


5-2020

Heterogeneous Nuclear Ribonucleoprotein K (hnRNP K) Overexpression and its Interaction with RUNX1 RNA in Acute Myeloid Leukemia

Marisa Aitken

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

 Part of the [Animal Experimentation and Research Commons](#), [Cancer Biology Commons](#), [Cell Biology Commons](#), [Genetics Commons](#), [Hemic and Lymphatic Diseases Commons](#), [Immune System Diseases Commons](#), and the [Immunopathology Commons](#)

Recommended Citation

Aitken, Marisa, "Heterogeneous Nuclear Ribonucleoprotein K (hnRNP K) Overexpression and its Interaction with RUNX1 RNA in Acute Myeloid Leukemia" (2020). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 994.

https://digitalcommons.library.tmc.edu/utgsbs_dissertations/994

This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN K (HNRNPK) OVEREXPRESSION
AND ITS INTERACTION WITH *RUNX1* IN ACUTE MYELOID LEUKEMIA

Marisa Janelle Lynne Aitken, B.S.

APPROVED:

Sean Post, Ph.D.

Advisory Professor

Scott Evans, M.D.

Marina Konopleva, M.D., Ph.D.

Dean Lee, M.D., Ph.D.

M. James You, M.D., Ph.D.

APPROVED:

Dean, The University of Texas

MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN K (HNRNPK) OVEREXPRESSION
AND ITS INTERACTION WITH *RUNX1* IN ACUTE MYELOID LEUKEMIA

A
DISSERTATION

Presented to the Faculty of
The University of Texas
MD Anderson Cancer Center UTHealth
Graduate School of Biomedical Sciences
in Partial Fulfillment
of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

By

Marisa Janelle Lynne Aitken, B.S.

Houston, Texas

May, 2020

Copyright

I did not formally copyright this work; however, it should be evident that the work here is mine.

Dedication

~

This work is dedicated...

to Dr. T., for encouraging me to pursue my curiosity,
to the patients who have allowed me to share in their journeys,
to Sam, for accepting the delayed gratification of my career path

~

Acknowledgements

It will become increasingly evident in this section that I am but a turtle on a post—I did not get here by myself.

I must thank my mentor, Sean Post, who has fostered an environment driven by self-motivation and scientific curiosity. Your ability to approach perplexing scientific scenarios with passion and rigor has fundamentally changed my understanding of this field.

I am immeasurably thankful to Prema Malaney, the most driven, talented, organized, and gracious postdoctoral fellow I have yet had the privilege to work with. Thank you for your incredible teamwork over the past several years, for the hundreds of plasmids you've cloned, and for being the first (and perhaps the last) individual to read this document in its entirety. Thank you for expanding beyond your job description to fill needed roles.

I am substantially indebted to Shelley Herbrich, who has graciously and effortlessly played the role of bioinformatician, biostatistician, science sister, mini-PI, confidant, bridesmaid, fact-checker, counselor, figure-generator, conference buddy, and innate immunologist. Words are insufficient to express my gratitude.

The members of the Post laboratory have been key players in the progress of this work. Thank you to: Xiaorui Zhang—for keeping our laboratory running, optimizing figures, incessantly editing, and keeping track of Seanisms. Lauren Chan—for being a trustworthy and reliable set of hands when mine were typing this document. Ruizhi Duan—for the FISHing trips. Miguel Gallardo—for generating preliminary data that enticed me to venture into the mouse world.

I also very much appreciate Riitta Nolo—learning from your unwavering belief that one should not be emotionally invested in the outcome an experiment has spared me much heartache.

An exuberant and heartfelt thank you:

To Patrick Zweidler-McKay—for instilling in me a love of hematology.

To Dean Lee—for teaching me how to write, empowering me to make difficult decisions, and providing continued, valuable feedback—even from afar.

To Scott Evans, Marina Konopleva, and M. James You—thank you for believing in my ability to develop a coherent story, providing useful guidance, and supporting my timeline.

To Michelle Barton and the members of her laboratory including Abhinav Jain, Sabrina Stratton, Vrutant Shah, Kendra Allton, and Lalit Patel—for consistently providing constructive feedback and encouragement over many years as I grew into this project.

To John and Charlene Kopchick—for their kind words, high expectations, and generous donations that directly funded major aspects of this work.

To Mr. Ma—for caring about my well-being and providing me with fresh, organic produce from his garden.

To Taghi Manshoury—for many laughs and many, many reagents.

To the members of Caspase Club, particularly Joya Chandra, Lisa Bouchier-Hayes, and Keri Schadler—for providing valuable outside perspectives, encouragement, and snacks.

To Chris Benton—for teaching me about the clinical aspects of this disease, allowing me to become a member of the clinic team, and providing a joke calendar that made 2019 particularly enjoyable.

To Nick Swanson, for reminding me why I study this disease.

To Carlos Bueso-Ramos and Rashmi Kanagal-Shamanna—for patiently explaining hematopathology.

To Wendy Schober and Nalini Patel—for facilitating my experiments by sorting the same cells repeatedly and for years.

To those affiliated with the MD/PhD program.

To the veterinary staff, particularly Angela Asch—for maintaining and caring for the mice used in my studies.

To my family—those of blood and those of choice—for being a fantastic cheerleading squad and fundamentally supporting my goals.

Finally—to my truest and best friend, Sam. Thank you for being the ears that hear the constant narrative, the eyes that see the sinusoidal emotions, and yet the heart that loves fiercely, completely, and without condition. You have been my ally in conflict and my comrade in adventure. I appreciate that you have incorporated the alphabet soup that is my research into your vocabulary. Thank you for cooking delicious meals, hiring a maid, funding our adventures, and for loving me without reservation, despite the fact that it will still be many years until I get a real job.

HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN K (HNRNPK) OVEREXPRESSION AND ITS INTERACTION WITH *RUNX1* IN ACUTE MYELOID LEUKEMIA

Marisa Janelle Lynne Aitken, B.S.

Advisory Professor: Sean Post, Ph.D.

Acute myeloid leukemia (AML) is an often devastating hematologic malignancy with 5-year overall survival lingering near 20%. Acquiring a deeper understanding of molecular underpinnings of leukemogenesis will provide a basis for developing more effective therapeutic strategies for patients with AML.

Here, we identified overexpression of hnRNP K as a recurrent abnormality in a subset (~20%) of AML patients. High levels of this RNA-binding protein associated with inferior clinical outcomes in *de novo* AML. Thus, to evaluate its putative oncogenic capacity in myeloid disease, we overexpressed hnRNP K in murine hematopoietic stem and progenitor cells isolated from fetal liver cells (FLCs). We revealed that hnRNP K-overexpression alters self-renewal capacity and differentiation potential of these cells *in vitro*. Such findings were recapitulated *in vivo*, as murine recipients of hnRNP K-overexpressing FLCs developed fatal myeloproliferative phenotypes.

To elucidate mechanisms by which hnRNP K overexpression causes myeloid neoplasia, we took an unbiased approach utilizing RNA-immunoprecipitation sequencing (fRIP-Seq). Among RNA transcripts interacting with hnRNP K was *RUNX1*—a pivotal transcriptional regulator of definitive hematopoiesis commonly mutated or translocated in AML. Consensus hnRNP K binding sites were identified in the 5' UTR and near the 3'

splice site in intron 5-6 of *RUNX1*. Fluorescence anisotropy studies confirmed these interactions were direct, and abrogated when hnRNP K binding sites within *RUNX1* were mutated. Manipulating hnRNP K expression in human cell lines and murine FLCs substantially altered *RUNX1* splicing surrounding exon 6. RNA-sequencing of FLCs confirmed these data, exposing *RUNX1* as a significantly differentially spliced entity in the context of hnRNP K overexpression. Importantly, the protein product of this spliced product (*RUNX1* Δ Ex6) exhibited disparate transcriptional activity in reporter assays compared to full-length *RUNX1*. Furthermore, we identified KH3 as the hnRNP K domain most critical for these splicing alterations; deletion of KH3 markedly abrogated hnRNP K overexpression phenotypes *in vitro*.

In sum, we established hnRNP K as an oncogene in myeloid leukemia that binds *RUNX1* RNA, altering its splicing and subsequent transcriptional activity. These findings shed light on a mechanism of myeloid leukemogenesis, paving the way for drug discovery efforts to improve outcomes for patients with this disease.

Table of Contents

Approval Sheet	i
Title Page	ii
Copyright	iii
Dedication	iv
Acknowledgements	v
Abstract.....	viii
Table of Contents.....	x
List of Figures	xiii
List of Tables.....	xv
Abbreviations	xvii
Chapter 1. Introduction.....	1
1.1 Acute myeloid leukemia.....	2
1.2 Heterogeneous nuclear ribonucleoprotein K (hnRNP K).....	4
1.3 RUNX1	5
1.4 RNA Splicing and its role in hematologic malignancies.....	7
1.5 hnRNP K and splicing.....	9
1.6 Regulation of RUNX1 by splicing.....	9
1.7 Conclusion	10
1.8 Hypothesis and research goals	10
1.9 Significance.....	11

Chapter 2. hnRNP K is overexpressed in acute myeloid leukemia and is associated with inferior outcomes	12
2.1 Introduction	13
2.2 Materials and Methods	15
2.3 Results	18
2.4 Discussion.....	28
Chapter 3. hnRNP K overexpression drives myeloproliferative disease in murine models.....	32
3.1 Introduction	33
3.2 Materials and Methods	34
3.3 Results	40
3.4 Discussion.....	55
Chapter 4. Examining the molecular basis of the oncogenic function of hnRNP K.....	58
4.1 Introduction	59
4.2 Materials/Methods	60
4.3 Results	65
4.4 Discussion.....	73
Chapter 5. Evaluating the functional consequence of the hnRNP K- <i>RUNX1</i> interaction	75
5.1 Introduction	76
5.2 Materials/Methods	79
5.3 Results	89
5.4 Discussion.....	109

Chapter 6. Discussion	112
References.....	121
Vita	166

List of Figures

Figure 1: Schematic of hnRNP K.....	4
Figure 2. <i>HNRNPK</i> mutations in AML at MD Anderson	18
Figure 3. <i>HNRNPK</i> RNA is overexpressed in a subset of AML.....	20
Figure 4. hnRNP K protein is overexpressed in a subset of AML	21
Figure 5. hnRNP K overexpression is associated with inferior outcomes	22
Figure 6. hnRNP K is elevated in more immature AML.....	23
Figure 7. hnRNP K overexpression is enriched in <i>NPM1</i> mutant cases.....	24
Figure 8. hnRNP K overexpression in the context of <i>NPM1</i> mutations is associated with decreased overall survival.....	25
Figure 9. <i>HNRNPK</i> locus is duplicated in a subset of AML	26
Figure 10. Additional <i>HNRNPK</i> loci exist as small supernumerary marker chromosomes	27
Figure 11. FLCs can overexpress hnRNP K.....	40
Figure 12. hnRNP K overexpression increases colony formation potential of FLCs.....	41
Figure 13. hnRNP K-overexpressing colonies have altered immunophenotypes.....	42
Figure 14. hnRNP K-overexpressing FLCs engraft in recipient NSG mice	44
Figure 15. hnRNP K overexpression is maintained <i>in vivo</i>	45
Figure 16. Survival of FLC recipients.....	46
Figure 17. Recipients of hnRNP K-overexpressing FLCs have bone marrow abnormalities	47
Figure 18. Recipients of hnRNP K-overexpressing FLCs develop splenomegaly	48
Figure 19. hnRNP K overexpression results in splenic extramedullary hematopoiesis	49
Figure 20. Hepatic extramedullary hematopoiesis in recipients of hnRNP K-overexpressing FLCs	51
Figure 21. Peripheral blood counts in mice transplanted with FLCs	52
Figure 22. Peripheral blood is abnormal in recipients of hnRNP K-overexpressing FLCs.....	54

Figure 23. Immunoprecipitation followed by mass spectrometry reveals hnRNP K interacting partners.....	66
Figure 24. hnRNP K binds to the <i>RUNX1</i> transcript	67
Figure 25. <i>RUNX1</i> mRNA interacts with hnRNP K	68
Figure 26. Predicted hnRNP K binding sites in <i>RUNX1</i>	69
Figure 27. hnRNP K binds stretches of poly(C) RNA derived from <i>RUNX1</i>	70
Figure 28. hnRNP K binds human and mouse <i>RUNX1</i>	71
Figure 29. <i>RUNX1</i> oligos thermostabilize hnRNP K	72
Figure 30. <i>Hnmpk</i> overexpression results in differential gene expression in murine FLCs	89
Figure 31. <i>Hnmpk</i> overexpression alters splicing in FLCs.....	90
Figure 32. <i>Runx1</i> is differentially spliced in hnRNP K overexpressing FLCs.....	91
Figure 33. hnRNP K overexpression leads to exclusion of <i>RUNX1</i> exon 6.....	93
Figure 34. Knockdown of hnRNP K in K562 increases <i>RUNX1</i> exon 6 inclusion.....	95
Figure 35. Addback of hnRNP K to shHNRNPK cells rescues <i>RUNX1</i> splicing.....	96
Figure 36. A smaller <i>RUNX1</i> isoform corresponds with hnRNP K expression at the protein level	97
Figure 37. <i>RUNX1</i> Δ Ex6 is present in AML patient samples	98
Figure 38. FLCs overexpressing hnRNP K have more <i>Runx1</i> Δ Ex6	99
Figure 39. <i>RUNX1</i> Δ Ex6 is more stable than full-length <i>RUNX1</i>	101
Figure 40. Full-length <i>RUNX1</i> and <i>RUNX1</i> Δ Ex6 have differential transcriptional activity.....	102
Figure 41. KH3 domain of hnRNP K mediates <i>RUNX1</i> exon 6 splicing	103
Figure 42. The KH3 domain of hnRNP K is required for an <i>in vitro</i> phenotype associated with hnRNP K overexpression.....	104
Figure 43. hnRNP K alterations affect global translation and that of <i>RUNX1</i>	107

List of Tables

Table 1. Co-occurring mutations and cytogenetic features of AML cases with mutant HNRNPK	19
Table 2. Antibodies used for IHC.....	38
Table 3. Fluorescence anisotropy oligos.	64
Table 4. Primers used in PCR.....	81
Table 5. Primers used for cloning.....	85

Abbreviations

3'ss	3' splice site
5'ss	5' splice site
6-FAM	6-Carboxyfluorescein
ABC	Avidin–Biotin Complex
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
AML1/ETO	Acute Myeloid Leukemia 1 Protein/Eight Twenty One Protein fusion
APL	Acute promyelocytic leukemia
bp	base pair
BRD4	Bromodomain Containing 4
BSA	bovine serum albumin
CBC	complete blood count
CBF-AML	core binding factor-acute myeloid leukemia
CBFB	Core-Binding Factor Subunit Beta
CD117	Cluster of Differentiation 117, same as c-kit
CD11b	Cluster of Differentiation 11b
CD14	Cluster of Differentiation 14
CD3	Cluster of Differentiation 3
CD34	Cluster of Differentiation 34
CD44	Cluster of Differentiation 44
CD45	Cluster of Differentiation 45
CDKN1A	Cyclin Dependent Kinase Inhibitor 1A
cDNA	complementary DNA
CEBPa	CCAAT Enhancer Binding Protein Alpha

CI	confidence interval
c-kit	KIT Proto-Oncogene, Receptor Tyrosine Kinase, same as CD117
CMV	cytomegalovirus
c-Myc	V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
CNIO	Centro Nacional de Investigaciones Oncológicas, National Centre for Cancer Research
C-terminus	carboxy terminus
Ctrl	control
DAB	3,3'-Diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DDX41	DEAD-Box Helicase 41
DLBCL	Diffuse Large B-cell Lymphoma
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DNMT3A	DNA Methyltransferase 3 Alpha
dNTP	deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescent protein
EV	empty vector
FA	Fluorescence Anisotropy
FAB	French-American-British
FBXW7	F-Box And WD Repeat Domain Containing 7
FDR	False discovery rate
FISH	Fluorescence in situ hybridization
FLCs	fetal liver cells
FLT3	Fms Related Receptor Tyrosine Kinase 3

FPDMM	familial platelet disorder with propensity to myeloid malignancy
fRIP	formaldehyde-fixed RNA Immunoprecipitation
G6PD	Glucose-6-Phosphate Dehydrogenase
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GATA2	GATA Binding Protein 2
GFP	green fluorescent protein
GO	gene ontology
Gr1	granulocytic marker
Gy	gray
H&E	Haemotoxylin and Eosin
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hgb	hemoglobin
hnRNP	Heterogeneous Nuclear Ribonucleoprotein
hnRNP K	Heterogeneous Nuclear Ribonucleoprotein K
HR	hazard ratio
HSC	Hematopoietic stem cells
HSP90	Heat Shock Protein 90
IDH1	Isocitrate Dehydrogenase (NADP(+)) 1
IDH2	Isocitrate Dehydrogenase (NADP(+)) 2
IGEPAL	Octylphenoxy poly(ethyleneoxy)ethanol
IgG	immunoglobulin
IGV	Integrative Genomics Viewer
IHC	immunohistochemistry
IL-3	Interleukin 3
IL-5	Interleukin 5
IL-6	Interleukin 6

IP	immunoprecipitation
IRES	internal ribosome entry site
IVIS	in vivo imaging systems
JAK2	Janus Kinase 2
KCl	potassium chloride
kDa	kilodalton
KH	K Homology domain
KI	K interactive domain
KNS	K nuclear shuttling domain
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LC-MS-MS	Liquid chromatography–mass spectrometry-mass spectrometry
LIC	large immature cells
LTR	Long terminal repeats
MDS	Myelodysplastic syndromes
MgCl ₂	Magnesium chloride
MPO	Myeloperoxidase
mRNA	Messenger Ribonucleic Acid
MSCV	Murine stem cell virus
MYH11	Myosin Heavy Chain 11
NaCl	sodium chloride
Ni-NTA	nickel-nitrilotriacetic acid
NP40	nonyl phenoxyethoxyethanol
NPM1	Nucleophosmin 1
NRAS	Neuroblastoma RAS Viral Oncogene Homolog
NSG	NOD scid gamma
OE	overexpression

ORF	Open reading frame
OS	overall survival
P2A	porcine teschovirus-1 2A peptide
PAGE	Polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCBP2	Poly(RC) Binding Protein 2
PCR	polymerase chain reaction
PGK	Phosphoglycerate Kinase 1
Plt	platelet
PML-RARA	Promyelocytic Leukemia/Retinoic Acid Receptor Alpha
PPIA	Peptidylprolyl Isomerase A
PVDF	Polyvinylidene fluoride
qRT-PCR	Quantitative Real-time polymerase chain reaction
RBC	red blood cell
RFP	red fluorescent protein
RHD	unt protein homology domain
RIP	RNA-immunoprecipitation
RNA	Ribonucleic Acid
RNA-Seq	Ribonucleic Acid sequencing
RPLP0	Ribosomal Protein Lateral Stalk Subunit P0
RPM	Revolutions per minute
RPMI	Roswell Park Memorial Institute medium
RPPA	Reverse Phase Protein Arrays
RUNX1	Runt Related Transcription Factor 1
RUNX1T1	RUNX1 Partner Transcriptional Co-Repressor 1
Sca1	Stem cells antigen-1

SCF	stem cell factor
SDS	sodium dodecyl sulfate
SF3B1	Splicing Factor 3b Subunit 1
shRNA	short hairpin RNA
SIN3A	SIN3 Transcription Regulator Family Member A
siRNA	Small interfering RNA
snRNP	small nuclear ribonucleoproteins
SNV	single-nucleotide variant
SR	Serine/arginine-rich (SR) proteins
SRSF2	Serine And Arginine Rich Splicing Factor 2
sSMC	small supernumerary marker chromosome
TAD	transactivation domain
t-AML	Therapy-related acute myeloid leukemia
Taq	Thermus aquaticus polymerase
TCEP	tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus
TP53	Tumor Protein P53
TPM	Transcripts per million
Tris	trisaminomethane
U2AF1	U2 Small Nuclear RNA Auxiliary Factor 1
UTR	untranslated region
VRC	vanadyl ribonucleoside complex
WBC	white blood cell
WT1	Wilms Tumor 1
ZRSR2	Zinc Finger CCCH-Type, RNA Binding Motif And Serine/Arginine Rich 2

Chapter 1

Introduction

1.1 Acute myeloid leukemia

Acute myeloid leukemia (AML) is a devastating hematologic malignancy wherein normal hematopoiesis is superseded by rapid proliferation of immature myeloid cells. In adults, AML is the most common acute leukemia, and the annual incidence in the United States ranges from 3-5 cases per 100,000 people.^{1,2} While the median age of diagnosis is 70 years, this disease is also found in children and young adults.^{3,4} AML is responsible for most leukemia-related deaths in the United States, where over 10,000 patients died from this disease in 2019 alone.⁵

Hematopoiesis, the process of blood development, occurs in the bone marrow. Through exquisitely orchestrated processes of self-renewal and differentiation, hematopoietic stem cells (HSCs) eventually give rise to all of the cellular components of blood.⁶ These include 1) red blood cells (RBCs), responsible for delivery of oxygen to tissues, 2) white blood cells (WBCs), used to combat infection, and 3) platelets, which aid in blood clotting, or hemostasis. In leukemia, aberrations in HSCs, or their slightly more mature progeny, result in unrestrained proliferation of immature WBCs. These rapidly dividing cells fill the bone marrow, usurping the space in which normal blood cells are made. As these malignant cells crowd out normal hematopoiesis, production of RBCs, WBCs, and platelets becomes severely compromised. This leads to anemia, leukopenia, and thrombocytopenia, respectively. These are responsible for the major clinical manifestations of the disease—all of which can be life threatening: infection (due to leukopenia, specifically neutropenia), bleeding (due to thrombocytopenia), and organ failure (due to anemia).⁷

AML derives its name from the lineage of cells from which it stems. White blood cells are grossly categorized into either myeloid or lymphoid lineages. Acute leukemias of lymphoid lineage are referred to as acute lymphoid leukemias (ALL), while those of myeloid lineage are acute myeloid leukemia (AML).

AML is characterized by several recurrent chromosomal and genetic aberrations. These provide insight not only into the underlying biology of the disease, but can also be used in some

instances to provide prognostic information and to guide treatment decisions. For example, a subtype of AML known as acute promyelocytic leukemia (APL) is defined by a translocation between chromosomes 15 and 17 (t(15;17)).^{8, 9} This chromosomal abnormality results in the formation of a fusion protein known as PML-RARA, which largely contributes to the disease.^{10, 11} Prior to the late 1980s, APL was a universally fatal subtype of AML.¹² Due to concerted efforts to understand the biology of PML-RARA, incredibly effective APL treatments have been tailored, and this disease is now almost universally curable (>97% 5-year overall survival).^{13, 14}

Another subtype of AML with generally fair prognosis includes those with recurrent cytogenetic abnormalities are associated with core-binding factor AML (CBF-AML). This subset of AML includes t(8;21) and inv(16), which lead to the creation of the fusion genes *RUNX1-RUNX1T1* and *CBFB-MYH11*, respectively.¹⁵⁻¹⁸ CBF-AML has a high complete response rate (80-90%) to standard chemotherapy with 3 days of anthracycline along with 7 days of cytarabine (3+7).¹⁹⁻²² In addition, 5-year overall survival (OS) in CBF-AML is more than 60% in younger patients.^{19, 20} Even in older patients with CBF-AML, 5-year OS is near 30%.²³

Unfortunately, not all subtypes of AML have seen this same progress. Taken as a whole, patients under age 60 have a 5-year survival of approximately 40%, while older patients have a much lower (~10%) 5-year survival.^{3, 24, 25}

Mutations in AML can also be used for prognostication and/or to tailor therapy. For example, *NPM1* mutations (in the absence of co-occurring mutant *FLT3*) or biallelic *CEBPA* mutations have been shown to confer a relatively fair prognosis.^{26, 27} Patients harboring other mutations, such as those in *FLT3*, *IDH1*, or *IDH2* can now benefit from the recently FDA-approved small molecule inhibitors of these mutant proteins.²⁸⁻³¹ While these new therapies are improving patient outcomes, many patients lack such targetable mutations.³²⁻³⁴ In addition, patients treated with these inhibitors are often still at risk for relapse.²⁸⁻³¹ This highlights the need to further understand the molecular underpinnings of AML such that alternative, effective therapeutic options can be developed.

