cell counts and FACS staining can be performed. After 3 months or development of pre-morbid symptoms, mice can be sacrificed for analysis by histology, FACS staining, and RNA-based studies of hematolymphoid tissues. We anticipate that BALR-6 will be differentially regulated in murine B-cell development, as it is in human B-cell progenitors (Chapter II). Furthermore, we anticipate that it has an important functional role in murine B-cell development.

BALR-6 in B-ALL pathogenesis *in vivo*

The most common translocation in B-ALL involves MLL being translocated to AF4[49]. We have shown in Chapter II and II that BALR-6 is important in MLL rearranged B-ALL and potentially regulated by the fusion protein. Several murine models have been developed to study B-ALL. However, a conditional knock-in mouse model has shown faithful recapitulation of the features of human B-ALL with MLL-AF4 translocations. These mice develop B-ALL and AML with a mean latency of 4-6 months following crossing with Mx1-Cre transgenic animals and poly-IC induction in a conditional knockout mouse[50]. With these mice, we can examine the expression of BALR-6 in the leukemia that develops. Then, we can use our retroviral transduction system to introduce siRNAs against murine BALR-6 into the knock-in marrow, followed by transplant into lethally irradiated wild type recipients. The recipient mice will be allowed to age (over 4 months) and leukemic transformation will be monitored by peripheral blood counts and careful observation. When the mice show signs of morbidity, or have aged enough, they will be sacrificed. A complete necropsy including gross and microscopic

pathology, FACS based analyses, and isolation of RNA and protein for expression analyses as described above. We expect that knockdown of BALR-6 will ameliorate the leukemic phenotype in mice, demonstrating its contribution to the leukemic transformation.

These are the first *in vivo*, *in vitro*, and cell culture studies of lncRNA contribution to B-ALL pathogenesis. Although this thesis has made significant strides to delineate the role of non-coding RNAs in B-ALL, there is still much to uncover. The story of RNA biology is one of continued growth. From understanding its simpler roles as messenger of information, to teasing out complex regulatory interactions with multiple effectors, our knowledge of its far-reaching influence upon the cell continues to expand.

References

1 Kapranov, P. *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**, 1484-1488, doi:10.1126/science.1138341 (2007).

2 Wilusz, J. E., Sunwoo, H. & Spector, D. L. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev* **23**, 1494-1504, doi:10.1101/gad.1800909 (2009).

3 Dey, B. K., Mueller, A. C. & Dutta, A. Long non-coding RNAs as emerging regulators of differentiation, development, and disease. *Transcription* **5**, e944014, doi:10.4161/21541272.2014.944014 (2014).

4 Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858, doi:10.1126/science.1064921 (2001).

5 Bertone, P. *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**, 2242-2246, doi:10.1126/science.1103388 (2004).

6 Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854 (1993).

7 Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T. & Jewell, D. MicroRNAs and other tiny endogenous RNAs in C. elegans. *Current biology : CB* **13**, 807-818 (2003).

8 Thai, T. H. *et al.* Regulation of the germinal center response by microRNA-155. *Science* **316**, 604-608, doi:10.1126/science.1141229 (2007).

9 O'Connell, R. M. *et al.* Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder. *The Journal of experimental medicine* **205**, 585-594, doi:10.1084/jem.20072108 (2008).

10 Rao, D. S. *et al.* MicroRNA-34a perturbs B lymphocyte development by repressing the forkhead box transcription factor Foxp1. *Immunity* **33**, 48-59, doi:10.1016/j.immuni.2010.06.013 (2010).

11 O'Connell, R. M. *et al.* MicroRNA-155 promotes autoimmune inflammation by enhancing inflammatory T cell development. *Immunity* **33**, 607-619, doi:10.1016/j.immuni.2010.09.009 (2010).

12 Boldin, M. P. *et al.* miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *The Journal of experimental medicine* **208**, 1189-1201, doi:10.1084/jem.20101823 (2011).

13 Contreras, J. R. *et al.* MicroRNA-146a modulates B-cell oncogenesis by regulating Egr1. *Oncotarget* **6**, 11023-11037 (2015).

14 Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13944-13949, doi:10.1073/pnas.0506654102 (2005).

15 Tam, W. & Dahlberg, J. E. miR-155/BIC as an oncogenic microRNA. *Genes, chromosomes & cancer* **45**, 211-212, doi:10.1002/gcc.20282 (2006).

16 Fernando, T. R., Rodriguez-Malave, N. I. & Rao, D. S. MicroRNAs in B cell development and malignancy. *J Hematol Oncol* **5**, 7, doi:10.1186/1756-8722-5-7 (2012).

17 Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223-227, doi:10.1038/nature07672 (2009).

18 Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. & Bartel, D. P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537-1550, doi:10.1016/j.cell.2011.11.055 (2011).

19 Lee, J. T. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. *Genes Dev* **23**, 1831-1842, doi:10.1101/gad.1811209 (2009).

20 Tripathi, V. *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39**, 925-938, doi:10.1016/j.molcel.2010.08.011 (2010).

21 Huarte, M. *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409-419, doi:10.1016/j.cell.2010.06.040 (2010).

22 Yoon, J.-H. *et al.* LincRNA-p21 Suppresses Target mRNA Translation. *Molecular Cell* **47**, 648-655, doi:http://dx.doi.org/10.1016/j.molcel.2012.06.027 (2012).

23 Lee, J. T. Epigenetic regulation by long noncoding RNAs. *Science* **338**, 1435-1439, doi:10.1126/science.1231776 (2012).

24 Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311-1323, doi:10.1016/j.cell.2007.05.022 (2007).

25 Wagner, L. A. *et al.* EGO, a novel, noncoding RNA gene, regulates eosinophil granule protein transcript expression. *Blood* **109**, 5191-5198, doi:10.1182/blood-2006-06-027987 (2007).

26 Dinger, M. E. *et al.* Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* **18**, 1433-1445, doi:10.1101/gr.078378.108 (2008).

27 Zhang, X. *et al.* A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. *Blood* **113**, 2526-2534, doi:10.1182/blood-2008-06-162164 (2009).

28 Hu, W., Yuan, B., Flygare, J. & Lodish, H. F. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev* **25**, 2573-2578, doi:10.1101/gad.178780.111 (2011).

29 Venkatraman, A. *et al.* Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature* **500**, 345-349, doi:10.1038/nature12303 (2013).

30 Yildirim, E. *et al.* Xist RNA is a potent suppressor of hematologic cancer in mice. *Cell* **152**, 727-742, doi:10.1016/j.cell.2013.01.034 (2013).

31 Fang, K. *et al.* A distinct set of long non-coding RNAs in childhood MLL-rearranged acute lymphoblastic leukemia: biology and epigenetic target. *Human Molecular Genetics* **23**, 3278-3288, doi:10.1093/hmg/ddu040 (2014).

32 Guo, G. *et al.* High expression of long non-coding RNA H19 is required for efficient tumorigenesis induced by Bcr-Abl oncogene. *FEBS Letters* **588**, 1780-1786, doi:http://dx.doi.org/10.1016/j.febslet.2014.03.038 (2014).

33 Hao, S. & Shao, Z. HOTAIR is upregulated in acute myeloid leukemia and that indicates a poor prognosis. *International Journal of Clinical and Experimental Pathology* **8**, 7223-7228 (2015).

34 Xing, C.-y. *et al.* Long non-coding RNA HOTAIR modulates c-KIT expression through sponging miR-193a in acute myeloid leukemia. *FEBS Letters* **589**, 1981-1987, doi:http://dx.doi.org/10.1016/j.febslet.2015.04.061 (2015).

35 Fernando, T. R. *et al.* LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia. *Molecular cancer research : MCR* **13**, 839-851, doi:10.1158/1541-7786.MCR-15-0006-T (2015).

36 Pui, C. H. *et al.* 11q23/MLL rearrangement confers a poor prognosis in infants with acute lymphoblastic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **12**, 909-915 (1994).

37 Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529-533, doi:http://www.nature.com/nature/journal/v478/n7370/abs/nature10509.html#supplementary-information (2011).

38 Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83-86, doi:10.1126/science.1091903 (2004).

39 Sheik Mohamed, J., Gaughwin, P. M., Lim, B., Robson, P. & Lipovich, L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* **16**, 324-337, doi:10.1261/rna.1441510 (2010).

40 Christoffersen, N. R. *et al.* p53-independent upregulation of miR-34a during oncogene-induced senescence represses MYC. *Cell death and differentiation* **17**, 236-245, doi:10.1038/cdd.2009.109 (2010).

41 Zhang, B. *et al.* The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep* **2**, 111-123, doi:10.1016/j.celrep.2012.06.003 (2012).

42 Garding, A. *et al.* Epigenetic upregulation of lncRNAs at 13q14.3 in leukemia is linked to the *In Cis* downregulation of a gene cluster that targets NF-kB. *PLoS Genet* **9**, e1003373, doi:10.1371/journal.pgen.1003373 (2013).

43 Han, H. J., Russo, J., Kohwi, Y. & Kohwi-Shigematsu, T. SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. *Nature* **452**, 187-193, doi:10.1038/nature06781 (2008).

44 Kohwi-Shigematsu, T. *et al.* Genome organizing function of SATB1 in tumor progression. *Semin Cancer Biol* **23**, 72-79, doi:10.1016/j.semcancer.2012.06.009 (2013).

45 Chen, Z. *et al.* SATB1 Promotes Pancreatic Cancer Growth and Invasion Depending on MYC Activation. *Dig Dis Sci* **60**, 3304-3317, doi:10.1007/s10620-015-3759-9 (2015).

46 Mir, R., Pradhan, S. J., Patil, P., Mulherkar, R. & Galande, S. Wnt/beta-catenin signaling regulated SATB1 promotes colorectal cancer tumorigenesis and progression. *Oncogene*, doi:10.1038/onc.2015.232 (2015).

47 Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529-533, doi:10.1038/nature10509 (2011).

48 O'Connell, R. M., Chaudhuri, A. A., Rao, D. S. & Baltimore, D. Inositol phosphatase SHIP1 is a primary target of miR-155. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7113-7118, doi:10.1073/pnas.0902636106 (2009).

49 Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* **7**, 823-833, doi:10.1038/nrc2253 (2007).

50 Krivtsov, A. V. *et al.* H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* **14**, 355-368, doi:10.1016/j.ccr.2008.10.001 (2008).

APPENDICES

APPENDIX I:

MicroRNAs in B cell Development and Malignancy

(reprint)

REVIEW Open Access

MicroRNAs in B cell development and malignancy

Thilini R Fernando[1†], Norma I Rodriguez-Malave[1,2†] and Dinesh S Rao[1,3,4,5*]

Abstract

MicroRNAs are small RNA molecules that regulate gene expression and play critical roles in B cell development and malignancy. miRNA expression is important globally, as B cell specific knockouts of Dicer show profound defects in B cell development; and is also critical at the level of specific miRNAs. In this review, we discuss miRNAs that are involved in normal B cell development in the bone marrow and during B cell activation and terminal differentiation in the periphery. Next, we turn to miRNAs that are dysregulated during diseases of B cells, including malignant diseases and autoimmunity. Further study of miRNAs and their targets will lead to a better understanding of B cell development, and should also lead to the development of novel therapeutic strategies against B cell diseases.

Introduction

In a relatively short time period, gene expression regulation by microRNAs (miRNAs) has changed the way that we view developmental and pathological processes. From initial discoveries in *C. elegans*, the identification of the novel small RNA biogenesis pathway and the identification of RNA interference, the field has moved rapidly [1-6]. The involvement of miRNAs in hematopoiesis has now been documented by numerous groups and they seem to regulate almost every aspect of hematopoietic development. In this review we focus on B cell development, where the importance of gene expression regulation has been appreciated for many years. miRNAs have emerged as critical regulators of gene expression and regulate many aspects of B cell development, and are dysregulated in B cell malignancies. Here, we review many of the studies that have been performed to delineate the roles of miRNAs in development and malignant transformation of B cells.

MicroRNA biogenesis

miRNAs are non-protein coding RNAs of about 19-23 nucleotides. They are post-transcriptional gene regulators that bind to partially complementary sequences in the 3' UTR on target messenger RNA transcripts, thereby causing downregulation of the target [7]. They were first discovered in 1993 in *C. elegans* by Victor Ambros, during a study of *lin-14*. They identified a small RNA product encoded by *lin-4* gene that is responsible for the downregulation of LIN-14 protein [2,3,8]. This central dogma of miRNA action has proven to stand the test of time, as miRNAs in most organisms are thought to behave similarly.

miRNAs can be grouped in to at least three categories depending on their genomic location: exonic miRNAs in non-coding genes, intronic miRNAs in non-coding genes and intronic miRNAs in protein-coding genes [9]. miRNAs are expressed as long primary RNA (pri-miRNA) as part of RNA polymerase II-driven transcript [10]. Therefore, it is possible that some miRNAs are co-regulated with their host gene as a part of transcriptional regulation during B cell development. The pri-miRNA is recognized by RNA binding protein DGCR8 and is processed by RNase III-type protein Drosha in the nucleus yielding a pre-mRNA [11,12]. Pre-miRNA is then exported to the cytoplasm by Exportin-5 where it is further processed by a second RNase III-type enzyme, Dicer, to produce a mature miRNA duplex [13]. The 19-25 nucleotide-long double stranded miRNA duplex is then unwound and incorporated into RNA-induced silencing complex (RISC), with strand selection based on thermodynamic properties. In the RISC, the miRNA binds to the target sequence in the 3' UTR via 6-8 nucleotide seed region and downregulates the expression of the target either by direct degradation or destabilization and eventually degradation of the target [14-16]. Since the repression is achieved by

* Correspondence: drao@mednet.ucla.edu
† Contributed equally
[1]Department of Pathology and Laboratory Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA
Full list of author information is available at the end of the article

complementary base pairing via a relatively short seed sequence, miRNAs are predicted to have multiple targets. A genome wide statistical analysis has shown that one miRNA can have hundreds of targets, indicating their critical role in post translation regulation [17]. It should be noted that recently, a Dicer-independent miRNA biogenesis pathway has also been reported. This pathway utilizes the catalytic activity of Argonaute2 (Ago2) [18-21]. miR-451 is the best characterized miRNA that is produced independently of Dicer and is involved in erythropoiesis. The unusual short stem structure of pre miR-451 promotes the binding and processing by Ago2 [19].

miRNAs have already found to influence immune cell differentiation. Recently, it was found that Dicer and miRNA play vital roles in both early and late B cell differentiation [22,23]. Deletions of individual miRNA genes are associated with several immune defects. In many instances, dysregulated expression of miRNAs has been seen in malignancies in the immune system, which we will discuss in detail later in the review.

B cell development

B cells are responsible for adaptive humoral immunity. B cell development is characterized by complex sequence of molecular events that is regulated by B - lineage transcription factors. It is evident that miRNAs play a major role in modulating the expression of these transcription factors and thereby the normal development of B cells. Conversely, dysregulation of miRNA expression is thought to be a key factor to the pathogenesis of B cell malignancies, including progenitor B cell-malignancies such as B-lymphoblastic leukemia (also referred to as B-Acute lymphoblastic leukemia or B-ALL) and mature B cell malignancies including several types of non-Hodgkin lymphoma. B cell development begins in the fetal liver and continues in the bone marrow of adult throughout the life (reviewed in [24,25]). The process of B cell formation starts in the bone marrow and ends in the peripheral secondary lymphoid organs such as the spleen (Figure 1). Here, we provide a primer on B cell development to orient the discussion on the role of miRNAs in B cells.

Bone marrow B cell development

In the bone marrow, cellular stages of development include the common lymphoid progenitor, pro-B cell, pre-B cell and immature-B cell stages. These stages of B cells are defined by the expression and the re-arrangement of functional B cell receptor (BCR)/immunoglobulin (Ig) genes. Complex and elegant mechanisms have evolved to generate a diverse repertoire of BCRs against the vast variety of antigen that we encounter in our lifetime. This diversity is achieved by V (D)J recombination of the immunoglobulin locus, which is a process of somatic recombination that brings together various gene segments within the heavy and light chain loci. The heavy chain is assembled from Variable (V), Diversity

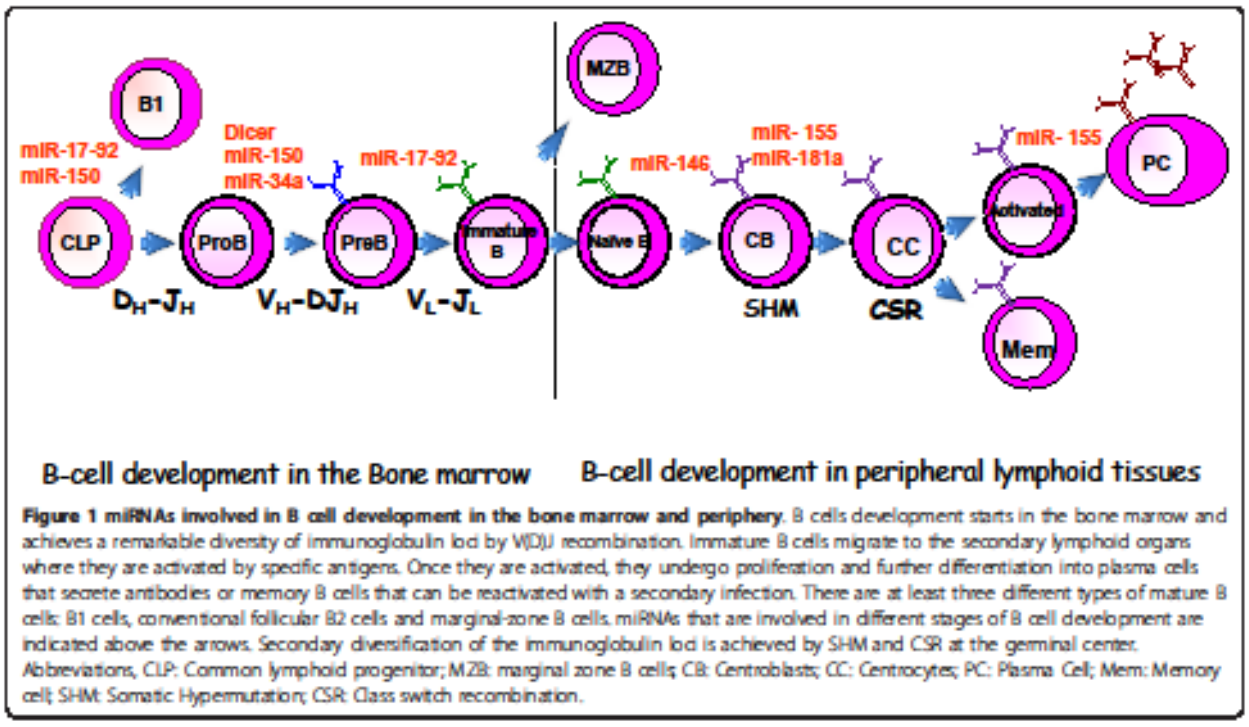


Figure 1 miRNAs involved in B cell development in the bone marrow and periphery. B cells development starts in the bone marrow and achieves a remarkable diversity of immunoglobulin loci by V(D)J recombination. Immature B cells migrate to the secondary lymphoid organs where they are activated by specific antigens. Once they are activated, they undergo proliferation and further differentiation into plasma cells that secrete antibodies or memory B cells that can be reactivated with a secondary infection. There are at least three different types of mature B cells: B1 cells, conventional follicular B2 cells and marginal-zone B cells. miRNAs that are involved in different stages of B cell development are indicated above the arrows. Secondary diversification of the immunoglobulin loci is achieved by SHM and CSR at the germinal center. Abbreviations, CLP: Common lymphoid progenitor; MZB: marginal zone B cells; CB: Centroblasts; CC: Centrocytes; PC: Plasma Cell; Mem: Memory cell; SHM: Somatic Hypermutation; CSR: Class switch recombination.

(D) and Joining (J) gene exons that somatically recombine with the Constant (C) region exons to generate unique immunoglobulins of differing antigenic specificity. On the other hand, the light chain variable region is composed of only a V and J segments. The numbers of these gene segments are highly variable in different species and the different segments seem to have evolved from gene duplications from ancestral V-gene exons. Each V-, D-, and J- gene segment is flanked by DNA sequence called recombination signaling sequence, which is recognized by RAG1 and RAG2 enzymes, which mediate recombination. Random selection of these segments during V(D)J recombination and junctional diversity introduced by addition or subtractions of nucleotides at the junctions of these segments enable the production of vast variety of Ig (reviewed in [26-32]).

The stages of B cell development have been defined by the steps in V(D)J recombination. The Pro-B cell stage is characterized by rearrangement of Ig heavy chain, which occurs first. D-J joining occurs first, following which the DJ segment is joined to a V segment. D-J rearrangement starts in the common lymphoid progenitor and occurs mainly in early pro-B cells. V-DJ rearrangement occurs in late pro-B cells. The assembled heavy chain is then expressed on the surface of the pre-B cells along with a surrogate light chain. The pro-B cell to pre-B cell transition is accompanied by cell proliferation. Rearrangement of the light chain by V to J joining occurs during the pre-B cell stage. Successful assembly of light chain leads to the expression of complete IgM molecule at the surface and if the rearrangement is successful, signal transduction from IgM binding allows for differentiation into an immature B cell. Immature B cells that are not self-reactive leave the bone marrow as transitional B cells. Self-reactive immature B cells will either undergo apoptosis (clonal deletion), generate a new B cell receptor by receptor editing, or become unresponsive to antigen (reviewed in [26-29,31,33]).

Several transcription factors including PU.1, STAT5, E2A (E12 and E47), EBF, Pax-5, IKZFI and FOXP1 together with cytokines (IL-7, SCF) and chemokines (CXCL12), which are provided by the stromal cells, regulate the commitment and maintenance of B cells (reviewed in [34,35]). Absence of IL-7 signaling leads to developmental arrest at the pre-pro-B cell stage, showing the essential role of IL-7 signaling in B cell development in mice [36]. Downstream of IL-7, deletion of both STAT5 and STAT3 lead to developmental arrest at the pro-B stage [37,38]. Transcriptional regulation of B cell development is complex and involves the interplay of several transcription factors, including E2A, EBF, PAX5, FOXP1 and IKZF1. E2A-null mutant mice were unable to generate mature B cells [39], while Pax-5 is required for commitment to the B cell lineage [40]. Deletion of EBF blocked the pro-B to pre-B transition [41]. It has also been shown that $Ebf^{-/-}$ hematopoietic cells do not express Pax-5 indicating that EBF acts upstream of Pax-5. This finding was further supported by the finding of an EBF-binding site in the Pax-5 promoter region [42]. Foxp1 is also an essential transcription factor for B cell development that is induced by E2A and in turn induces expression of the Rag enzymes. Deletion or knockdown of Foxp1 resulted in a reduction of B cell specific gene expression and interrupted the transition from pro-B cell to pre-B cell [35,43]. Other transcriptional regulators of B cell development include Ikzf1, which seems to play an important role in early lymphoid commitment.

These repeated cycles of DNA damage and repair may explain the reason for Pro-B and pre-B stages being more susceptible to oncogenic transformation. Also of interest, almost every transcriptional regulator of B cell development is disrupted in B-lineage malignancy. E2A has found in chromosomal translocations associated with B-ALL, and Pax-5 deletions are common in B-ALL [44-48]. Also in genome wide analysis, Pax-5 and EBF have been shown to be associated with B ALL [46]. Foxp1 translocation and overexpression is noted in mature B cell neoplasms, while IKZF1 is disrupted in pre-B-ALL [49-60]. Hence, the study of B cell development also informs the understanding of B cell malignancy pathogenesis.

B cell development in the periphery

Mature B cells in the periphery are generally divided into B1 and B2 cells. B2 cells are conventional B cells that are derived from the bone marrow, undergo V(D)J recombination, and are part of the adaptive immune response (reviewed in [26]). A second set of recently described B cells, called B1 cells, are characterized by a limited immunoglobulin repertoire, and are part of the innate immune system (reviewed in [61]). These are best described in the mouse lymphoid system, where they express distinct sets of markers and are located mainly in the peritoneum and in the spleen. The lineage origin and relationship with conventional B2 cells is not entirely clear. The rest of our discussion will focus on B2 cells.

The majority of the cells in the spleen and in the circulation are thought to be B2 cells, and we focus the rest of the discussion in this section of such cells. B cells that have not encountered their specific antigen are called naïve B cells (CD 27-, IgD+). When a B cell binds to an antigen, it enters the germinal center of peripheral lymphoid tissues and eventually differentiates into plasma cells and memory cells. Plasma cells are a terminally differentiated, highly specialized B cell that

secretes massive quantities of Ig, and whose differentiation is mediated by activation of plasma cell transcription factors such as Blimp1 and Xbp1 [62]. Memory B cells are long-lived and can be re-activated during a secondary infection. The entry of the activated B cells into primary lymphoid follicles results in the formation of germinal centers (GCs). Follicles with germinal centers have three distinct zones, namely dark, light and mantle zone. Rapidly dividing B cells which are called centroblasts form the dark zone of the GC. CBs eventually differentiate into non cycling centrocytes which make up the light zone of the GC. In the GC, secondary diversification of the immunoglobulin repertoire is achieved through somatic hypermutation and class switch recombination [63]. During this remarkable process, CBs can refine the specificity of their antigen receptor by somatic hypermutation of IgV gene, via the introduction of point mutations. If this process results in enhancement of the binding affinity for antigen, the B cell is selected for, and it survives to differentiate further. Also during the germinal center reaction, class switching from IgM to any other immunoglobulin can occur. Here, the variable portion of the heavy chain (VH exon) is brought adjacent to with different immunoglobulin constant regions (CH exons). This process allows making antibodies with different effector function. Remarkably, both of these processes-somatic hypermutation and class-switch recombination-are under the control of the same enzyme, adenosine-induced deaminase (AID) [64]. CC expressing Igs with enhanced affinity may eventually be released as memory B cells (CD 27+) from the GC. These memory cells are long lived and have potential to become antibody secreting cells during secondary infection (reviewed in [29,32,65-67]).

MicroRNAs in bone marrow B cell development

As can be seen from the preceding discussion, bone marrow B cell development is carefully orchestrated, as only one gene locus is rearranged at a given time in a fixed sequence. Ig rearrangement is mediated by the sequential action of a gene regulatory network composed of transcription factors and growth factor receptors. miRNAs are known to act as post-transcriptional regulators of gene expression and it therefore stands to reason that they play a critical role in this network. The importance of miRNAs was first established by a seminal study that delineated a role for miRNAs in hematopoietic lineage choice selection. In this study, the authors determined that miR-181 (now known as miR-181a) was expressed most highly in B cells and that its overexpression in hematopoietic stem and progenitor cells led to increased output of B cells [68,69]. Further studies have shown additional miRNAs of importance in B cell development; in this section, we will focus on the role of miR-150, miR-34a and the miR-17-92 cluster in antigen independent B cell development at the bone marrow.

A general role for miRNAs in B cell development has also been established (Figure 1). The conditional knockout of Dicer in early B cells led to a developmental arrest at the pro-B cell to pre-B cell transition, and also caused an effect on antibody diversification [70]. Gene-expression profiling from Dicer-deficient cells indicated that Bim, a known miR-17-92 cluster target, was upregulated in the mice. Functionally, B cell development was partially rescued by concurrent Bim ablation in Dicer-deficient mice. The implication of these studies was that miR-17-92, via repression of Bim, was the key player that was missing in Dicer-deficient B cell development. The miR-17-92 cluster is located on chromosome 13 and encodes six different miRNAs. The cluster is highly expressed in progenitor B cells and expression diminishes as cells mature. Ectopic expression of the cluster in mice resulted in expansion and activation of all lymphocyte populations in the periphery [71]. Compound heterozygous mutations of two target genes of miR-17-92, *Bim* and *Pten*, resulted in an accumulation of activated lymphocytes, indicating that partial repression of two targets may explain the majority of the miR-17-92-induced phenotype [71]. Taken together these data indicate that the miR-17-92 cluster plays a critical role in proliferation control in B cells, in B cell development and Ig rearrangement.

A second B cell-relevant miRNA, miR-150, is highly expressed in progenitor B cells and levels decrease at the pro-B cell to pre-B cell transition. miR-150 targets the c-Myb transcription factor in B cell development [72]. Confirming targeting, B cells that are deficient in miR-150 showed higher levels of c-Myb, while over-expression of miR-150 in transgenic mice caused reduced levels of c-Myb. Over-expression of miR-150 in mouse HSC led to a defect at the pro-B cell to pre-B cell transition [72]. Mice with targeted deletion of miR-150 had more B-1 cells in the spleen and peritoneal cavity and fewer B-2 cells, although they appeared phenotypically normal [73]. Knock-out mice for miR-150 at baseline contained higher serum concentrations of Ig classes, especially IgA, likely due to an expansion of B-1 cells [72]. Similarly, mice that were haploinsufficient for c-Myb had fewer mature B cells in the spleen and fewer B-1 cells, consistent with what was seen with the miR-150 transgenic mice [74] These data indicate a critical role for miR-150 during B cell development.

Along with miR-150, miR-34a is highly expressed in progenitor cells and downregulated and the pro-B cell to pre-B cell transition. Ig rearrangement has multiple check points dependent on TP53 [75]. TP53 targets miR-34a which in turn targets genes involved in cell

cycle regulation, cell proliferation and apoptosis [76,77]. Among its targets are the anti-apoptotic protein, BCL2, and the transcription factor Foxp1 [43,78]. Mice with constitutive expression of miR-34a showed a block at the pro-B cell to pre-B cell transition with a reduction in mature B cells [43]. This arrest resulted from the inhibition of Foxp1 which is required for early B cell development. These findings elucidate a crucial role for miR-34a regulation at early B cell development.

MicroRNAs in spleen and periphery B cell development

As in bone marrow B cell development, miRNAs as a whole, as well as specific miRNAs, have now been appreciated to play important roles in peripheral or antigen-dependent B cell development (Figure 1). At the global level, Dicer ablation in mature B cells (as opposed to early precursor B cells) using CD21-Cre resulted in an increase in marginal zone B cells and a decrease in follicular B cells [23]. Mice deficient for Dicer in mature B cells had an increased titer of autoimmune immunoglobulins with frank autoimmune disease in a proportion of the female mice. The mechanistic basis of these findings remains to be determined, but this study suggested that a miRNA may be responsible of regulating Bruton's tyrosine kinase. However, there are certainly some miRNAs that play major roles in B cell development; here we will focus on the role of miR-155, miR-146a and miR-181a in B cell development in the spleen and periphery.

In normal lymphopoiesis, miR-155 is expressed in moderate level in HSCs, at high levels in the germinal center and at much lower levels in mature B cells [79-81]. Expression of miR-155 is rapidly induced in B cells after engagement of the antigen receptor and exposure to inflammatory mediators [82,83]. Mice lacking miR-155 showed normal steady state immune cell populations; however, mice had a defective humoral response when immunized [84]. This response involved impaired germinal center formation and led to low antibody class switching to IgG in a B cell-intrinsic manner. The targets responsible for this appear to be multiple but likely include PU.1, SHIP1, and possibly AID [85-87]). The inhibition of the latter target is interesting, because the phenotype that is observed in the miR-155 deficient mice is one of decreased class-switching, whereas derepression of AID might be expected to cause increased class switching. The targeting of AID by miR-155 was extensively studied by mutating the binding site for miR-155 in the AID 3'UTR and these studies determined that disruption of the interaction did indeed lead to increased class-switching, and hence the overall effect of miR-155 likely includes additional targets [86,87]. Overall, these data are consistent with miR-155 playing an important role in regulating antigen-dependent B cell development. More recently, a second miRNA has been identified to regulate CSR- miR-181b overexpression in B cells was found to reduce the CSR rates, possibly by also downregulating AID [88].

The miR-146 family has distinct expression patterns amongst various hematopoietic lineages and is involved in maintaining lineage identity in lymphocytes. Vertebrates have two genomic copies of the miRNA, miR-146a and miR-146b, although the latter is likely a pseudogene [89]. miR-146a is induced by Toll-like receptor 4 and latent membrane protein 1 activation, and is NF-κB dependent [83,90]. miR-146a targets IRAK1 and TRAF6, two adapter proteins involved in Toll like receptor and interleukin 1 receptor signaling [91]. The role of miR-146a in B cells remains to be definitively determined, but overall B cell numbers are lower in miR-146a-deficient mice as the mice have a myeloproliferative disorder. Curiously however, many mice show dramatic follicular hyperplasia and active germinal centers with increased B cell function [92].

Some other miRNAs, including miR-125a, miR-125b, miR-99b and let-7e transcripts are preferentially expressed by the actively dividing centroblasts in germinal centers. In functional assays, miRNA-125b overexpression inhibited the differentiation of primary B cells [93]. Hence, it can be seen that several miRNAs show important roles during antigen-dependent B cell development.

MicroRNAs in B cell lymphoma and leukemia

The expression of miRNAs in a particular cell type (miRNome) can vary between normal and diseased tissues. The relationship between miRNAs and cancer was first appreciated when loss of miR-15a/16-1 was discovered in chronic lymphocytic leukemia (CLL) [94]. Also, the discovery that miRNAs were located in cancer-associated genomic regions (CAGRs) furthered studies of the miRNome in a vast number of cancers, from solid tumors to hematological malignancies [95,96]. The function of miRNAs depends greatly on the cellular context for they can act as tumor suppressor genes or oncogenes depending on the genes that are expressed in a given cell. In this section we will focus on miRNAs that have been implicated in B cell lymphoma and leukemia (Figure 2). We will focus our discussion on those miRNAs that have some functional role in B cell development or lymphoma; it is beyond the scope of this review to list the plethora of profiling studies that exist in the literature.

Dysregulation of miR-155 and the miR-15a/16 and miR-17-92 clusters has been implicated in B-ALL [97]. miR-155 overexpression has been observed in certain subtypes of ALL and its over-expression in mice gives rise to B cell lymphoproliferative disease by targeting

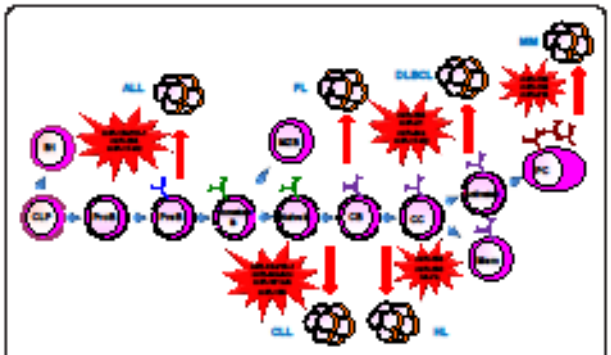


Figure 2 miRNA dysregulation leads to B cell malignancies. Dysregulation of key miRNAs at different stages in B cell development can cause malignant transformation and expansion. ALL: B-lymphoblastic lymphoma; CLL: Chronic lymphocytic leukemia; HL: Hodgkin lymphoma; NHL: Non-Hodgkin lymphoma; DLBCL: Diffuse large B cell lymphoma (a type of NHL); FL: Follicular lymphoma (a type of NHL); MM: Multiple Myeloma.

the SHIP1 inositol phosphatase [98-100]. miR-16 targets anti-apoptotic genes BCL-2, MCL1 and CDK46, thereby acting as a putative tumor suppressor [101]. The miR-17-92 cluster is upregulated in ALL due to copy number amplification and direct upregulation, and transgenic overexpression of this cluster leads to a lymphoproliferative disease in mice, as we have previously discussed (see preceding section on miRNAs and bone marrow B cell development [71,97]). Interestingly, overexpression of a single miRNA, miR-21, can lead to a high-grade B-ALL in mice and similar to protein-coding oncogenes, demonstrates the phenomenon of "oncogene addiction" [102].

In CLL miR-15a/16-1 dysregulation is observed along with dysregulation of miR-34a, miR-34b, miR-34c, miR-181b, miR-181a and miR-155. In fact miR-34a expression has recently been validated as a prognostic marker in CLL in a fairly large clinical study [103,104]. It is interesting to note that one of the most frequent abnormalities in CLL is 13q14 deletion, and the search for candidate tumor suppressor genes in the deleted region had not been successful in the pre-miRNA era [105]. The minimal deleted region (MDR) within 13q14 contains a long non-coding RNA (lncRNA) named DLEU2 [106,107]. Carlo Croce and his colleagues found that the miR-15a/16- cluster was located within intron 4 of DLEU2 [95]. This identification of the first tumor suppressor miRNA was followed by extensive study and delineation of multiple targets [103]. However, the formal assessment of tumor suppressor function was completed much more recently. In an exacting study, Dalla-Favera and colleagues created conditional knockout mice with deletions of the minimal deleted region (MDR), deleting both DLEU2 and the miRNA, or of the miRNA cluster only [108]. Remarkably, the mice with deletion of the miRNA only showed a pre-leukemic expansion of B cells, while the mice with deletion of the MDR developed a CLL-like disease. In this study, the authors showed that there were likely multiple miRNA targets responsible for the phenotypes observed, including some that had been previously identified.

Dysregulation of miRNAs has also been described in Hodgkin's Lymphoma (HL) and Non-Hodgkin's Lymphoma (NHL). In NHL miR-155, miR-21, miR-34a and the miR-17-92 clusters have been implicated. miR-155 carries prognostic implications as high expression of the miRNA is typical of ABC-DLBCL which has a 5 year survival rate [109]. *In vivo* studies demonstrate an oncogenic role for miR-21 in B-lymphomagenesis [102]. Mice expressing miR-21 showed a pre-B malignant lymphoid-like phenotype and inactivating miR-21 caused regression of the malignancy. The miR-17-92 cluster is upregulated in approximately 65% of lymphomas [110]. Furthermore, let-7a, miR-150 and miR-155 are found dysregulated in HL. Let-7a is upregulated causing low expression of PRDM1/Blimp, presumably interrupting plasma cell differentiation [111]. Downregulation of miR-150 and upregulation of miR-155 is common in HL. In normal lymphopoiesis there are high levels of miR-155 in the germinal center, where HL has its origin, suggesting that miR-155 upregulation in HL is due to an abnormal block of lymphocyte differentiation at the germinal center stage [24].

Lastly, p53-mediated miRNA regulation has been found to be important in multiple myeloma (MM), a neoplasm of mature plasma cells. In an effort to understand the miRNA effectors of p53 in this context, Pichiorri and colleagues defined a set of p53-regulated miRNAs, which include miR-192, miR-194, and miR-215 that are downregulated in a subset of newly diagnosed MM. These miRNAs target and downregulate MDM2, a negative regulator of p53. Hence, the expression of these miRNAs reinforces the activity of p53, and the authors found that enforced expression of these miRNAs had a negative effect on the growth of MM cells. Therapeutic possibilities are suggested by the effects of miRNA reconstitution in tempering MM cell growth [112].

The preceding discussion should establish the contribution of miRNAs in lymphoid tumorigenesis. Although some molecular effectors of the miRNnome are known, much remains to be discovered. miRNAs are likely to be of use as diagnostic biomarkers for cancer and as prognostic indicators. Additional work to uncover the roles of miRNAs as therapeutic agents remains to be completed, where a major limitation remains delivery of small RNAs into lymphoid cells.

MicroRNAs in autoimmunity

Strong responses to self-antigens are thought to be the basis of autoimmune diseases. Many autoimmune diseases are heavily dependent on T cells, but B cells are almost certainly involved, for example, in the secretion of autoantibodies. Indeed mice with a conditional deletion of Dicer in mature B cells develop abnormal B cell subsets, have high autoantibody titers, and female mice develop autoimmune disease with end-organ damage. Several specific miRNAs have been found to be dysregulated in variety of autoimmune disease, and many have a role in T-cell function [113].

It has been found that miR-146a null mice develop a severe autoimmune disorder characterized by enlarged spleen and lymph nodes. These null mice produced about 60 fold higher amounts of autoantibodies against double standard DNA than wild type mice. Autoimmune phenotype in miR-146a null mice is consistent with the finding of elevated amount of activated T cells in the periphery, but may also be dependent on increased activation of B cells [89]. miR-146a also plays a role in the pathogenesis of Systemic lupus erythematosus (SLE). It represses the function of IFN (type one interferon), a factor that is important in SLE, by repressing the target genes such as TRAF6/IRAK1, STAT1 and TLR7 or TLR9 [114-117].

It has been found out that the generation and function of regulatory T-cells (T reg) in autoimmunity is dependent on Dicer dependent miRNA biosynthesis pathway. Mice that have conditional deletion of Dicer in T reg cells showed early onset of autoimmunity which is similar to the observed phenotype in Foxp3 mutant mice that completely lack T reg cells [118,119]. Later study showed that Foxp3 regulate the expression of miR-155 in T reg cells and deficiency of miR-155 results in decreased number, proliferation and fitness of T reg cells compared to wild type [120]. In a similar set of experiments, miR-146a-deficient hematopoietic cells failed to rescue Foxp3-deficient T-cell-mediated autoimmunity [121].

Although the role of miRNAs in B cell-mediated autoimmunity is less firmly established, it is likely that further discoveries in B cells are forthcoming. Finally, it has been shown in a pilot study that miRNA can be potentially used as biomarkers for diagnosis and prognosis of autoimmune diseases such as rheumatic diseases [122].

Conclusions

The advances in understanding the biological and pathological roles of miRNAs in B cells have been tremendous in the last few years. Despite this progress, there are many questions that remain. The first is how extensive are the networks that are controlled by miRNAs in B cells? Although some studies, including our own, have shown that a single or few targets may be critical at a given developmental stage, it remains to be delineated whether this is generally true or if there are multiple targets that a miRNA regulates. A second major question, which remains largely unexplored in mammalian systems, is how miRNA degradation is regulated. This will help define how gene expression programs may be regulated at one stage but not another by a given miRNA. Lastly, the utilization of miRNA-based therapeutics in B cell malignancies and inflammatory conditions is an area of active research. There are several avenues of promising work that suggests that we will be able to leverage miRNA-based pathways in treating these diseases, but current challenges include delivery into specific cell-types [123]. Research into viral vector-based delivery and into chemically modified small RNA sequences are particularly promising, and are likely to be the next frontiers.

Acknowledgements

In this review we focused on those miRNAs that have some functional role in B cell development or lymphoma, thus we acknowledge that we may not have included many papers describing profiling studies on miRNAs. We would like to thank Jorge Contreras for his help and support in the completion of this article. TF is a recipient of Developmental Hematology Training Grant T32HL086345-05 from the National Institute of Health (NIH). NIRM is a recipient of the Eugene V. Cota-Robles Fellowship from UCLA and of the Graduate Research Fellowship Program from the National Science Foundation (NSF). DSR is a Kimmel Scholar of the Sidney Kimmel Foundation for Cancer Research and received a career development award from the NIH (5 K08-CA133251).

Author details

[1]Department of Pathology and Laboratory Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA. [2]Cellular and Molecular Pathology Ph. D. Program, Department of Pathology and Laboratory Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA. [3]Jonsson Comprehensive Cancer Center, UCLA, 650 Charles E. Young Drive South, Factor 8-684, Los Angeles, CA 90095, USA. [4]Broad Stem Cell Research Center, UCLA, 650 Charles E. Young Drive South, Factor 12-272, Los Angeles, CA 90095, USA. [5]Division of Biology, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91106, USA.

Authors' contributions

TF and NIRM participated in its design and coordination and drafted the manuscript. DSR conceived of the review, and participated in its design and coordination and drafted the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 16 February 2012 Accepted: 8 March 2012
Published: 8 March 2012

References

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.** *Nature* 1998, 391:806-811.
2. Lee RC, Feinbaum RL, Ambros V: **The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.** *Cell* 1993, 75:843-854.