1.2 Heterogeneous nuclear ribonucleoprotein K (hnRNP K)

The gene encoding hnRNP K (*HNRNPK*) resides on chromosome 9q21.32. Deletions of chromosome 9q (del(9q)) are recurrent cytogenetic abnormalities observed in approximately 2% of patients with AML.³⁵⁻³⁷ Scattered mutations in this gene have also been identified in AML, albeit at an exceedingly low frequency (~1%).³²⁻³⁴ While these mutations have not been formally characterized, it has been suggested that they result in haploinsufficiency of hnRNP K.³⁸

Functionally, hnRNP K is a renaissance (wo)man of proteins. It has been implicated in a wide variety of normal cellular functions, including signal transduction³⁹⁻⁴³, chromatin remodeling⁴⁴⁻⁴⁷, transcription⁴⁸⁻⁵⁴, RNA splicing⁵⁵⁻⁵⁹, mRNA stability⁶⁰⁻⁶², and translation^{54, 63-69}. Not surprisingly, homozygous *Hnrnpk* knockout is embryonic lethal in mice, supporting its role as an essential gene.⁵³ This broad functionality of hnRNP K is largely due to its composite protein domains. A nuclear localization signal (NLS) and a nuclear shuttling domain (KNS) allow hnRNP K to move bi-directionally between the cytoplasm and nucleus.^{70,71} The K-interactive (KI) domain is largely responsible for the interaction between hnRNP K and other proteins.^{72, 73} In addition, three K homology (KH) domains are located throughout the protein, which recognize RNA and single-stranded DNA.⁷⁴ KH1 and KH2 are located at the amino terminus, while KH3 resides at the carboxy terminus—separated by a large unstructured region, including the KI domain.⁷²⁻⁷⁴ These domains are depicted in Figure 1.

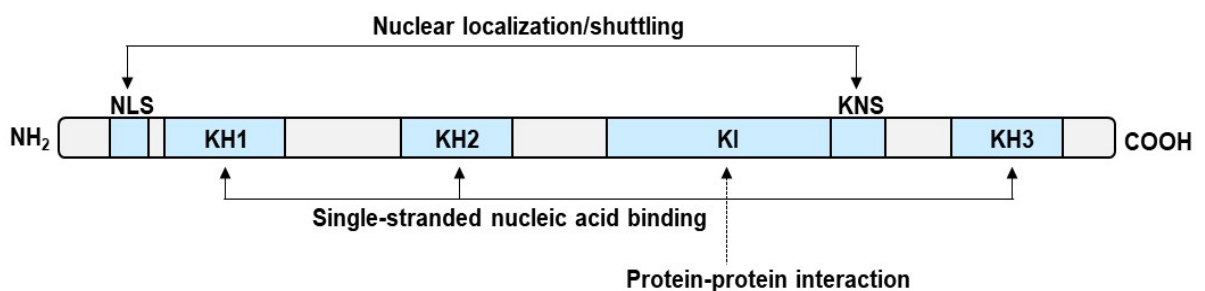


Figure 1: Schematic of hnRNP K: Each domain is indicated. Diagram is not to scale. NLS: nuclear localization signal; KH: K homology domain; KI: K interactive domain; KNS: K nuclear shuttling.

hnRNP K has been implicated in a variety of malignancies. For example, our group developed a mouse model to demonstrate that *HNRNPK* is the haploinsufficient tumor suppressor at the recurrently deleted 9q locus in AML.⁵³ In contrast to this observation, several clinical studies have observed that increased hnRNP K expression is correlated with advanced disease and poor clinical outcomes in several malignancies including breast, pancreatic, and colorectal cancers, among others.⁷⁵⁻⁸² Indeed, we recently identified overexpression of hnRNP K associated with substantially inferior overall survival in a subset of patients with diffuse large B-cell lymphoma (DLBCL), regardless of subtype.⁸³ Critically, when we generated a mouse model of B-cell specific hnRNP K overexpression, mice developed B-cell lymphoma—thus formally demonstrating that hnRNP K is an oncogene in B-cell malignancies when overexpressed.⁸³

1.3 RUNX1

RUNX1 (also called AML1, CBF α 2, or PEBP2 α B) is a pivotal hematopoietic transcription factor.⁸⁴ Homozygous *Runx1* knockout is embryonic lethal in mice due to absence of definitive hematopoiesis and propensity to hemorrhage.⁸⁵ *Runx1* heterozygotes have diminished erythroid and myeloid progenitor cells.⁸⁵ When *Runx1* is conditionally deleted in adult mice, significant hematopoietic alterations are observed, including inefficient production of platelets and common lymphocyte progenitors.⁸⁶ These mice also display a myeloproliferative phenotype.⁸⁶

RUNX1 is located on chromosome 21 in humans and chromosome 16 in mouse. Two alternative promoters—distal (P1) and proximal (P2)—control expression of *RUNX1* in vertebrates.⁸⁷ Isoforms arising from P1 and P2 have different 5' untranslated regions (UTRs) and coding sequences at the amino terminus, but are otherwise identical in their full-length forms.⁸⁷ *RUNX1C* is the major isoform transcribed from P1.⁸⁷ Utilization of P2 results in the formation of *RUNX1B*.⁸⁷ In primates, alternative exon usage also results in *RUNX1A*, which

lacks the 203 N-terminal amino acids present in *RUNX1B*.⁸⁷⁻⁹⁰ In mice, a functional ortholog of *RUNX1A* is generated by use of an extended terminal exon, referred to as *Runx1bEx6e*.⁸⁸

All of these major *RUNX1* isoforms share a conserved runt-homology domain (RHD), which is responsible for sequence-specific DNA recognition and binding, as well as heterodimerization with CBF β .^{91, 92} The longer isoforms (*RUNX1B* and *RUNX1C*) also share a transactivation domain (TAD) and inhibitory domain, where protein-protein interactions occur that can influence the transcriptional activity of *RUNX1*.^{93, 94}

While all *RUNX1* isoforms have been identified throughout various temporal stages in hematopoiesis, *RUNX1B* predominates in the adult.⁹⁵ *RUNX1C* is mostly present early in development, as definitive HSCs emerge⁹⁵, but is re-expressed in adult B-cells, where it inhibits proliferation⁹⁶. In contrast to the embryonic lethality observed in complete *Runx1* knockout mice, knockout of only the P1 locus (encoding this C isoform) does not result in an overt phenotype.^{86, 97} Throughout this work, we refer to *RUNX1B* as simply “*RUNX1*” unless otherwise specified.

Overexpression of human *RUNX1A* in mouse models leads to lymphoid leukemia⁹⁸, expansion of hematopoietic stem and progenitor cell populations⁹⁹, and enhanced engraftment into recipient mice.^{99, 100} In striking contrast, overexpression of the longer isoforms of *RUNX1* leads to altered hematopoietic differentiation, favoring monocytes over granulocytes¹⁰¹, decreased engraftment potential^{95, 99}, and p53-dependent senescence^{102, 103}.

In leukemias of various lineages, *RUNX1* is a common target of chromosomal translocations or mutations. Indeed, *RUNX1* was originally identified in the context of the t(8;21) translocation in AML (thus its alias, *AML1*).¹⁰⁴ Other recurrent translocations affecting *RUNX1* include t(3;21) in myelodysplastic syndrome (MDS) and therapy-related AML (t-AML)¹⁰⁵, and t(12;21) in pediatric ALL¹⁰⁶, though dozens more have been reported¹⁰⁷. Somatic point mutations in *RUNX1* are also evident in AML, ALL, and MDS^{108, 109}, and are associated with high-risk disease¹¹⁰⁻¹¹⁵. Furthermore, germline *RUNX1* mutations are associated with familial platelet disorder with associated myeloid malignancy (FPDMM).¹¹⁶

1.4 RNA Splicing and its role in hematologic malignancies

Though only approximately 20,000 protein-coding genes have been identified in the human genome^{117, 118}, many more proteins have been observed in cells^{119, 120}. In fact, the vast majority of these genes are thought to give rise to multiple isoforms.^{121, 122} Such entities are occasionally referred to as “spliceforms”, as it is now widely accepted that alternative splicing is the major mechanism behind this phenomenon.

Early descriptions and imaginations of alternative splicing were documented in the 1970s based on the observed piece-like structure of bacterial genes.^{123, 124} (These references are worth reading, as they offer intriguing historic insight via fascinating speculations on a phenomenon that was, at the time, almost completely enigmatic.) It is now appreciated that splicing is required for the maturation of the vast majority of mRNAs, and occurs in the nucleus where non-coding introns are excised from pre-mRNA to allow for a continuous stretch of protein-coding exons in a mature transcript prior to translation.^{121, 122} Use of alternative splice sites results in selective use of exons.¹²⁵ This contributes to the immense proteomic diversity observed in cells.^{119, 120}

Large molecular complexes known as the major and minor spliceosome mediate the RNA splicing process.¹²⁵⁻¹²⁷ These are comprised of several small nuclear ribonucleoprotein complexes (snRNPs; consisting of small RNAs complexed with proteins): U1, U2, U4, U5, and U6 in the major spliceosome¹²⁵, and U5, U11, U12, U4, and U6 in the minor spliceosome^{126, 128}. As transcription occurs, the U1 snRNP binds the 5' splice site (5'ss; located between the upstream exon and intron).^{125, 126, 129} The 3'ss is soon bound by U2 auxiliary factors (U2AFs).^{126, 129} Shortly thereafter, the U2 snRNP binds sequence motifs at the 3' splice site (3'ss; flanking the intron and downstream exon).^{125, 126} Finally, U4/U5/U6 snRNPs are recruited, the fully assembled spliceosome assumes an active conformation, and two sequential transesterification reactions occur, resulting in intron excision and ligation of adjacent exons.¹²⁵

Promoters and repressors of splicing can further impact splicing outcomes. Serine-arginine-rich proteins (SR proteins) and heterogeneous nuclear ribonucleoproteins (hnRNPs) are

canonically classified as promoters and repressors, respectively.¹³⁰⁻¹³⁴ However, their impact on splicing can vary depending on a variety of factors, including the relative location within the pre-mRNA in which they bind, cellular context, expression level or mutation of the splicing factor, etc.¹³⁵⁻¹³⁸

Alterations in RNA splicing have been increasingly recognized in numerous solid and hematologic malignancies. In particular, mutations in genes encoding splicing factors are common in MDS and AML.¹³⁹⁻¹⁴¹ The most common of these mutated splicing factors are serine/arginine-rich splicing factor 2 (SRSF2), splicing factor 3B subunit 1 (SF3B1), U2 small nuclear auxiliary factor 1 (U2AF1), and zinc finger RNA-binding motif and serine/arginine-rich 2 (ZRSR2), all of which contribute to recognition of the 3'ss.¹³⁹⁻¹⁴¹ These mutations are invariably mutually exclusive with one another. With the exception of *ZRSR2*, the observed mutations are heterozygous and affect specific amino acid residues.³²⁻³⁴ Mutations in *ZRSR2* confer loss of function, whereas mutations in the remainder of these genes confer gain of function and/or altered functionality.¹⁴² Individual splicing factor mutations have been correlated with specific subtypes of myeloid malignancies¹⁴³, indicating that differentially altered splicing can result in varied phenotypic consequences. For example, *Srsf2*^{P95H} mice develop features consistent with MDS, such as leukopenia, morphologic dysplasia, and increased numbers of myeloid progenitor cells.¹⁴⁴ While *U2af1*^{S34F} mice do not develop the same dysplasia, though leukopenia and increased myeloid progenitor cells are evident.¹⁴⁵

Even in the absence of these splicing mutations, aberrant splicing in myeloid malignancies has been described.¹⁴⁶ This is likely due in part to the fact that dysregulated splicing can give rise to oncogenic protein isoforms.^{147, 148} Understanding the factors involved in aberrant splicing as well as the consequences of aberrant protein isoform expression are thus important in more deeply understanding disease.

1.5 hnRNP K and splicing

hnRNP K has been shown to influence the splicing of several genes.^{56, 58, 59, 149-151} This is in part due to its interactions with splicing factors, such as SRSF7 (also called 9G8) and SRSF3 (also called SRp20).⁴⁶ In hepatocellular carcinoma, hnRNP K has been shown to alter splicing of *G6PD* in opposite directions depending on its interacting protein partners.¹⁵²

As a poly(C) binding protein, hnRNP K can bind polypyrimidine tracts in the proximity of both 3' and 5' splice sites.^{56, 74, 153} In the case of chicken β -tropomyosin pre-mRNA, hnRNP K promotes splicing of an alternative exon 6A.⁵⁵ Likewise, inclusion of *MRPL33* exon 3 is dependent on elevated levels of hnRNP K.¹⁵⁰ Contrastingly, hnRNP K acts as a splicing inhibitor in the case of Bcl-X_s, preventing production of this pro-apoptotic entity.⁵⁸ Thus, even in the case of direct RNA-binding, hnRNP K can have variable effects on pre-mRNA splicing.¹⁵⁴

Indirectly, hnRNP K can alter splicing by regulating expression of splicing factors. This is exemplified in gastric cancers, where hnRNP K transcriptionally upregulates of *SRSF1*, which in turn alters splicing of *CD44*.¹⁵¹

Together, these observations demonstrate that hnRNP K can directly and indirectly influence splicing in a context-dependent manner.

1.6 Regulation of RUNX1 by splicing

Like the majority of protein-coding genes with more than one exon, *RUNX1* pre-mRNA undergoes splicing. As discussed in section 1.3, the isoforms *RUNX1A* and *RUNX1B* arise due to alternative splicing. However, the mechanisms underlying the specifics of this splicing have not been fully elucidated. Another isoform of *RUNX1* lacks an internal 64 amino acid residues (corresponding to exon 6), and has been identified in mouse and human.^{88, 93, 149, 155} In ovarian cancer, this isoform (herein called *RUNX1 Δ Ex6*) has been correlated with inferior outcomes.¹⁵⁵ Poly(C) binding proteins, including PCBP2 and hnRNP K have been shown to play a role in this

splicing event.^{56, 149} However, the existence, regulation, and function of *RUNX1*ΔEx6 has not been characterized in AML.

1.7 Conclusion

Acute myeloid leukemia is a heterogeneous group of hematologic malignancies. With few exceptions, outcomes for patients with this disease are suboptimal. We have outlined several features of AML that can be used as prognostic and therapeutic guides. Despite newly approved therapeutic agents, many patients remain ineligible for these drugs—mandating further drug development via a deeper understanding of leukemogenic mechanisms. To this end, we discussed a role of the highly multifunctional protein hnRNP K in AML and reviewed the biology of the critical transcription factor, *RUNX1*, in this disease. Furthermore, we discussed pre-mRNA splicing biology and the aberrant splicing observed in AML and other myeloid malignancies. We presented the role of hnRNP K in splicing, and finally, discussed alternative splicing of *RUNX1*.

1.8 Hypothesis and research goals

This dissertation begins with an investigation into the expression and clinical impact of hnRNP K in AML. While haploinsufficiency of *HNRNPK* has been described, along with scarce mutations in this gene, no studies have systematically evaluated hnRNP K expression in cases of AML. In Chapter 3, we formally test the hypothesis that hnRNP K is an uncharacterized oncogene in AML by developing and characterizing a mouse model of hnRNP K-overexpression. In Chapter 4, we begin to evaluate a mechanism of the oncogenic functions of hnRNP K by utilizing unbiased screening methods. In the fifth and final experimental chapter, we characterize the interaction between hnRNP K and *RUNX1* RNA *in vitro* and *in vivo*.

Our working hypothesis is that hnRNP K, when overexpressed, is an uncharacterized oncogene in AML that functions in part through its post-transcriptional interactions with *RUNX1*.

To this end, I endeavored to:

1. Evaluate the hnRNP K expression in AML and characterize its clinical impact
2. Describe the hematologic consequences of hnRNP K overexpression *in vivo*
3. Elucidate the molecular basis of the oncogenic function of hnRNP K
4. Examine the functional consequences of the hnRNP K-*RUNX1* interaction

1.9 Significance

This work will provide great insight into the role of hnRNP K overexpression in myeloid biology and leukemogenesis. By understanding the interaction between *RUNX1* and hnRNP K, future studies will be able to explore therapeutic opportunities to alter this interaction and thereby dramatically improve outcomes for patients with hnRNP K-overexpressing AML.

Chapter 2

hnRNP K is overexpressed in acute myeloid leukemia and is associated with inferior outcomes

2.1 Introduction

Acute myeloid leukemia (AML) encompasses a constellation of diseases characterized by recurrent genetic and/or cytogenetic abnormalities. Outcomes for patients with AML have historically been quite dismal. For example, in the 1970s, 5-year overall survival (OS) for patients diagnosed with AML and treated at MD Anderson Cancer Center was a mere 13%.¹² With improvements in chemotherapeutic regimen composition and dosing, as well as supportive care, this number improved to 49% in 2015.¹² To further improve upon these outcomes, the hematology community has expended incredible efforts to understand the driving events underlying AML development such that better therapeutic options can be created, thereby prolonging the quantity and quality of patient life. Indeed, many of the mutated proteins identified in AML (FLT3, IDH1, IDH2, JAK2) can now be targeted with small molecular inhibitors that have either been FDA-approved or are in late phases of clinical trials.³³

In 2013, The Cancer Genome Atlas published a landmark study delineating the most commonly mutated genes in *de novo* AML.³² This list of genes has been validated and expanded upon by others.^{33,34} Among these genes is *HNRNPK*, which encodes for the RNA-binding protein heterogeneous nuclear ribonucleoprotein K (hnRNP K).^{32,33} *HNRNPK* mutations were described to occur in approximately 1% of AML patients in these studies.³²⁻³⁴ In addition to the small proportion of patients with *HNRNPK* mutations, the mutations that were described did not cluster in a single hotspot of the gene, nor did they aggregate in a portion of the gene encoding for a single functional domain. This motivated us to evaluate our own patient population at MD Anderson Cancer Center, where targeted exonic *HNRNPK* sequencing has been a routine clinical test since 2016.¹⁵⁶

HNRNPK resides on chromosome 9q21.32, which is specifically deleted in ~2% of patients with AML, and leads to haploinsufficiency.³⁵⁻³⁷ Our group previously demonstrated that decreased expression of *HNRNPK* causes myeloid malignancy in a mouse model, thus confirming its role as a haploinsufficient tumor suppressor.⁵³ This finding highlights that gene

dosage, perhaps independent of mutation status, may also contribute to AML development. Therefore, we evaluated *HNRNPK* expression levels in AML patients.

To assess whether *HNRNPK* aberrations beyond genomic aberrations may be evident in these patient populations, we evaluated RNA and protein expression and correlated these results with patient outcomes. Taken together, our data demonstrate that hnRNP K is highly expressed in a subset of patients with AML and these patients have inferior clinical outcomes compared to patients with lower levels of hnRNP K.

2.2 Materials and Methods

HNRNPK Sequencing: Genomic DNA was extracted from bone marrow specimens with an Autopure extractor (Qiagen, Valencia, CA, USA). Libraries were prepared from genomic DNA using hybridization capture-based enrichment of regions of interest. For *HNRNPK*, this included exons 3-7 (amino acids 1-95) and 7-17 (amino acids 100-465). A next generation sequencing platform was used to perform bidirectional paired-end sequencing to detect single nucleotide variants (SNVs) and insertions or deletions up to 52 base pairs (Illumina Inc., San Diego, CA, USA). GRCh37/hg19 was the genomic reference sequence used. Total sequencing coverage depth was at least 250 reads. The lower limit of detection was 5% for clinical reporting.¹⁵⁶

Fluorescence in situ hybridization (FISH) and Giemsa staining: Bone marrow cells harvested from patients were treated with 0.1µg/mL colcemid for 30 minutes and resuspended in a hypotonic solution for 30 minutes, after which Carnoy's fixative solution (3:1 methanol:glacial acetic acid) was added. Cells were then placed onto slides, and kept at 60°C overnight. The next day, slides were placed in 0.05% trypsin-EDTA for 1 minute and 45 seconds, then sequentially washed in isoton diluent and Gurr's buffer (pH 6.8), Giemsa stained for 45 seconds, and rinsed in water. G-banding was evaluated via microscopy. For FISH hybridization, slides were aged and dehydrated in an ethanol gradient before denaturation took place at 73°C for 5 minutes followed by hybridization for 24 hours at 37°C. The RP11-19G1 probe was developed to recognize chromosome 9p (control) and the *HNRNPK* probe (RP11-101L4) to the 9q21.32 locus. DAPI was used as a counterstain.

qRT-PCR: Bone marrow or peripheral blood samples were subjected to adequate red blood cell lysis with BD Pharm Lyse lysing solution (BD Biosciences, San Jose, CA, USA) and RNA extracted and purified using phenol/chloroform.¹⁵⁷ Samples were treated with DNase for 30

minutes at 37°C, and quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). 1µg of RNA was reverse transcribed using iScript (BioRad, Hercules, CA, USA). qRT-PCR was performed using iTaq Universal SYBR Green Supermix as per manufacturer's instructions (BioRad, Hercules, CA, USA) using an ABI StepOnePlus Real Time PCR System. Expression of *HNRNPK* was evaluated using primers as follows: Forward: 5' GCAGGAGGAATTATTGGGGTC 3', Reverse: 5' TGCACTCTACAACCCTATCGG 3'. Changes in expression were determined by comparing *HNRNPK* expression to the housekeeping control *RPLP0*. Primers for *RPLP0* Forward: 5' CCTTCTCCTTTGGGCTGGTCATCCA 3' and reverse: 5' CAGACACTGGCAACATTGCGGACAC 3'. Individual samples were assayed in triplicate. Calculations were performed using the Pfaffl method comparing expression changes between target genes and housekeeping control.¹⁵⁸

Immunohistochemistry: Formalin-fixed paraffin-embedded bone marrow biopsies were deparaffinized in xylene and rehydrated in an alcohol gradient. Antigen retrieval was performed using 10mM sodium citrate and 0.05% Tween 20 (pH 6.0) in a steam chamber for 45 minutes. Slides were incubated in 3% hydrogen peroxide/methanol solution to deactivate endogenous peroxidases and subsequently incubated with anti-hnRNP K antibody (Abcam, Cambridge, MA, USA, ab18195, 3C2, 1:3000 dilution) at 4°C overnight in a humidity chamber. Slides were washed with 0.1% Tween-PBS before biotinylated secondary antibody was added at room temperature for 30 minutes. Antibody-protein interactions were visualized with Vectastain Elite ABC and DAB peroxidase substrate kits (Vector Laboratories, Burlingame, CA, USA). Nuclear fast red was used as a counterstain.

Reverse phase protein array: Methodology is fully described elsewhere.¹⁵⁹⁻¹⁶¹ Briefly, bone marrow protein samples were printed in replicate in serial dilutions onto slides along with controls for normalization and expression. A panel of 230 strictly validated primary antibodies was used

to probe the slides, and a stable dye is precipitated after addition of a secondary antibody. A complete description of the antibodies is included elsewhere.¹⁶¹ The primary antibody against hnRNP K was from Santa Cruz Biotechnology (Dallas, TX, USA, sc-28380, D6). Stained slides were analyzed with Microvigene software (Version 3.4, Vigene Tech, Carlisle, MA, USA). Data is publicly available at www.leukemiaatlas.org.¹⁶² Acute promyelocytic (FAB M3) cases were excluded.

Data normalization and processing: SuperCurve algorithms¹⁶³ was used to generate a single expression value for each sample from the series of dilutions. The average expression level of each replicate sample was used for downstream analysis. Topographical normalization procedures¹⁶⁴ and loading controls¹⁶⁵ were included. Protein expression levels were described as relative to the median of CD34+ bone marrow specimens from healthy donors.

Statistical analyses: Overexpression of hnRNP K protein was defined as any expression greater than or equal to one standard deviation above the median expression for normal, healthy CD34+ bone marrow. Correlations between these expression levels and clinical features were assessed using Wilcoxon rank-sum test. Survival curves were generated using the Kaplan-Meier method.

2.3 Results

HNRNPK mutations are rare in AML

Since implementation of the 81-gene targeted-sequencing panel that tests for mutations in leukemia-associated genes, samples from 238 unique AML patients have been sequenced at MD Anderson. Of these patients, *HNRNPK* mutations were identified in seven individuals (2.9% of cases). A schematic of these mutations and their location within the coding sequence are depicted in Figure 2. These mutations were not clustered in a single region of the gene, and were distributed throughout regions encoding for several functional domains of the hnRNP K protein.

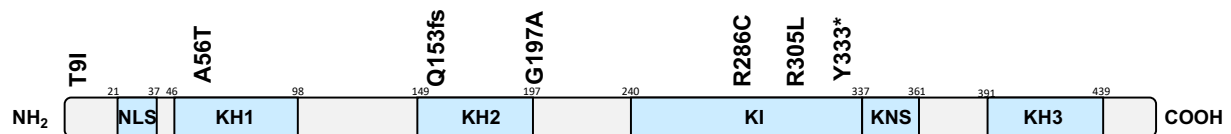


Figure 2. *HNRNPK* mutations in AML at MD Anderson. Schematic of hnRNP K with each domain indicated. Numbers above the domains correspond to the amino acid residue that limit each domain. Mutations identified in AML patients at MD Anderson are depicted on the domain diagram. Each mutation was identified once within our patient cohort. Diagram is not to scale.

Splicing factor mutations are overrepresented in cases with mutant HNRNPK

Table 1 includes the cytogenetics and co-occurring mutations for each case with an identified *HNRNPK* mutation. Due to the very small sample size, it is challenging to draw definitive conclusions about enrichment of co-occurring mutations or cytogenetic abnormalities. However, 5/7 (71.4%) patients with an *HNRNPK* mutation also had a co-occurring mutation in a splicing factor gene, such as *U2AF1*, *SRSF2*, or *SF3B1*. This is more than would be expected, since the incidence of such mutations in AML ranges from 4-6% for *U2AF1*, 1-12% for *SRSF2*, and 1-4% for *SF3B1*, and these mutations are virtually always mutually exclusive of one another.³²⁻³⁴

HNRNPK mutation	Amino acid change	Other mutations	Cytogenetics
26C>T	T9I	IDH2 R140Q NPM1 W288fs SRSF2 P95H	46,XY
166G>A	A56T	KRAS G12C NRAS Q61K RUNX1 R166Q SF3B1 K666E	46,XX,t(3;12)(q26.2;p13),del(7)(q22q36)[18]/46,idem[cp2]
457del	Q153fs	FLT3 D835H TP53 R248G RUNX1 D332fs	47,XY,+8[12]/48,idem,+9[8]
590G>C	G197A	KRAS G12V NRAS G12D U2AF1 S34F WT1 H465L	46,XY
856C>T	R286C	DDX41 R525H DDX41 R159* SRSF2 P95H	46,XY,del(2)(p16)[1]/46,XY,t(8;19)(p21;q13.3)[1]/46,XY[18]
914G>T	R305L	TP53 A119G FBXW7 R505C GATA2 S24*	43~45,Y,-X,der1(t1;5)(q21;q13),-2,-3,del(4)(q22), der(5)del(5)(p14)t(1;5),t(6;7)(p23;q11.2),del(7)(p15),-9,-16, add(21)(q22),+2~5mar[cp20]
998dupA	Y333*	DNMT3A R887fs NPM1 W288fs NRAS G12A SF3B1 K666N	46,XY

Table 1. Co-occurring mutations and cytogenetic features of AML cases with mutant HNRNPK. Table lists the mutations observed in *HNRNPK*, the resultant amino acid change in hnRNP K, co-occurring mutations detected at diagnosis, and cytogenetic features. Mutations in genes encoding splicing factors are indicated in red.

HNRNPK RNA is elevated in a subset of AML patients

In the context of our previous work demonstrating that low expression of hnRNP K could result in myeloid malignancy, we next sought to evaluate whether expression of hnRNP K was altered in patients with AML without 9q21.32 deletions. We evaluated expression of *HNRNPK* RNA by qRT-PCR on bone marrow samples from patients with AML compared to CD34+ cells

from healthy donor bone marrow (Figure 3A). While the difference in median *HNRNP*K expression between AML patients and healthy donors was not statistically significant ($p=0.16$), a subset of patients (5/18; 27.8%) had demonstrably higher *HNRNP*K expression than healthy controls. When analyzed separately, this difference was statistically significant ($p=0.02$, Figure 3B). To confirm these data, we performed qRT-PCR on a separate cohort of AML bone marrow samples from patients with cytogenetically normal disease from a different institution (Centro Nacional de Investigaciones Oncológicas [CNIO]; Madrid, Spain; Figure 3C). Again, a subset of patients (3/15; 20%) had clearly elevated *HNRNP*K expression even though the difference in median *HNRNP*K expression between groups did not differ.

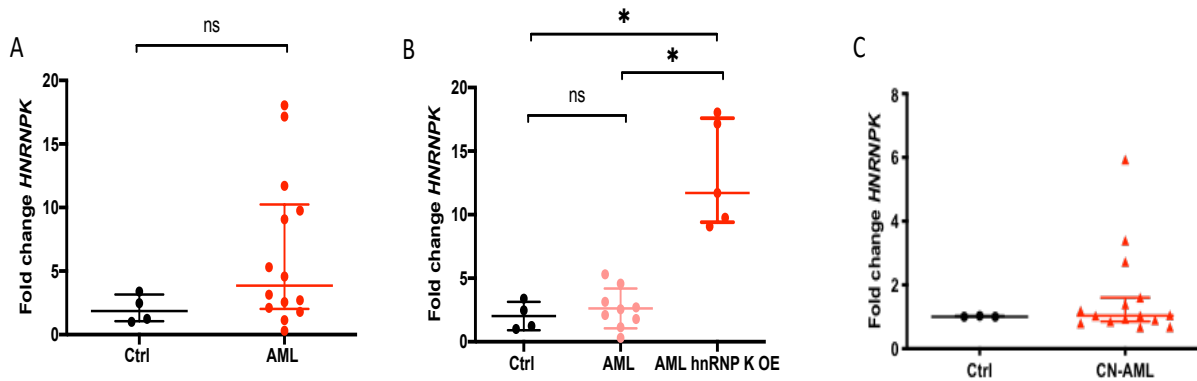


Figure 3. *HNRNP*K RNA is overexpressed in a subset of AML. qRT-PCR for *HNRNP*K expression in patients with AML from MD Anderson Cancer Center (A-B) or CNIO (C). Each point represents an individual patient sample. Lines are drawn at the median with tails extending to the interquartile range. Each dot indicates an individual patient sample.

hnRNP K protein is elevated in a subset of AML

To evaluate whether this increase in RNA expression corresponded to an increase in hnRNP K protein expression, we analyzed a reverse phase protein array (RPPA) from a different cohort of patients previously treated at our institution. This compared hnRNP K expression in bone marrow from patients with protein expression in CD34+ bone marrow cells from healthy

donors. Strikingly, median hnRNP K expression was significantly higher in AML samples compared with healthy donors ($p=0.0056$, Figure 4A). We confirmed this altered protein expression using hnRNP K immunohistochemistry (IHC) from samples used in the RPPA (Figure 4B). Indeed, cases that were deemed to have high hnRNP K expression by RPPA stained strongly for hnRNP K by IHC. Likewise, samples with low expression of hnRNP K by RPPA had only weak hnRNP K staining by IHC.

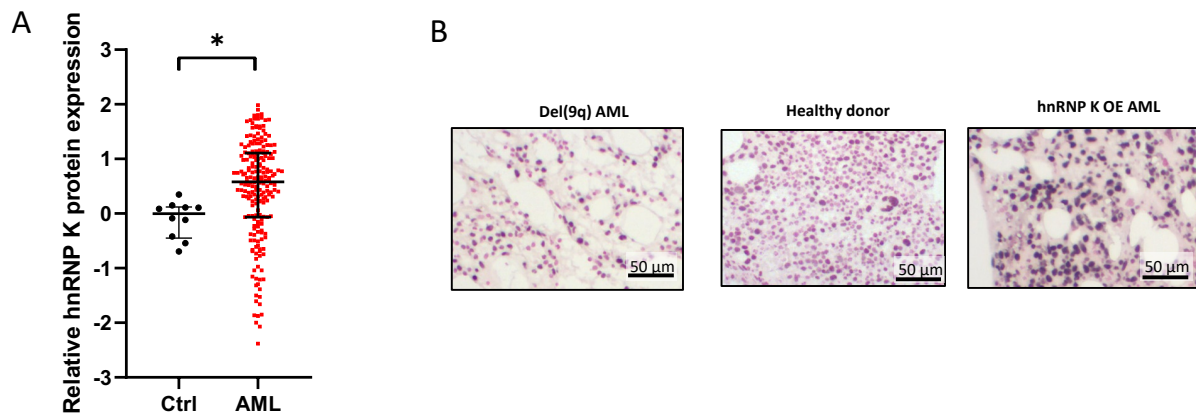


Figure 4. hnRNP K protein is overexpressed in a subset of AML. A. Normalized hnRNP K protein expression data from RPPA. Line is drawn at median expression with bars reaching within one interquartile region. Each dot represents a sample from an individual patient. B. hnRNP K IHC on representative samples from the RPPA dataset represented in panel A.

Elevated hnRNP K expression is associated with inferior clinical outcomes in AML

To determine whether increased hnRNP K expression was associated with altered clinical outcomes, we performed Kaplan-Meier analysis on patients stratified according to hnRNP K expression by RPPA. For these analyses, patients with hnRNP K expression greater than or equal to one standard deviation above the median of healthy control bone marrow were considered to have high hnRNP K expression. With this cutoff, 45 patients (22%) overexpressed hnRNP K, and 160 had normal or low hnRNP K expression. Patients with elevated hnRNP K

had statistically significant decreases in both remission duration (HR 2.1; 95% CI 1.2-3.6; Figure 5A) and overall survival (24.3 months versus 48.7 months; HR 1.9; 95% CI 1.3-2.7; Figure 5B).

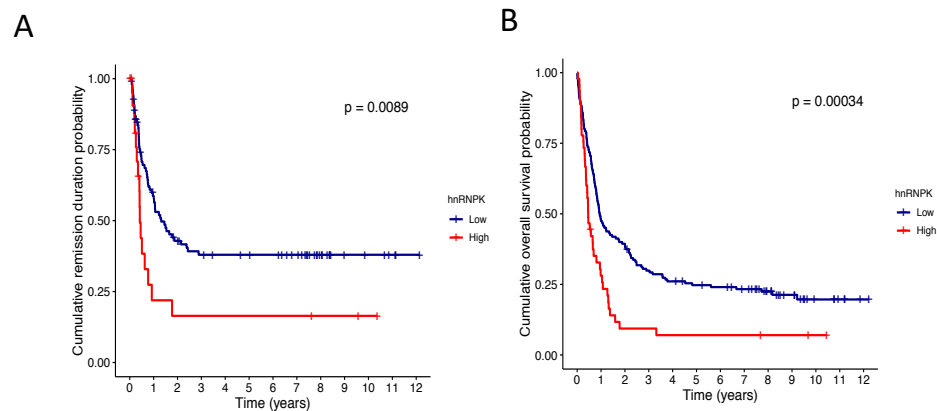


Figure 5. hnRNP K overexpression is associated with inferior outcomes. Kaplan-Meier analysis of remission duration (A) or overall survival (B) in patients with AML stratified according to hnRNP K protein expression. Red lines indicate patients with hnRNP K overexpression (greater than one standard deviation above the median of healthy controls). Blue lines indicate the remainder of the group with normal/low hnRNP K expression.

Elevated hnRNP K expression is enriched in immature AML subtypes

To evaluate whether hnRNP K overexpression was enriched in AML of particular morphology, we evaluated hnRNP K protein expression in cases according to French-American-British (FAB) classification. Median hnRNP K expression was higher in M0 and M1 cases compared with M2, M4, or M5 (Figure 6).

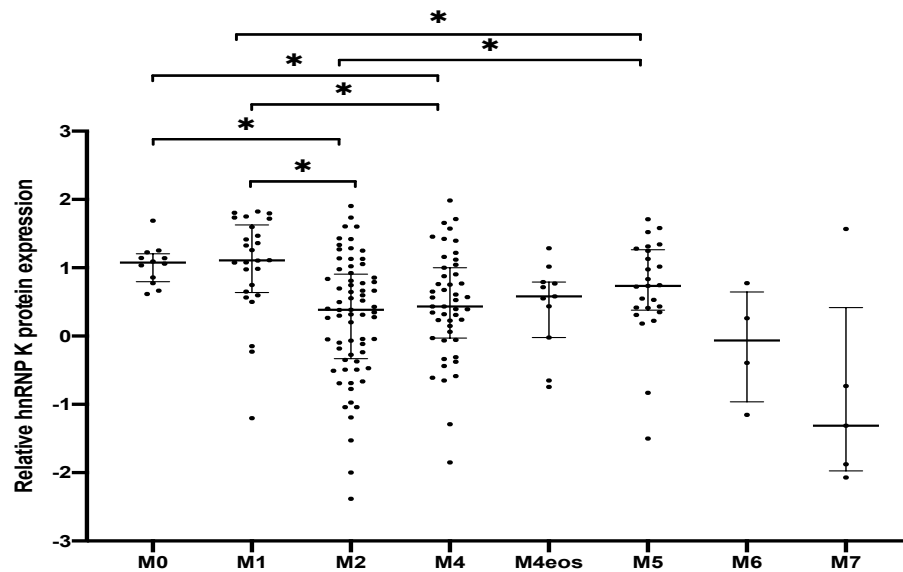


Figure 6. hnRNP K is elevated in more immature AML. Normalized hnRNP K protein expression data as assessed by RPPA and categorized according to FAB subtype. Lines are drawn at the median with tails reaching the interquartile range.

Elevated hnRNP K expression is correlated with mutations in NPM1 and FLT3

To evaluate the specific context in which hnRNP K is highly expressed in AML, we sought to determine whether high hnRNP K expression co-occurred with common mutations. Mutations in *FLT3* in the absence of a concomitant *NPM1* mutation did not exhibit an appreciable difference in hnRNP K expression (Figure 7). Interestingly, an *NPM1* mutation in the absence of *FLT3* mutation did have a statistically significant increase in hnRNP K expression. In line with this observation, cases with both *NPM1* and *FLT3* mutations also had a statistically significant increase in hnRNP K expression (Figure 7). Mutations in *TP53*, *NRAS*, *KRAS*, *RUNX1*, and *IDH1/2* did not correlate with changes in hnRNP K expression (data not shown).

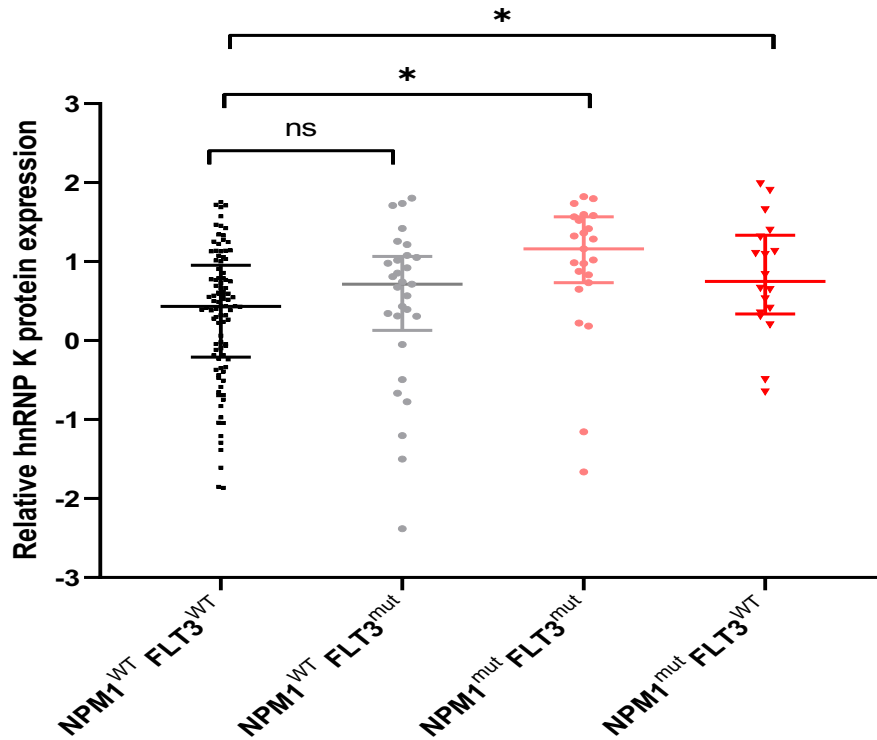


Figure 7. hnRNP K overexpression is enriched in NPM1 mutant cases. Normalized hnRNP K expression stratified according to mutational status of NPM1 and FLT3. Lines are drawn at the median and tails reach to within one interquartile range.

High hnRNP K expression in the context of NPM1 mutations is associated with inferior outcomes

To assess the impact of hnRNP K expression in the context of these mutations, we performed Kaplan-Meier analysis on patients based on hnRNP K status. Patients with any mutation in *NPM1* who also had greater than median hnRNP K expression had a statistically significant decrease in overall survival ($p=0.022$, Figure 8A). Since *NPM1* mutations in the context of wildtype *FLT3* have a relatively fair prognosis²⁷, we also specifically evaluated this subset of patients. Patients harboring *NPM1*^{Mut}*FLT3*^{WT} and higher than median hnRNP K expression had inferior overall survival, though this was not statistically significant, likely due to limited sample size (Figure 8B). Given these findings, evaluating hnRNP K expression in this specific subset of patients may offer prognostic guidance for this otherwise favorable subgroup.

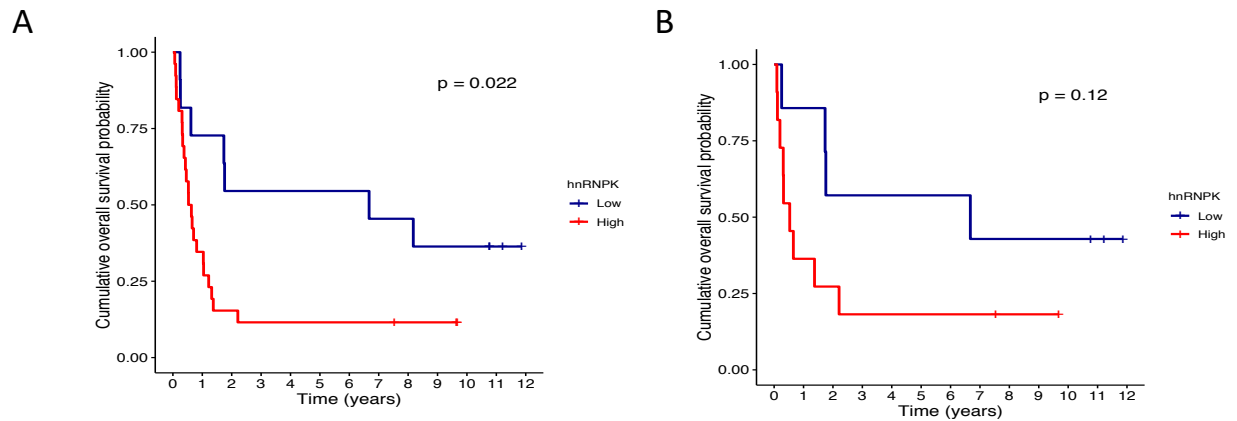


Figure 8. hnRNP K overexpression in the context of NPM1 mutations is associated with decreased overall survival. Kaplan-Meier analyses of overall survival of AML patients with mutant *NPM1* (A) or mutant *NPM1* and wildtype *FLT3* (B) stratified according to hnRNP K expression status. High expressors are shown in red.

The HNRNPK locus is amplified in a subset of AML and corresponds with increased hnRNP K protein.

To understand a possible mechanism of this hnRNP K overexpression, we designed a fluorescence in situ hybridization (FISH) probe that hybridizes to the *HNRNPK* locus. As expected, when we analyzed AML cases harboring chromosome 9q deletions, we only observed one copy of *HNRNPK* that was detectable by FISH (Figure 9A). Strikingly, we found numerous AML cases not harboring del(9q) had an extra copy of the *HNRNPK* locus (Figure 9A). Critically, when we performed IHC on samples with extra copies of the *HNRNPK* locus, the protein product, hnRNP K, was highly abundant (Figure 9B). Therefore, one mechanism of hnRNP K overexpression may be additional copies of the *HNRNPK* locus.

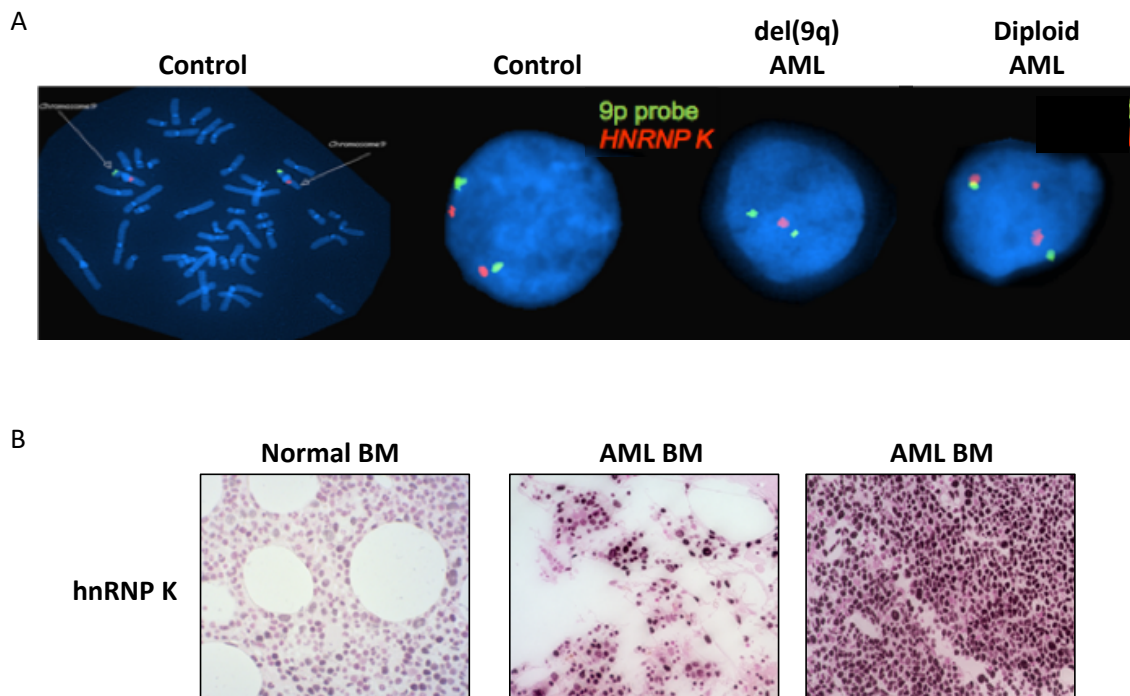


Figure 9. *HNRNPK* locus is duplicated in a subset of AML. A. (Left) Metaphase FISH using *HNRNPK* specific probe (red) and a control 9p probe (green). Other images represent nuclei of interphase cells from a healthy human donor, AML with del(9q), or diploid AML. Cells were probed with *HNRNPK* (red) or control 9p probe (green) and counterstained with DAPI (blue). B. hnRNP K IHC on healthy human bone marrow (left) or AML bone marrow with three copies of *HNRNPK* as identified by FISH.

Additional HNRNPK loci are present as small supernumerary marker chromosomes

Given the presence of additional copies of *HNRNPK* in several AML cases, and the propensity of translocations to occur in hematologic malignancies, we wanted to examine whether *HNRNPK* may partake in a translocation event. When we performed FISH on AML samples alongside Giemsa staining, we were unable to locate a consistent chromosome with which *HNRNPK* was associated. Instead, *HNRNPK* appeared in locations that did not stain with Giemsa, but only with DAPI—meeting the criteria of a small supernumerary marker chromosome

(sSMC; Figure 10).¹⁶⁶ This indicates that *HNRNP*K duplications do not occur in the context of translocations, but are present as sSMCs.