3. Ruvkun G, Giusto J: **The Caenorhabditis elegans heterochronic gene lin-14 encodes a nuclear protein that forms a temporal developmental switch.** *Nature* 1989, **338**:313-319.
4. Bernstein E, Caudy AA, Hammond SM, Hannon GJ: **Role for a bidentate ribonuclease in the initiation step of RNA interference.** *Nature* 2001, **409**:363-366.
5. Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ: **Argonaute2, a link between genetic and biochemical analyses of RNAi.** *Science (New York, NY)* 2001, **293**:1146-1150.
6. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH: **Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans.** *Genes Dev* 2001, **15**:2654-2659.
7. Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D: **MicroRNAs and other tiny endogenous RNAs in C. elegans.** *Curr Biol* 2003, **13**:807-818.
8. Wightman B, Ha I, Ruvkun G: **Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.** *Cell* 1993, **75**:855-862.
9. Kim VN, Han J, Siomi MC: **Biogenesis of small RNAs in animals.** *Nat Rev Mol Cell Biol* 2009, **10**:126-139.
10. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN: **MicroRNA genes are transcribed by RNA polymerase II.** *EMBO J* 2004, **23**:4051-4060.
11. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN: **The nuclear RNase III Drosha initiates microRNA processing.** *Nature* 2003, **425**:415-419.
12. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R: **The Microprocessor complex mediates the genesis of microRNAs.** *Nature* 2004, **432**:235.
13. Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U: **Nuclear export of microRNA precursors.** *Science (New York, NY)* 2004, **303**:95-98.
14. Bartel DP: **MicroRNAs: target recognition and regulatory functions.** *Cell* 2009, **136**:215-233.
15. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP: **MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing.** *Mol Cell* 2007, **27**:91-105.
16. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.** *Cell* 2005, **120**:15-20.
17. Brennecke J, Stark A, Russell RB, Cohen SM: **Principles of microRNA-target recognition.** *PLoS biology* 2005, **3**:e85.
18. Czech B, Hannon GJ: **Small RNA sorting: matchmaking for Argonautes.** *Nat Rev Genet* 2011, **12**:19-31.
19. Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, *et al*: **A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity.** *Science (New York, NY)* 2010, **328**:1694-1698.
20. Carmell MA, Xuan Z, Zhang MQ, Hannon GJ: **The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis.** *Genes Dev* 2002, **16**:2733-2742.
21. Sasaki T, Shiohama A, Minoshima S, Shimizu N: **Identification of eight members of the Argonaute family in the human genome small star, filled.** *Genomics* 2003, **82**:323-330.
22. Xu S, Guo K, Zeng Q, Huo J, Lam KP: **The RNase III enzyme Dicer is essential for germinal center B-cell formation.** *Blood* 2012, **119**:767-776.
23. Belver L, de Yebenes VG, Ramiro AR: **MicroRNAs prevent the generation of autoreactive antibodies.** *Immunity* 2010, **33**:713-722.
24. Fabbri M, Croce CM: **Role of microRNAs in lymphoid biology and disease.** *Current Opinion in Hematology* 2011, **18**:266-272.
25. Shaffer AL, Lin KI, Kuo TC, Yu X, Hurt EM, Rosenwald A, Giltnane JM, Yang L, Zhao H, Calame K, Staudt LM: **Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program.** *Immunity* 2002, **17**:51-62.
26. Hardy RR, Hayakawa K: **B cell development pathways.** *Annu Rev Immunol* 2001, **19**:595-621.
27. Perez-Vera P, Reyes-Leon A, Fuentes-Panana EM: **Signaling proteins and transcription factors in normal and malignant early B cell development.** *Bone marrow research* 2011, **2011**:502751.
28. Baltimore D, Boldin MP, O'Connell RM, Rao DS, Taganov KD: **MicroRNAs: new regulators of immune cell development and function.** *Nature Immunology* 2008, **9**:839-845.
29. Nagasawa T: **Microenvironmental niches in the bone marrow required for B-cell development.** *Nature reviews Immunology* 2006, **6**:107-116.
30. Arawaka S, Wada M, Goto S, Karube H, Sakamoto M, Ren CH, Koyama S, Nagasawa H, Kimura H, Kawanami T, *et al*: **The role of G-protein-coupled receptor kinase 5 in pathogenesis of sporadic Parkinson's disease.** *The Journal of neuroscience: the official journal of the Society for Neuroscience* 2006, **26**:9227-9238.
31. Nussenzweig RS: **Human trials of malaria vaccine.** *Science (New York, NY)* 1987, **236**:763.
32. Murphy K: **Janeway's Immunobiology.** *Book Janeway's Immunobiology*. 8 edition. City: Garland Science, Taylor & Francis Group, LLC; 2012.
33. Nemazee D, Buerki K: **Clonal deletion of autoreactive B lymphocytes in bone marrow chimeras.** *Proc Natl Acad Sci USA* 1989, **86**:8039-8043.
34. Nutt SL, Kee BL: **The transcriptional regulation of B cell lineage commitment.** *Immunity* 2007, **26**:715-725.
35. Hu H, Wang B, Borde M, Nardone J, Maika S, Allred L, Tucker PW, Rao A: **Foxp1 is an essential transcriptional regulator of B cell development.** *Nature Immunology* 2006, **7**:819-826.
36. Kikuchi K, Kondo M: **Developmental switch of mouse hematopoietic stem cells from fetal to adult type occurs in bone marrow after birth.** *Proc Natl Acad Sci USA* 2006, **103**:17852-17857.
37. Yao Z, Cui Y, Watford WT, Bream JH, Yamaoka K, Hissong BD, Li D, Durum SK, Jiang Q, Bhandoola A, *et al*: **Stat5a/b are essential for normal lymphoid development and differentiation.** *Proc Natl Acad Sci USA* 2006, **103**:1000-1005.
38. Hoelbl A, Kovacic B, Kerenyi MA, Simma O, Warsch W, Cui Y, Beug H, Hennighausen L, Moriggl R, Sexl V: **Clarifying the role of Stat5 in lymphoid development and Abelson-induced transformation.** *Blood* 2006, **107**:4898-4906.
39. Bain G, Maandag EC, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC, Krop I, Schlissel MS, Feeney AJ, van Roon M, *et al*: **E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements.** *Cell* 1994, **79**:885-892.
40. Cobaleda C, Schebesta A, Delogu A, Busslinger M: **Pax5: the guardian of B cell identity and function.** *Nature Immunology* 2007, **8**:463-470.
41. Lin YC, Jhunjhunwala S, Benner C, Heinz S, Welinder E, Mansson R, Sigvardsson M, Hagman J, Espinoza CA, Dutkowski J, *et al*: **A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate.** *Nature Immunology* 2010, **11**:635-643.
42. O'Riordan M, Grosschedl R: **Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A.** *Immunity* 1999, **11**:21-31.
43. Rao DS, O'Connell RM, Chaudhuri AA, Garcia-Flores Y, Geiger TL, Baltimore D: **MicroRNA-34a perturbs B lymphocyte development by repressing the forkhead box transcription factor Foxp1.** *Immunity* 2010, **33**:48-59.
44. Look AT: **Oncogenic transcription factors in the human acute leukemias.** *Science* 1997, **278**:1059-1064.
45. Smith E, Sigvardsson M: **The roles of transcription factors in B lymphocyte commitment, development, and transformation.** *J Leukoc Biol* 2004, **75**:973-981.
46. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, Girtman K, Mathew S, Ma J, Pounds SB, *et al*: **Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia.** *Nature* 2007, **446**:758-764.
47. Familiades J, Bousquet M, Lafage-Pochitaloff M, Bene MC, Beldjord K, De Vos J, Dastugue N, Coyaud E, Struski S, Quelen C, *et al*: **PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: a GRAALL study.** *Leukemia* 2009, **23**:1989-1998.
48. Bhojwani D, Howard SC, Pui CH: **High-risk childhood acute lymphoblastic leukemia.** *Clin Lymphoma Myeloma* 2009, **9**(Suppl 3):S222-230.
49. Mullighan CG, Miller CB, Radtke I, Phillips LA, Dalton J, Ma J, White D, Hughes TP, Le Beau MM, Pui CH, *et al*: **BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros.** *Nature* 2008, **453**:110-114.
50. Sun L, Heerema N, Crotty L, Wu X, Navara C, Vassilev A, Sensel M, Reaman GH, Uckun FM: **Expression of dominant-negative and mutant isoforms of the antileukemic transcription factor Ikaros in infant acute lymphoblastic leukemia.** *Proc Natl Acad Sci USA* 1999, **96**:680-685.
51. Nakase K, Ishimaru F, Avitahl N, Dansako H, Matsuo K, Fujii K, Sezaki N, Nakayama H, Yano T, Fukuda S, *et al*: **Dominant negative isoform of the Ikaros gene in patients with adult B-cell acute lymphoblastic leukemia.** *Cancer Res* 2000, **60**:4062-4065.

52. Olivero S, Maroc C, Bellard E, Gabert J, Nietfeld W, Chabannon C, Tonnelle C: **Detection of different Ikaros isoforms in human leukaemias using real-time quantitative polymerase chain reaction.** *Br J Haematol* 2000, **110**:826-830.
53. Nishii K, Katayama N, Miwa H, Shikami M, Usui E, Masuya M, Araki H, Lorenzo F, Ogawa T, Kyo T, *et al*: **Non-DNA-binding Ikaros isoform gene expressed in adult B-precursor acute lymphoblastic leukemia.** *Leukemia* 2002, **16**:1285-1292.
54. Takanashi M, Yagi T, Imamura T, Tabata Y, Morimoto A, Hibi S, Ishii E, Imashuku S: **Expression of the Ikaros gene family in childhood acute lymphoblastic leukaemia.** *Br J Haematol* 2002, **117**:525-530.
55. Meleshko AN, Movchan LV, Belevtsev MV, Savitskaja TV: **Relative expression of different Ikaros isoforms in childhood acute leukemia.** *Blood Cells Mol Dis* 2008, **41**:278-283.
56. Mullighan CG, Downing JR: **Global genomic characterization of acute lymphoblastic leukemia.** *Semin Hematol* 2009, **46**:3-15.
57. Georgopoulos K, Bigby M, Wang JH, Molnar A, Wu P, Winandy S, Sharpe A: **The Ikaros gene is required for the development of all lymphoid lineages.** *Cell* 1994, **79**:143-156.
58. Morgan B, Sun L, Avitahl N, Andrikopoulos K, Ikeda T, Gonzales E, Wu P, Neben S, Georgopoulos K: **Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation.** *EMBO J* 1997, **16**:2004-2013.
59. Streubel B, Vinatzer U, Lamprecht A, Raderer M, Chott A: **T(3;14)(p14.1;q32) involving IGH and FOXP1 is a novel recurrent chromosomal aberration in MALT lymphoma.** *Leukemia* 2005, **19**:652-658.
60. Wlodarska I, Veyt E, De Paepe P, Vandenberghe P, Nooijen P, Theate I, Michaux L, Sagaert X, Marynen P, Hagemeijer A, De Wolf-Peeters C: **FOXP1, a gene highly expressed in a subset of diffuse large B-cell lymphoma, is recurrently targeted by genomic aberrations.** *Leukemia* 2005, **19**:1299-1305.
61. Montecino-Rodriguez E, Dorshkind K: **B-1 B cell development in the fetus and adult.** *Immunity* 2012, **36**:13-21.
62. Shapiro-Shelef M, Calame K: **Regulation of plasma-cell development.** *Nature reviews Immunology* 2005, **5**:230-242.
63. Klein U, Tu Y, Stolovitzky GA, Keller JL, Haddad J Jr, Miljkovic V, Cattoretti G, Califano A, Dalla-Favera R: **Transcriptional analysis of the B cell germinal center reaction.** *Proc Natl Acad Sci USA* 2003, **100**:2639-2644.
64. Odegard VH, Schatz DG: **Targeting of somatic hypermutation.** *Nature reviews Immunology* 2006, **6**:573-583.
65. McHeyzer-Williams LJ, McHeyzer-Williams MG: **Antigen-specific memory B cell development.** *Annu Rev Immunol* 2005, **23**:487-513.
66. Klein U, Casola S, Cattoretti G, Shen Q, Lia M, Mo T, Ludwig T, Rajewsky K, Dalla-Favera R: **Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination.** *Nature Immunology* 2006, **7**:773-782.
67. Cambier JC, Gauld SB, Merrell KT, Vilen BJ: **B-cell anergy: from transgenic models to naturally occurring anergic B cells?** *Nat Rev Immunol* 2007, **7**:633-643.
68. Bartel DP, Chen CZ: **Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs.** *Nat Rev Genet* 2004, **5**:396-400.
69. Chen CZ, Li L, Lodish HF, Bartel DP: **MicroRNAs modulate hematopoietic lineage differentiation.** *Science (New York, NY* 2004, **303**:83-86.
70. Koralov SB, Muljo SA, Galler GR, Krek A, Chakraborty T, Kanellopoulou C, Jensen K, Cobb BS, Merkenschlager M, Rajewsky N, Rajewsky K: **Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage.** *Cell* 2008, **132**:860-874.
71. Xiao C, Srinivasan L, Calado DP, Patterson HC, Zhang B, Wang J, Henderson JM, Kutok JL, Rajewsky K: **Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes.** *Nature Immunology* 2008, **9**:405-414.
72. Xiao C, Calado DP, Galler G, Thai TH, Patterson HC, Wang J, Rajewsky N, Bender TP, Rajewsky K: **MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb.** *Cell* 2007, **131**:146-159.
73. Zhou B, Wang S, Mayr C, Bartel DP, Lodish HF: **miR-150, a microRNA expressed in mature B and T cells, blocks early B cell development when expressed prematurely.** *Proc Natl Acad Sci USA* 2007, **104**:7080-7085.
74. Xiao C, Rajewsky K: **MicroRNA control in the immune system: basic principles.** *Cell* 2009, **136**:26-36.
75. Meffre E, Casellas R, Nussenzweig MC: **Antibody regulation of B cell development.** *Nature Immunology* 2000, **1**:379-385.
76. He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, *et al*: **A microRNA component of the p53 tumour suppressor network.** *Nature* 2007, **447**:1130-1134.
77. Chang T-C, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KwangÂ H, Feldmann G, Yamakuchi M, Ferlito M, Lowenstein CJ, *et al*: **Transactivation of miR-34a by p53 Broadly Influences Gene Expression and Promotes Apoptosis.** *Molecular Cell* 2007, **26**:745-752.
78. Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, Zhai Y, Giordano TJ, Qin ZS, Moore BB, *et al*: **p53-Mediated Activation of miRNA34 Candidate Tumor-Suppressor Genes.** *Curr Biol* 2007, **17**:1298-1307.
79. O'Connell RM, Chaudhuri AA, Rao DS, Gibson WS, Balazs AB, Baltimore D: **MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output.** *Proc Natl Acad Sci USA* 2010, **107**:14235-14240.
80. Georgantas RW, Hildreth R, Morisot S, Alder J, Liu CG, Heimfeld S, Calin GA, Croce CM, Civin CI: **CD34+ hematopoietic stem-progenitor cell microRNA expression and function: a circuit diagram of differentiation control.** *Proc Natl Acad Sci USA* 2007, **104**:2750-2755.
81. Gibcus JH, Tan LP, Harms G, Schakel RN, de Jong D, Blokzijl T, Moller P, Poppema S, Kroesen BJ, van den Berg A: **Hodgkin lymphoma cell lines are characterized by a specific miRNA expression profile.** *Neoplasia (New York, NY* 2009, **11**:167-176.
82. O'Connell RM, Taganov KD, Boldin MP, Cheng G, Baltimore D: **MicroRNA-155 is induced during the macrophage inflammatory response.** *Proc Natl Acad Sci USA* 2007, **104**:1604-1609.
83. Taganov KD, Boldin MP, Chang KJ, Baltimore D: **NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses.** *Proc Natl Acad Sci USA* 2006, **103**:12481-12486.
84. Rodriguez A, Vigorito E, Clare S, Warren MV, Couttet P, Soond DR, van Dongen S, Grocock RJ, Das PP, Miska EA, *et al*: **Requirement of bic/microRNA-155 for normal immune function.** *Science (New York, NY* 2007, **316**:608-611.
85. Vigorito E, Perks KL, Abreu-Goodger C, Bunting S, Xiang Z, Kohlhaas S, Das PP, Miska EA, Rodriguez A, Bradley A, *et al*: **microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells.** *Immunity* 2007, **27**:847-859.
86. Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R, Papavasiliou FN: **MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase.** *Immunity* 2008, **28**:621-629.
87. Dorsett Y, McBride KM, Jankovic M, Gazumyan A, Thai TH, Robbiani DF, Di Virgilio M, San-Martin BR, Heidkamp G, Schwickert TA, *et al*: **MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation.** *Immunity* 2008, **28**:630-638.
88. de Yebenes VG, Belver L, Pisano DG, Gonzalez S, Villasante A, Croce C, He L, Ramiro AR: **miR-181b negatively regulates activation-induced cytidine deaminase in B cells.** *The Journal of Experimental Medicine* 2008, **205**:2199-2206.
89. Boldin MP, Taganov KD, Rao DS, Yang L, Zhao JL, Kalwani M, Garcia-Flores Y, Luong M, Devrekanli A, Xu J, *et al*: **miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice.** *The Journal of Experimental Medicine* 2011, **208**:1189-1201.
90. Cameron JE, Yin Q, Fewell C, Lacey M, McBride J, Wang X, Lin Z, Schaefer BC, Flemington EK: **Epstein-Barr virus latent membrane protein 1 induces cellular MicroRNA miR-146a, a modulator of lymphocyte signaling pathways.** *J Virol* 2008, **82**:1946-1958.
91. Kawai T, Akira S: **Signaling to NF-kappaB by Toll-like receptors.** *Trends Mol Med* 2007, **13**:460-469.
92. Zhao JL, Rao DS, Boldin MP, Taganov KD, O'Connell RM, Baltimore D: **NF-kappaB dysregulation in microRNA-146a-deficient mice drives the development of myeloid malignancies.** *Proc Natl Acad Sci USA* 2011, **108**:9184-9189.
93. Gururajan M, Haga CL, Das S, Leu CM, Hodson D, Josson S, Turner M, Cooper MD: **MicroRNA 125b inhibition of B cell differentiation in germinal centers.** *Int Immunol* 2010, **22**:583-592.
94. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, *et al*: **A MicroRNA Signature Associated with**

Prognosis and Progression in Chronic Lymphocytic Leukemia. *The New England Journal of Medicine* 2005, 353:1793-1801.
95. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, *et al*: **Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.** *Proc Natl Acad Sci USA* 2002, 99:15524-15529.
96. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM: **Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, 101:2999-3004.
97. Mi S, Li Z, Chen P, He C, Cao D, Elkahloun A, Lu J, Pelloso LA, Wunderlich M, Huang H, *et al*: **Aberrant overexpression and function of the miR-17-92 cluster in MLL-rearranged acute leukemia.** *Proc Natl Acad Sci USA* 2010, 107:3710-3715.
98. Costinean S, Zanesi N, Pekarsky Y, Tili E, Volinia S, Heerema N, Croce CM: **Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice.** *Proc Natl Acad Sci USA* 2006, 103:7024-7029.
99. 99. Costinean S, Sandhu SK, Pedersen IM, Tili E, Trotta R, Perrotti D, Ciarlariello D, Neviani P, Harb J, Kauffman LR, *et al*: **Src homology 2 domain-containing inositol-5-phosphatase and CCAAT enhancer-binding protein beta are targeted by miR-155 in B cells of Emicro-MIR-155 transgenic mice.** *Blood* 2009, 114:1374-1382.
100. Garzon R, Volinia S, Liu CG, Fernandez-Cymering C, Palumbo T, Pichiorri F, Fabbri M, Coombes K, Alder H, Nakamura T, *et al*: **MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia.** *Blood* 2008, 111:3183-3189.
101. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M, *et al*: **miR-15 and miR-16 induce apoptosis by targeting BCL2.** *Proc Natl Acad Sci USA* 2005, 102:13944-13949.
102. Medina PP, Nolde M, Slack FJ: **OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma.** *Nature* 2010, 467:86-90.
103. Fabbri M, Bottoni A, Shimizu M, Spizzo R, Nicoloso MS, Rossi S, Barbarotto E, Cimmino A, Adair B, Wojcik SE, *et al*: **Association of a microRNA/TP53 feedback circuitry with pathogenesis and outcome of B-cell chronic lymphocytic leukemia.** *JAMA* 2011, 305:59-67.
104. Zenz T, Mohr J, Eldering E, Kater AP, Buhler A, Kienle D, Winkler D, Durig J, van Oers MH, Mertens D, *et al*: **miR-34a as part of the resistance network in chronic lymphocytic leukemia.** *Blood* 2009, 113:3801-3808.
105. Chiorazzi N, Rai KR, Ferrarini M: **Chronic lymphocytic leukemia.** *N Engl J Med* 2005, 352:804-815.
106. Liu Y, Corcoran M, Rasool O, Ivanova G, Ibbotson R, Grander D, Iyengar A, Baranova A, Kashuba V, Merup M, *et al*: **Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia.** *Oncogene* 1997, 15:2463-2473.
107. 107. Migliazza A, Bosch F, Komatsu H, Cayanis E, Martinotti S, Toniato E, Guccione E, Qu X, Chien M, Murty VV, *et al*: **Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia.** *Blood* 2001, 97:2098-2104.
108. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, Ambesi-Impiombato A, Califano A, Migliazza A, Bhagat G, Dalla-Favera R: **The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia.** *Cancer Cell* 2010, 17:28-40.
109. Lawrie CH, Soneji S, Marafioti T, Cooper CD, Palazzo S, Paterson JC, Cattan H, Enver T, Mager R, Boultwood J, *et al*: **MicroRNA expression distinguishes between germinal center B cell-like and activated B cell-like subtypes of diffuse large B cell lymphoma.** *International Journal of Cancer* 2007, 121:1156-1161.
110. Ota A, Tagawa H, Karnan S, Tsuzuki S, Karpas A, Kira S, Yoshida Y, Seto M: **Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma.** *Cancer Res* 2004, 64:3087-3095.
111. 111. Nie K, Gomez M, Landgraf P, Garcia JF, Liu Y, Tan LH, Chadburn A, Tuschl T, Knowles DM, Tam W: **MicroRNA-mediated down-regulation of PRDM1/Blimp-1 in Hodgkin/Reed-Sternberg cells: a potential pathogenetic lesion in Hodgkin lymphomas.** *The American Journal of Pathology* 2008, 173:242-252.
112. 112. Pichiorri F, Suh SS, Rocci A, De Luca L, Taccioli C, Santhanam R, Zhou W, Benson DM Jr, Hofmainster C, Alder H, *et al*: **Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development.** *Cancer Cell* 2010, 18:367-381.
113. Iborra M, Bernuzzi F, Invernizzi P, Danese S: **MicroRNAs in autoimmunity and inflammatory bowel disease: Crucial regulators in immune response.** *Autoimmunity reviews* 2010, 11:305-314.
114. Rromblom L, Alcaranta ML, Alm GV: **The type I interferon system in systemic lupus erythematosus.** *Arthritis and rheumatism* 2006, 54:408-420.
115. Pascual V, Farkas L, Banchereau J: **Systemic lupus erythematosus: all roads lead to type I interferons.** *Curr Opin Immunol* 2006, 18:676-682.
116. Crow MK: **Type I interferon in systemic lupus erythematosus.** *Current topics in microbiology and immunology* 2007, 316:359-386.
117. Tang Y, Luo X, Cui H, Ni X, Yuan M, Guo Y, Huang X, Zhou H, de Vries N, Tak PP, *et al*: **MicroRNA-146A contributes to abnormal activation of the type I interferon pathway in human lupus by targeting the key signaling proteins.** *Arthritis and rheumatism* 2009, 60:1065-1075.
118. Rudensky AY: **Regulatory T cells and Foxp3.** *Immunol Rev* 2011, 241:260-268.
119. Liston A, Lu LF, O'Carroll D, Tarakhovsky A, Rudensky AY: **Dicer-dependent microRNA pathway safeguards regulatory T cell function.** *The Journal of experimental medicine* 2008, 205:1993-2004.
120. Lu LF, Thai TH, Calado DP, Chaudhry A, Kubo M, Tanaka K, Loeb GB, Lee H, Yoshimura A, Rajewsky K, Rudensky AY: **Foxp3-dependent microRNA155 confers competitive fitness to regulatory T cells by targeting SOCS1 protein.** *Immunity* 2009, 30:80-91.
121. Lu LF, Boldin MP, Chaudhry A, Lin LL, Taganov KD, Hanada T, Yoshimura A, Baltimore D, Rudensky AY: **Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses.** *Cell* 2010, 142:914-929.
122. Alevizos I, Illei GG: **MicroRNAs as biomarkers in rheumatic diseases.** *Nature reviews Rheumatology* 2010, 6:391-398.
123. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M: **Silencing of microRNAs in vivo with 'antagomirs'.** *Nature* 2005, 438:685-689.

doi:10.1186/1756-8722-5-7
Cite this article as: Fernando *et al*.: MicroRNAs in B cell development and malignancy. *Journal of Hematology & Oncology* 2012 5:7.