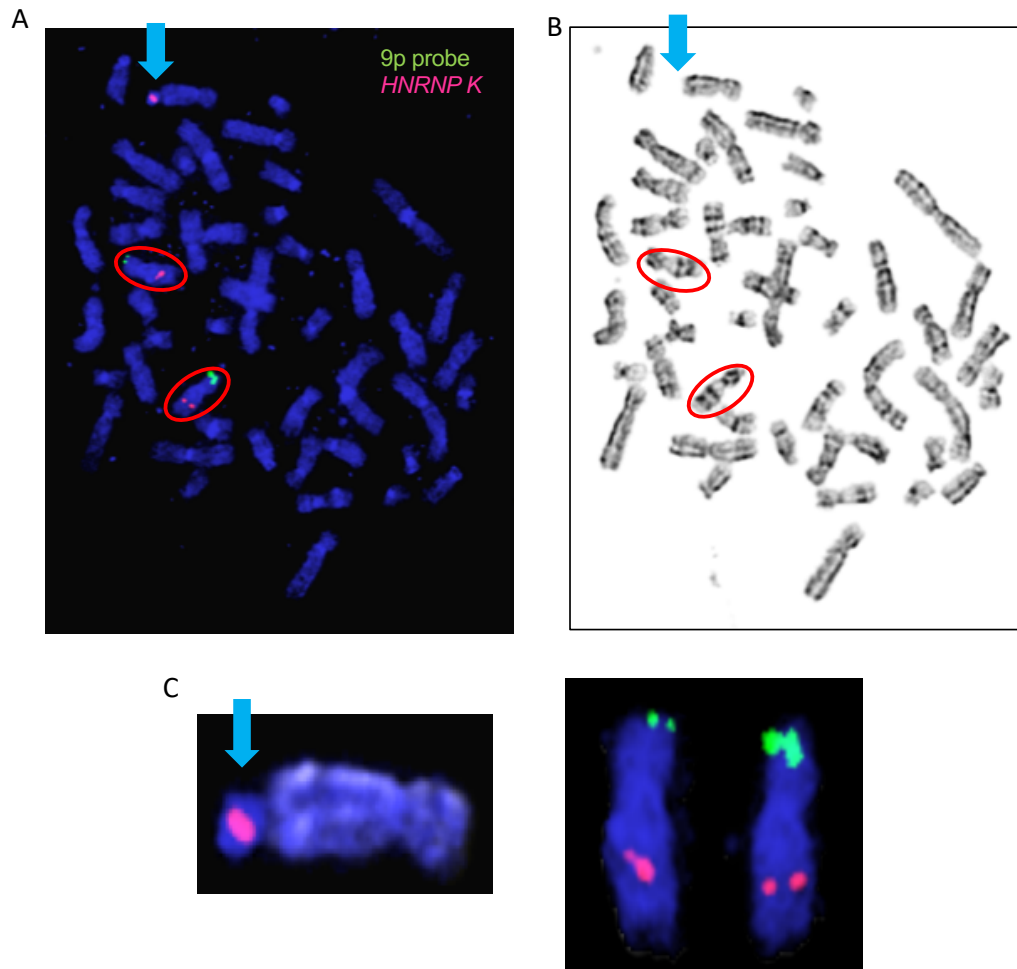


Figure 10. Additional *HNRNP*K loci exist as small supernumerary marker chromosomes. AML patient bone marrow metaphase subjected to (A) FISH with *HNRNP*K shown in red and a control 9p probe in green or (B) Giemsa staining. Red circles are drawn around chromosome 9. Blue arrow indicates presence of an additional *HNRNP*K locus. C. Magnified view of the additional *HNRNP*K locus (left) and the two chromosomes 9 (right).

2.4 Discussion

These studies elucidate that hnRNP K is overexpressed in a subset of patients with AML at both the RNA and protein level. Importantly, the overexpressed hnRNP K is unlikely to be mutated, since mutations in *HNRNPK* are rare events—occurring in less than 3% of patients, and the overexpression occurs in a much larger subset of patients (~20%) (Figures 3, 4). This corroborates previous reports of *HNRNPK* mutations, where incidences mingled near a mere 1%.³²⁻³⁴

The *HNRNPK* mutations we found in AML patients at MD Anderson were largely unique. A single AML patient has been described as having an A56G mutation, similar to the A56T mutation we identified.³⁴ Of interest, the Y333* mutation has been described in a patient with AML³³, but has also been observed in a patient with Au-Kline syndrome—a multiple malformation syndrome resulting from germline mutations or deletions of *HNRNPK*.^{38, 167} Perhaps then, it is not surprising that this mutation has been suggested to undergo nonsense-mediated decay, and therefore mimic *HNRNPK* haploinsufficiency.³⁸ However, the functional consequences of this mutation and the others we identified have not been formally characterized. Despite the rarity of these mutations, future work evaluating their functional consequences may be an impactful means to more thoroughly understand the biologic role of hnRNP K in multiple tissue types.

Despite having a very small cohort, we noticed that *HNRNPK* mutations in AML appeared to coexist with splicing factor mutations (*SRSF2*, *U2AF1*, or *SF3B1*). This is an intriguing observation since mutations in these genes are only estimated to occur in approximately 15% of AML.^{32-34, 168} As hnRNP K has itself been implicated in splicing^{56, 149, 154, 169} (see Chapters 4-5), future studies investigating these relationships would be of particular interest.

Data presented in this chapter indicate that *HNRNPK* transcript is elevated in a subset of AML patients (Figure 3). This observation was mirrored in a cohort of patients from CNIO in Spain (Figure 3C). While both patient cohorts for qRT-PCR analyses were relatively small, similar observations between these samples from completely separate institutions bolsters the

conclusion that *HNRNPK* is overexpressed in ~20% of AML. These findings were greatly substantiated by the observation that hnRNP K protein was also elevated in AML samples compared to healthy controls (Figure 4). While some may therefore conclude that the majority of AML cases “overexpress” hnRNP K (as one interpretation of the term “overexpression” is anything above “normal expression”¹⁷⁰), we chose to be more cautious in our definition of this admittedly subjective term. As the median of hnRNP K expression in AML was already greater than that of healthy controls, we chose to increase our threshold to one standard deviation above this measure. With this cutoff, we identified cases with substantially increased hnRNP K, a finding that was further validated by hnRNP K IHC (Figure 4B). Therefore, we used this as the definition of hnRNP K overexpression.

When AML cases were stratified according to hnRNP K overexpression, we observed statistically significant decreases in remission duration and overall survival (Figure 5). It is relevant to note that the RPPA data, and therefore these cutoffs, were generated from a cohort of patients treated prior to 2015.¹⁶² Almost all of these patients were treated with similar chemotherapy-based regimens, devoid of hypomethylating agents, targeted small molecules, immunotherapies (including antibodies), or cellular therapies. It is thus possible that differences in these clinical outcomes associated with hnRNP K expression are more reflective of the underlying disease biology, not merely an artifact of response to more sophisticated therapeutic agents. On the other hand, with the widespread use of these agents, hnRNP K overexpression as a single stratification criterion may no longer be sufficient to separate outcomes. Prospective studies would be relevant to further understand the impact of hnRNP K overexpression on patient outcomes in the context of the most up to date therapeutic options.

We found that AML cases with more immature phenotypes by FAB subclass (M0 or M1) had increased hnRNP K expression compared to more mature cases (Figure 6). This is only partly consistent with microarray data in healthy human hematopoiesis, where *HNRNPK* expression is relatively high in hematopoietic stem cell subsets and decreases slightly in cells

that are fully committed to monocytic or megakaryocytic differentiation.¹⁷¹ In our data, cases of M5 (monocytic) and M7 (megakaryocytic) AML had lower hnRNP K levels than more undifferentiated cases (M0 and M1). Unlike the observations in healthy hematopoiesis, where *HNRNPK* expression did not substantially decrease early in differentiation¹⁷¹, we observed that M2 and M4(eos) cases had a relatively steep drop off in hnRNP K expression.

To more specifically identify subpopulations of patients with elevated hnRNP K expression, we evaluated hnRNP K expression in cases with mutated *NPM1* and/or *FLT3*. hnRNP K expression did appear to be increased in the setting of *NPM1* and *FLT3* mutations (Figure 7). Patients harboring mutated *NPM1* with concomitant hnRNP K overexpression had decreased overall survival compared to those with lower hnRNP K expression. This is particularly important in the subgroup of patients with mutated *NPM1* in the context of wildtype *FLT3*, as such patients are normally considered to have a favorable prognosis.²⁷ Strikingly, these patients had dramatically worse overall survival when hnRNP K was overexpressed. Future studies evaluating the interaction between mutant *NPM1* and hnRNP K overexpression are certainly warranted.

The mechanism of the hnRNP K overexpression described in this chapter remains unclear; however, we did identify the presence of the *HNRNPK* locus as a small supernumerary marker chromosome (sSMC) that corresponded to high hnRNP K protein expression by IHC. When so-called marker chromosomes occur in AML, these can predict adverse prognosis¹⁷²; however, these are identified in routine cytogenetic testing. The sSMCs we identified are only detected by FISH, and would be missed in routine cytogenetic testing. Therefore, the FISH probe we developed could have clinical potential as it would identify these sSMCs and also readily identify patients with AML harboring del(9q) encompassing this locus. Prospective studies with this probe may be of clinical utility. In addition, future investigations into the genomic contents of these sSMCs would be of great interest, as it may provide insight into the endogenous regulation of *HNRNPK* expression in disease states.

Small clinical studies have correlated high hnRNP K expression with inferior clinical course in several solid tumor types, including melanoma, breast, nasopharyngeal, pancreatic, hepatocellular, colorectal, bladder, and prostate cancers.⁷⁵⁻⁸² Importantly, we have recently identified that overexpression of hnRNP K can have oncogenic activity in the setting of B-cell lymphoma.⁸³ Thus, it is reasonable to hypothesize that high expression of hnRNP K may be pathogenic in the setting of myeloid malignancy. Chapter 3 will address this hypothesis in detail.

Chapter 3

hnRNP K overexpression drives myeloproliferative disease in murine models

3.1 Introduction

In Chapter 2, we determined that hnRNP K was overexpressed in a subset of patients with AML, and this overexpression was associated with inferior clinical outcomes. These observations alluded to the notion that hnRNP K may act as an oncogene when overexpressed in AML. While limited clinical studies in solid tumors have seen hnRNP K overexpression⁷⁵⁻⁸², data regarding a role for hnRNP K in hematologic malignancies is sparse.

AML has been described as a disease arising from aberrations in hematopoietic stem and progenitor cells.¹⁷³⁻¹⁷⁵ Thus, an *in vivo* model used to evaluate an oncogenic role for hnRNP K should harbor hnRNP K overexpression in these immature cells. We approached generating a mouse model of such using retroviral transduction of *Hnrnpk* into murine hematopoietic stem and progenitor cells followed by transplantation into recipient mice.

Throughout a portion of mammalian development *in utero*, the fetal liver is the primary site of hematopoiesis. In mice, beginning at embryonic day 11.5 (E11.5), hematopoietic stem cells (HSCs) begin to seed the fetal liver, and by E12.5, the fetal liver becomes the predominant site of HSC expansion.^{176, 177} Here, HSC number increases between E13.5-16.5, then declines steadily until HSCs shift to bone marrow at E17.5.¹⁷⁸ Throughout this time, progenitor cells are also abundant in the fetal liver.¹⁷⁹ Murine fetal liver cells (FLCs) from E13.5-16.5 therefore serve as a rich source of hematopoietic stem and progenitor cells. Furthermore, FLCs can be virally transduced to stably alter gene expression.^{180, 181} Transplantation of these genetically altered FLCs into recipient mice is a rapid and efficient way to restrict expression of the gene of interest to the hematopoietic system.^{180, 181}

3.2 Materials and Methods

Plasmids: Retroviruses were made in an MSCV backbone and were modified from those described previously.¹⁸⁰ MSCV-AML1/ETO-IRES-GFP plasmid was obtained from Addgene (plasmid #60832).¹⁸⁰ The green fluorescent protein (GFP) coding sequence was replaced with firefly luciferase ORF obtained from a luciferase-pcDNA plasmid (Addgene plasmid #18964).¹⁸² For generation of empty vectors, AML1/ETO was excised using EcoRI and BamH1 and plasmid re-ligated. To generate MSCV-HNRNPK-IRES-GFP and MSCV-HNRNPK-IRES-Luciferase, AML1/ETO was replaced with HNRNPK that was PCR amplified from 293T cells.

Retroviral production: HEK293T cells at 50% confluency in T75 flasks were transfected with 4 µg of transfer vector (MSCV-(HNRNPK)-IRES-GFP/Luciferase) and 4 µg of pCL-Eco packaging vector (Addgene plasmid #12371)¹⁸³ with 12 µL of JetPrime DNA transfection reagent (Polyplus Transfection, New York, NY, USA). Fresh stem cell medium without cytokines (see below) was supplied after 4-8 hours and cells were incubated at 32°C for 48-72 hours. High-titer retroviral supernatant was collected and passed through a 0.45 µm filter.

Fetal liver cell (FLC) isolation: The University of Texas MD Anderson Cancer Center Animal Care and Use Committee approved all mouse experiments performed in these studies under protocols 0000787-RN01 and 0000787-RN02. Pregnant wildtype CD45.2+ C57/Bl6 females were euthanized by exposure to CO₂ at day 13.5-16.5 of gestation. Fetal livers were sterilely dissected and gently disrupted on a 70 µm filter to obtain a single-cell suspension. The collected cells were briefly subjected to a red blood cell (RBC) lysis with BD Pharm Lyse buffer (BD Biosciences, San Jose, CA, USA). After RBC lysis, cells were resuspended in stem cell medium at ~3x10⁶ cells/mL and incubated at 37°C overnight prior to retroviral transduction. Stem cell medium contained 37% DMEM (Corning, Corning, NY, USA), 37% Iscove's modified Dulbecco's Medium (Corning, Corning, NY, USA), 20% fetal bovine serum, 2% L-glutamine (200mM; Corning, Corning, NY,

USA), 100 U/mL penicillin/streptomycin (Sigma Aldrich, St. Louis, MO, USA), 5×10^{-5} M 2-mercaptoethanol (Sigma Aldrich, St. Louis, MO, USA), recombinant murine interleukin-3 (IL-3; 0.2 ng/mL), interleukin-6 (IL-6; 2 ng/mL), and stem cell factor (SCF; 20 ng/mL; all cytokines from Stem Cell Technologies, Vancouver, Canada).

FLC transduction and sorting: Approximately 5×10^6 FLCs were resuspended in 2 mL of high-titer retroviral supernatant supplemented with 12 μ g/mL polybrene and cytokines (IL-3; 0.2 ng/mL, IL-6 2 ng/mL, and SCF 20 ng/mL). Plates were then spun at 600xg for 90 minutes at room temperature and incubated at 32°C for 48-72 hours. After transduction, cells infected with GFP-containing constructs were sorted for GFP positivity using the MoFlo Astrios cell sorter (Beckman Coulter, Brea, CA, USA) at the MD Anderson Cancer Center North Campus Flow Cytometry Core Facility.

qRT-PCR: Bone marrow and spleen samples were subjected to adequate red blood cell lysis with BD Pharm Lyse lysing solution (BD Biosciences, San Jose, CA, USA). For all samples, RNA was extracted and purified using phenol/chloroform.¹⁵⁷ Samples were treated with DNase for 30 minutes at 37°C, and quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). 1 μ g of RNA was reverse transcribed using iScript (BioRad, Hercules, CA, USA). qRT-PCR was performed using iTaq Universal SYBR Green Supermix as per manufacturer's instructions (BioRad, Hercules, CA, USA) using an ABI StepOnePlus Real Time PCR System. Changes in expression were determined by comparing expression to the housekeeping control *Rplp0*. Expression of *Hnmpk* was evaluated using primers as follows: Forward (sp240): 5'-GAAGATATGGAAGAGGAGCAAGCC-3', Reverse (sp242): 5'-CAAGGTAGGGATGATTTTCTTC-3' (sp240-242), *Rplp0* Forward: 5'-CCCTGAAGTGCTCGACATCA-3', *Rplp0* Reverse: 5'-TGCGGACACCCTCCAGAA-3'.

Individual samples were assayed in triplicate, and calculations were performed using the Pfaffl method comparing expression changes between target genes and housekeeping control.¹⁵⁸

Western blotting: Cells were homogenized in NP40 lysis buffer containing protease and phosphatase inhibitors (Millipore Sigma, Burlington, MA, USA). Soluble proteins were boiled in Laemmli buffer, resolved on a 10% SDS-PAGE gel, and transferred to a PVDF membrane. Membranes were blocked with 5% milk for one hour at room temperature and incubated with primary antibody at 4°C overnight while rocking. Primary antibodies were hnRNP K (3C2, 1:1000, Abcam, Cambridge, MA, USA) and β -actin (AC-15, 1:2000, Santa Cruz, Biotechnology, Dallas, TX, USA). Membranes were incubated with secondary antibody and antibody-protein interactions were visualized using enhanced chemiluminescence (GE Healthcare, Chicago, IL, USA) or BCIP/NBT color development substrate (VWR International, Radnor, PA, USA).

Colony formation assay: GFP sorted FLCs were cultured in quadruplicate wells of a 12-well plate in methylcellulose medium with cytokines IL-3, IL-6, erythropoietin, and stem cell factor (Methocult GF M3434, StemCell Technologies, Vancouver, Canada). 50,000 cells were plated per well and colonies were counted after 7 days. Colonies were gently disrupted in PBS and cells counted manually with a hemacytometer using trypan blue dye exclusion, or subjected to cytopsin or flow cytometry. Cytopsin were stained with Wright-Giemsa.

Flow cytometry: Cells were harvested and washed with PBS supplemented with 5% bovine serum albumin (BSA; Sigma Aldrich, St. Louis, MO, USA). Prior to antibody staining, cells were treated with murine Fc block (TruStain FcX, BioLegend, San Diego, CA, USA) at room temperature for 15 minutes. Cells were incubated with fluorescently labeled antibodies in the dark at room temperature for 30 minutes-1 hour. Antibodies used: Gr1 [RB6-8C5], CD11b [M1/70;], CD117 [2B8], CD45 [30F11] (all from BD Biosciences, East Rutherford, NJ, USA), and

Sca-1 [D7; eBioscience, San Diego, CA, USA]. Flow cytometry was performed in the MDACC North Campus Flow Cytometry Core Facility using a Gallios flow cytometer (Beckman Coulter, Brea, CA, USA). Data was analyzed using FlowJo Software (Beckton Dickinson, Franklin Lakes, NJ, USA).

FLC transplantation: NOD-scid-IL2R-gamma (NSG) mice were purchased from the MDACC Department of Experimental Radiation Oncology. NSG mice were sub-lethally irradiated with 2.5 Gy administered as a single dose 4-24 hours prior to transplantation. At least 40,000 (up to 1×10^6) cells suspended in sterile PBS were injected into the retroorbital sinus of mice under isoflurane-induced anesthesia. Engraftment was evaluated using *in vivo* bioluminescent imaging (see below) or flow cytometry analysis of peripheral blood. In the latter case, blood was collected from the retro-orbital sinus of isoflurane-anesthetized mice on a monthly basis.

In vivo bioluminescent imaging: Mice were injected intraperitoneally with D-Luciferin, anesthetized with isoflurane, and imaged using an IVIS system (Xenogen/Caliper Life Sciences, Alameda, CA, USA). Imaging was performed on a weekly basis beginning one week after transplantation to monitor and ensure engraftment.

Peripheral blood analysis: Blood was collected from the retro-orbital sinus of isoflurane-anesthetized mice into EDTA-coated tubes. Complete blood count analysis was performed with an ABX Pentra analyzer (Horiba, Kyoto, Japan). Peripheral blood smears were stained with Wright-Giemsa.

Tissue procurement: At time of necropsy, spleen, liver, and sternum were collected from each mouse. Tissues were fixed in 10% neutral-buffered formalin and processed by the Research Histopathology Facility at MD Anderson Cancer Center where paraffin-embedded blocks were

made and sectioned. Each section was stained with standard hematoxylin/eosin staining. Cells were also collected from femurs and remaining spleen in a single cell suspension, subjected to a gentle RBC lysis, and processed for western blotting or qRT-PCR as described above.

Immunohistochemistry: Formalin-fixed paraffin-embedded tissues were deparaffinized in xylene and rehydrated in an alcohol gradient. Antigen retrieval was performed using 10mM sodium citrate and 0.05% Tween 20 (pH 6.0) in a steam chamber for 45 minutes. Slides were incubated in 3% hydrogen peroxide/methanol solution to deactivate endogenous peroxidases. Incubation with primary antibody was at 4°C overnight in a humidity chamber. Slides were washed with 0.1% Tween-PBS before biotinylated secondary antibody was added at room temperature for 30 minutes. Antibody-protein interactions were visualized with Vectastain Elite ABC and DAB peroxidase substrate kits (Vector Laboratories, Burlingame, CA, USA). For all antibodies, DAB time was 3 minutes. Nuclear fast red was used as a counterstain. Antibodies used for IHC are included in Table 2.

Antibody	Clone	Secondary antibody	Vendor	Dilution
hnRNP K	3C2	mouse	Abcam	1:3000
CD34	MEC14.7	rat	ThermoFisher	1:100
CD117	2B8	rat	ThermoFisher	1:100
CD3	SP162	rabbit	Abcam	1:150
MPO	polyclonal; ab9535	rabbit	Abcam	1:100
CD14	4B4F12	mouse	Abcam	1:100

Table 2. Antibodies used for IHC. Antibodies used for immunohistochemistry. With the exception of MPO, antibodies were monoclonal; therefore, the clone is listed in lieu of a catalog number.

Survival analysis: Kaplan-Meier curves were generated using Prism software (GraphPad Software, San Diego, CA). Time between transplantation of FLCs and death or veterinarian-mandated euthanasia was used for analyses.

3.3 Results

hnRNP K can be overexpressed in murine FLCs

To evaluate whether the FLC system was suitable for modeling hnRNP K-overexpression, we first infected FLCs with an hnRNP K-overexpression construct. Compared to empty vector controls, these cells exhibited clear hnRNP K overexpression at both the RNA and protein levels (Figure 11). This indicated that FLCs were capable of overexpressing hnRNP K, making these a suitable model system.

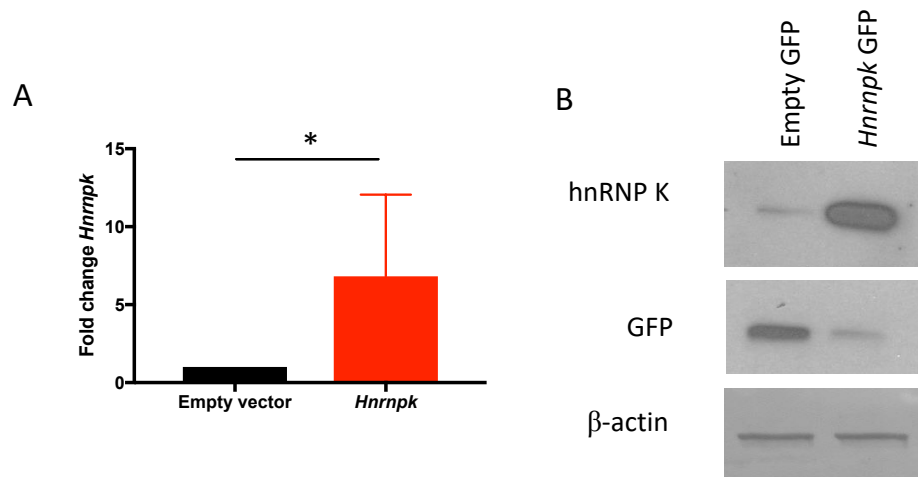


Figure 11. FLCs can overexpress hnRNP K. A. Bar graph of *Hnrnpk* qRT-PCR results, normalized to *Rplp0*, in FLCs infected with an hnRNP K-overexpression vector (red) or empty vector control (black). B. Representative western blot of FLCs infected with either hnRNP K-overexpression vector or empty GFP control. Cells were infected 48 hours prior to harvest.

hnRNP K overexpression results in altered colony forming potential

To evaluate whether overexpression of hnRNP K had a biologic effect on FLCs, we performed colony formation assays. These assays can be used to assess self-renewal capacity and differentiation potential of cells.¹⁸⁴ When placed in colony formation assays, FLCs

overexpressing hnRNP K formed more colonies than empty vector controls (Figure 12A). Of note, colonies from hnRNP K-overexpressing FLCs were also visibly larger than empty vector controls (Figure 12B), suggesting that hnRNP K-overexpression provides a proliferative advantage to these cells.

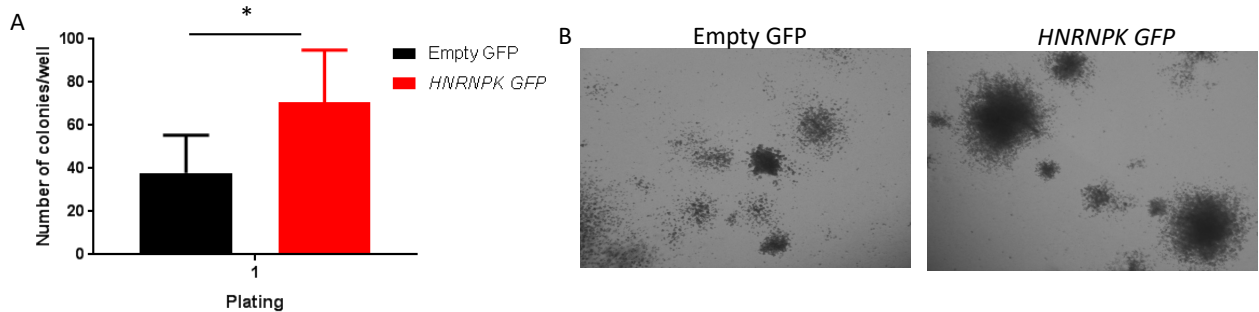


Figure 12. hnRNP K overexpression increases colony formation potential of FLCs. A. Bar graph representing number of colonies formed per well in FLCs overexpressing hnRNP K (red) or empty vector controls (black). B. Representative brightfield image of colonies from panel A.

hnRNP K overexpression alters colony phenotype

Given the increase in colony formation in hnRNP K-overexpressing cells, we sought to examine the properties of the colonies that formed. Morphologic differences were observable in Wright-Giemsa stained cytopins from colonies arising from hnRNP K-overexpressing cells compared to empty vector controls (Figure 13A). In particular, cells from empty vector controls appeared relatively uniform, while colonies overexpressing hnRNP K were more varied in cell size and morphologic features. To evaluate these differences in a more definitive manner, we examined cell surface marker expression by flow cytometry. These experiments revealed that hnRNP K-overexpressing colonies had fewer mature myeloid cells (CD11b+Gr1+) compared to controls (Figure 13B). Consistent with this lack of mature myeloid cells, hnRNP K-overexpressing colonies were enriched in immature c-kit+Sca1+ cells (Figure 13C). As hematopoietic stem cells differentiate, c-kit and Sca1 expression is downregulated.¹⁸⁵⁻¹⁸⁸ Thus,

these data suggest that hnRNP K-overexpression results in maintenance of a stem-like phenotype, leading to decreased percentage of mature myeloid cells, and an increase in colony formation and self-renewal capacity. This could also represent a differentiation block, which is a phenomenon often observed when leukemic cells are placed into colony formation assays.¹⁸⁹

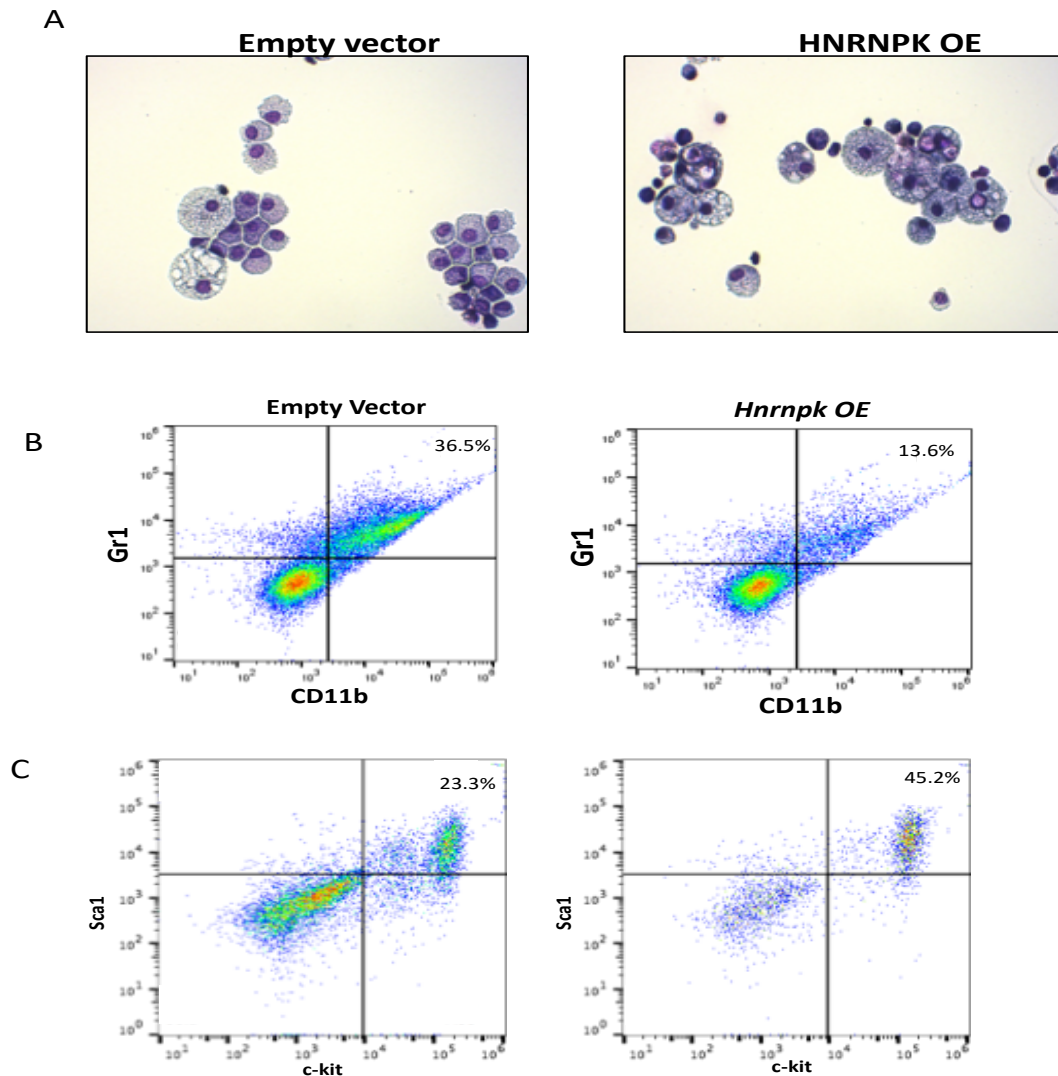


Figure 13. hnRNP K-overexpressing colonies have altered immunophenotypes. A. Wright-Giemsa stain of cytopins from representative colonies. B. Dot plot of Gr1 and CD11b expression from representative colonies. Inset numbers indicate the percentage of cells that were Gr1+ and CD11b+. C. Dot plot of Sca1 and c-kit expression from representative colonies.

Inset numbers indicate the percentage of cells that were Sca1+ckit+. Flow cytometry data is gated on CD45+ cells.

hnRNP K-overexpressing FLCs engraft in recipient mice

Since colony formation assays indicated that hnRNP K-overexpression altered the differentiation and self-renewal properties of FLCs, we next asked whether this would translate to a similar phenotype *in vivo*. Thus, we injected luciferase-containing FLCs intravenously into sub-lethally irradiated NSG mice. Serial *in vivo* imaging revealed luciferase activity in hematopoietic tissues (bone marrow and spleen) of transplanted mice (Figure 14A), indicating successful engraftment. Interestingly, recipients of hnRNP K-overexpressing FLCs frequently had faster engraftment—a finding consistent with the increase in colony formation seen in Figure 13.

Since luciferase-containing FLCs could be successfully transplanted into NSG mice regardless of hnRNP K expression, we next transplanted FLCs transduced with GFP-containing constructs. The presence of GFP allowed us to sort and inject only cells that had successfully been infected with the viral constructs, thus ensuring a more homogeneous population. Indeed, mice injected with GFP-containing FLCs had measurable GFP positivity in peripheral blood that was stable over time (Figure 14B).

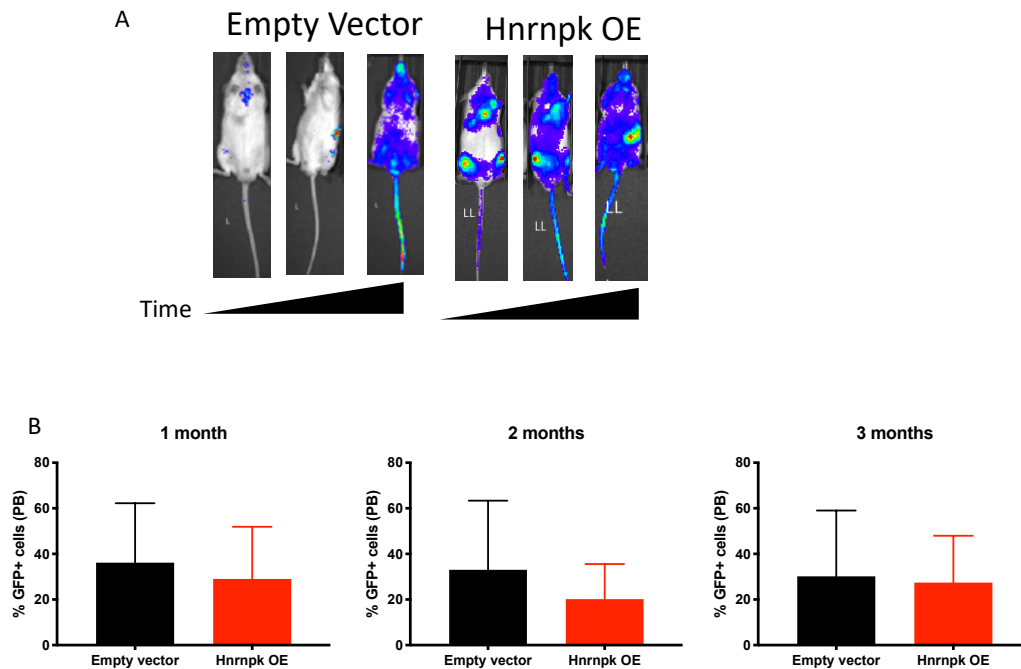


Figure 14. hnRNP K-overexpressing FLCs engraft in recipient NSG mice. A. *In vivo* bioluminescent imaging of representative mice injected with either empty-Luciferase FLCs (left) or hnRNP K-overexpressing-Luciferase FLCs (right). Images are from one representative mouse from each group at matching timepoints over the course of 6 weeks. B. Bar graphs of GFP positive cells in peripheral blood from mice injected with either empty-GFP FLCs (black) or hnRNP K-overexpressing-GFP FLCs (red).

hnRNP K overexpression is maintained in vivo

To ensure that the transplanted FLCs maintained hnRNP K-overexpression as intended, we assessed cells from spleen and bone marrow of recipient mice. qRT-PCR revealed consistently increased *Hnrnpk* expression in mice receiving hnRNP K-overexpressing FLCs compared to empty vector controls (Figure 15A). Increased hnRNP K expression was also seen at the protein level in spleen (Figure 15B). These findings support the notion that hnRNP K-overexpressing FLCs can engraft in recipient mice and be maintained in an *in vivo* setting.

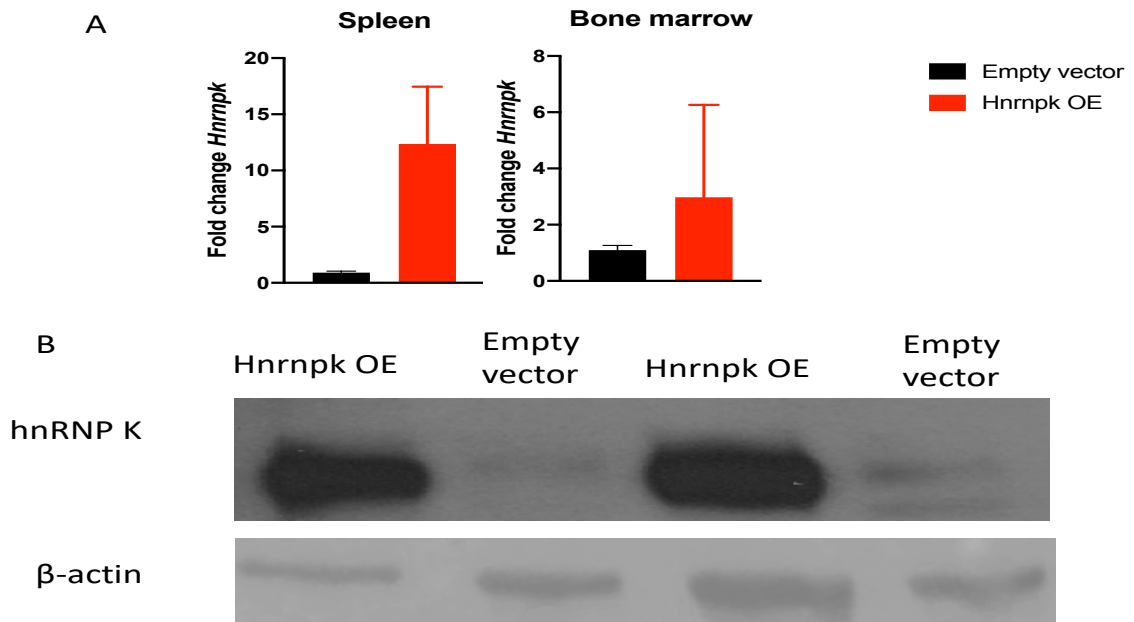


Figure 15. hnRNP K overexpression is maintained *in vivo*. A. *Hnrnpk* qRT-PCR, normalized to *Rplp0*, from spleen and bone marrow of mice transplanted with empty vector FLCs or hnRNP K-overexpressing FLCs (n=4). B. Western blot from spleens of representative mice. Each lane is an individual animal. All samples were collected 4 weeks after transplant.

Recipients of hnRNP K-overexpressing FLCs have a survival disadvantage

Our data have shown that hnRNP K-overexpressing FLCs engraft in recipient mice and maintain elevated hnRNP K expression (Figures 14, 15). Next, we asked whether this affected the outcomes of the mice. Strikingly, mice that received hnRNP K-overexpressing FLCs had substantially shortened overall survival compared to recipients of empty vector controls (Figure 16). Median survival for recipients of hnRNP K-overexpressing FLCs was 7.2 weeks, while median survival for recipients of empty vector controls was not reached ($p=0.02$). Notably, mice were flagged for euthanasia due to hyperpnea, emaciation, and failure to thrive.

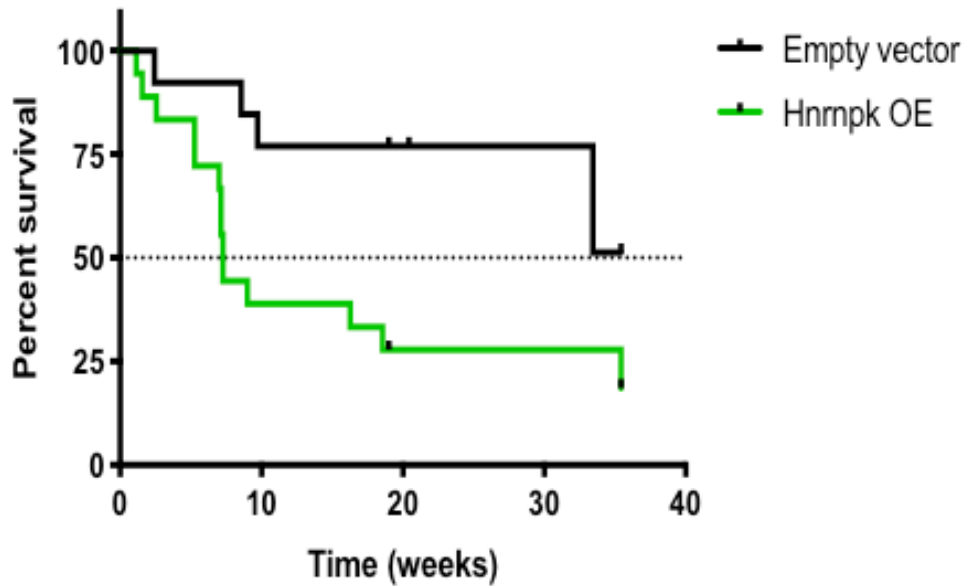


Figure 16. Survival of FLC recipients. Kaplan-Meier curve of overall survival of mice transplanted with FLCs containing an empty vector control (black; n=13) or hnRNP K-overexpression construct (green; n=20).

Bone marrow is abnormal in mice receiving hnRNP K overexpressing FLCs

We next sought to understand the basis of the shortened survival observed in mice that received hnRNP K-overexpressing FLCs. Since the transplanted FLCs were intended to reconstitute the bone marrow of recipient mice, we first examined this compartment. Mice that received FLCs overexpressing hnRNP K had more cellular bone marrow compared to empty vector controls (Figure 17). These cells were substantially skewed toward those with morphologic features consistent with myeloid differentiation. In addition, mild eosinophilia was observed.

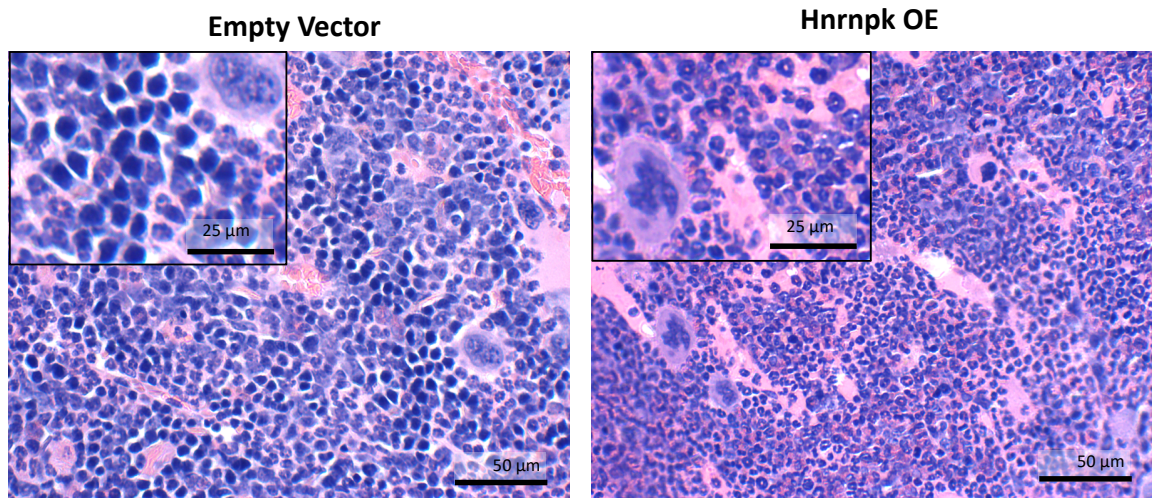


Figure 17. Recipients of hnRNP K-overexpressing FLCs have bone marrow abnormalities. H&E staining of sternal bone marrow from representative mice transplanted with empty vector control FLCs (left) or hnRNP K-overexpressing FLCs (right).

Recipients of hnRNP K overexpressing FLCs exhibit splenomegaly

To evaluate whether these bone marrow abnormalities affected other organs, we next examined the spleen, since splenic abnormalities can reflect hematopoietic pathology.¹⁹⁰ We observed that recipients of empty vector control FLCs had spleens that were approximately the size and weight of adult wildtype C57Bl6 animals¹⁹¹ (Figure 18)—consistent with efficient engraftment of FLCs derived from C57Bl6 mice (Figure 14). However, mice that received hnRNP K-overexpressing FLCs exhibited modest splenomegaly (Figure 18), which can be indicative of hematopoietic disease.¹⁹⁰

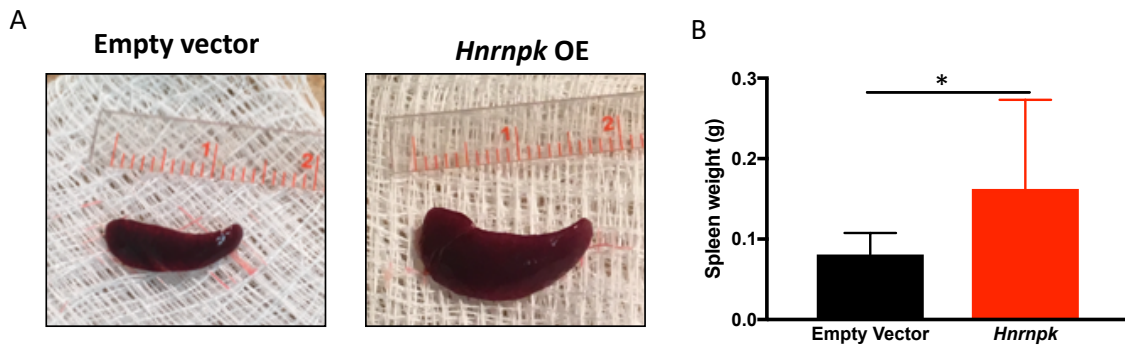


Figure 18. Recipients of hnRNP K-overexpressing FLCs develop splenomegaly. A. Gross images of spleens from representative mice transplanted with either empty vector FLCs (left) or hnRNP K-overexpressing FLCs (right). B. Quantitation of spleen weights.