APPENDIX II:

MicroRNA-146a Modulates B-cell Oncogenesis by Regulating Egr1

(reprint)

MicroRNA-146a modulates B-cell oncogenesis by regulating Egr1

Jorge R. Contreras[1,2], Jayanth Kumar Palanichamy[1], Tiffany M. Tran[1], Thilini R. Fernando[1], Norma I. Rodriguez-Malave[1,2], Neha Goswami[1], Valerie A. Arboleda[1], David Casero[1], Dinesh S. Rao[1,3,4]

[1]Department of Pathology and Laboratory Medicine, UCLA

[2]Cellular and Molecular Pathology Ph.D. Program, UCLA

[3]Jonsson Comprehensive Cancer Center, UCLA

[4]Broad Stem Cell Research Center, UCLA

Correspondence to:
Dinesh S. Rao, **e-mail:** drao@mednet.ucla.edu

Keywords: microRNA, B-cell, lymphoma, leukemia, c-Myc

Received: January 26, 2015 *Accepted:* February 24, 2015 *Published:* April 13, 2015

ABSTRACT

miR-146a is a NF-κB induced microRNA that serves as a feedback regulator of this critical pathway. In mice, deficiency of miR-146a results in hematolymphoid cancer at advanced ages as a consequence of constitutive NF-κB activity. In this study, we queried whether the deficiency of miR-146a contributes to B-cell oncogenesis. Combining miR-146a deficiency with transgenic expression of c-Myc led to the development of highly aggressive B-cell malignancies. Mice transgenic for c-Myc and deficient for miR-146a were characterized by significantly shortened survival, increased lymph node involvement, differential involvement of the spleen and a mature B-cell phenotype. High-throughput sequencing of the tumors revealed significant dysregulation of approximately 250 genes. Amongst these, the transcription factor Egr1 was consistently upregulated in mice deficient for miR-146a. Interestingly, transcriptional targets of Egr1 were enriched in both the high-throughput dataset and in a larger set of miR-146a-deficient tumors. miR-146a overexpression led to downregulation of Egr1 and downstream targets with concomitant decrease in cell growth. Direct targeting of the human EGR1 by miR-146a was seen by luciferase assay. Together our findings illuminate a bona fide role for miR-146a in the modulation of B-cell oncogenesis and reveal the importance of understanding microRNA function in a cell- and disease-specific context.

INTRODUCTION

MicroRNAs (miRNAs) are a class of small non-coding RNAs, 21–22 nucleotides in length, which have physiological roles in many developmental systems [1]. miRNAs primarily act through post-transcriptional repression of target mRNAs via short complementary sequences in the 3'untranslated region (UTR) of mRNA transcripts [2, 3]. It has been reported that nearly 2000 miRNAs exist in the human genome and more than half of protein-coding genes are potential targets for miRNAs [4]. Both oncogenic and tumor suppressive miRNAs have been described in oncogenesis, acting via repression of tumor-suppressive and growth-promoting targets, respectively [5–8]. It is important to note, however, that miRNA regulation of gene expression is highly context-dependent: they regulate a cell type-specific transcriptome generated by a set of oncogenic or developmental transcriptional regulators. Hence, uncovering the oncogenic role of a miRNA requires the study of lineage specific transcriptional dysregulation.

miR-146a was discovered as a transcriptional target of the NF-κB pathway acting as a negative feedback regulator of this pathway and repressing some key components, such as *Traf6* and *Irak1* [9–13]. In line with its function in the NF-κB pathway, miR-146a deficiency in mice results in the development of a hyper inflammatory phenotype characterized by myeloid proliferation, lymphoid

hyperplasia, T-cell hyper activation and autoantibody production [9, 11, 14, 15]. Subsequently, aged knockout mice develop myeloid and lymphoid malignancies [9, 11]. These phenotypes are characterized by a dependence on constitutive NF-κB activity, as demonstrated by the correction of many phenotypes by deletion of elements of NF-κB signaling or downstream mediators [11, 16].

Constitutive NF-κB activity is a hallmark of many different types of cancer including B-cell malignancies [17]. The activated B-cell type of diffuse large B-cell lymphoma (ABC-DLBCL), which demonstrates constitutive NF-κB activation, is more aggressive and leads to worse outcomes in patients. Currently, several components of the NF-κB pathway have been found mutated in DLBCL, producing activation of NF-κB [18, 19]. The role of miR-146a as a negative regulator of this critical pathway, along with the development of B-cell malignancies in knockout mice, suggest that loss of miR-146a via undefined mechanisms may represent a pathogenetic event in B-cell malignancies that contributes to constitutive NF-κB activity.

In addition to being regulated by NF-κB, miR-146a has been shown to be positively regulated by the potent oncogene, *c-Myc*, in a melanoma cell line [20]. In contrast, primary samples of B-cell lymphoma with high levels of *c-Myc* expression show dramatic downregulation of miR-146a expression, and additional studies demonstrate negative regulation of miR-146a by c-Myc [21–23]. This led us to question the role that miR-146a plays in *c-Myc*-mediated oncogenesis in the B-cell lineage. Since *c-Myc* is a powerful transcriptional regulator with a specific transcriptome, we hypothesized that miR-146a mediated effects on the *c-Myc* gene expression program would reveal unique cancer relevant pathways. To test our hypotheses, we intercrossed the Eμ-Myc mouse with miR-146a-deficient animals. We found that miR-146a deficiency accelerates oncogenesis, decreases survival, and alters the differentiation stage of the tumors that are formed in the resulting mice. Histopathologic and flow cytometric analyses revealed a distinctive pattern of involvement in miR-146a-deficient animals. Mechanistically, few genes were significantly differentially regulated between wild-type and miR-146a deficient, *c-Myc* driven tumors. Of these, *Egr1* and its downstream mediators were identified as a novel pathway regulated by miR-146a in B-cells. Our findings promise to open up a new area of research and demonstrate a tumor suppressive function for miR-146a in B-cell oncogenesis.

RESULTS

miR-146a deficiency decreases survival of Eμ-Myc transgenic mice

Given the proposed roles for miR-146a in tumor suppression and negative feedback regulation of the NF-κB pathway, we examined whether miR-146a deficiency would synergize with *c-Myc* during B-cell oncogenesis. miR-146a deficient and Eμ-Myc transgenic mice were bred to yield cohorts of mice that carried the Eμ-Myc transgene with wild-type, heterozygous or homozygous knockout alleles of miR-146a (Figure 1a–1b). Most tumors that formed in Eμ-Myc mice showed a lymphoblastic morphology with numerous mitotic figures and apoptotic bodies on H&E sections (Figure 1c–1e). Conversely, tumors from the miR-146a-deficient mice demonstrated a more heterogeneous appearance. Many tumors had lymphoblastic morphology, but others showed a plasmacytoid appearance, including eosinophilic cytoplasmic concretions, suggestive of immunoglobulin deposits (Figure 1e inset shows cells with immunoglobulin concretions). Eμ-Myc miR-146a$^{+/-}$ mice did not have a significant reduction in their survival (Figure 1f). On the other hand, homozygous deficiency caused a decrease in survival from 104.5 days to 82.5 days (Figure 1f). Gender differences were noted, with female miR-146a$^{-/-}$ mice showing significant differences in survival, while males only showed a trend towards reduced survival (Supplementary Figure 1a–1d). Finally, virtually all mortality in both sets of mice was attributable to tumor formation (data not shown).

miR-146a deficient tumors demonstrate differential anatomic patterns of involvement

Anatomically, tumors in both sets of mice showed differential patterns of involvement of hematopoietic and lymphoid organs, with virtually all mice showing thymic involvement. 31% of Eμ-Myc miR-146a$^{+/+}$ did not show any lymph node involvement, whereas all of the miR-146a$^{-/-}$ did show involvement (Figure 2a). While the majority of mice in both groups showed small numbers of circulating tumor cells in the peripheral blood, 6/10 Eμ-Myc miR-146a$^{-/-}$ mice examined showed frank leukemia (defined as a white blood cell count of greater than 30,000/μL) (Figure 2b–2d). This was in contrast to the lower numbers of Eμ-Myc miR-146a$^{+/+}$ mice that demonstrated leukemia by blood counts (4/14). Amongst mice with predominantly solid tumors, miR-146a deficiency caused a statistically significant increase in peripheral blood CD11b+ myeloid cells but not in B220+ B-cells, CD3ε+ T-cells, hemoglobin or platelets (Supplementary Figure 2a–2f). This may represent the propensity of miR-146a-deficient hematopoietic progenitors to produce increased numbers of myeloid cells. Bone marrow analysis of these mice found similar proportions of myeloid cells, erythroid cells, and B- lymphocytes (Supplementary Figure 2g–2i).

Mice in both groups demonstrated enlarged spleens, with average weights of approximately 400 mg (SupplementaryFigure 3a). In Eμ-Myc miR-146a$^{+/+}$ mice there was involvement of the white pulp with contiguous spread between the lymphoid follicles (Figure 2e, dotted area). High power views showed the malignant cells in both the white and red pulp (Supplementary Figure 3b–3c). On the other hand, Eμ-Myc miR-146a$^{-/-}$ mice showed

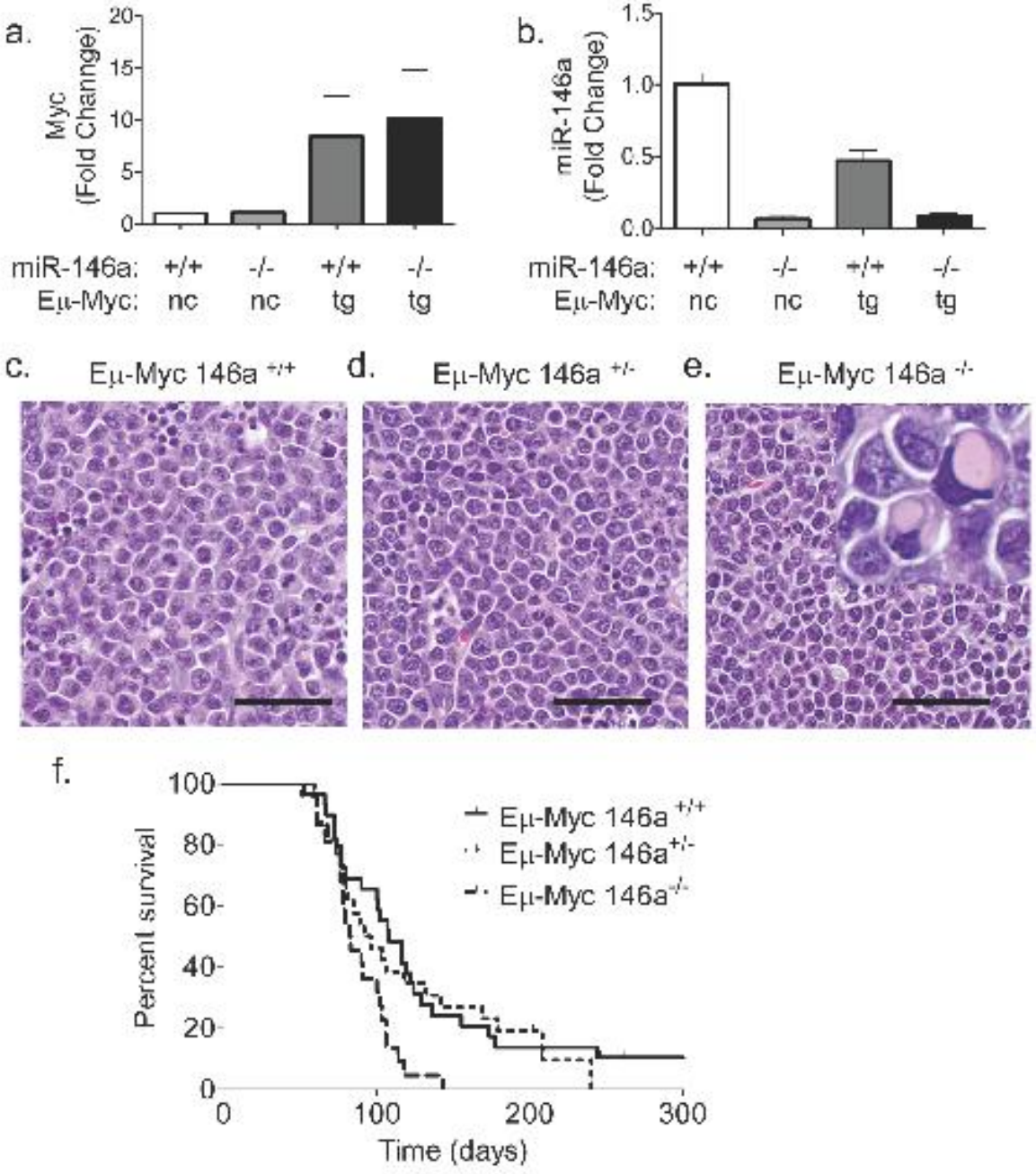


Figure 1: miR-146a deficiency causes increased mortality in Eμ-Myc mice. **(a)** RT-qPCR for c-Myc was performed on splenic B-cells from WT and miR-146a$^{+/-}$ mice or tumor samples from Eμ-Myc miR-146a$^{+/+}$and Eμ-Myc miR-146a$^{+/-}$ animals (n = 3, 3, 12, 11 respectively) (nc: non-carrier; tg: transgene). **(b)** RT-qPCR for miR-146a was performed from the same samples as for c-Myc. **(c–e)** Hematoxylin and Eosin (H&E) stained sections of lymph node tumors derived from Eμ-Myc miR-146a$^{+/+}$, Eμ-Myc miR-146a$^{+/-}$and Eμ-Myc miR-146a$^{-/-}$mice, respectively. The inset in (e) shows a subset of cells with immunoglobulin concretions. Scale bar for Figures 1c-e, 40 μm. **(f)** Kaplan Meier survival curve of mice with Eμ-Myc oncogene and wild-type, heterozygous, or homozygous deficiency of miR-146a (n = 26 for Eμ-Myc miR-146a$^{+/+}$(Solid line in graph), n = 23 for Eμ-Myc miR-146a$^{+/-}$ (Dotted line on graph), n = 22 for Eμ-Myc miR-146a$^{-/-}$ (Dashed line on graph);w.t. vs. het comparison, Log-Rank Test, p = 0.6725; w.t. vs. k.o. comparison, Log-Rank Test, p = 0.0027).

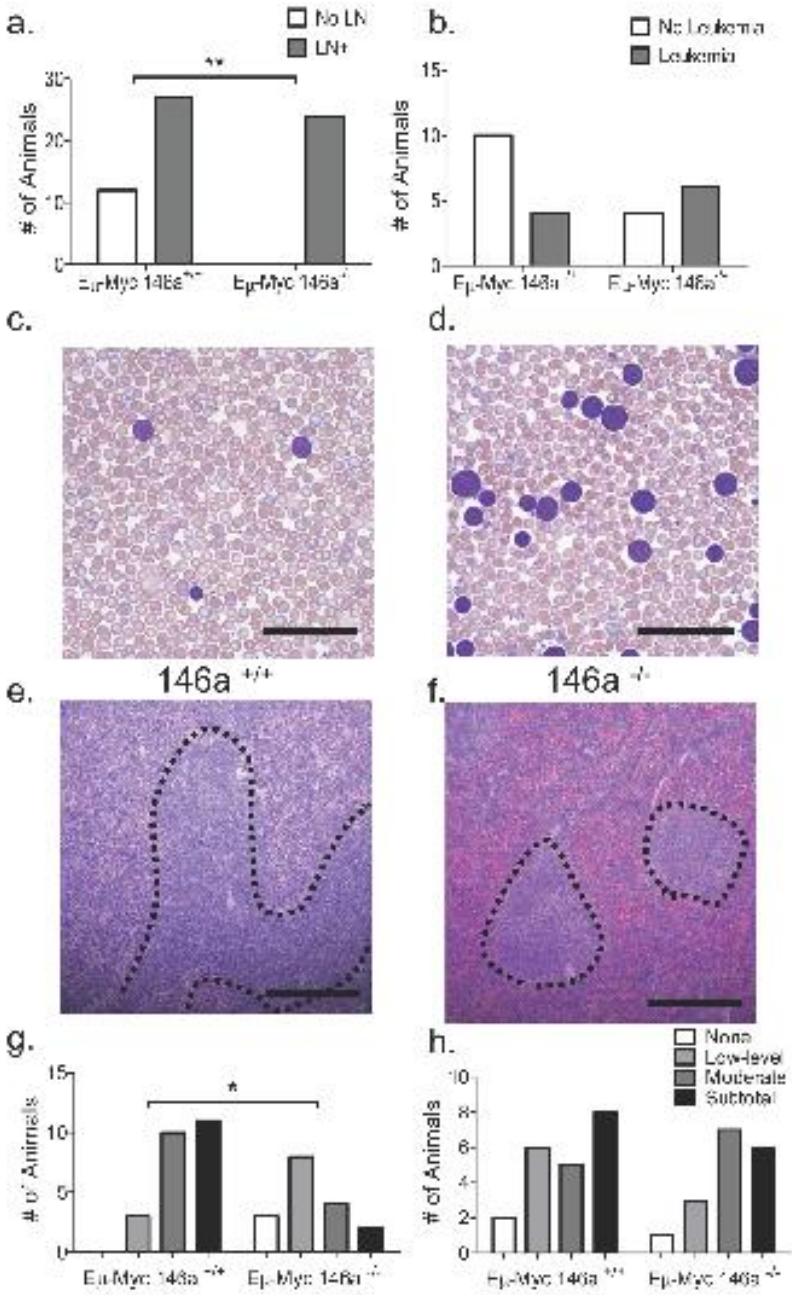


Figure 2: miR-146a deficient tumors show a differential pattern of anatomic involvement. (a) Macroscopic lymph node involvement is significantly increased in miR-146a$^{-/-}$ deficient mice ($n = 39$ for Eμ-Myc miR-146a$^{+/+}$ and $n = 24$ for Eμ-Myc miR-146a$^{-/-}$; Fisher's exact test, $p = 0.002$). **(b)** Quantitation of the incidence of leukemia, defined as a peripheral white blood cell count of greater than 30, 000/μL, in these mice, shows a trend towards statistical significance between the groups ($n = 14$ for Eμ-Myc miR-146a$^{+/+}$ and $n = 10$ for Eμ-Myc miR-146a$^{-/-}$; Fisher's exact test, $p = 0.21$). **(c–d)** Wright stained peripheral blood smears from mice with miR-146a-sufficient and deficient Eμ-Myc driven tumors. Scale bar, 40 μm. **(e)** H&E stained section of Eμ-Myc miR-146$^{+/+}$spleen, low power image. Dotted lines delineate expanded white pulp. Scale bar, 400 μm. **(f)** Low power view of an H&E stained section showing relative sparing of the white pulp (dotted lines) in Eμ-Myc miR-146a$^{-/-}$ mice. Scale bar, 400 μm. **(g)** Quantitation of splenic white pulp involvement in Eμ-Myc miR-146a$^{+/+}$and Eμ-Myc miR-146a$^{-/-}$ mice on an ordinal 4-point scale going from no involvement to subtotal involvement of the spleen ($n = 24$ for Eμ-Myc miR-146a$^{+/+}$$n = 17$ for Eμ-Myc miR-146a$^{-/-}$; Chi Square Test, $p = 0.004$). **(h)** Quantitation of splenic red pulp involvement in Eμ-Myc miR-146a$^{+/+}$and Eμ-Myc miR-146a$^{-/-}$ mice on an ordinal 4-point scale going from no involvement to subtotal involvement of the spleen ($n = 24$ for Eμ-Myc miR-146a$^{+/+}$$n = 17$ for Eμ-Myc miR-146a$^{-/-}$; Chi Square Test, $p = 0.671$).

extensive involvement of the red pulp with relative sparing of the white pulp (Figure 2f). Using a semi-quantitative 4-point scale to grade involvement, we found that white pulp involvement was significantly higher in Eμ-Myc miR-146a$^{+/+}$ mice compared to knockout mice (Figure 2g). Red pulp involvement was not different between the two groups (Figure 2h). Despite these differential patterns of involvement, the relative numbers of B-cells, T-cells and myeloid cells in the spleen were equivalent between the two groups of mice (Supplementary Figure 3d–3k). Together, the data suggest similar overall infiltration of the spleen (given similar weights and cellular composition), but a predilection for the red pulp when miR-146a is deficient, suggesting that the deficiency of miR-146a may change the homing properties of the malignant B-cells. The patterns of involvement are somewhat reminiscent of certain subtypes of B-cell lymphoma/leukemia that show peripheral blood involvement and red pulp involvement in the spleen, but are not correlated with NF-κB activity or histologic subtype in humans.

miR-146a deficient tumors demonstrate a mature B-cell phenotype

To further characterize the increased mortality seen in the miR-146a-deficient mice, we undertook immunophenotypic analyses. The tumors in both sets of mice were predominantly of B-cell phenotype (Figure 3a–3b). Similarly, Eμ-Myc mice with a heterozygous deficiency for miR-146a also developed B-cell tumors (Supplementary Figure 4a–4d). To examine the stage of differentiation, we examined expression of IgM, finding that greater than 70% of tumors from Eμ-Myc miR-146a$^{+/+}$ mice were IgM-. In contrast, only 42% of tumors from Eμ-Myc miR-146a$^{-/-}$ mice were IgM- (Figure 3c–3d). Amongst IgM- tumors, several were plasmacytic, shown by morphology and staining for CD138 (Figure 3e). Next, we dichotomized the data by mean fluorescent intensity (MFI), finding that CD138+ tumors were more frequently seen in the miR-146a-deficient background (Figure 3f). When we combined positivity for CD138 and IgM, most miR-146a-deficient tumors showed a mature B-cell phenotype (either IgM+ or CD138+) whereas miR-146a-sufficient tumors were negative for both IgM and CD138 (Figure 3). Tumor cells from miR-146a-deficient mice showed lower expression of memory B-cell/ activation related antigen, CD80 (Supplementary Figure 4e–4f), but similar expression of CD44 (Supplementary Figure 4g–4h). To further characterize the stage of B-cell differentiation in these tumors, we performed RT-qPCR to quantitate the expression of genes involved in B-cell differentiation. We found that transcripts for Blimp1, CD43, Bcl6 and Igh6 were all more highly expressed in tumors from miR-146a-deficient mice (Figure 3h–3k). Interestingly, female mice, which showed a statistically significant difference in survival, showed similar trends in their immunophenotypic profiles (Supplementary Figure 5a–5f) as well as in gene expression of maturation-related B-cell transcripts (Supplementary Figure 5g–5j) when compared to the group overall. Together these findings indicate that miR-146a deficient tumors are composed of malignant B-cells that derive from a different stage of differentiation than tumors sufficient for miR-146a.

miR-146a-deficient tumors show a limited difference in transcriptome expression, including many putative targets of miR-146a

To define a mechanistic basis for miR-146a deficient B-lymphomagenesis, we performed RNA-sequencing on four miR-146a sufficient and two miR-146a deficient tumors. Based on this comparison, we arrived at a list of 249 genes that were differentially regulated with an adjusted *p*-value of 0.05 or lower (Figure 4a). We then searched the dataset for miR-146a targets predicted by TargetScan [2, 3]. Of the differentially regulated genes, 53 genes are predicted to be miR-146a targets (Figure 4b). When we examined the genes that were upregulated, 29 out of 140 genes were predicted miR-146a targets (Figure 4c), and this did not represent a statistical enrichment. Next, we confirmed some of the findings by RT-qPCR in the larger set of tumor samples that we had collected. Four of the top ten genes from RNA sequencing had significantly different expression levels in the tumors when assayed by qPCR. These genes include *Jhy* and *Camk2b* (Figure 4d and 4f). *Jhy* is a recently described novel gene with no known function in oncogenesis or hematopoiesis; while *Camk2b* has a previously described putative role in epithelial cancer [24]. The gene *Dtx3*, which showed differential regulation by RNA-sequencing, was not differentially expressed in the larger set of tumor samples (Figure 4h). Other genes that were differentially expressed included the putative target *Egr1*, with *Nrp2* showing a trend towards differential expression (Figure 4e and 4g). A third predicted target, *Axl*, failed to show differential regulation by qPCR in this larger set of samples (Figure 4i). Hence, the transcriptome data provided us with a starting point for understanding tumorigenesis, uncovering putative miR-146a targets in the setting of B-cell oncogenesis.

The transcriptome regulated by EGR1 is differentially regulated in miR-146a-deficient tumors

The early growth response-1 gene (*Egr1*), has previously described functions in hematopoietic differentiation [25, 26]. Given that *Egr1* is overexpressed in miR-146a deficient tumors, we undertook an analysis to determine whether the *Egr1* transcriptome is differentially regulated. Using a publically available ChIP-Seq dataset, we gathered a list of EGR1 transcription factor binding

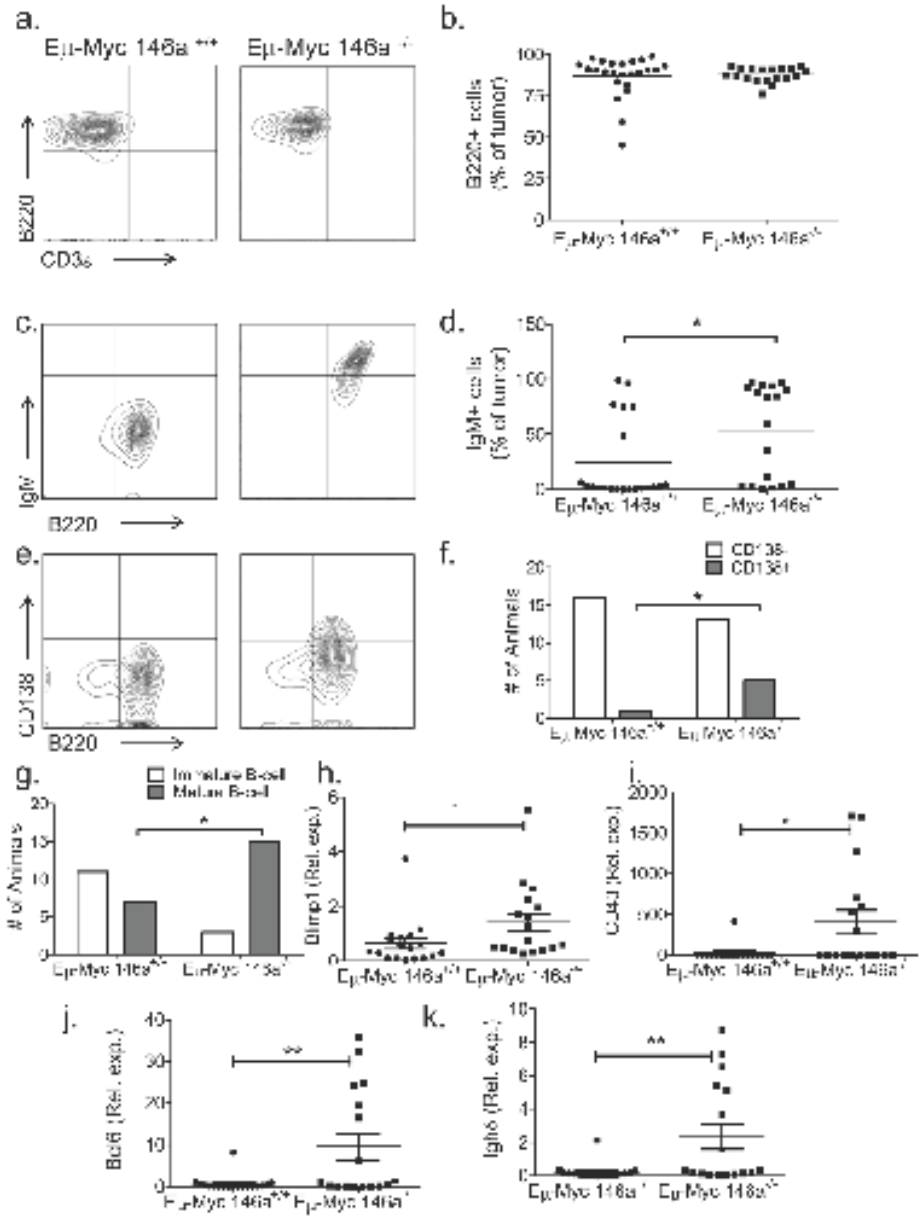


Figure 3: miR-146a deficiency causes a mature B-cell phenotype in Eμ-Myc mice. **(a)** Representative FACS plot showing staining for B220 and CD3ε in Eμ-Myc miR-146a$^{+/+}$ and Eμ-Myc miR-146a$^{-/-}$mice. **(b)** FACS shows that both groups of mice demonstrate B-cell tumors (n = 24 for Eμ-Myc miR-146a$^{+/+}$ and n = 19 for Eμ-Myc miR-146a$^{-/-}$; t-test p = 0.794). **(c)** Representative FACS plots stained for B220 and IgM from tumors derived from Eμ-Myc miR-146a$^{+/+}$ and Eμ-Myc miR-146a$^{-/-}$ mice. **(d)** Percentage of IgM positive cells in tumors from both cohorts showing increased IgM positivity in Eμ-Myc miR-146a$^{-/-}$ tumors (n = 25 for Eμ-Myc miR-146a$^{+/+}$ and n = 17 for Eμ-Myc miR-146a$^{-/-}$; t-test, p = 0.03). **(e)** Representative FACS plots stained for B220 and CD138 from tumors derived from Eμ-Myc miR-146a$^{+/+}$ and Eμ-Myc miR-146a$^{-/-}$ mice. **(f)** Dichotomized CD138 expression data (see methods for details on dichotomization), showing animals with CD138+ versus CD138- tumors (n = 18 for Eμ-Myc miR-146a$^{+/+}$ and n = 18 for Eμ-Myc miR-146a$^{-/-}$; Chi-square test, one-sided p = 0.04). **(g)** Tumors were dichotomized as being either immature (double negative for CD138 and IgM) or mature (having expression of either marker) (n = 18 for Eμ-Myc miR-146a$^{+/+}$ and n = 18 for Eμ-Myc miR-146a$^{-/-}$; Chi-square test, p = 0.0062). **(h–k)** RT-qPCR data for Blimp1, CD43, Bcl6, and Ighδ in tumors from Eμ-Myc miR-146a$^{+/+}$ (n = 18)and Eμ-Myc miR-146a$^{-/-}$ mice (n = 17). All comparisons showed statistically significant differences by T-test (p = 0.05 for Blimp1 (h), p = 0.014 for CD43 (i), p = 0.0070 for Bcl6 (j), and p = 0.0067 for Ighδ(k).

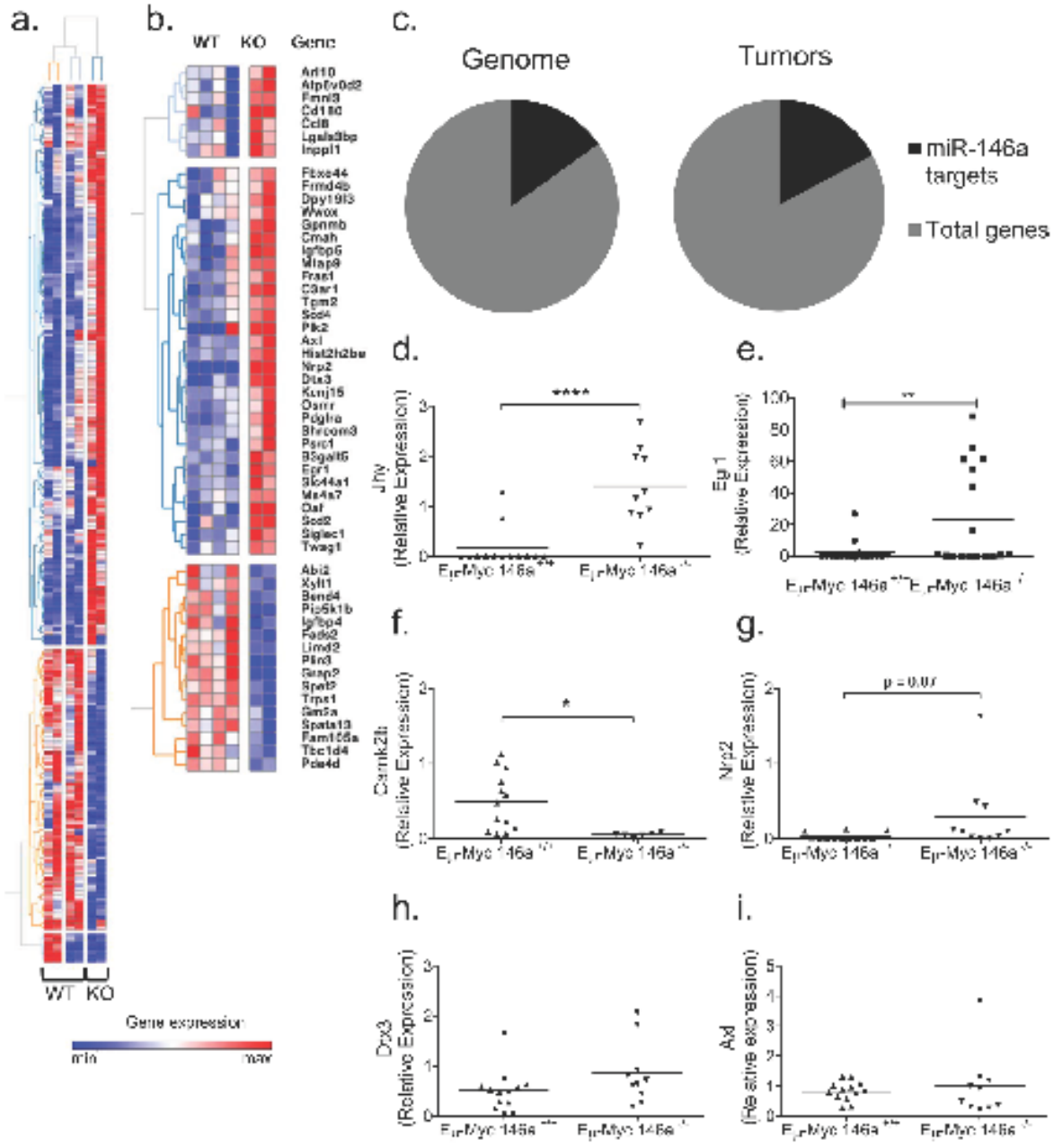


Figure 4: Gene expression analyses of Eμ-Myc driven tumors with miR-146a deficiency. **(a)** Genes differentially expressed between Eμ-Myc miR-146a$^{+/+}$ and Eμ-Myc miR-146a$^{-/-}$ tumors. **(b)** Differentially expressed genes with miR-146a sites in their UTR as predicted by the TargetScan algorithm. The heat map color scale represents, for each gene, the relative expression level using the average mean gene expression as a reference. **(c)** Graphical representation of the percentage of the genome (left) predicted to be targeted by miR-146a (2773 predicted targets out of 18, 393 annotated UTRs), compared with the percentage of upregulated genes in the tumor dataset (right) that are predicted miR-146a targets (29 predicted targets out of 169 upregulated genes). No statistically significant enrichment was found (Chi-square test). **(d–i)** RT-qPCR of genes found to be differentially regulated by RNA-sequencing analysis, including *Jhy* (d; *t*-test, $p < 0.0001$), *Egr1* (e; *t*-test, $p = 0.0086$), *Camk2b* (f; *t*-test, $p = 0.01$), *Nrp2* (g; *t*-test, $p = 0.0743$), *Dtx3* (h; *t*-test, $p = 0.127$) and *Axl* (i; *t*-test, $p = 0.616$). The three genes on the left (*Jhy, Camk2b and Dtx3*) were amongst the most differentially regulated genes between the miR-146a sufficient and deficient tumors by RNA sequencing. The three genes on the right (*Egr1, Nrp2 and Axl*) represent putative targets of miR-146a (For RT-qPCR analysis $n = 13$ for Eμ-Myc miR-146a$^{+/+}$and $n = 10$ for Eμ-Myc miR-146a$^{-/-}$ for panels d, g, h, i; $n = 13$ for Eμ-Myc miR-146a$^{+/+}$and $n = 6$ for Eμ-Myc miR-146a$^{-/-}$ for panel f; $n = 18$ for Eμ-Myc miR-146a$^{+/+}$ and $n = 17$ for Eμ-Myc miR-146a$^{-/-}$ for panel e).

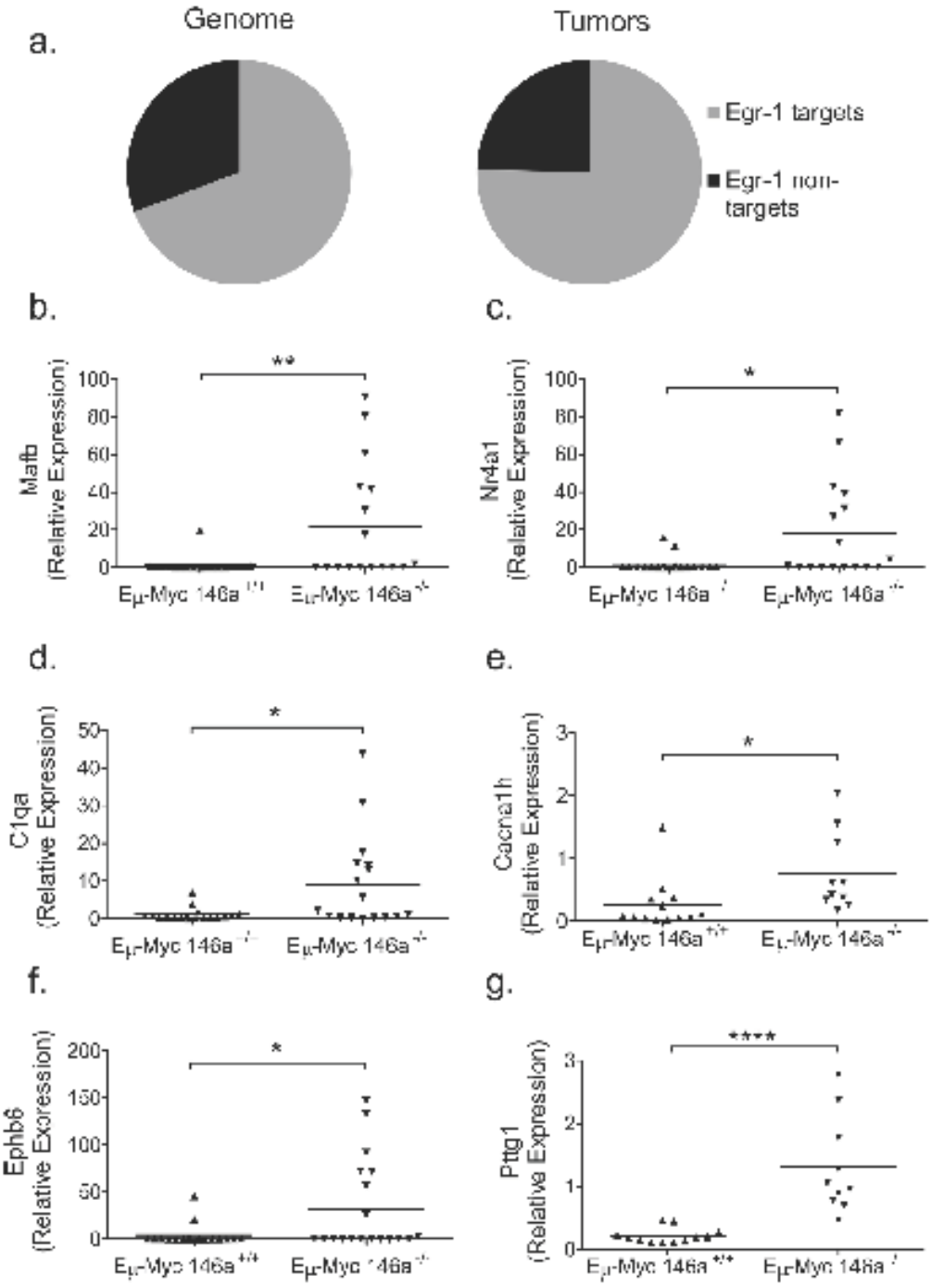


Figure 5: Genes with EGR1 transcription factor binding sites are enriched in Eμ-Myc driven tumors with miR-146a deficiency. (a) Graphical representation of the percentage of the genome (left) containing EGR1 TFBS (11136 TFBS out of 16288 genes that were conserved between human and mouse), compared with the percentage of differentially regulated genes in the tumor dataset (right) that show EGR1 TFBS (153 with TFBS out of 203 differentially regulated genes that are conserved between mouse and human). This difference was found to be statistically significant (Chi-square test, one-sided $p = 0.0353$). **(b–g)** RT-qPCR confirmation of differentially regulated genes that show EGR1 TFBS, including *Maqb* (b; *t*-test, $p = 0.0093$), *Nr4a1* (c, *t*-test, $p = 0.0126$), *C1qa* (d; *t*-test, $p = .0101$), *Cacna1 h* (e; *t*-test, $p = .0308$), *Ephb6* (f; *t*-test, $p = .0248$), *Pttg1* (g; *t*-test, $p < 0.0001$). For these RT-qPCR analyses, $n = 18$ for Eμ-Myc miR-146a$^{+/+}$and $n = 17$ for Eμ-Myc miR-146a$^{-/-}$.

sites (TFBS) around known protein coding genes in three different human cell lines (K562, GM12878, and H1-hESC) [27]. This dataset was compared to the list of differentially regulated genes in miR-146a deficient tumors that have human homologs (Supplementary Figure 6a–6b). Remarkably, genes that show TFBS for EGR1 were statistically overrepresented in the differentially regulated gene set from miR-146a deficient tumors (Figure 5a). We then confirmed several targets of EGR1 that (i) were differentially regulated in the RNA-sequencing dataset and (ii)had been previously shown in the literature to be EGR1 targets or had EGR1 binding sites based on the ChIP-Seq datasets. Several genes that are important in hematopoiesis and/or cancer were profiled in the larger set of tumors and showed a significantly differential regulation. These genes included *Maf*b (Figure 5b), *Nr4a1* (Figure 5c), *C1qa* (Figure 5d), *Cacna1 h* (Figure 5e), *Ephb6* (Figure 5f) and *Pttg1* (Figure 5g and Supplementary Figure 6c). Changes in gene expression in Egr1 and a subset of its targets were conserved in the subset of tumors derived from female mice, hinting that these molecular changes may underlie the increased lethality in the knockout mice (Supplementary Figure 5k–5n). Together, these findings indicate that miR-146a-regulated *Egr1* may represent a critical target that leads to the elaboration of a gene expression signature and the more aggressive phenotype observed during miR-146a-deficient, Eμ-Myc-mediated oncogenesis.

Egr1 is regulated by miR-146a and overexpression of miR-146a has an anti-growth effect on B-cell lymphoma cell lines

To elucidate whether miR-146a targets *Egr1*, we examined the 3' untranslated region (UTR) of the cDNA transcript. In the human *EGR1* sequence, there is a miR-146a 7-mer binding site located at position 111–117 of the 3' UTR (Figure 6a). The DNA sequence surrounding this area is somewhat conserved between the human and the mouse, but the complete 7-mer site is not present in the mouse (Supplementary Figure 6d). To examine direct targeting, we cloned a 996 bp segment of the human *EGR1* 3'UTR into the pmiRGlo vector. Co-transfection of a miR-146a over- expression vector along with the luciferase-*EGR1* 3'UTR fusion construct showed significant repression of luciferase activity, compared to the empty vector, similar to that observed for *Traf6*. Mutation of the binding site for miR-146a in the EGR1 3' UTR derepressed luciferase expression. A similar repression was not consistently observed for the murine Egr1 3'UTR (Figure 6b). Stable overexpression of miR-146a using a retroviral vector in the murine leukemia cell lines, 70Z/3 and WEHI-231 led to a repression of *Egr1* at both the transcript and protein levels (Figure 6c–6e, 6i–6j, 6l). In addition, overexpression of miR-146a led to decreased growth of both cell lines at baseline and following serum starvation (Figure 6f–6g, 6m–6n). Moreover, we observed repression of the EGR1 target, *Nr4a1* (Figure 6h, 6k), implicating the same sequence of regulation with miR-146a overexpression as that observed with miR-146a deficiency in the tumors. Moreover, miR-146a overexpressing cells showed a downregulation of Blimp1 and Bcl6 (Supplementary Figure 6i–6l), in line with the observations made in the tumors. In the human DLBCL cell line, SUDHL2, miR-146a overexpression led to repression of EGR1 as well as the expected target of miR-146a, TRAF6 (Figure 6o). Nrp2 was validated as an additional target of miR-146a (Supplementary Figure 6e–6h). These findings imply a role for miR-146a in the regulation of B-cell leukemia/lymphoma cell growth and demonstrate that in the human, miR-146a directly targets EGR1 via canonical 3' UTR-mediated targeting. Hence, miR-146a overexpression and knockout results in significant effects on Egr1 and downstream gene expression, suggesting a conserved regulatory module in the human and mouse.