Recipients of hnRNP K-overexpressing FLCs have disrupted splenic architecture

We next sought to understand the basis of the splenomegaly in mice harboring hnRNP K-overexpression. While recipients of control FLCs had splenic structure mimicking wildtype C57Bl6 mice, those that received hnRNP K-overexpressing FLCs had markedly disrupted splenic architecture (Figure 19). Critically, hnRNP K expression was still abundant in these spleens (Figure 19A), supporting the notion that this aberrant hematopoiesis was driven by hnRNP K-overexpression. Spleens from hnRNP K-overexpressing mice had an increase in the number of cells expressing the immature hematopoietic stem cell markers CD34 and CD117 (Figures 19B, C). The presence of these cell types indicated that extramedullary hematopoiesis, a pathologic process, was likely to be a factor underlying the observed splenomegaly.¹⁹⁰

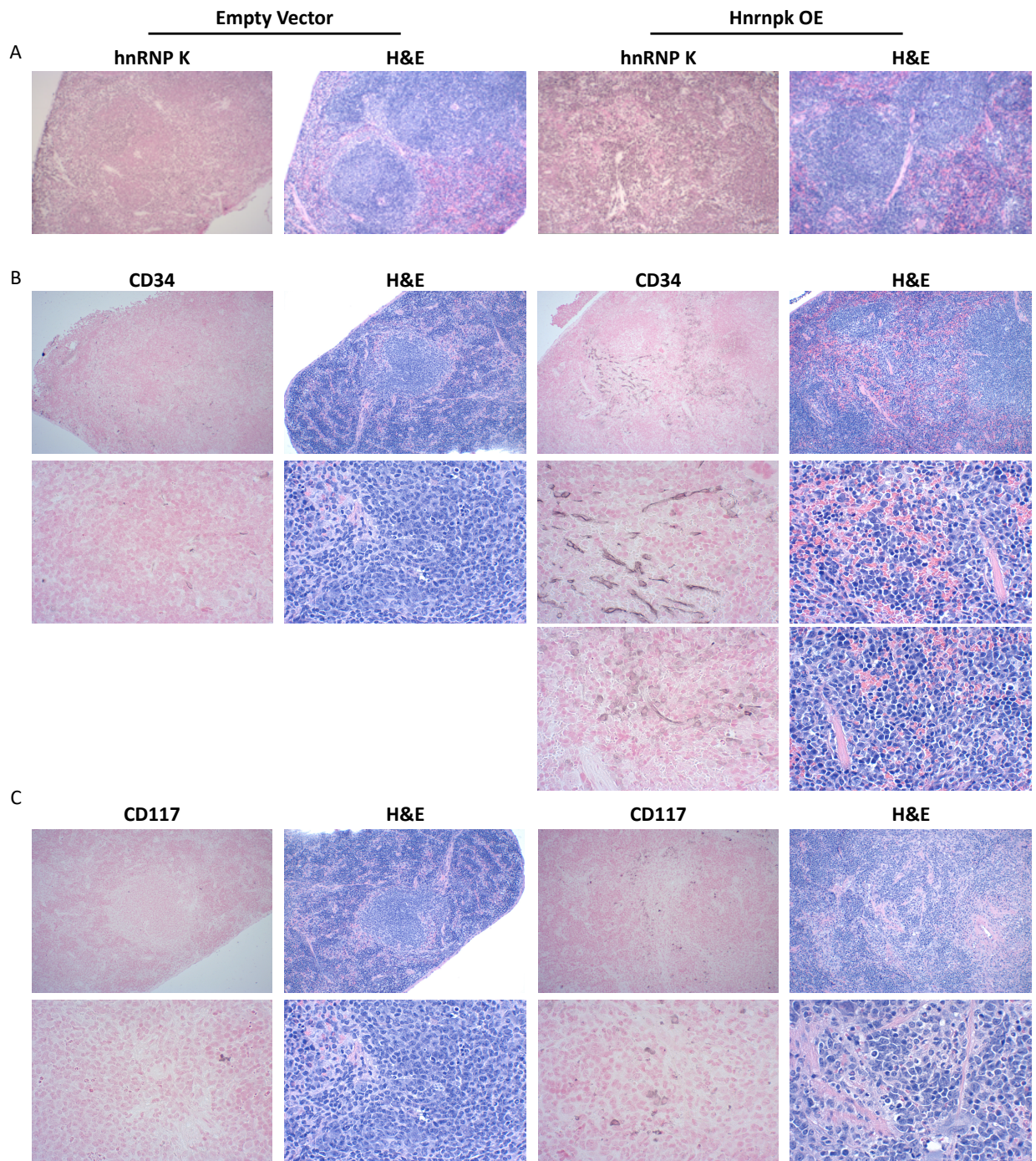


Figure 19. hnRNP K overexpression results in splenic extramedullary hematopoiesis. A. IHC staining of hnRNP K (A), CD34 (B), or CD117 (C) and corresponding H&E of spleen from representative mice transplanted with empty vector FLCs (left) or hnRNP K overexpressing FLCs (right). Additional images in panel B from the *Hnrnpk* OE mouse are included.

Recipients of hnRNP K overexpressing FLCs have hepatic leukocyte infiltrates

Given the splenomegaly and altered splenic architecture in mice receiving hnRNP K-overexpressing FLCs, we next evaluated the liver, as this can be another site of extramedullary hematopoiesis.¹⁹⁰ We observed a striking infiltration of cells into the hepatic parenchyma of mice receiving hnRNP K-overexpressing FLCs, but these cells were not present in mice transplanted with control FLCs (Figure 20A). These cells stained largely negative for CD3, precluding the notion that these might be a reactive infiltration of T cells (Figure 20B). A small subset of cells stained positive for CD117 (Figure 20C), indicative of extramedullary hematopoiesis. Interestingly, a large portion of these cells were CD14 positive, suggesting a monocytic lineage (Figure 20D). Likewise, a subset of cells were also strongly positive for MPO (Figure 20E), often utilized as a defining characteristic of myeloid lineage cells.¹⁹² Due to the absence of infiltrating leukocytes, livers from recipients of control FLCs were completely negative for these markers. Taken together, these findings demonstrate that hnRNP K-overexpression promotes hepatic extramedullary hematopoiesis with a myeloid-lineage bias.

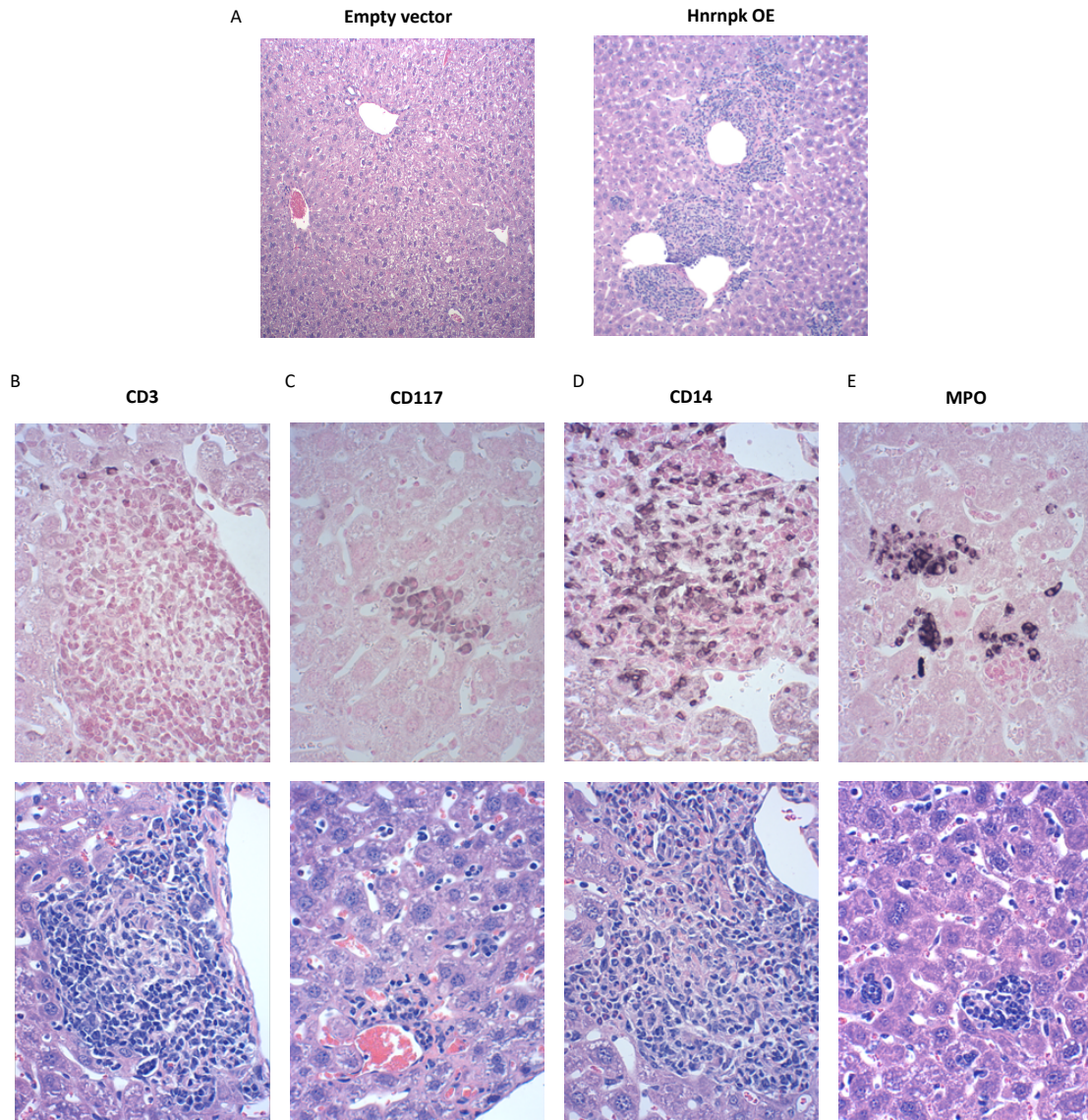


Figure 20. Hepatic extramedullary hematopoiesis in recipients of hnRNP K-overexpressing FLCs. A. H&E staining of liver from a representative mouse transplanted with empty vector control FLCs (left) or hnRNP K-overexpressing FLCs (right). B-E. Higher magnification immunohistochemistry in livers of representative mice transplanted with hnRNP K-overexpressing FLCs. Corresponding H&E staining is in the bottom panels.

Mice receiving hnRNP K-overexpressing FLCs have abnormal blood counts

We next evaluated how this aberrant hematopoiesis affected the components of peripheral blood. Recipients of hnRNP K-overexpressing FLCs showed an increase in total white blood cell counts (Figure 21A). Likewise, a decrease in segmented cells (neutrophils) was evident in these recipients compared with empty vector controls (Figure 21B). A slight increase in lymphocytes was evident in hnRNP K-overexpressing FLC recipients (Figure 21C). Eosinophil percentages were also significantly increased in the context of hnRNP K overexpression (Figure 21D). Monocyte percentages were unchanged between groups (Figure 21E). While not statistically significant, the percentage of large immature cells (LICs), trended towards increased in recipients of hnRNP K-overexpressing FLCs (Figure 21F). Differences in platelet count and hemoglobin were not evident between groups (Figures 21G). These observations indicate that hnRNP K-overexpression affects predominantly leukocytes, while largely sparing red blood cell and platelet production.

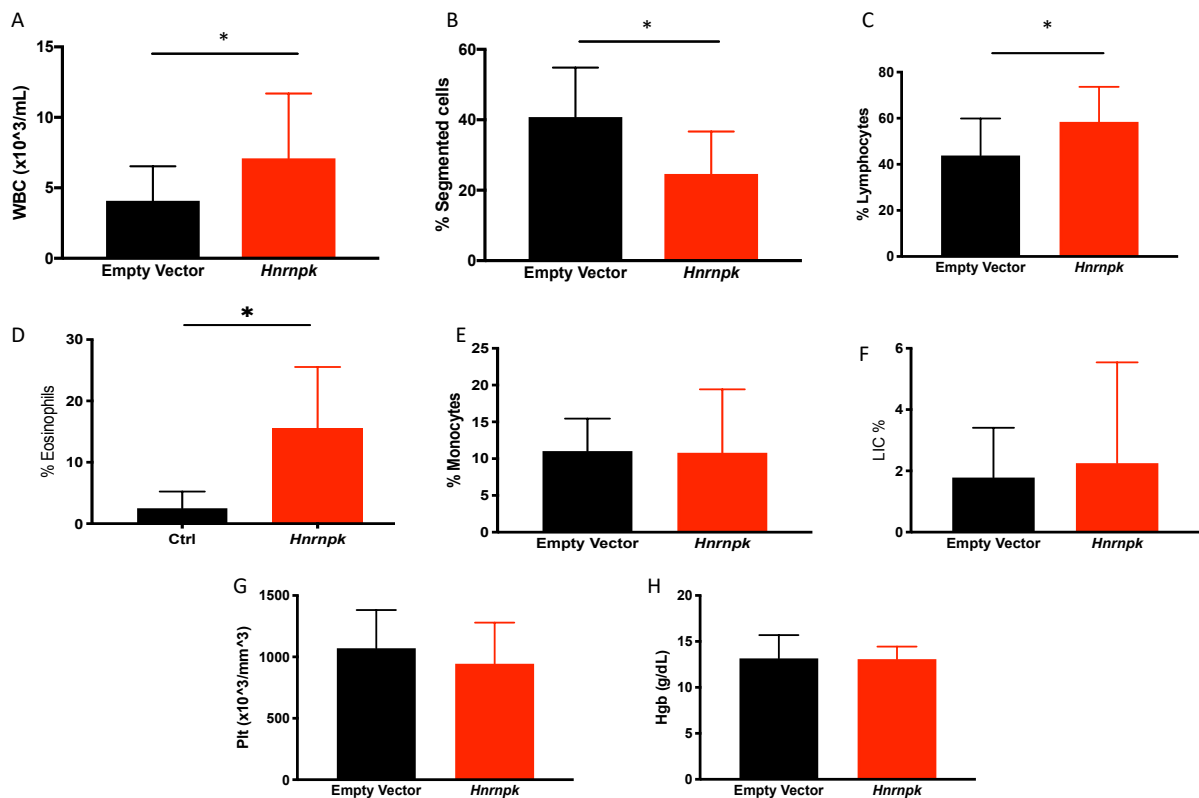


Figure 21. Peripheral blood counts in mice transplanted with FLCs. Bar graphs of A. White blood cell count (WBC) B. Segmented cell percentage C. Lymphocyte percentage D. Eosinophil

percentage E. Monocyte percentage F. Large immature cell (LIC) percentage G. Platelet (Plt) count H. Hemoglobin (Hgb) in recipients of hnRNP K-overexpressing FLCs (red) or empty vector control FLCs (black).

Peripheral blood smears are abnormal in recipients of hnRNP K-overexpressing FLCs

To visualize and confirm findings from our CBC analyses, we examined peripheral blood smears from our mice. Recipients of hnRNP K-overexpressing FLCs frequently had large numbers of WBCs, with numerous leukocytes present in a high-powered field. We observed myeloid cells in various levels of maturation (Figures 22B-E)—indicative of bone marrow stress. In addition, several mice had large numbers of highly abnormal leukocytes appearing to be myeloid in origin (Figures 22F-H). In line with the CBC data, circulating eosinophils were also present (Figure 22I).

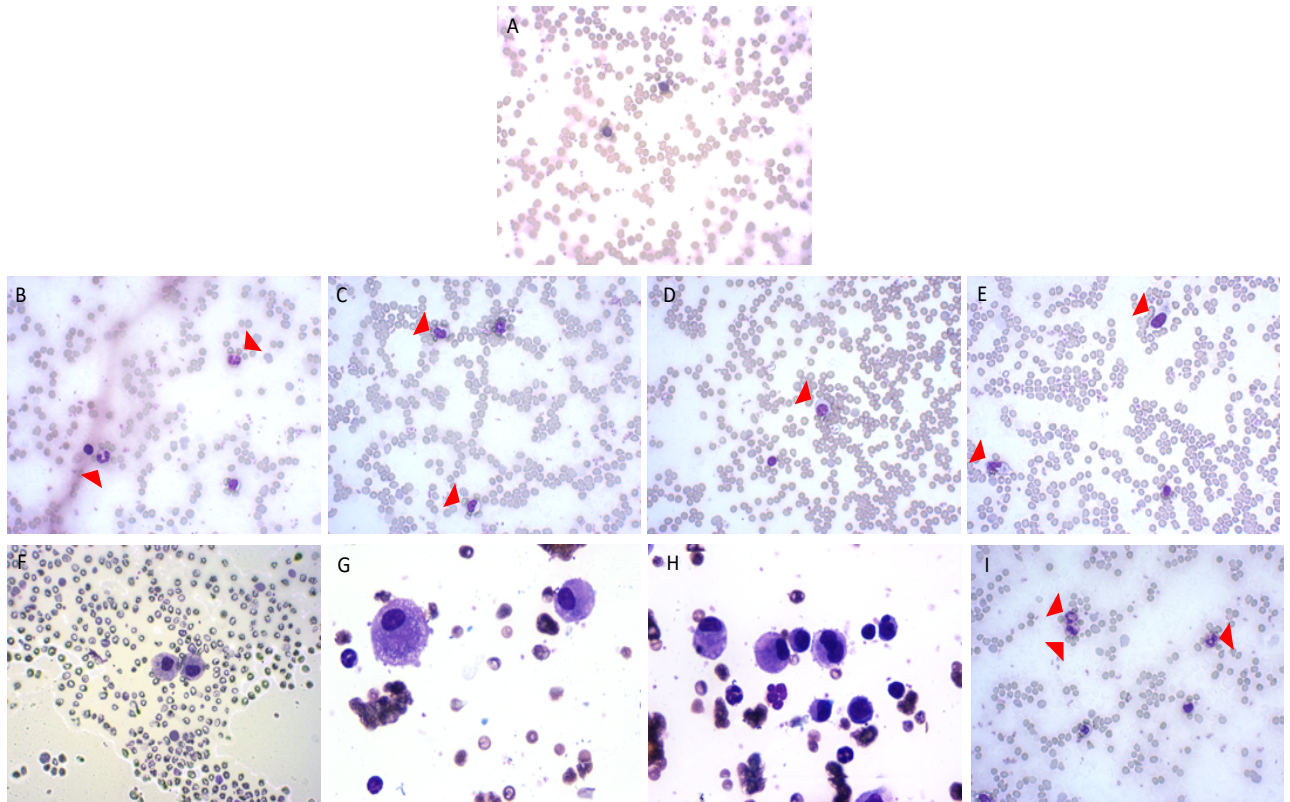


Figure 22. Peripheral blood is abnormal in recipients of hnRNP K-overexpressing FLCs. Wright-Giemsa staining of peripheral blood smears from an empty vector control (A) or recipients of hnRNP K-overexpressing FLCs (B-I). Red arrows indicate cells with morphologic features consistent with bands (B), metamyelocytes (C), myelocytes (D), or promyelocytes (E). F-H highlight other circulating abnormal leukocytes. I. Red arrows indicate eosinophils.

3.4 Discussion

Based on our clinical observations of increased hnRNP K expression in AML and its correlation with inferior clinical outcomes (Chapter 2), we proposed that hnRNP K may have oncogenic capability when overexpressed. To formally test this hypothesis, we transplanted mice with hematopoietic stem and progenitor cells (from the fetal liver) that were stably transduced to overexpress hnRNP K. Recipients of these cells had substantially shortened survival (Figure 16) and grossly aberrant hematopoiesis, characterized by bone marrow harboring substantial morphologic myeloid hyperplasia, extramedullary hematopoiesis in the spleen and liver, and circulating immature myeloid cells (Figures 17-22). These findings support the notion that hnRNP K overexpression drives myeloproliferative disease in mice.

hnRNP K-overexpression in FLCs led to increased colony formation in methylcellulose (Figure 12), suggesting that hnRNP K can alter a cell's ability to self-renew and to differentiate. Thus, our findings insinuate that hnRNP K overexpression may influence the stemness of hematopoietic cells. It would then follow that compared to wildtype hematopoietic stem cells, fewer hnRNP K-overexpressing cells would be required to reconstitute hematopoiesis in a lethally-irradiated recipient mouse. While we did not perform these limiting dilution studies, these future pursuits would be of interest.

Despite not being routinely reported in the literature, we used an empty vector control to abrogate alterations in FLC characteristics due solely to transduction. Since the focus of this chapter was on the influence of hnRNP K overexpression in a mouse model, however, we did not fully characterize the impact of altered hnRNP K levels in the initial FLCs. Therefore, we cannot definitively conclude that the hnRNP K-overexpressing FLCs transplanted into recipient mice were immunophenotypically identical to the empty vector controls. However, we performed our experiments with numerous primary transplant models and generated very similar results. To characterize these models in greater depth, as well as to more quantitatively assess their oncogenic properties, secondary transplants are currently underway. Preliminary data indicate

cells from several primary recipients of hnRNP K-overexpressing FLCs have been successfully transplanted into secondary recipients.

In these studies, we transplanted FLCs into immunocompromised NSG mice. While use of these mice exponentially diminishes the probability of transplant rejection, due to a grossly absent adaptive immune system, this is a suboptimal model for evaluating other aspects of biology, particularly host immune response.^{193, 194} Since these cells are isolated from C57Bl6 mice, mice of this immunocompetent background are also suitable transplant recipients for these FLCs.¹⁸⁰ These mice are particularly useful models for evaluating treatment response, including to immunotherapies.^{195, 196} Therefore, future studies transplanting hnRNP K-overexpressing FLCs into immunocompetent mice would be worthwhile.

The myeloproliferative phenotype observed with hnRNP K overexpression in our FLC transduction and transplantation model is particularly intriguing for several reasons. We have recently described a mouse model wherein hnRNP K is overexpressed in the B-cell compartment.⁸³ These mice develop highly penetrant, transplantable B-cell lymphomas that faithfully recapitulate clinical characteristics of human diffuse large B-cell lymphoma (DLBCL), where we observed high expression of hnRNP K.⁸³ Interestingly, none of the mice that received hnRNP K-overexpressing FLCs in our current study developed lymphoma. Since these experiments placed hnRNP K-overexpression in hematopoietic stem cells, this suggests that high hnRNP K expression favors myeloid differentiation. Consistent with this, microarray data in normal human hematopoiesis revealed RNA expression of *HNRNPK* was substantially higher in progenitor cells with a myeloid bias (common myeloid progenitors, granulocyte-monocyte progenitors, or megakaryocyte-erythroid progenitors) than in lymphoid-committed progenitor cells.¹⁷¹ Other groups have also confirmed this, with *HNRNPK* being observed at much higher levels in myeloid-committed progenitors compared to lymphoid-biased cells.¹⁹⁷

This notion that hnRNP K overexpression causes a myeloproliferative phenotype in our current mouse model is also captivating given that a strikingly similar phenotype is observed in

mice that are haploinsufficient for *Hnrnpk*.⁵³ These mice also have shortened survival, mediated most frequently by a phenotype involving myeloid hyperplasia, though a subset of *Hnrnpk*^{+/-} animals also develop lymphoma or other tumors.⁵³ Such an observation strongly suggests that normal hematopoiesis, particularly myelopoiesis, requires exquisitely tight regulation of hnRNP K expression.

In particular, an interesting similarity between these two mouse models was eosinophilia (Figures 21, 22). While myeloid malignancy can cause this phenomenon, more common causes of eosinophilia include allergic reaction or parasitic infection.¹⁹⁸ In our current study, we did not observe any obvious infections in the mice. While the recipients of empty vector control FLCs did not exhibit increased peripheral blood eosinophils, it is possible that hnRNP K-overexpressing FLCs could confer an immunologic deficit to recipient mice, thus predisposing this group to parasitic infections. However, the absence of visible signs of parasitic infection in these mice, such as diarrhea, abdominal swelling, or substantial weight loss, diminish interest in this possibility. In humans, rare cases of T-cell malignancies secreting cytokines (mostly IL-5) that stimulate eosinophil production have also been described as a mechanism of eosinophilia.¹⁹⁹ We did not evaluate IL-5 levels in our mice; however, it is possible that hnRNP K-overexpression alters secretion of cytokines, including those that promote eosinophil development. Given that increased eosinophils were evident in recipients of hnRNP K-overexpressing FLCs and *Hnrnpk*^{+/-} mice⁵³, it is possible that this is a manifestation of the aberrant myeloid development driven by hnRNP K. Of note, clinical eosinophilia is observed in cases of AML with core-binding factor abnormalities, usually in cases of inv(16) or t(16;16), and occasionally t(8;21).^{200, 201} These chromosomal aberrations result in altered RUNX1 transcriptional programs—a topic that will be addressed in Chapters 4 and 5 as a possible mechanism of hnRNP K-mediated oncogenesis.

Chapter 4

Examining the molecular basis of the oncogenic function of hnRNP K

4.1 Introduction

In acute myeloid leukemia (AML), hnRNP K overexpression occurs in approximately 20% of cases (Chapter 2). This increased expression is correlated with decreased remission durations and overall survival (Chapter 2). To formally test the hypothesis, then, that hnRNP K is a yet uncharacterized oncogene in AML, we developed a mouse model wherein hnRNP K is overexpressed in hematopoietic stem cells (Chapter 3). These mice developed hematologic malignancies characterized by splenomegaly, abnormal blood counts, hepatic infiltration of immature myeloid cells, and early death (Chapter 3). In this chapter, we aimed to define the mechanisms by which overexpression of hnRNP K causes such disease by examining the molecular basis of the oncogenic function of hnRNP K.

hnRNP K is a ssDNA/RNA-binding protein implicated in a myriad of biologic processes. These include signal transduction³⁹⁻⁴³, chromatin remodeling⁴⁴⁻⁴⁷, transcription⁴⁸⁻⁵⁴, RNA splicing⁵⁵⁻⁵⁹, mRNA stability⁶⁰⁻⁶², and translation^{54, 63-69}. In our recent model of B-cell lymphomas, hnRNP K overexpression caused malignant transformation by stabilizing *MYC* RNA and promoting translation of c-Myc.⁸³ Contrastingly, *HNRNPK* haploinsufficient mice developed myeloid neoplasms via transcriptional deregulation of *Cdkn1a*, *Cebpa*, and *Cebpb*.⁵³

The multifunctionality of hnRNP K is due in large part to its composite domains. With both a nuclear localization signal (NLS) and a nuclear shuttling domain (KNS), hnRNP K can move bi-directionally between nucleus and cytoplasm.^{70, 71} This protein contains three K homology (KH) domains, which recognize single-stranded nucleic acid—either DNA or RNA.⁷⁴ In addition, an unstructured region (K-interactive [KI]) resides between the terminal KH domains, and is where most known protein interactions occur.^{72, 73} This domain allows hnRNP K to be part of multi-protein complexes, which may affect its nucleic acid binding activities.^{41, 73}

In this chapter, we sought to determine which of hnRNP K's functions was most critical to the development of myeloid neoplasia in the context of hnRNP K overexpression.