DISCUSSION

In this manuscript we describe the modulation of tumorigenesis by the NF-κB induced tumor suppressor microRNA, miR-146a. miR-146a plays a very important role in immune cells and seems to be critical in modulating feedback inhibition of the NF-κB pathway. Its role in T-cells, myeloid cells and hematopoietic stem cells is well-established, with deletion of this miRNA leading to T-cell hyper activation, myeloid hyperplasia and tumors, and stem cell exhaustion [9, 11, 14, 16]. The role of miR-146a in the developmental sequence of B-cells is less understood. In young miR-146a deficient mice, B-cell development appears to proceed normally, but by the age of six months, lymphoid follicles in the spleen and other lymphoid tissue demonstrate hyperplasia [9, 28]. Following this phase, myeloproliferative disease becomes the dominant phenotype and B-cell numbers drop as the mice age. Nonetheless, aged miR-146a deficient mice show an increased incidence of B-cell malignancies. Interestingly, these tumors show a predilection for the lymph nodes, similar to what we have observed here with Eμ-Myc driven tumors.

miR-146a-deficient Eμ-Myc transgenic mice develop mature B-cell neoplasms with IgM and/or CD138 expression, leading to a higher proportion of lymph node tumors and leukemia in the peripheral blood. There is some heterogeneity in the proportion of Eμ-Myc mice reported to develop IgM+ tumors in the literature [29, 30], but our results have been consistently in the 20–30% range. The immunophenotypic differences, along with concordant gene expression changes (e.g., Blimp1 and Bcl6), indicate that miR-146a deficiency may alter the stage of B-cell development that is most susceptible to transformation by *c-Myc*. This is an interesting observation as B-cell neoplasms in humans that have increased levels of *c-Myc* can also derive from different stages of development (for example, B-lymphoblastic leukemia, Burkitt's lymphoma,

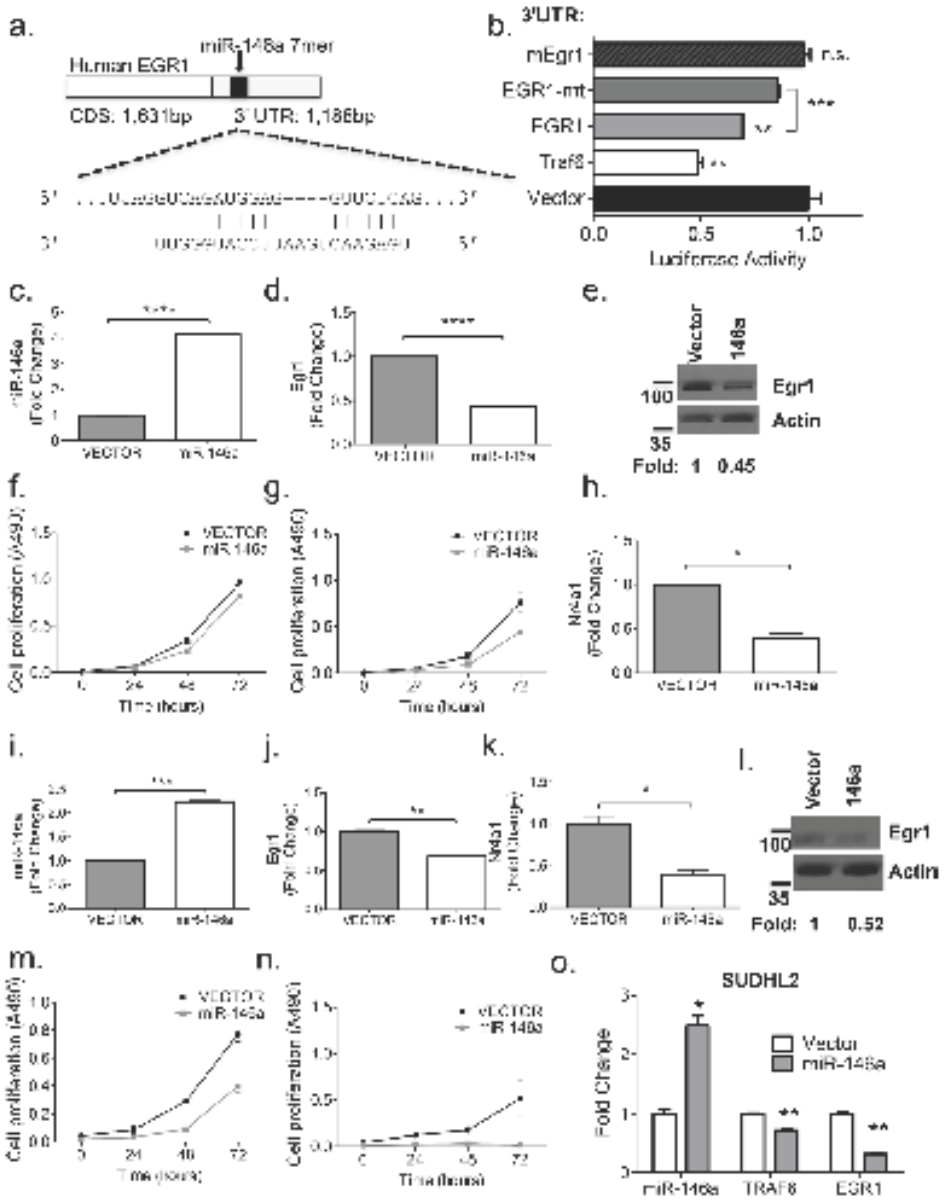


Figure 6: Egr1 is regulated by miR-146a. (a) The human EGR1 cDNA contains a 1, 188 bp 3'UTR that contains an intact miR-146a 7-mer binding site. Shown is a schematic of the binding between EGR1 and miR-146a. **(b)** Luciferase assays quantitating repression with MGP/miR-146a co-transfection relative to MGP alone for each of the UTRs depicted. Each measurement is representative of firefly luciferase normalized to renilla luciferase, and was performed in duplicate, with the experiment was repeated at least three times (*T*-test; Traf6 v Vector, $p = 0.0013$; EGR1 vs. vector, $p = 0.0077$; EGR1 vs. mutant EGR1, $p = 0.0002$). **(c–d)** RT-qPCR analyses of miR-146a and *Egr1*, respectively in the murine 70Z/3 cell line. **(e)** Western Blot analysis confirms EGR1 repression with miR-146a over expression in the 70Z/3 cell line. Fold repression was computed using ImageJ software. **(f–g)**. Cell proliferation (MTS) assays were performed using 70Z/3 cells transduced with either empty vector (MGP) or miR-146a over expressing vector (MGP-miR-146a). Basal growth is shown in (f), while growth in samples following 24 hours of serum starvation is shown in (g). **(h)** RT-qPCR of the EGR1 target gene, *Nr4a1*, shows repression in miR-146a overexpressing cells. **(i–k)**. RT-qPCR analyses of miR-146a, *Egr1* and *Nr4a1*, respectively, in the murine WEHI-231 cell line. **(l)** Western Blot analysis, as in (e), using the WEHI-231 cell line **(m–n)**. Cell proliferation (MTS) assays were used to measure basal growth and growth following 4 hours of serum starvation **(o)** miR-146a overexpression in a human DLBCL cell line, SUDHL2 results in the repression of TRAF6 and EGR1. All comparisons were made with *T*-test for Figure 6c-o, with the following legend: $*p < 0.05$; $**p < 0.005$; $***p < 0.0005$; $****p < 0.0001$.

DLBCL, and plasma cell myeloma) [31–35]. Here, the expression of a miRNA in an experimental model of Myc-mediated oncogenesis does alter the stage of B-cell oncogenesis. Although the relevance to human disease remains to be established, it is interesting to speculate that miRNA expression may have an important role in defining the cellular composition of lymphoma from a given driver mutation.

An important question raised by this study is whether the observed phenotypes occur as a consequence of cell-intrinsic or cell-extrinsic mechanisms. The primary tumor sites showed primarily B-cells, and histologically, the tumors appeared to be quite homogeneous. In miR-146a deficient mice, T-cell activation is thought to occur as a consequence of repeated bouts of subclinical infection and inflammation without the "recalibrating" effects of miR-146a expression. We do not think this is a likely cause for the augmentation of oncogenesis, since the development of tumors under specific pathogen free conditions do not occur early in life in miR-146a singly-deficient mice. Heterozygotes did not show increased mortality in the presence of Eμ-Myc, whereas miR-146a heterozygosity alone causes inflammatory changes [9]. Hence, it is likely that increased tumorigenesis in our mice occurs primarily as a consequence of B-cell intrinsic mechanisms. However, we cannot entirely exclude a cell-extrinsic process driven by benign but hyperactivated T-cells. Future studies to address this issue will include producing B-cell specific knockouts and knock-ins of miR-146a to study disease progression, but are beyond the scope of the current study.

In an effort to further characterize tumorigenesis in these mice, we undertook gene expression analysis by high-throughput sequencing. We have found that a small set of genes are significantly differentially regulated between miR-146a sufficient and deficient Eμ-Myc tumors. While the functional analysis did not reveal an overall pattern to the differentially regulated gene set, the individual genes do seem to be important in various aspects of tumorigenesis (Supplementary Table 2) [36, 37]. Amongst the differentially regulated genes, many have roles in oncogenesis and B-cell development. Perhaps the most interesting gene to be identified by our analysis is *Egr1*, a factor known to promote differentiation in the hematopoietic lineage. EGR1 transcriptionally induces a range of genes, and the differentially regulated gene set in miR-146a deficient mice was enriched for these targets. Indeed, some of the most differentially regulated genes in our dataset were previously described targets of EGR1 or putative targets as defined by the presence of transcription factor binding sites. Critically, miR-146a regulates *Egr1*, and provides an explanation for the observed phenotypic differences in the tumors from these mice. However, we must note that direct targeting was only seen with the human EGR1 3'UTR, and hence the mechanism of this regulation in the mouse may be indirect. This could include non-canonical mechanisms of miRNA targeting (such as in the 5'UTR) and/or indirect regulation.

miR-146a overexpression changes the growth of murine B-cell lines, suggesting the importance of the EGR1-mediated transcription program in maintaining growth of these cells. Notably, miR-146a overexpression also led to the repression of certain mRNAs that are important in B-cell differentiation including Blimp1 and Bcl6, once again supporting the notion of miR-146a playing a role in the maturation stage of the tumor cells. This is line with prior reports showing that expression of an Egr1 transgene supported the development of progenitor cells into mature, IgM-expressing B-cells [26]. Interestingly, some of the genes that contain TFBS for EGR1 are also regulated by miR-146a (for example, *Nrp2*), suggesting that miR-146a may target several points in the same pathway during B-cell oncogenesis. Downstream, genes without a defined role in B-cell neoplasms were also identified. *Pttg1* is overexpressed in a wide variety of endocrine and non-endocrine tumors, modulates tumor invasiveness and recurrence in several systems, and has functions in chromatid separation and cell cycle progression [38]. *Jhy* is another gene we identified whose deficiency causes juvenile hydrocephalus in mice [39]. It will be of great interest to study how miR-146a deficiency causes differential regulation of these novel genes and what their roles are in normal and malignant B-lymphopoiesis.

Our findings also point to the cell-type specific nature of miRNA mediated regulation. The targets uncovered in a malignant B-cell are different than those found in an activated T-cell or a myeloid cell. For example, our findings suggest that *Traf6* and *Irak1*, which are highly important in the elaboration of myeloid phenotypes, may not be as important in B-cell oncogenesis, particularly that induced by *c-Myc*, as these genes were not differentially regulated in the tumors that we examined (data not shown). These findings highlight the need for experimental work in carefully defined physiological and pathological systems to comprehensively understand miRNA function.

In summary, we show that concurrent *c-Myc* overexpression coupled with the absence of a *bona fide* tumor suppressor miRNA leads to more aggressive tumor due to a small set of genes that are regulated directly or indirectly by miR-146a. . Our novel set of targets may indicate that miR-146a regulates components of signaling networks other than the NF-κB inflammatory pathway. Hence, our work opens the door to new areas of investigation in B-cell oncogenesis and miRNA biology.

MATERIALS AND METHODS

Mice

miR-146a-deficient (miR-146a$^{-/-}$) mice were developed as previously described [9, 11, 16]. Eμ-Myc mice were purchased from Jackson laboratories and housed under pathogen free conditions at the University of California, Los Angeles [40]. Eμ-Myc and miR-146a$^{-/-}$mice

were bred to obtain Eμ-Myc miR-146a$^{+/-}$ mice with further miR-146a$^{-/-}$ intercross producing Eμ-Myc miR-146a$^{-/-}$ mice. Mice were monitored for tumors and sacrificed when they became pre-moribund indicated by the following criteria: tumors larger than 1.5cm, emaciation, or any other signs of distress. All mouse studies were approved by the UCLA Office of Animal Research Oversight.

Flow cytometry

Blood, bone marrow, spleen, and lymph node tumors were collected from the mice under sterile conditions. Single cell suspensions were lysed in red blood cell lysis buffer. Fluorochrome conjugated antibodies against B220, CD3ε, CD11b, Ter119, CD19, IgM, CD80, CD138, CD44, CD21, CD23, and CD5 were used for staining (all antibodies obtained from Biolegend). Flow cytometry was performed on a FACSAria and analysis performed using FlowJo software. Dichotomization of flow cytometric measurements was accomplished by visual inspection of the data and identification of clusters within the data. These were then validated by comparison of the means and averages of the two clusters. For CD138, this was accomplished by examining the Mean Fluorescence Intensity and determining that the low expression cluster had a mean MFI, 154.0 ± 13.96 (N = 30) and 557.6 ± 71.87 (N = 7) for the high CD138 samples (p <0.0001 for this comparison).

Histopathology

Organs were collected after necropsy and fixed in 10% neutral buffered formalin. These were then embedded in paraffin, processed for hematoxylin and eosin staining by the Translational Pathology Core Laboratory at UCLA. Histopathologic analysis was performed by a board certified hematopathologist (D.S.R). The degree of splenic involvement was scored on a 4-point scale for red and white pulp involvement. Analysis of dichotomized or ordinal-type histopathologic data was accomplished by the use of Fisher's Exact Test.

Statistical analyses

Figures are graphed as mean with the standard deviation of the mean (SD) for continuous numerical data. Bar graphs are employed to show dichotomized or ordinal-type histopathologic data. Student's t-test, Fisher's exact test, Chi square test, and Kaplan-Meier survival analysis were performed using GraphPad Prism software, applied to each experiment as described in the figure legends.

RNA-sequencing and analysis

Total RNA was extracted from tumors using Trizol combined with Qiagen miRNEasy mini kit with additional on column DNAse I digestion. Following isolation of RNA, cDNA libraries were built using the Illumina(San Diego, CA) TrueSeq RNA Sample Preparation kit V2 (RS-122–2001). An Agilent Bioanalyzer was used to determine RNA quality (RIN > 8) prior to sequencing. RNA-Seq libraries were sequenced on an Illumina HiSeq 2000 (single-end 100bp). Raw sequence files were obtained using Illumina's proprietary software and are available at NCBI's Gene Expression Omnibus resource (GEO Series accession number GSE67113; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67113) resource. RNA-Seq reads were aligned using STAR v2.3.0 [41]. The GRCm38 assembly (mm10) of the mouse genome and the junction database from Ensembl's gene annotation (release 71) were used as reference for STAR. The count matrix for genes in Ensembl's genome annotation (excluding rRNAs, Mt_rRNAs and Mt_tRNAs) was generated with HTSeq-count v0.5.4p3 (http://www-huber.embl.de/users/anders/HTSeq/) and normalized using the geometric mean across samples [42]. DESeq v1.14.0 [42] was used to classify genes as differentially expressed (Benjamini-Hochberg adjusted p-value < 0.05). Moderate fold changes between conditions were obtained from variance-stabilized data [42]. Functional annotation of differentially expressed genes was generated through the use of DAVID [36, 37]. Hierarchical gene clustering was performed with GENE-E (http://www.broadinstitute.org/cancer/software/GENE-E/). To display the heatmap, the expression levels were re-scaled so that, for each gene, the limits of the color scale correspond to the minimum and maximum expression levels across all samples.

EGR1 transcription factor binding site analysis

Publically available ENCODE data for EGR1 Transcription Factor Binding Site ChIP-Seq Uniform Peak analysis was downloaded from the UCSC Genome Browser for the K562, H1-hESC, and GM12878 cell lines [27]. For each line, the locations for all EGR1 Transcription Factor Binding Sites (TFBS) were grouped based on the closest known gene based using the UCSC Main (hg19) ccds gene list. The genes with one or more TFBS were compared to the mouse (mm10) RNASeq data set to identify genes that were differentially expressed in the miR-146a deficient tumors and also had at least one EGR1 TFBS in close proximity to the gene (defined as 3 kb). A Chi-Square test was performed with one degree of freedom to compare the relative frequency of EGR1 TFBS in the differentially expressed dataset (Observed) with the frequency across the genome (Expected). Only the mouse genes with a human homolog (total of 16288 genes) were used.

RT-qPCR

RNA collected from the murine tumors was reverse transcribed using qScript reagent and PerfeCTa SYBR Green FastMix reagent (Quanta Biosciences) or TaqMan MicroRNA Assay (Life Technologies). Primer sequences used are listed in Supplementary Table 1.

Western blot

Tumor cell suspensions were lysed in RIPA buffer (Boston BioProducts) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific). Equal amounts of protein lysate (as quantified by using bicinchoninic acid protein assay, BCA (Thermo Scientific)) were electrophoresed on a 5–12% SDS–PAGE and electroblotted onto a nitrocellulose membrane. Antibodies used were c-MYC Rabbit polyclonal (#9402), EGR1 (44D5) Rabbit monoclonal (all antibodies from Cell Signaling), and ß-Actin (AC15) mouse monoclonal antibody (Sigma Aldrich). Secondary HRP-conjugated antibodies were purchased from Santa Cruz Biotechnology.

MTS assay

Cell proliferation was measured using the Promega Cell Titer 96 Aqueous Non-Radioactive Cell Proliferation Assay kit. After addition of reagent according to the manufacturer's protocol, cells were incubated at 37°, 5% CO_2 for 4 hours and absorbance was measured at 490 nm.

Luciferase assays

A 996-bp segment of the human *EGR1* 3'UTR containing the miR-146a site was cloned into the pmiRGlo dual luciferase vector (Promega). A similar cloning strategy was used to clone murine Egr1 3'UTR and the Nrp2 UTR (see Supplementary Table 1). For mutation of the miR-146a binding site, we utilized site-directed mutagenesis as previously described using the primers shown in Supplementary Table 1 [43]. Co-transfections were performed with Lipofectamine 2000 (Life Technologies) as per the manufacturer's instructions. Cells were lysed after 24 hours, substrate was added and luminescence was measured on a Glomax-Multi Jr (Promega).

Genotyping for miR-146a mice and c-Myc mice

Mice were genotyped for miR-146a deletion and Eμ-Myc presence using DNA extracted from tail samples. Genotyping for miR-146a deletion was done as described previously [11]. Primers are listed in Supplementary Table 1.

ACKNOWLEDGMENTS

We thank members of the Rao lab for helpful discussions regarding the research. This work was supported by an R01CA166450–01 and Career Development Award K08CA133521 from the National Institutes of Health (DSR). DSR was a Sidney Kimmel Scholar supported by the Sidney Kimmel Foundation for Cancer Research (Translational Award SKF-11–013). JRC was supported by the Eugene V. Cota-Robles Grant and by a training award from the institutional Tumor Immunology training Grant (NIH T32CA009120). TRF was supported by Tumor Biology Training Grant NIH T32CA009056 from the National Institute of Health. NIRM was supported by the Eugene V. Cota-Robles Fellowship from UCLA and the Graduate Research Fellowship Program from the National Science Foundation. Flow cytometry was performed in the UCLA Jonsson Comprehensive Cancer Center (JCCC) and Center for AIDS Research Flow Cytometry Core Facility that is supported by National Institutes of Health awards AI-28697, and award number P30CA016042, the JCCC, the UCLA AIDS Institute, and the David Geffen School of Medicine at UCLA. We would like to acknowledge the ENCODE project and in particular the EGR1 Transcription Factor Binding Site Chip-Seq data, generated and analyzed by the Hudson Alpha Institute in Huntsville, AL.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

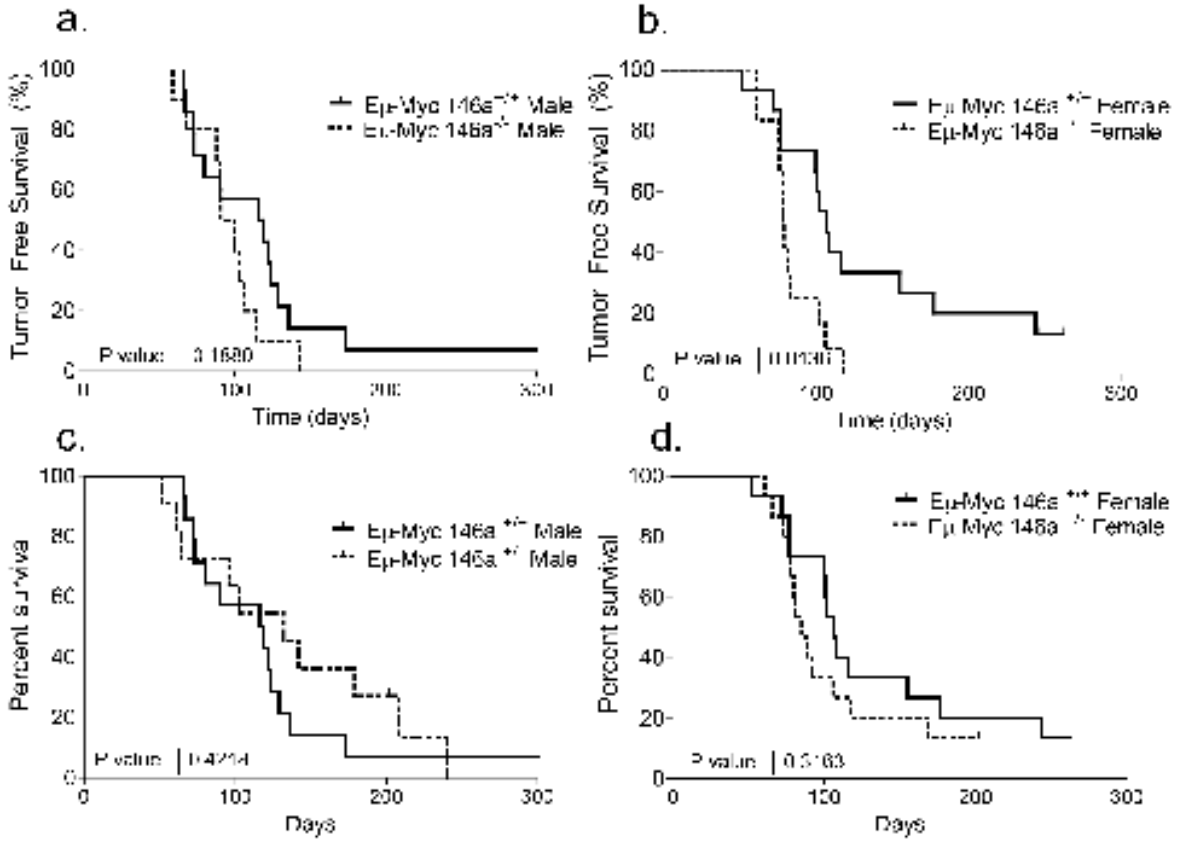
REFERENCES

1. O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D. Physiological and pathological roles for microRNAs in the immune system. Nature reviews Immunology. 2010; 10:111–122.
2. Bartel DP, Chen CZ. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nature reviews Genetics. 2004; 5:396–400.
3. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009; 136:215–233.
4. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2006; 34:D140–144.
5. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2002; 99:15524–15529.
6. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, Ambesi-Impiombato A, Califano A, Migliazza A, Bhagat G, Dalla-Favera R. The DLEU2/miR-15a/16–1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell. 2010; 17:28–40.
7. O'Connell RM, Rao DS, Chaudhuri AA, Boldin MP, Taganov KD, Nicoll J, Paquette RL, Baltimore D. Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder. The Journal of experimental medicine. 2008; 205:585–594.
8. O'Connell RM, Chaudhuri AA, Rao DS, Baltimore D. Inositol phosphatase SHIP1 is a primary target of miR-155. Proc Natl Acad Sci U S A. 2009; 106:7113–7118.

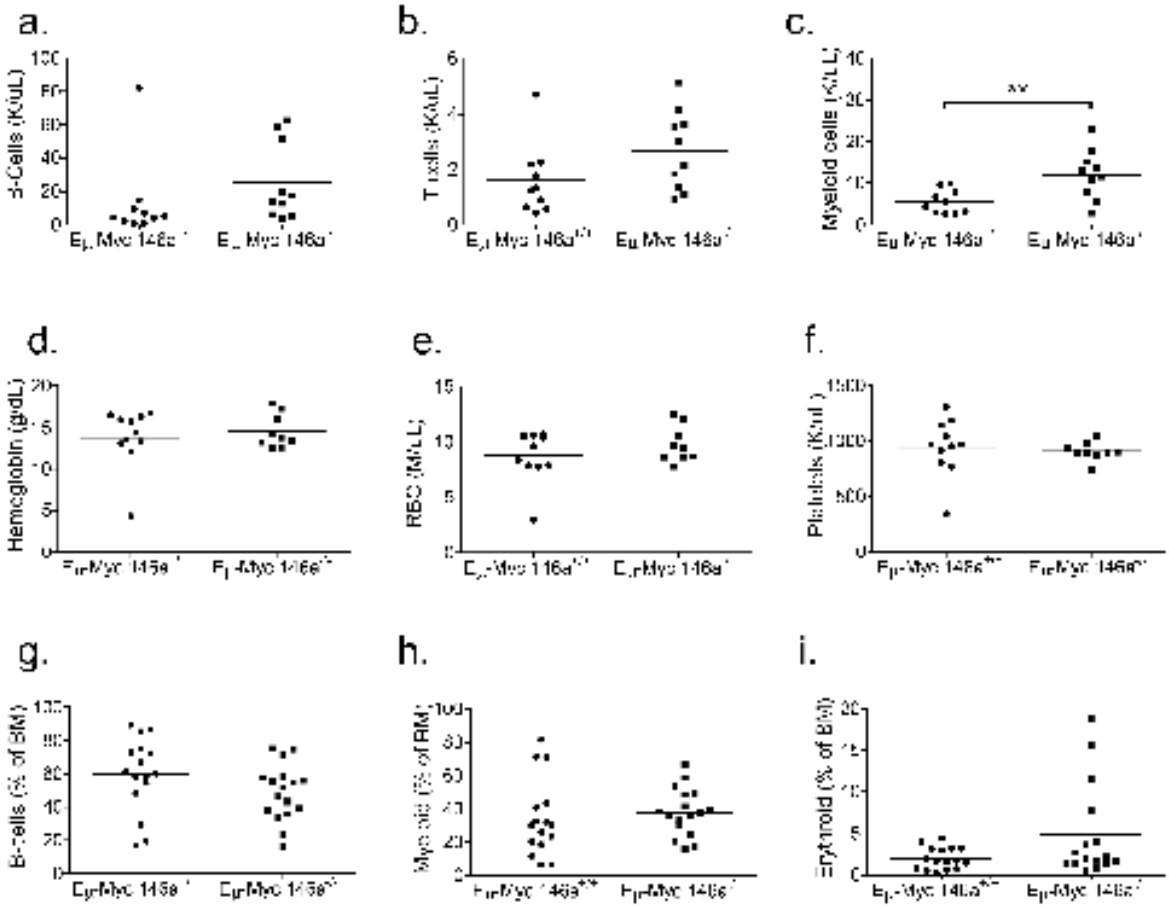
9. Boldin MP, Taganov KD, Rao DS, Yang L, Zhao JL, Kalwani M, Garcia-Flores Y, Luong M, Devrekanli A, Xu J, Sun G, Tay J, Linsley PS, Baltimore D. miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. J Exp Med. 2011.
10. Taganov KD, Boldin MP, Chang KJ, Baltimore D. NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proc Natl Acad Sci U S A. 2006; 103:12481–12486.
11. Zhao JL, Rao DS, Boldin MP, Taganov KD, O'Connell RM, Baltimore D. NF-kappaB dysregulation in microRNA-146a-deficient mice drives the development of myeloid malignancies. Proc Natl Acad Sci U S A. 2011; 108:9184–9189.
12. Starczynowski DT, Kuchenbauer F, Argiropoulos B, Sung S, Morin R, Muranyi A, Hirst M, Hogge D, Marra M, Wells RA, Buckstein R, Lam W, Humphries RK, Karsan A. Identification of miR-145 and miR-146a as mediators of the 5q- syndrome phenotype. Nat Med. 2010; 16:49–58.
13. Starczynowski DT, Kuchenbauer F, Wegrzyn J, Rouhi A, Petriv O, Hansen CL, Humphries RK, Karsan A. MicroRNA-146a disrupts hematopoietic differentiation and survival. Exp Hematol. 2010.
14. Yang L, Boldin MP, Yu Y, Liu CS, Ea CK, Ramakrishnan P, Taganov KD, Zhao JL, Baltimore D. miR-146a controls the resolution of T cell responses in mice. The Journal of experimental medicine. 2012; 209:1655–1670.
15. Lu LF, Boldin MP, Chaudhry A, Lin LL, Taganov KD, Hanada T, Yoshimura A, Baltimore D, Rudensky AY. Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses. Cell. 2010; 142:914–929.
16. Zhao JL, Rao DS, O'Connell RM, Garcia-Flores Y, Baltimore D. MicroRNA-146a acts as a guardian of the quality and longevity of hematopoietic stem cells in mice. eLife. 2013; 2:e00537.
17. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J, Jr., Lu L, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000; 403:503–511.
18. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, Kohlhammer H, Xu W, Yang Y, Zhao H, Shaffer AL, Romesser P, Wright G, Powell J, Rosenwald A, Muller-Hermelink HK, et al. Oncogenically active MYD88 mutations in human lymphoma. Nature. 2011; 470:115–119.
19. Compagno M, Lim WK, Grunn A, Nandula SV, Brahmachary M, Shen Q, Bertoni F, Ponzoni M, Scandurra M, Califano A, Bhagat G, Chadburn A, Dalla-Favera R, Pasqualucci L. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. Nature. 2009; 459:717–721.
20. Forloni M, Dogra SK, Dong Y, Conte D, Jr., Ou J, Zhu LJ, Deng A, Mahalingam M, Green MR, Wajapeyee N. miR-146a promotes the initiation and progression of melanoma by activating Notch signaling. eLife. 2014; 3:e01460.
21. Robertus JL, Kluiver J, Weggemans C, Harms G, Reijmers RM, Swart Y, Kok K, Rosati S, Schuuring E, van Imhoff G, Pals ST, Kluin P, van den Berg A. MiRNA profiling in B non-Hodgkin lymphoma: a MYC-related miRNA profile characterizes Burkitt lymphoma. Br J Haematol. 2010; 149:896–899.
22. Chang TC, Yu D, Lee YS, Wentzel EA, Arking DE, West KM, Dang CV, Thomas-Tikhonenko A, Mendell JT. Widespread microRNA repression by Myc contributes to tumorigenesis. Nat Genet. 2008; 40:43–50.
23. Chang TC, Zeitels LR, Hwang HW, Chivukula RR, Wentzel EA, Dews M, Jung J, Gao P, Dang CV, Beer MA, Thomas-Tikhonenko A, Mendell JT. Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation. Proc Natl Acad Sci U S A. 2009; 106:3384–3389.
24. Johansson FK, Goransson H, Westermark B. Expression analysis of genes involved in brain tumor progression driven by retroviral insertional mutagenesis in mice. Oncogene. 2005; 24:3896–3905.
25. Kharbanda S, Nakamura T, Stone R, Hass R, Bernstein S, Datta R, Sukhatme VP, Kufe D. Expression of the early growth response 1 and 2 zinc finger genes during induction of monocytic differentiation. J Clin Invest. 1991; 88:571–577.
26. Dinkel A, Warnatz K, Ledermann B, Rolink A, Zipfel PF, Burki K, Eibel H. The transcription factor early growth response 1 (Egr-1) advances differentiation of pre-B and immature B cells. The Journal of experimental medicine. 1998; 188:2215–2224.
27. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74.
28. Hu R, Kagele DA, Huffaker TB, Runtsch MC, Alexander M, Liu J, Bake E, Su W, Williams MA, Rao DS, Moller T, Garden GA, Round JL, O'Connell RM. miR-155 promotes T follicular helper cell accumulation during chronic, low-grade inflammation. Immunity. 2014; 41:605–619.
29. Frenzel A, Labi V, Chmelewskij W, Ploner C, Geley S, Fiegl H, Tzankov A, Villunger A. Suppression of B-cell lymphomagenesis by the BH3-only proteins Bmf and Bad. Blood. 2010; 115:995–1005.
30. Nemajerova A, Petrenko O, Trumper L, Palacios G, Moll UM. Loss of p73 promotes dissemination of Myc-induced B cell lymphomas in mice. J Clin Invest. 2010; 120:2070–2080.
31. Kuehl WM, Brents LA, Chesi M, Huppi K, Bergsagel PL. Dysregulation of c-myc in multiple myeloma. Current topics in microbiology and immunology. 1997; 224:277–282.
32. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci U S A. 1982; 79:7824–7827.

33. Adams JM, Gerondakis S, Webb E, Corcoran LM, Cory S. Cellular myc oncogene is altered by chromosome translocation to an immunoglobulin locus in murine plasmacytomas and is rearranged similarly in human Burkitt lymphomas. Proc Natl Acad Sci U S A. 1983; 80:1982–1986.
34. Battey J, Moulding C, Taub R, Murphy W, Stewart T, Potter H, Lenoir G, Leder P. The human c-myc oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell. 1983; 34:779–787.
35. Rabbitts TH, Hamlyn PH, Baer R. Altered nucleotide sequences of a translocated c-myc gene in Burkitt lymphoma. Nature. 1983; 306:760–765.
36. Huang da W, Sherman BT, Zheng X, Yang J, Imamichi T, Stephens R, Lempicki RA. Extracting biological meaning from large gene lists with DAVID. Curr Protoc Bioinformatics. 2009; . Chapter 13:Unit 13 11.
37. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4:44–57.
38. Salehi F, Kovacs K, Scheithauer BW, Lloyd RV, Cusimano M. Pituitary tumor-transforming gene in endocrine and other neoplasms: a review and update. Endocr Relat Cancer. 2008; 15:721–743.
39. Appelbe OK, Bollman B, Attarwala A, Triebes LA, Muniz-Talavera H, Curry DJ, Schmidt JV. Disruption of the mouse Jhy gene causes abnormal ciliary microtubule patterning and juvenile hydrocephalus. Dev Biol. 2013; 382:172–185.
40. Adams JM, Harris AW, Pinkert CA, Corcoran LM, Alexander WS, Cory S, Palmiter RD, Brinster RL. The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature. 1985; 318:533–538.
41. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21.
42. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010; 11:R106.
43. Rao DS, O'Connell RM, Chaudhuri AA, Garcia-Flores Y, Geiger TL, Baltimore D. MicroRNA-34a perturbs B lymphocyte development by repressing the forkhead box transcription factor Foxp1. Immunity. 2010; 33:48–59.

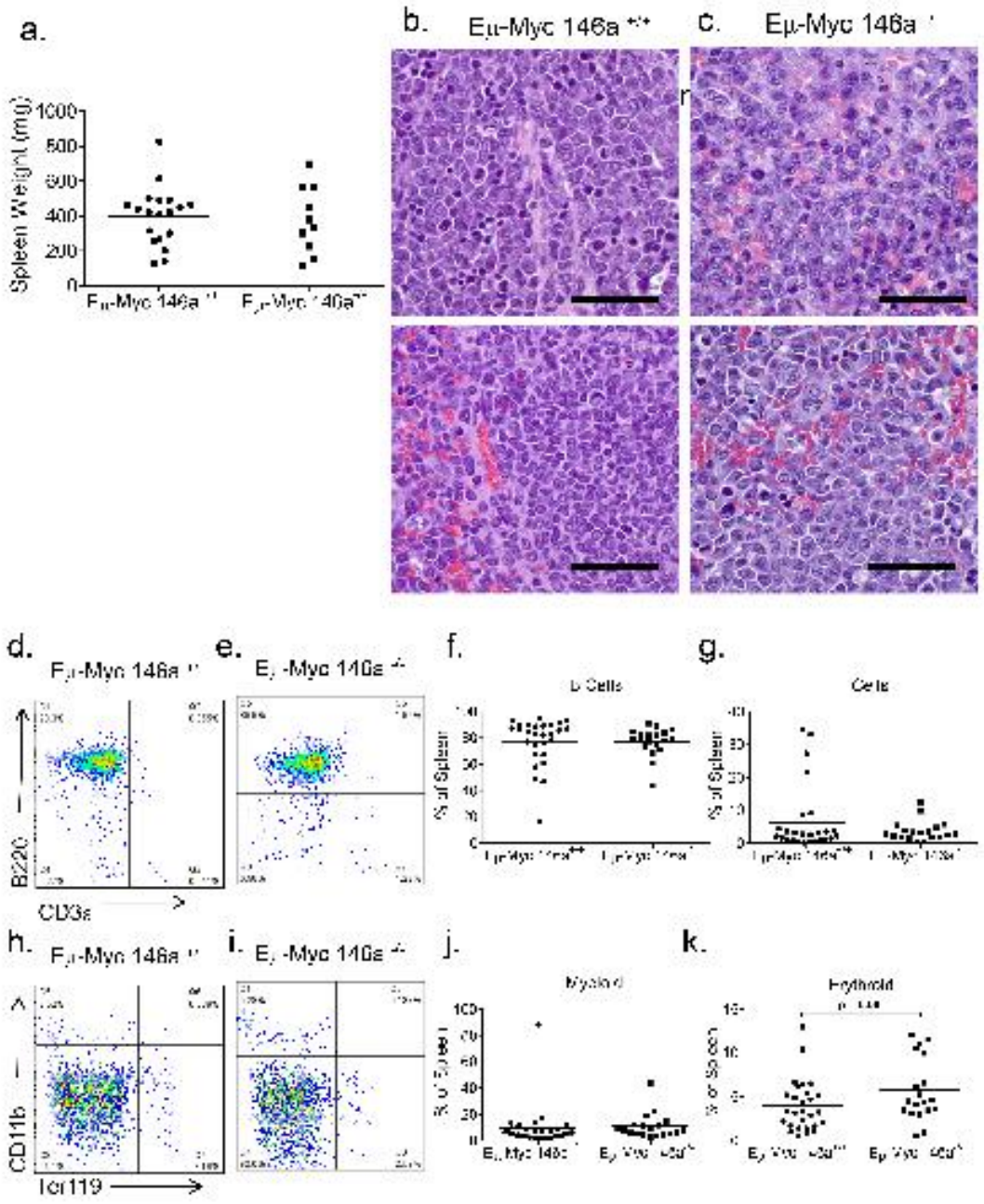
SUPPLEMENTARY FIGURES AND TABLES



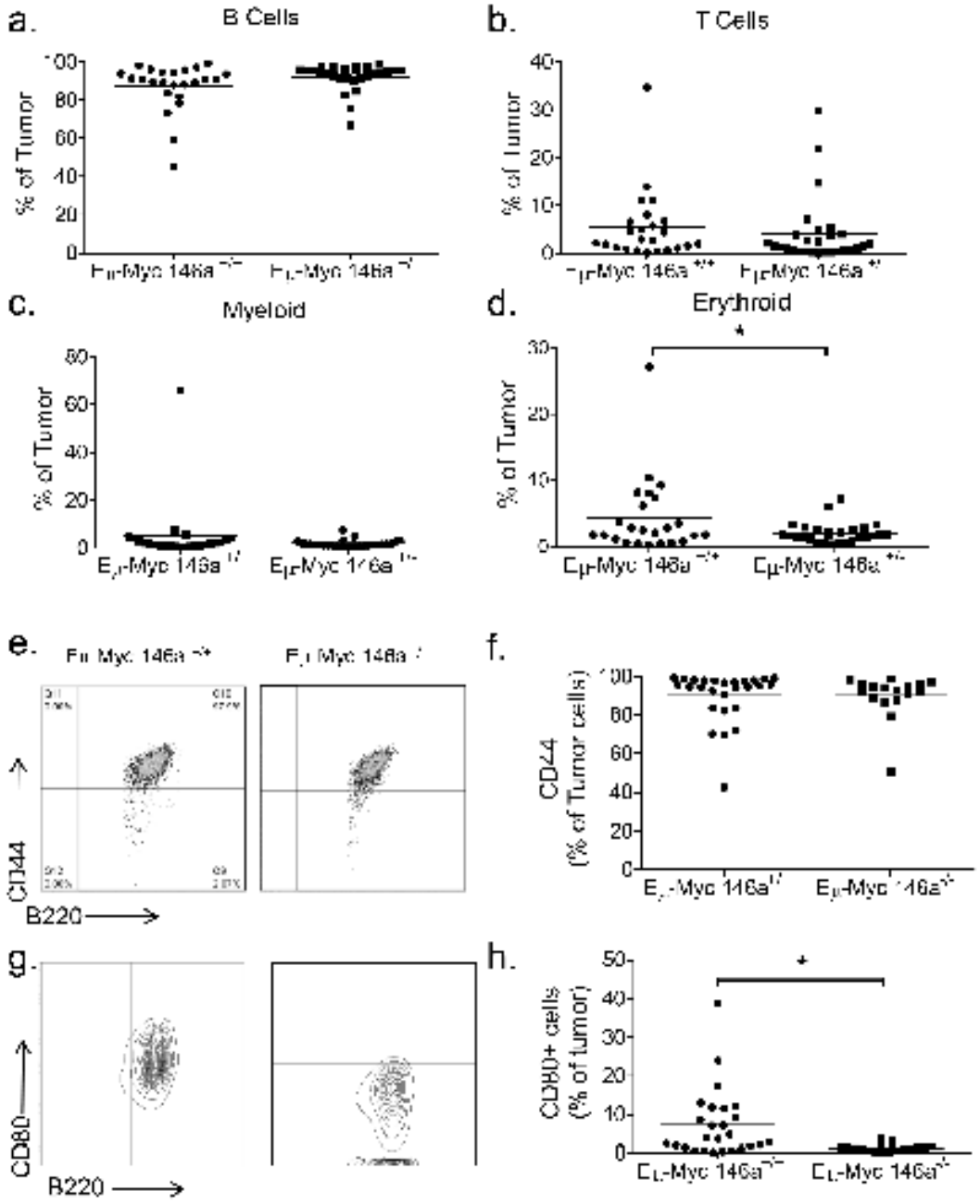
Supplementary Figure S1: miR-146a deficiency causes increased mortality in female Eμ-Myc mice. (a) Tumor-free survival curve of male mice with Eμ-Myc oncogene and either wild-type or homozygous deficiency of miR-146a ($n = 14$ for Eμ-Myc miR-146a$^{+/+}$ $n = 10$ for Eμ-Myc miR-146a$^{-/-}$; Log-Rank Test, $p = 0.1880$). **(b)** Tumor-free survival curve of female mice with Eμ-Myc oncogene and either wild-type or homozygous deficiency of miR-146a ($n = 15$ for Eμ-Myc miR-146a$^{+/+}$, $n = 12$ for Eμ-Myc miR-146a$^{-/-}$; Log-rank Test, $p = 0.0136$). **(c)** Tumor-free survival curve of male mice with Eμ-Myc oncogene and either wild-type or heterozygous deficiency of miR-146a ($n = 14$ for Eμ-Myc miR-146a$^{+/+}$, $n = 11$ for Eμ-Myc miR-146a$^{+/-}$; Log-Rank Test, $p = 0.4218$). **(d)** Tumor free survival curve of female mice with Eμ-Myc oncogene and either wild-type or heterozygous deficiency of miR-146a ($n = 15$ for Eμ-Myc miR-146a$^{+/+}$, $n = 15$ for Eμ-Myc miR-146a$^{+/-}$; Log-Rank Test, $p = 0.3163$).