4.2 Materials/Methods

Immunoprecipitation-mass spectrometry: OCIAML3 cells were transduced with pCDH-510-rtTA-EF1-puro²⁰² and pCDH-TRE-HNRNPK-EF1-CopGFP prior to selection in puromycin and fluorescence-activated cell sorting for GFP-positive cells. Cells transduced with pCDH-TRE-EF1-CopGFP plasmid served as negative (empty vector) controls. hnRNP K overexpression was induced by exposing cells to 1 µg/mL of doxycycline for 32 hours. Nuclear and cytoplasmic fractions were separated by dounce homogenization as described.²⁰³ Immunoprecipitation was performed with antibodies against hnRNP K (D6, Santa Cruz Biotechnology, Dallas, TX, USA) or IgG (ab18413, Abcam, Cambridge, MA, USA). Proteins immunoprecipitated out of solution were resolved on a 4-15% SDS-PAGE gel and silver stained (catalog# 24612, Thermo Fisher, Waltham, MA, USA). Bands of visibly different intensities were excised and digested in the gel with 200ng of sequencing grade trypsin (Promega, Madison, WI, USA) at 37°C for 18 hours. The extracted bands were then analyzed by high-sensitivity liquid chromatography with tandem mass spectrometry (LC-MS-MS) with an orbital ion-trap mass spectrometer (Orbitrap Elite, Thermo Fisher Scientific, Waltham, MA, USA). Proteins were subsequently identified by searching the fragment spectra against the SWISS-PROT database²⁰⁴ (EBI) using Mascot²⁰⁵ (Matrix Science, Boston, MA, USA) or Sequest²⁰⁶ (Thermo Fisher Scientific, Waltham, MA, USA). Further details are available elsewhere.⁸³

Formaldehyde-RNA immunoprecipitation (fRIP): fRIP was performed using a modified protocol described previously.²⁰⁷ 100 million OCIAML3 cells were resuspended to a final concentration of 1x10⁶ cells/mL in RPMI 1640 media without FBS or antibiotics. Cross-linking was performed via addition of formaldehyde to a final concentration of 0.25% and incubated at room temperature. After 20 minutes, glycine at a final concentration of 125mM was used to quench the reaction. Fixed cells were pelleted and resuspended in polysome lysis buffer (100mM KCl, 5mM MgCl₂, 10mM HEPES (pH 7.0), 0.5% NP40, 1mM DTT, 100 units/mL RNaseOut (Thermo

Fisher Scientific, Waltham, MA, USA), 400 μ M vanadyl ribonucleoside complex (VRC), and protease inhibitor cocktail (Roche, Basel, Switzerland)) and left at -80°C overnight. After thawing on ice, the suspension was lysed with a digital sonifier at 10% amplitude in 1 second bursts followed by 4 seconds of rest for a total of 90 seconds. Lysates were subsequently cleared by high-speed centrifugation at 4°C and pre-cleared using protein A and protein G beads in a 1:1 ratio. Beads coated with hnRNP K antibodies were suspended in NT2 buffer (50 mM Tris-HCl (pH 7.4), 150mM NaCl, 1mM MgCl₂, 0.05% IGEPAL) supplemented with RNaseOut, VRC, DTT, and EDTA. Next, 10% of the pre-cleared lysate removed to serve as the “input” sample, and the remainder was added to the bead slurry. The bead-lysate mixture was incubated while tumbling at room temperature. After 2 hours, the mixture was washed four times with NT2 buffer. Crosslinking was reversed in the lysates by addition of NaCl (final concentration 200 mM) and proteinase K (final concentration 20mg/mL), incubated at 42°C for one hour, then moved to 65°C for an additional hour. Remaining RNA was isolated using Trizol (Sigma-Aldrich, St. Louis, MO, USA), purified, and concentrated with Zymo RNA Clean and Concentrator Kit, as per manufacturer instructions (Zymo Research, Irvine, CA, USA). Finally, RNA was reverse transcribed to cDNA and subjected to single-read sequencing on an Illumina HiSeq 2000 at 36 nucleotides per read at the MD Anderson Sequencing Core. All samples were performed in triplicate. These data is now publicly available under GSE126479 in the Gene Expression Omnibus.⁸³

fRIP analysis: Raw reads were trimmed to a minimum length of 25 base pairs using FLEXBAR.²⁰⁸ BOWTIE2²⁰⁹ was used to align trimmed reads to Genome Reference Consortium Human Build 38 (GRCh38),²¹⁰ and outputted as .sam files, which were converted to .bam files using SAMtools.²¹¹ This algorithm was then used to index, sort, and filter unique read counts per alignment. Differential representation between input and immunoprecipitated (IP) samples was done with Cufflinks, Cuffmerge, and Cuffdiff.^{212, 213} Using Cuffdiff, q-values were calculated

comparing input to IP samples. Peaks were visualized in Integrative genomics viewer (IGV).²¹⁴ Output was sorted by q-value, indexed, and subsetted by causal implication in cancers using datasets of known tumor suppressors and oncogenes.^{215, 216} Further details with analysis specifications are available elsewhere.⁸³

Native RNA immunoprecipitation (RIP): OCIAML3 cells were lysed and processed according to manufacturer's instructions using the Magna-RIP RNA-binding protein immunoprecipitation kit (Millipore, Burlington, MA, USA). In brief, cells were lysed in RIP lysis buffer and incubated with magnetic beads and 5 μ g of either IgG or hnRNP K antibodies (3C2, Abcam, Cambridge, MA, USA) at 4°C for three hours. After washing, RNA-protein complexes were subsequently disrupted by incubating with the presence of proteinase K. RNA was then extracted, purified, and DNase treated at 37°C for one hour. Reverse transcription of RNA was performed as with other PCR assays using iScript cDNA synthesis kit (BioRad, Hercules, CA, USA). qRT-PCR was performed with iTaq Universal SYBR green using an ABI StepOnePlus Real Time PCR machine.

Identifying putative hnRNP K binding sites: A computer-based algorithm was developed to scan the *RUNX1* transcript for possible hnRNP K binding sites. Putative hnRNP K consensus RNA-binding sequences^{67, 217} contained either ≥ 2 (U/C)CCC motifs within 19 nucleotides or a motif consisting of: CCAUCN₂₋₇(A/U)CCC(A/U)N₇₋₁₈UCA(C/U)C, where N is any nucleic acid. Further details are available elsewhere.⁸³

Recombinant hnRNP K protein expression and purification: Using ligation-independent cloning²¹⁸, the DNA sequence encoding full-length hnRNP K was placed into the pNic28-Bsa4 plasmid with a TEV protease cleavage site immediately upstream of *HNRNPK*. *E. coli* transformed with this plasmid were grown at overnight at 37°C in auto-induction media. After

centrifugation, cell pellets were stored at -80°C overnight, thawed, and resuspended in 50mM HEPES, 300mM NaCl, 10% glycerol, 0.5mM TCEP, complete protease inhibitors, 0.5mg/mL lysozyme, and 20 $\mu\text{g}/\text{mL}$ DNase at pH 8.0 and lysed by sonication. Lysate was then centrifuged at 18,000 rpm for one hour at 4°C . Full-length hnRNP K protein was purified from the clarified lysate with Ni-NTA resin in 50mM HEPES with 300mM NaCl at pH 8.0. To remove contaminating proteins, the resin was washed with 50mM HEPES, 300mM NaCl, and 20mM imidazole. Recombinant hnRNP K protein was eluted with 500mM imidazole. His6-TEV protease was used to cleaved the N-terminal His-tag during dialysis against 50mM HEPES, 150mM NaCl, and 0.5mM TCEP at pH 7.5. This solution was passed through the Ni-NTA column to remove the cleaved His-tag as well as the remaining TEV protease. Size exclusion chromatography was then performed using a Superdex 200 column equilibrated with 20mM HEPES, 200mM NaCl, 5% glycerol, and 0.5mM TCEP at pH 7.5. After these steps, SDS-PAGE analysis indicated the recombinant full-length hnRNP K protein was $>95\%$ pure.

Fluorescence anisotropy: RNA oligonucleotides with a 5' 6-FAM label were purchased from Sigma Aldrich (St. Louis, MO, USA) and used at a final concentration of 2nM. Sequences are included in Table 3. Recombinant hnRNP K protein, serially diluted in PBS to range from 0.1nM to 10 μM , was added to the oligos. Fluorescence anisotropy (FA) values were measured with excitation wavelength 485nm and emission wavelength detected at 528nm on a Synergy Neo multi-mode plate reader (BioTek, Winooski, VT, USA). Fluorescein at 10nM was used as an internal standard. All readings were done at room temperature and performed in at least triplicate. FA data was analyzed with Prism 8. Data was fit to the following equation: $FA = FA_i + B_{\max} * [\text{oligo}] / (K_d + [\text{oligo}])$ where initial FA is represented by FA_i and the overall change in FA is represented by B_{\max} . Here it is relevant to note that I adore penguins as much as Sean Post loves B_{\max} .

Human	
Name	Sequence
hRUNX1 1bs	UCCCCUC
hRUNX1 2bs	UCCCCUCCCC
hRUNX1 3bs	UCCCCUCCCCUCCC
hRUNX1 5' UTR	CGCCCCCCCCACCCCCCGCAGUAAUAAAGGCCCCUGA
hRUNX1 5' UTR (mut)	CGC GCGCGCGCACGCGCC GCAGUAAUAAAGGC GCC UGA
hRUNX1 int5-6	UCUCUUCCCUCCCUCCUCCCUCCCCCAU
hRUNX1 int5-6 (mut)	UCU GUUCGCUCGCUCGUUCGCUCGCG CAU
Mouse	
Name	Sequence
mRunx1 intron 5-6	UCCUCCUCCCUUCCCCUCCCGGUCCUA
mRunx1 intron 5-6 (mut)	UCCUCCUC GPUUCGCCUCGCGGUCG CUA

Table 3. Fluorescence anisotropy oligos. Human and mouse sequences of RNA oligos used for fluorescence anisotropy assays. Sequences are listed from 5' to 3'. Red letters highlight where mutations were made.

Thermal shift assay: The thermal shift assay was performed in a 96-well plate using SYPRO orange dye in a StepOne Plus Real Time PCR System (Applied Biosystems, Foster City, CA, USA). The reaction was prepared in PBS buffer and contained 5 μ M of recombinant hnRNP K protein, SYPRO orange dye at a final concentration of 2X and DNA oligos at 1 μ M, 2.5 μ M, and 5 μ M. Oligos used were the DNA equivalents to those used in the fluorescence anisotropy assays. The total reaction volume was 20 μ L and 4 technical replicates were done for each sample. Negative control samples containing no protein were also run on the same plate. The samples were heated from 25°C to 99°C and the fluorescence measured at each temperature increment. The first derivative of fluorescence was calculated at each temperature and the temperature corresponding to the minima was designated as the melting temperature of the sample.

4.3 Results

hnRNP K interacts with RNA processing machinery and ribosomal subunits

Given the plethora of cellular processes in which hnRNP K has been implicated, we performed hnRNP K immunoprecipitation (IP) followed by mass spectrometry to identify proteins with which hnRNP K interacts. In a human AML cell line, OCIAML3, compared to IgG IP, hnRNP K IP pulled down multiple proteins that were clearly visible after silver staining (Figure 23A). This was true in both cytoplasmic and nuclear fractions of OCIAML3 cells. Endogenous cytoplasmic hnRNP K was found to interact with numerous ribosomal proteins (Figure 23B). In the nucleus, hnRNP K interacting partners were enriched for RNA-binding proteins in addition to ribosomal proteins (Figure 23B). Gene ontology (GO) analyses²¹⁹ from each fraction indicate hnRNP K interacts with proteins involved in translation and RNA binding/modification (Figures 23C,D). Together, this strongly indicates that hnRNP K is involved in post-transcriptional processes.

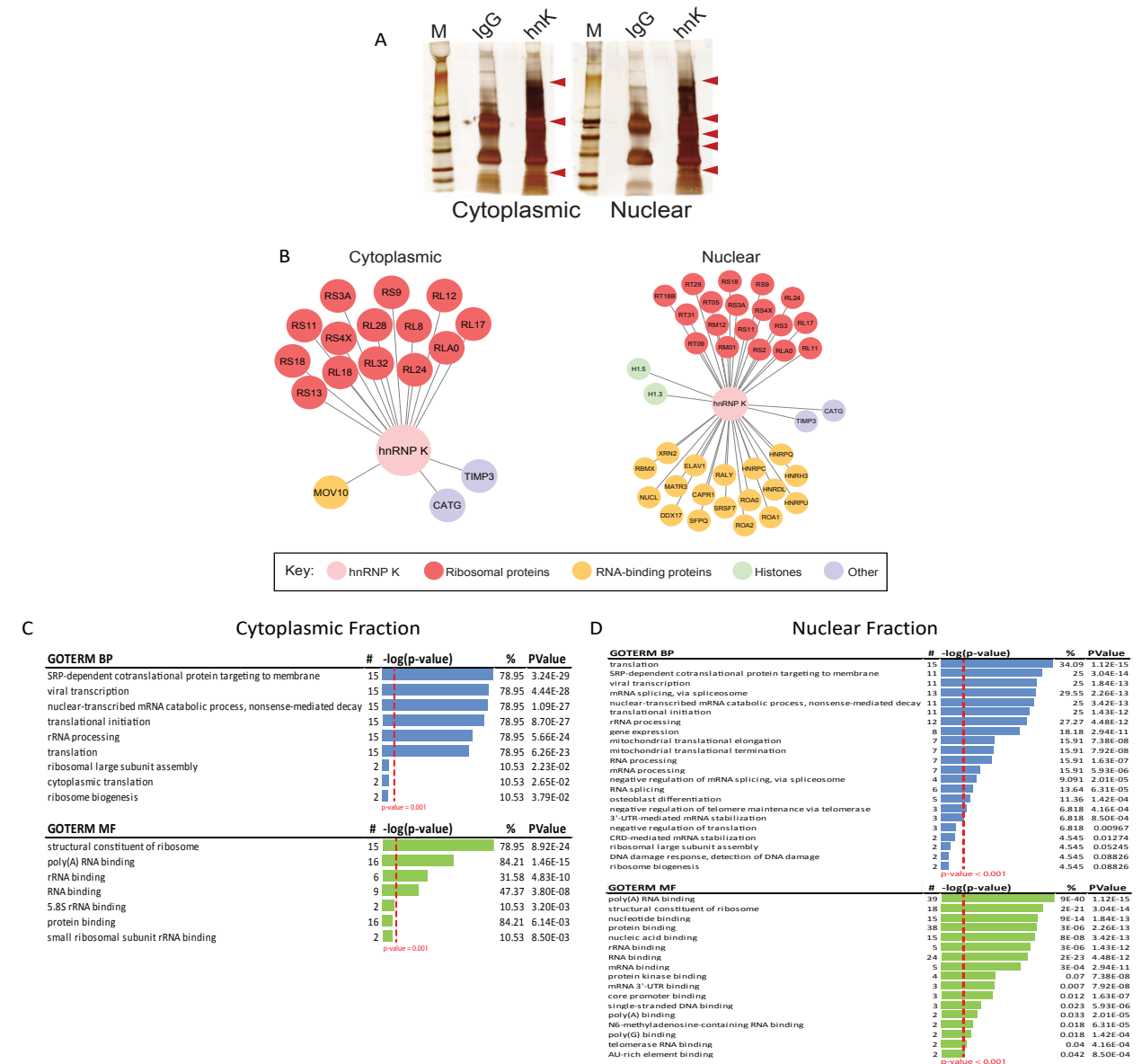


Figure 23. Immunoprecipitation followed by mass spectrometry reveals hnRNP K interacting partners. A. Silver stain of SDS-PAGE blots from OCIAML3 cells fractionated to cytoplasmic (left) and nuclear (right) compartments. IPs were performed with either IgG or hnRNP K (hnK). Red arrows indicate bands that were excised and sent for mass-spectrometry. B. Proteins co-immunoprecipitated with hnRNP K were visualized with Cytoscape.²²⁰ C-D. Gene ontology (GO) analyses of proteins immunoprecipitated with hnRNP K in cytoplasmic (C) or nuclear (D) fractions. Blue bars divide proteins by biologic process and green bars divide proteins by molecular function. Bars indicate statistical significance of proteins falling into the

associated categories. Figure partially reproduced with permission from Gallardo & Malaney et al., JNCI, 2019.⁸³

hnRNP K interacts with numerous RNA transcripts, including RUNX1

As indicated by the hnRNP K IP-MS, hnRNP K interacts with numerous proteins that bind RNA and/or are involved in RNA processing. To understand how modulation of RNA species might contribute to leukemogenesis, we sought to identify the RNA targets with which hnRNP K interacts. Formaldehyde-RNA immunoprecipitation followed by sequencing revealed that hnRNP K was associated with numerous transcripts, including many involved in cancer progression. When evaluated in the context of known tumor suppressors or oncogenes, one of the most significant hnRNP K-associated transcripts was that of *RUNX1*, which a critical transcription factor for both normal and malignant hematopoiesis²²¹ (Figure 24).

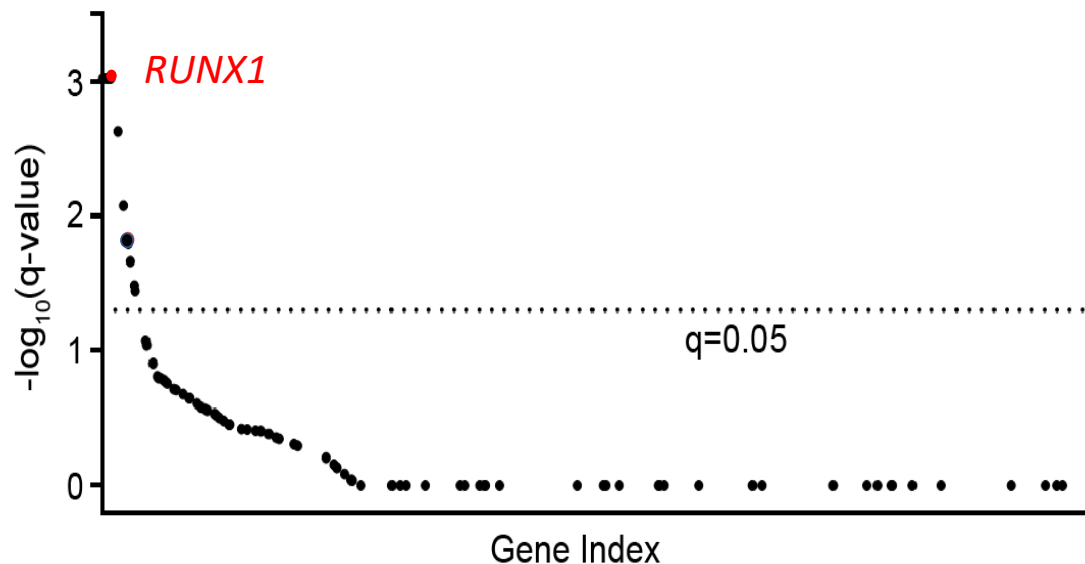


Figure 24. hnRNP K binds to the *RUNX1* transcript. Depiction of hnRNP K-associated transcripts as determined by fRIP analyses. Results are subsetted based on a list of known oncogenes/tumor suppressors. The dotted line indicates a cutoff representing a q-value of 0.05. Figure modified with permission from Gallardo & Malaney et al., JNCI, 2019.⁸³

hnRNP K interacts with RUNX1 transcript

To validate the finding from fRIP-Seq that hnRNP K interacts with *RUNX1*, we performed native RIP followed by PCR. Indeed, *RUNX1* mRNA was substantially enriched after hnRNP K IP compared to input (Figure 25). This supports the notion that hnRNP K interacts with *RUNX1* in human AML cells.

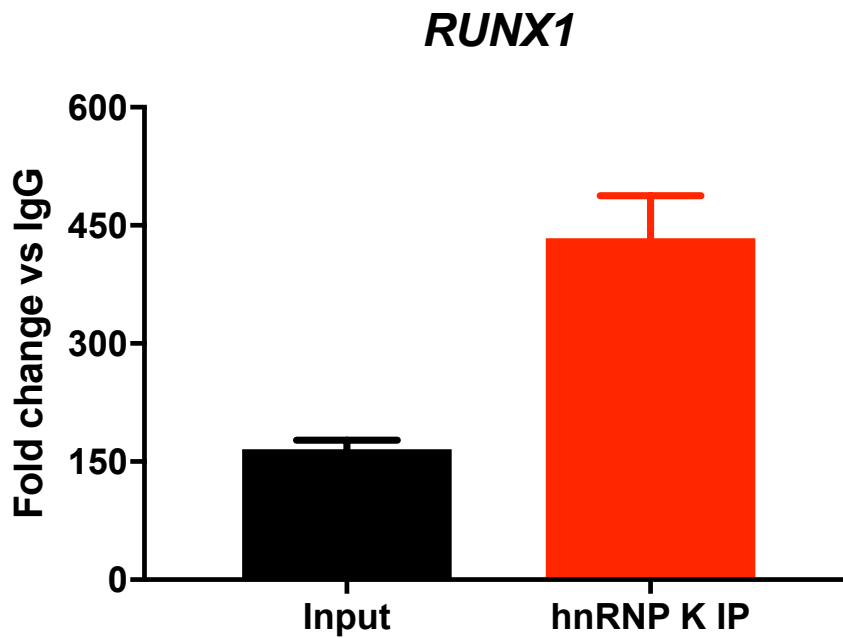


Figure 25. *RUNX1* mRNA interacts with hnRNP K. Bar graph illustrating fold change of *RUNX1* RNA in input sample of OCIAML3 cells compared with hnRNP K IP. Data is represented as fold change of input (black) or hnRNP K IP (red) as compared to IgG IP.

RUNX1 RNA harbors putative hnRNP K binding sites

Given our fRIP and RIP data indicating that hnRNP K binds human *RUNX1*, we sought to determine whether this occurred in a sequence-specific manner. Knowing that hnRNP K is a poly-C binding protein²²², we used a computer program to scan the human *RUNX1* transcript for multiple stretches of at least three cytosine residues. We identified two putative hnRNP K binding sites in the human *RUNX1* transcript (Figure 26). One binding site was located in the 5' untranslated region (UTR) of the majority of *RUNX1* isoforms. No analogous hnRNP K binding

site was identified in murine *Runx1*. In contrast, a second consensus hnRNP K binding site was identified at an intron-exon boundary in human *RUNX1*; this site was also found in murine *Runx1* isoforms (Figure 26). In the majority of *RUNX1* isoforms, this site exists at the boundary of intron 3 and exon 4. Of note, we have referred to this site as intron 5-6, since this is its location in the longest isoform of *RUNX1* (*RUNX1C*; ENST00000300305.7).