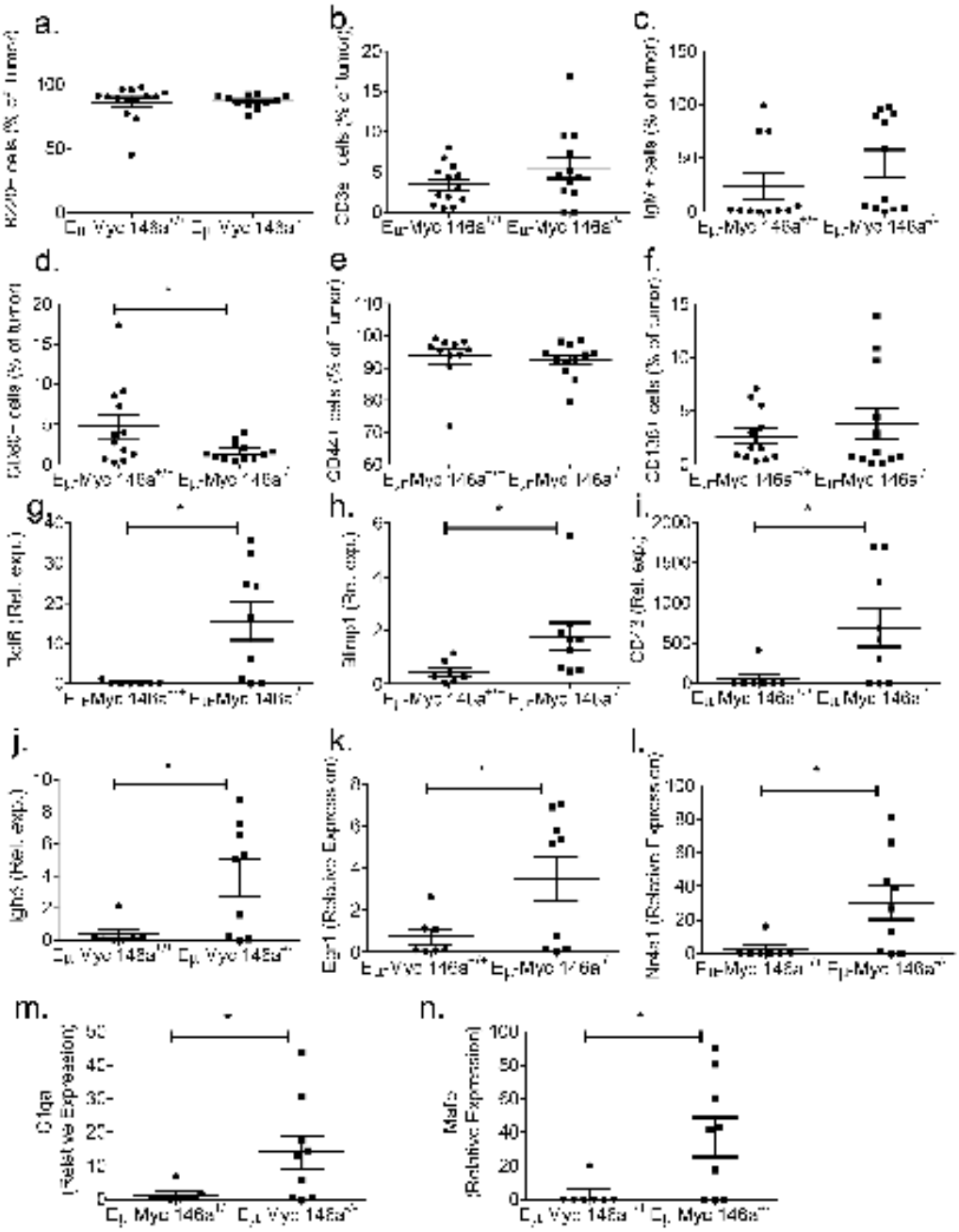
Supplementary Figure S2: Blood and bone marrow composition in Eμ-Myc animals at the time of death. (a–c) Quantitation of absolute number of B-cells (a), T-cells (b), and Myeloid cells (c) based on CBC and FACS analysis of blood at time of death (n = 10 Eμ-MycmiR-146a$^{+/+}$, and n = 10 Eμ-Myc miR-146a$^{-/-}$; t-test p = 0.2719, 0.0904 and 0.0052 respectively). **(d–i)** Hemoglobin levels, red blood cell, and platelet counts at time of death (n = 11 Eμ-MycmiR-146a$^{+/+}$, and n = 9 Eμ-Myc miR-146a$^{-/-}$; t-test, p = 0.5663, 0.3150 and 0.6521 respectively). Percentage of B-cells (g), Myeloid (h), and Erythroid (i) cells in the bone marrow of mice at time of death (n = 16 Eμ-Myc miR-146a$^{+/+}$, and n = 17 Eμ-Myc miR-146a$^{-/-}$; t-test, p = 0.1250, 0.5529 and 0.0635 respectively).



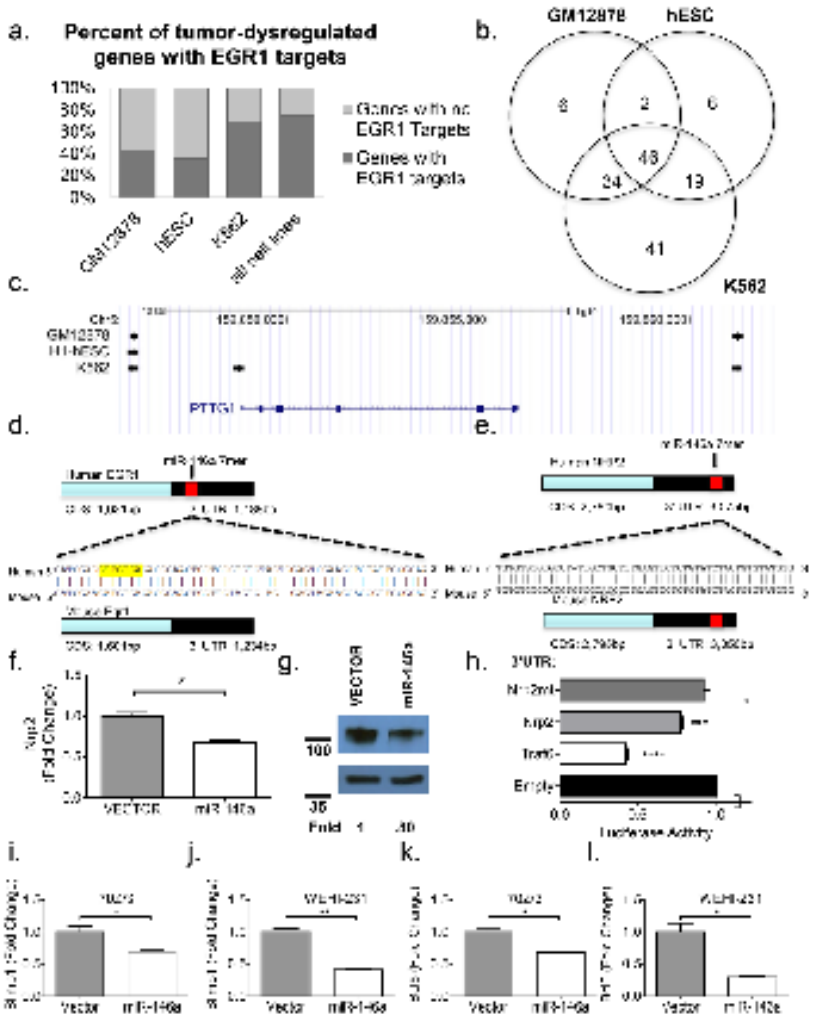
Supplementary Figure S3: Analysis of spleens from Eμ-Myc animals deficient for miR-146a. **(a)** Spleen weights at the time of death for all animals where data was available ($n = 19$ Eμ-Myc miR-146a$^{+/+}$ and $n = 10$ Eμ-Myc 146a$^{-/-}$; t-test, $p = 0.7662$). **(b–c)** Representative high power images of paraffin embedded H&E stained tumor samples showing red (top panels), and white (bottom panels) pulp of splenic tumors from Eμ-Myc animals sufficient or deficient for miR-146a. **(d–e, h–i)** Representative FACS plots from Eμ-Myc miR-146a$^{+/+}$ and Eμ-MycmiR-146a$^{-/-}$ spleens stained with B220 and CD3e, or CD11b and Ter119, respectively. **(f–g, j–k)** Quantitation of the percentage of B-lymphocytes, T-lymphocytes, Myeloid or Erythroid cells in the spleens of Eμ-Myc miR-146a$^{+/+}$ and Eμ-MycmiR-146a$^{-/-}$ mice, based on FACS ($n = 28$, Eμ-Myc miR-146a$^{+/+}$, and $n = 19$, Eμ-Myc miR-146a$^{-/-}$; t-test, $p = 0.9505$, 0.2568, 0.7114, and 0.0600 respectively).



Supplementary Figure S4: Immunophenotypic properties of tumors in mice with heterozygous and homozygous deficiency of miR-146a. (a–d) B-, T-, myeloid, and erythroid cells, respectively, quantitated by FACS in tumors from mice carrying the Eμ-Myc oncogene and wild-type or heterozygous for miR-146a ($n = 24$ for Eμ-Myc miR-146a$^{+/+}$ and $n = 28$ for Eμ-Myc miR-146a$^{+/-}$; t-test $p = 0.0704$, 0.5164 and 0.2368, 0.0427 respectively). **(e)** Representative FACS plots from Eμ-Myc 146a$^{+/+}$ or Eμ-Myc 146a$^{-/-}$ tumors stained with CD44 and B220. **(f)** Quantitation of all mice where data was available for the percent of CD44 positive cells in tumors from Eμ-Myc 146a$^{+/+}$ or Eμ-Myc 146a$^{-/-}$ ($n = 27$ Eμ-Myc 146a$^{+/+}$ and $n = 17$ Eμ-Myc 146a$^{-/-}$; T-Test, $p = 0.9922$). **(g)** Representative FACS plots from Eμ-Myc 146a$^{+/+}$ or Eμ-Myc 146a$^{-/-}$ tumors stained with CD80 and B220. **(h)** Quantitation of all mice where data was available for the percent of CD44 positive cells in tumors from Eμ-Myc 146a$^{+/+}$ or Eμ-Myc 146a$^{-/-}$ ($n = 27$ Eμ-Myc 146a$^{+/+}$ and $n = 17$ Eμ-Myc 146a$^{-/-}$; T-Test, $p = 0.0069$).



Supplementary Figure S5: Analysis of immunophenotypic and gene expression features of tumors from female Eμ-Myc 146a$^{-/-}$ mice. (a–f) Quantitation of B220+ (a), CD3e+ (b), IgM+ (c), CD80+ (d), CD44+ (e) and CD138+ (f) cells from female mice that carry the Eμ-Myc transgene and are either sufficient or deficient for miR-146a. Only CD80 expression is significantly different (*T*-test, $p = 0.0487$). This may be due to the reduced numbers of animals available for the analysis ($n = 12$ Eμ-Myc 146a$^{+/+}$ and $n = 12$ Eμ-Myc 146a$^{-/-}$). **(g–j)** RT-qPCR analyses of B-cell maturation associated transcripts from female mice that carry the Eμ-Myc transgene and are either sufficient or deficient for miR-146a. All comparisons showed a statistically significant difference ($p < 0.05$ for all comparisons between the mice; $n = 7$ Eμ-Myc 146a$^{+/+}$ and $n = 9$ Eμ-Myc 146a$^{-/-}$). **(k–n)**. RT-qPCR analyses of Egr1 and putative Egr1-regulated transcripts from female mice that carry the Eμ-Myc transgene and are either sufficient or deficient for miR-146a. All comparisons showed a statistically significant difference ($p < 0.05$ for all comparisons between the mice; $n = 7$ Eμ-Myc 146a$^{+/+}$ and $n = 9$ Eμ-Myc 146a$^{-/-}$).



Supplementary Figure S6: High throughput analyses for EGR1 TFBS and demonstration of a novel miR-146a target, *Nrp2*. **(a)** Plot showing the numbers of genes that were differentially regulated in miR-146a deficient tumors that have EGR1 binding sites in the three different cell lines. Note that the K562 cell line had the highest number of binding sites. The cell lines displayed a wide range in the total number of EGR1 transcription factor binding sites (TFBS). The K562 cell line had a total of 36,367 sites representing 12,741 genes, the GM12878 line had 16,530 sites representing 9,170 genes, and the H1-hESC line had only 8, 818 sites representing 6,349 genes. There exists a wide range in the total number of EGR1 TFBS sites as well as the total genes with EGR1 binding sites. In the K562 cell line, 62% of total protein coding genes are putatively regulated by EGR1. In the H1-hESC line 31% of genes have EGR1 TFBS and in GM12878 line 45% of genes have EGR1 TFBS sites. When all three cell lines are combined, there are a total of 62,071 sites representing 13,944 unique genes. **(b)** Intersection of differentially regulated genes from the RNA-Seq dataset overlaid with the genes containing EGR1 TFBS obtained from ChIP-Seq data of the three different cell lines. **(c)** *Pttg1* is a differentially regulated gene in miR-146a deficient tumors. It is found on chromosome 5 and has EGR1 transcription factor binding sites based on the reanalyzed ChIP-Seq data presented above. **(d)** Schematic representation of human and mouse *Egr1* gene showing the miR-146a binding site in the 3′UTR. **(e)** Schematic representation of human and mouse *Nrp2* gene showing the miR-146a binding site in the 3′UTR. **(f)** RT-qPCR analysis of *Nrp2* in 70Z/3 cell lines either expressing MGP or MGP-miR-146a vector (*t*-test, $p = 0.0148$). **(g)** Western blot analysis for NRP2 after miR-146a shows reduction in the protein levels in 70Z/3 cells when compared to the reference gene actin (upper panel: Nrp2 and lower panel, β-actin). Shown below are fold repression computed using ImageJ software. **(h)** Luciferase assays showing repression seen with MGP/miR-146a co-transfection relative to MGP alone for each of the UTRs depicted. Each measurement is representative of firefly luciferase normalized to renilla luciferase, and was performed in duplicate, with the experiment was repeated at least three times (*T*-test; Traf6 v Vector, $p < 0.0001$; Nrp2 vs. vector, $p = 0.0005$; Nrp2 vs. mutant Nrp2, $p = 0.033$). **(i–l)**. RT-qPCR analyses of Blimp1 and Bcl6 in 70Z/3 and WEHI-231 cells that are overexpressing miR-146a. All comparisons showed statistically significant downregulation of these genes in miR-146a overexpressing cell lines (*$p < 0.05$; **$p < 0.005$).

Supplementary Table S1: RT-qPCR primers and genotyping primers used. Listed are the primers used for RNAseq data validation, genotyping miR-146a and Eμ-Myc allele, and cloning. 5'*P** indicates phosphorylated 5' end.

RT-qPCR primers	Direction	Sequence
Jhy	FOW	5' GGTGCCGGCAGGATGAATAA 3'
	REV	5' AAGTTGGTGTGATGGACGGG 3'
Cacna1h	FOW	5' ATGCTTGGGAACGTGCTTCTT 3'
	REV	5' GTCTGGTAGTATGGCCGCAA 3'
Camk2b	FOW	5' TGGTGGAACAAGCCAAGAGTTT 3'
	REV	5' GAGGGAGAGATCCTTTGGGG 3'
Myo18b	FOW	5' AGAACAATGGAGTCCGCTGG 3'
	REV	5' GCTGGCTGTGGATCTTCTGT 3'
Pttg1	FOW	5' CCCTCCAACCAAAACAGCC 3'
	REV	5' TCCCTTACCAGATTCCCATGAT 3'
Axl	FOW	5' GTGGTTTCCAGACAACCTACG 3'
	REV	5' CGGATGTGATACGGGGTGTG 3'
Egr1	FOW	5' TTGTGGCCTGAACCCCTTTT 3'
	REV	5' AGATGGGACTGCTGTCGTTG 3'
Oaf	FOW	5' GAAGGGGCAGAGTCAGTTCC 3'
	REV	5' GTTTTTCTGCCGGAGCTTGG 3'
Nrp2	FOW	5' GCTGGCTACATCACTTCCCC 3'
	REV	5' CAATCCACTCACAGTTCTGGTG 3'
Dtx3	FOW	5' ACCCAATGTCATCACTTGGAAC 3'
	REV	5' CCTCTTGCACCCTAGTCAGGT 3'
Mafb	FOW	5'TGGATGGCGAGCAACTACC3'
	REV	5'CCAGGTCATCGTGAGTCACA3'
Nr4a	FOW	5'TTGAGTTCGGCAAGCCTACC3'
	REV	5'GTGTACCCGTCCATGAAGGTG3'
C1qa	FOW	5'AAAGGCAATCCAGGCAATATCA3'
	REV	5'TGGTTCTGGTATGGACTCTCC3'
Bcl6	FOW	5'CCGGCACGCTAGTGATGTT3'
	REV	5'TGTCTTATGGGCTCTAAACTGCT3'
IgD	FOW	5'CTTAGCTGCCGAGAGGGATG3'
	REV	5'ACACTGTGCTCGAAGGTGTT3'
IgM	FOW	5'AACATTGCTGGCAGGGGTAG3'
	REV	5'ACCAGAGGTTGTCCCTCCTT3'
CD43	FOW	5'GACCCACTTCCTTTCCCCCT3'
	REV	5'CGTACCCAGCAAGATCATACCC3'

RT-qPCR primers	Direction	Sequence
Bcl6	FOW	5'CCGGCACGCTAGTGATGTT3'
	REV	5'TGTCTTATGGGCTCTAAACTGCT3'
Blimp1	FOW	5'TTCTCTTGGAAAAACGTGTGGG3'
	REV	5'GGAGCCGGAGCTAGACTTG3'
Traf6	FOW	5'GCACAAGTGCCCAGTTGAC3'
	REV	5'TGCAAAATTGTCGGGAAACAGT3'
EGR1	FOW	5'GGTCAGTGGCCTAGTGAGC3'
	REV	5'GTGCCGCTGAGTAAATGGGA3'
mmu-miR-146a		5' UGAGAACUGAAUUCCAUGGGUU 3'
Cloning primers		
Egr1 CDS	FOW	5'AGCTAGA-AGATCT-TTCTCCAGCTCGCTGGTCC3'
	REV	5'AGCATCT-CTCGAG-TTCCTGCCTCTCCCTTTGCT3'
Egr1-3'-UTR	FOW	5'TAATCTGGTTTAAACGAGCTCTGGAAGATCTCAGAGCCAAG3'
	REV	5'TCGAATCCCTGCAGGCTCGAGGAACTTCATGTTCATAACATAC AAAAA3'
Egr1-3'-UTR-Mutant	FOW	5'P*ATGTCCACTGGACTGTCACCTC3'
	REV	5'P*GGCTGTTTCAGGCAGCTGAAG3'
EGR1-3'-UTR	FOW	5'TAATCTGGTTTAAACGAGCTCGAGGAGATGGCCATAGGAGA3'
	REV	5'TCGAATC-CCTGCAGGCTCGAGTACAAAAATCGCCGCCTACT3'
EGR1-3'-UTR-Mutant	FOW	5'P*GATGGAGCTGGACTGGAGCCAA3'
	REV	5'P*TGACCTAAGAGGAACCCTCC3'
Nrp2-3'-UTR	FOW	5'TAATCTGGTTTAAACGAGCTCACTGTGGTGGCCAAGTGAAT3'
	REV	5'TCGAATCCCTGCAGGCTCGAGCAGCACTGAGTCCCACGTTA3'
Nrp2-3'-UTR-Mutant	FOW	5'P*ACCCTTGCTGGACTGTGTATCT3'
	REV	5'P*GCGACACACACACACACACA3'
Traf6-3'-UTR	FOW	5'TAATCTGGTTTAAAGAGCTCTGAAAATCACCACTGCCTGT3'
	REV	5'TCGAATCCCTGCAGGCTCGAGGGATCCCCTCTGCTTCCTTA3'
Traf6-3'-UTR-Mutant	FOW	5'P*GGTGTTGCTGGACTGTTTAGTT3'
	REV	5'P*AGAGCGGTAACTTCTCTACTG3'
Bcl6-3'-UTR	FOW	5'TAATCTG- GTTTAAAC-GAGCTC-CCAGCCCCTTCTCAGAATC3'
	REV	5'TCGAATC-CCTGCAGG-CTCGAG CAACGCACTAATGCAGTTTAGA3'
Genotyping primers		
miR-146a WT	FOW	5' CTTGGACCAGCAGTCCTCTTGATGCACCTT 3'
miR-146a KO	FOW	5' ATCGCGGCCGCTTTAAGTGTAGAGAGGGGGTCAAGTA 3'
	REV	5' ATTGCTCAGCGGTGCTGTCCATCTGCACGA 3'
Eμ-Myc	FOW	5' ACCCAGGCTAAGAAGGCAAT 3'
	REV	5' GCTCCGGGGTGTAAACAGTA 3'

Supplementary Table S2: Functional annotation results for gene expression data from miR-146a deficient tumors. Genes that were differentially expressed between tumors from Eμ-Myc 146a$^{+/+}$ and for Eμ-Myc 146a$^{-/-}$ were used as the input in the DAVID Functional Annotation Tool for Functional Annotation Analysis. Listed are keywords associated with subgroups of genes.

Term	Count	%	*P*-Value
disulfide bond	67	28.5106383	4.68E–11
signal	74	31.4893617	1.45E–10
Secreted	47	20	2.36E–10
glycoprotein	79	33.61702128	1.03E–08
innate immunity	11	4.680851064	1.82E–08
complement pathway	7	2.978723404	4.80E–07
immune response	12	5.106382979	1.23E–05
collagen	8	3.404255319	6.11E–05
transmembrane protein	15	6.382978723	1.35E–04
inflammatory response	7	2.978723404	3.21E–04
Growth factor binding	4	1.70212766	6.67E–04
chemotaxis	6	2.553191489	8.23E–04
Immunoglobulin domain	15	6.382978723	8.25E–04
inflammation	4	1.70212766	9.79E–04
extracellular matrix	10	4.255319149	9.99E–04
gpi-anchor	7	2.978723404	0.003569296
cell adhesion	12	5.106382979	0.005673897
hydroxylation	5	2.127659574	0.006928978
thiolester bond	3	1.276595745	0.007127014
immunoglobulin c region	3	1.276595745	0.007127014
sulfation	4	1.70212766	0.007435725
phosphoprotein	93	39.57446809	0.007720982
ATP	7	2.978723404	0.009174195
membrane	81	34.46808511	0.01649583
Fatty acid biosynthesis	4	1.70212766	0.01701409
cell membrane	31	13.19148936	0.018684946
tumor suppressor	5	2.127659574	0.020455856
calmodulin-binding	5	2.127659574	0.033890063
sh3 domain	7	2.978723404	0.034520031
ubl conjugation	12	5.106382979	0.047435024
lipoprotein	13	5.531914894	0.047618299
duplication	5	2.127659574	0.054099645
oxidoreductase	12	5.106382979	0.077770979

(*Continued*)

Animal	Genotype	Gender	Bcl6	Blimp1	CD43	IgM	IgD
2222	ko	f	6.066195	1.687667	290.3511	0.311407	1.574364
1958	ko	f	0.133845	0.609388	0.570343	0.496057	0.051101
2320	ko	m	0.043995	0.332472	0.113839	0.673455	0.321443
2581	ko	m	0.044904	0.251925	0.063658	0.500618	0.190218
2590	ko	f	0.092365	0.496068	0.172707	0.88351	0.284531
2843	ko	m	0.203623	0.472634	0.582692	1.103613	0.227785
2949	ko	f	1.158255	0.482188	0.921204	0.36023	0.031063
2589	ko	m	0.070043	0.372983	0.223328	0.789975	0.045549