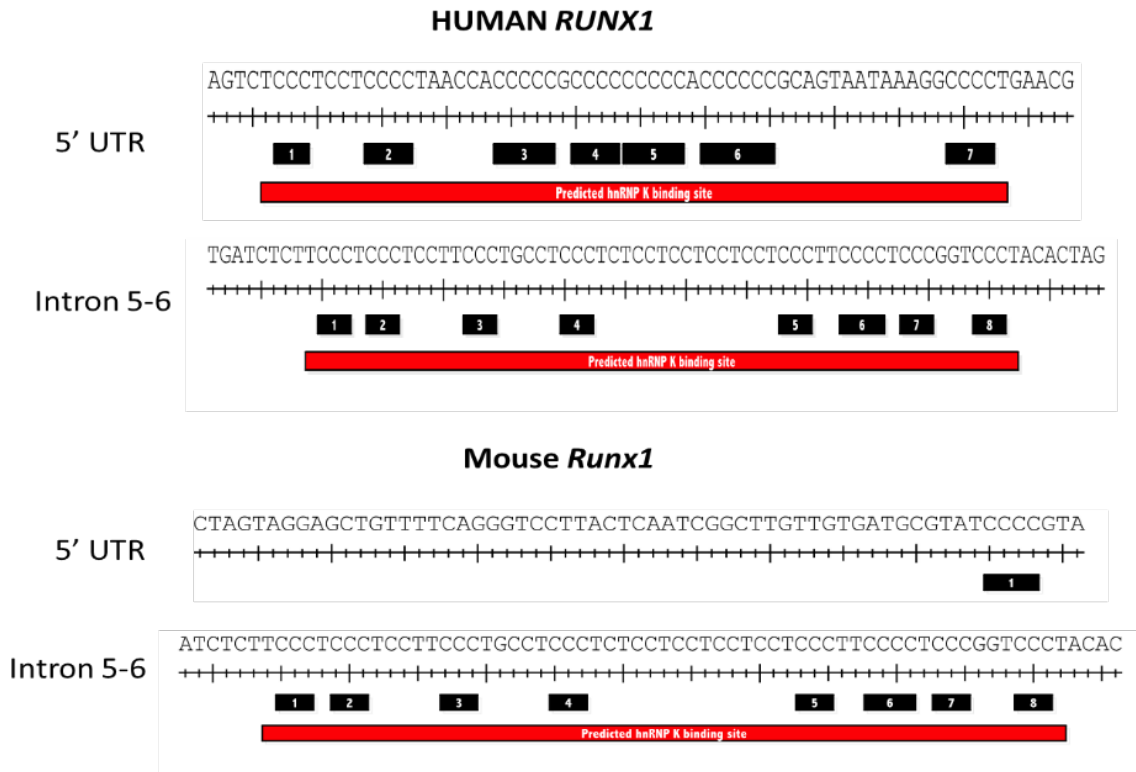


Figure 26. Predicted hnRNP K binding sites in *RUNX1*. Poly(C) regions in the *RUNX1* transcript are denoted by black boxes. A cluster of 2 poly(C) regions within 19 nucleotides or 4 poly(C) regions within 38 nucleotides constitutes the hnRNP K consensus binding sequence, as reported in the literature^{67, 217}; these are indicated in red.

hnRNP K binds tightly to poly(C) containing RNA

To evaluate whether hnRNP K was capable of directly binding the predicted sites in *RUNX1*, we performed fluorescence anisotropy (FA) studies using recombinant hnRNP K

protein. In these assays, a 6-FAM labelled *RUNX1* RNA containing the predicted hnRNP K binding site was placed into solution with recombinant full-length hnRNP K protein. In the absence of protein binding, the fluorescently labelled oligonucleotide rotates rapidly in solution, being measured as low anisotropy. Upon protein binding, the rotation of the oligonucleotide in solution is reduced, and is measured as high anisotropy.^{223, 224} Oligos, derived from the *RUNX1* transcript, containing 1, 2 or 3 poly(C) stretches were used in the FA assays. Multiple stretches of poly(C) residues enhanced hnRNP K/RNA binding, (Figure 27), indicating more efficient binding with oligos containing multiple stretches of poly(C) residues.

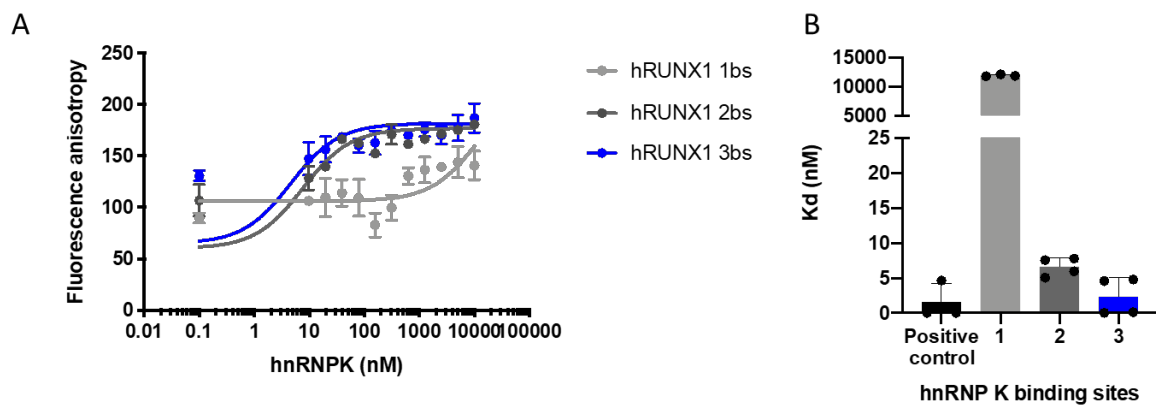


Figure 27. hnRNP K binds stretches of poly(C) RNA derived from *RUNX1*. A. Fluorescence anisotropy curves as a function of increasing hnRNP K concentration. RNA oligos contained increasing numbers of poly(C) regions derived from the *RUNX1* intron 5-6 sequence. B. Summary Kd values for data presented in panel A.

hnRNP K specifically binds RUNX1 RNA

It is clear that hnRNP K stringently binds portions of RNA derived from the *RUNX1* transcript. To evaluate whether hnRNP K was capable of binding larger portions of *RUNX1* that may form secondary structures, we performed fluorescence anisotropy studies with longer portions of the endogenous *RUNX1* RNA. At all predicted hnRNP K binding sites—5' UTR in human and intron 5-6 in human and mouse—hnRNP K was found to bind strongly (Figure 28).

Critically, in each instance, mutations in that disrupted any triple C sequence markedly abrogated hnRNP K binding (Figure 28). This demonstrates that hnRNP K is capable of binding *RUNX1* RNA in a stringent, sequence-specific manner.

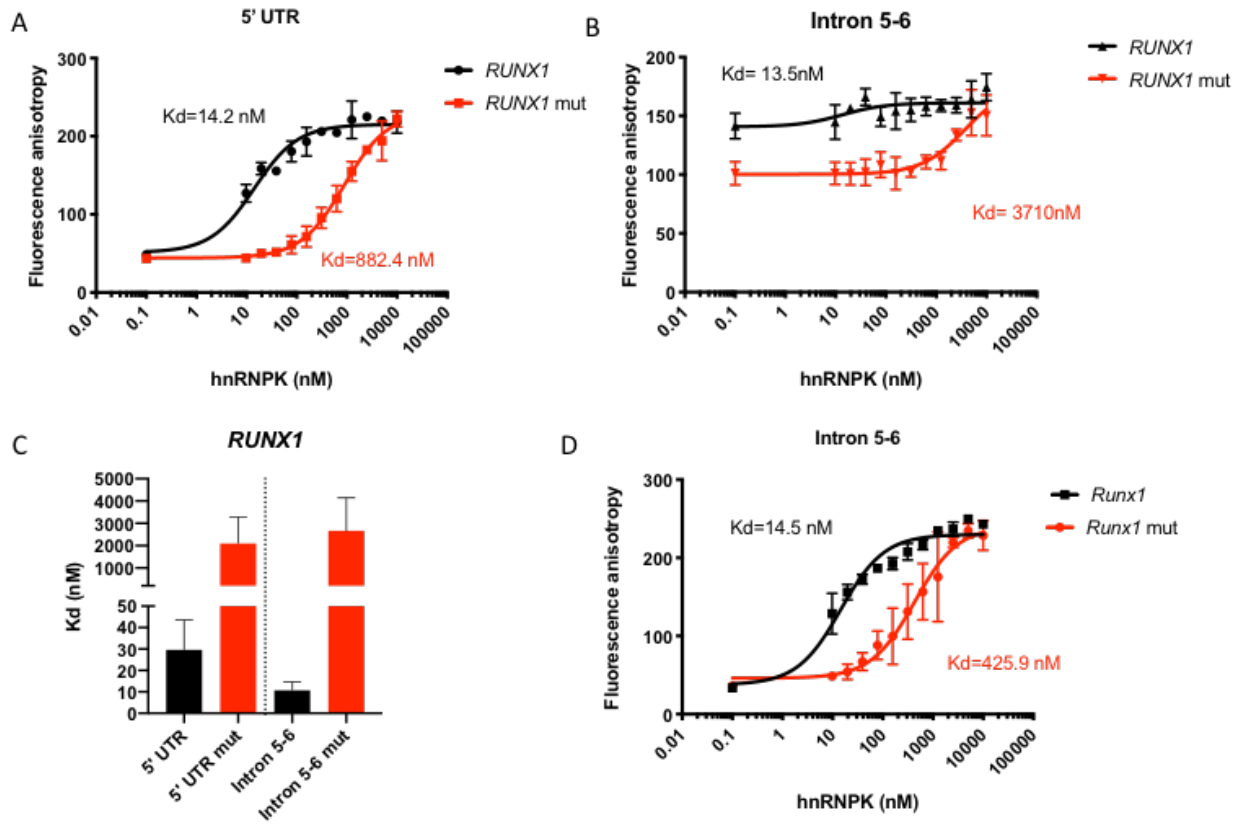


Figure 28. hnRNP K binds human and mouse *RUNX1*. Curves representing fluorescence anisotropy as a function of increasing hnRNP K protein concentration. Oligos spanning the predicted hnRNP K binding sequence in: A. the 5' UTR of human *RUNX1* or B. intron 5-6 of human *RUNX1*. C. Summary K_d values of data from panels A and B. D. Oligos spanning the predicted hnRNP K binding sequence spanning intron 5-6 of mouse *Runx1*.

Oligos derived from *RUNX1* bind to and thermostabilize hnRNP K

As an orthogonal method to evaluate binding of hnRNP K to *RUNX1* oligos, we performed thermal shift assays. In these experiments, recombinant full-length hnRNP K is incubated with SYPRO orange—a dye that becomes fluorescent upon binding hydrophobic amino acid residues.

As the reaction is heated and protein begins to unfold, these hydrophobic residues are exposed, and SYPRO orange binds. By measuring SYPRO orange fluorescence over a gradient of temperatures, the melting temperature of a protein can be calculated. An increase in the melting temperature of a protein, upon addition of any entity, such as a drug or a DNA oligo, indicates an interaction between the protein and that entity.²²⁵

When recombinant full-length hnRNP K was incubated with oligos derived from *RUNX1*—either at the 5' UTR or at intron 5-6, a significant increase in the melting temperature of hnRNP K was observed (Figure 29). Critically, these increases were abrogated in the presence of mutated oligos, further suggesting that hnRNP K binds *RUNX1* RNA in a sequence-specific manner.

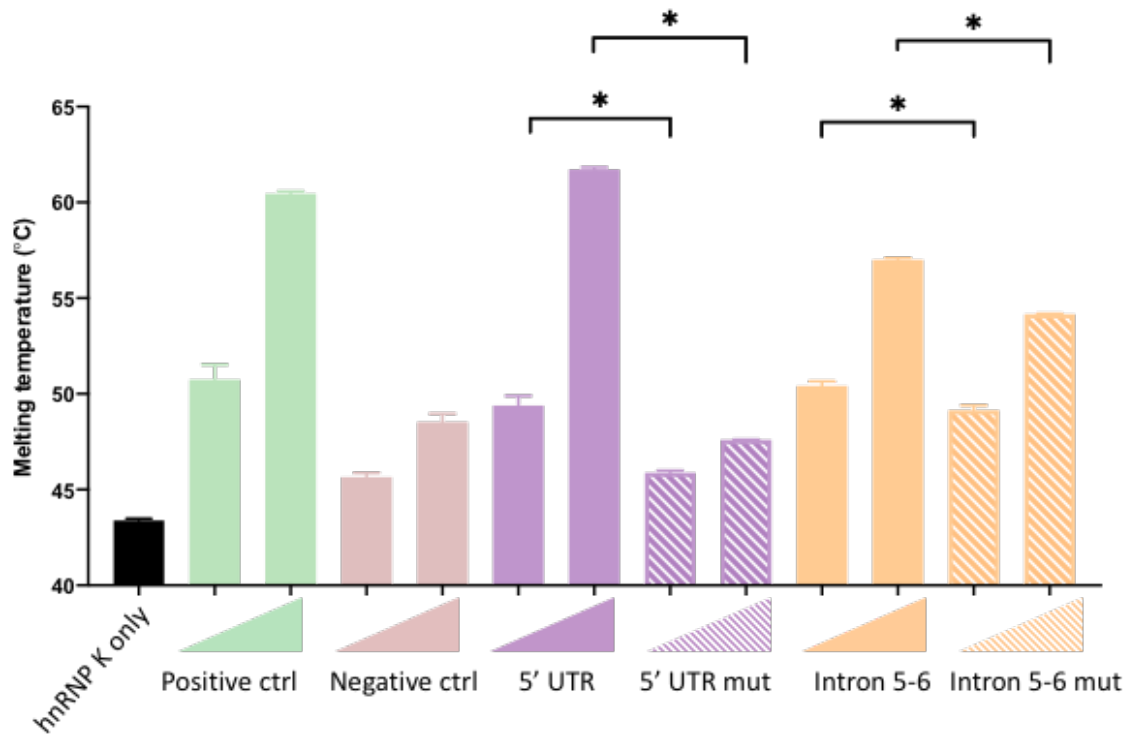


Figure 29. *RUNX1* oligos thermostabilize hnRNP K. Bar graph representing the melting temperature of hnRNP K upon interaction with increasing amounts of the indicated oligos derived from human *RUNX1*. For each oligo, concentrations were 2.5 μM and 5 μM.

4.4 Discussion

In Chapters 2 and 3, we determined that hnRNP K is overexpressed in AML, and is a bona fide oncogene when overexpressed in a mouse model. In this chapter, we aimed to begin identifying a mechanism by which hnRNP K exerts its oncogenic activity—a formidable undertaking considering most descriptions of hnRNP K begin with a litany of its described biologic functions.²²⁶ We therefore conducted hnRNP K IP-MS to provide direction as to which of these biologic functions were most worthy of pursuit. In line with published literature, hnRNP K interacted with a long list of proteins.²²⁷ However, the hnRNP K-interacting proteins we identified in the context of AML were predominantly RNA-binding proteins and ribosomal subunits—entities involved in post-transcriptional cellular processes.

We began our evaluation of hnRNP K as an RNA-binding protein with fRIP-Seq. In this study, we identified *RUNX1* as one of the top hnRNP K-interacting transcripts. This finding was particularly intriguing as *RUNX1* is a critical transcription factor in both normal and malignant hematopoiesis. Definitive hematopoiesis does not occur in *Runx1* knockout mice, resulting in embryonic lethality, and heterozygous loss of *Runx1* also results in aberrant hematopoiesis.⁸⁵ In leukemias of various lineages, *RUNX1* is frequently translocated¹⁰⁴⁻¹⁰⁶, mutated^{109, 116, 228}, or aberrantly expressed.^{98, 229-231} Therefore, an interaction between hnRNP K and an entity so integrally involved in leukemogenesis was of great interest. Of note, *RUNX1* was not identified as an hnRNP K-interacting protein in our IP-MS data. Rather, the interaction was found between hnRNP K protein and *RUNX1* RNA.

We identified two hnRNP K binding sites in *RUNX1* (Figure 26). To our knowledge, the 5' UTR site has not been described, though others recently reported hnRNP K binding at the 3' UTR *in vitro*.²³² An hnRNP K-*Runx1* interaction has been described at the intron 5-6 junction during neuronal differentiation in rats.⁵⁶ Indeed, this intron 5-6 sequence is largely conserved between species¹⁴⁹, which is also supported by our studies that hnRNP K stringently binds the intron 5-6 junction in both human and mouse (Figures 28, 29). Disruptions in poly(C) stretches

in the *RUNX1* transcript dramatically reduce hnRNP K binding (Figures 28, 29). Together, this is compelling evidence that hnRNP K interacts with *RUNX1* RNA in a sequence-specific manner.

hnRNP K is overexpressed in AML (Chapter 2), a disease which often has aberrancies in *RUNX1*.²²¹ In this chapter, we presented compelling evidence that hnRNP K stringently and specifically interacts with *RUNX1* RNA. Hence, one mechanism of hnRNP K's oncogenicity may stem from its ability to interact with *RUNX1*. This hypothesis will be addressed in detail in Chapter 5.

Chapter 5

Evaluating the functional consequence of the hnRNP K-*RUNX1* interaction

5.1 Introduction

hnRNP K is a versatile protein with roles in copious biologic processes including chromatin remodeling⁴⁴⁻⁴⁷, signal transduction³⁹⁻⁴³, transcription⁴⁸⁻⁵⁴, RNA splicing⁵⁵⁻⁵⁹, mRNA stability⁶⁰⁻⁶², and translation^{54, 63-69}. These functions are mediated in large part through its capacity to bind single-stranded nucleic acid, either DNA or RNA, via its KH domains.^{74, 226} In Chapter 4, we observed that hnRNP K stringently and specifically bound human and mouse *RUNX1* RNA near an intron-exon boundary. Therefore, in this chapter, we examined the possibility that hnRNP K has a role in regulating *RUNX1* alternative splicing.

Splicing is a requisite step for the maturation of almost all human mRNAs, where non-coding introns are removed before translation occurs.^{121, 122} Further, a single pre-mRNA can give rise to numerous distinctive mature mRNAs via differential utilization of splice sites—thus including or excluding various exons—a process referred to as alternative splicing.¹²⁵ Such a process therefore allows a single gene to encode for multiple protein products. Indeed, alternative splicing is the major process underlying the disparity between the vast proteomic diversity in cells and a much more limited array of protein-coding genes.¹¹⁷⁻¹²⁰

Splicing occurs via processes involving macromolecular complexes referred to as the major and minor spliceosome.¹²⁵⁻¹²⁷ Small nuclear ribonucleoprotein complexes (snRNPs) U1, U2, U4, U5, and U6 comprise the major spliceosome.¹²⁵ The minor spliceosome consists of U5, U11, U12, U4atac, and U6atac, which are functionally analogous to the components of the major spliceosome.^{126, 128} Sequence motifs at the 5' splice site (5'ss; flanking the upstream exon and intron) and the 3' splice site (3'ss; flanking the intron and downstream exon) are required for intron excision to occur.¹²⁵ The 5'ss is recognized and bound by the U1 snRNP, while the U2 snRNP recognizes and binds the 3'ss.^{125, 126} U2 auxiliary factors (U2AFs) also bind the 3'ss and facilitate binding of the U2 snRNP. Next, the U4/U5/U6 tri-snRNP is recruited, and the fully assembled spliceosome splices the pre-mRNA via two sequential transesterification reactions.¹²⁵

Splice sites can be considered constitutive or alternative, depending on whether they are always or only sometimes used, respectively.^{125, 142} In either case, the same spliceosomal machinery catalyzes the splicing reaction. However, *trans*-acting splicing factors often mediate recruitment efficiency of the spliceosome to alternative splice sites.¹⁴² Such *trans*-acting factors are most often RNA-binding proteins that recognize sequence motifs that enhance/promote or silence/repress splicing.^{233, 234} The canonical splicing enhancers are the serine/arginine-rich proteins (SR proteins, such as SRSF2).¹⁴² In contrast, canonical splicing repressors are heterogeneous nuclear ribonucleoproteins (hnRNPs, such as hnRNP K).^{142, 235} However, context largely dictates whether a protein, even one from one of these canonical classes, will act as a repressor or enhancer of splicing.¹⁴²

Depending on where in a pre-mRNA a protein binds, as well as the capability of that protein to interact with other proteins, one protein may have differential splicing effects around a particular site.¹⁴² For example, hnRNP K binding to *RUNX1* RNA near the 3'ss of intron 5-6 could lead to competition with normal binding of U2AFs and/or the U2 snRNP. Indeed, in a rat cell line, hnRNP K competed with U2AF2 (also known as U2AF65) for binding to the poly(C) tract immediately upstream of 3' splice sites of several genes.⁵⁶ Depending, then, on the interaction between hnRNP K and the U2 snRNP, hnRNP K may act as a splicing repressor.

The role of splicing in hematologic malignancies has become increasingly recognized. Mutations in splicing factors such as *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2* have been identified in numerous leukemias, including AML.^{32, 34, 236, 237} Intriguingly, altered splicing events can occur in AML even in the absence of mutations in splicing machinery.¹⁴⁶ In one study of several tumor types, AML samples had the highest rate of alternative splicing events, even when only analyzing samples without splicing factor mutations.¹⁴⁶ Furthermore, dysregulation of the alternative splicing process can give rise to protein isoforms with functions resulting in oncogenic insults—even without concomitant genomic aberrations.^{147, 148} Thus, our understanding of these processes is imperative to reveal the underpinnings of leukemogenesis.

RUNX1 is a key transcription factor in normal and malignant hematopoiesis.⁸⁵ In leukemias, *RUNX1* is often mutated^{109, 116, 228}, translocated¹⁰⁴⁻¹⁰⁶, or aberrantly expressed.^{98, 229-231} Like the majority of protein-coding genes with more than one exon, *RUNX1* is spliced. Indeed, alternative splicing of *RUNX1* results in an isoform lacking the C-terminus transactivation domain, known as *RUNX1A*. Another alternatively spliced, less understood isoform of *RUNX1*, lacking an internal 64 amino acid residues (corresponding to exon 6), has been identified in mouse and human.^{88, 93, 149, 155} However, the mechanisms underlying the production and functional relevance of these isoforms have not been fully elucidated.

hnRNP K has been shown to influence splicing of several genes, including *RUNX1*.^{56, 58, 59, 149-151} However, it is unclear what influence, if any, hnRNP K overexpression has on splicing of *RUNX1*, particularly in the context of AML. Additionally, the physiologic relevance of the hnRNP K-*RUNX1* interaction has not been elucidated. In this chapter, we therefore aimed to functionally characterize the interaction between hnRNP K and *RUNX1*.

5.2 Materials/Methods

Viral production: HEK293T cells at 50% confluency in T75 flasks were transfected with 4 µg of transfer vector (see plasmids section) and 4 µg of packaging vector. For virus used to infect FLCs, the packing vector used was pCL-Eco (Addgene plasmid #12371).¹⁸³ For virus used to infect human cell lines, packaging was done with 4 µg of pCMV-dR8.2 (Addgene plasmid #8455)²³⁸, and 800 ng of pCMV-VSV-G (Addgene plasmid #8454)²³⁸. This DNA was combined with 12 µL of JetPrime DNA transfection reagent (Polyplus Transfection, New York, NY, USA). Fresh medium was supplied after 6-8 hours (stem cell medium without cytokines for FLCs; DMEM with 10% tetracycline-free FBS for human cell lines), and cells were incubated at 32°C for 48-72 hours. High-titer viral supernatant was collected and passed through a 0.45 µm filter.

Fetal liver cell (FLC) isolation: The University of Texas MD Anderson Cancer Center Animal Care and Use Committee approved all mouse experiments performed in these studies under protocols 0000787-RN01 and 0000787-RN02. Pregnant wildtype CD45.2+ C57/Bl6 females were euthanized by exposure to CO₂ at day 13.5 of gestation. Fetal livers were sterilely dissected and gently disrupted on a 70 µm filter to obtain a single-cell suspension. The collected cells were briefly subjected to a red blood cell (RBC) lysis with BD Pharm Lyse buffer (BD Biosciences, San Jose, CA, USA). After RBC lysis, cells were resuspended in stem cell medium at ~3x10⁶ cells/mL and incubated at 37°C overnight prior to retroviral transduction. Stem cell medium contained 37% DMEM (Corning Inc, Corning, NY, USA), 37% Iscove's modified Dulbecco's Medium (Corning Inc, Corning, NY, USA), 20% fetal bovine serum, 2% L-glutamine (200mM; Corning Inc, Corning, NY, USA), 100 U/mL penicillin/streptomycin (Sigma-Aldrich, St. Louis, MO, USA), 5x10⁻⁵ M 2-mercaptoethanol (Sigma-Aldrich, St. Louis, MO, USA), recombinant murine interleukin-3 (0.2 ng/mL), interleukin-6 (2 ng/mL), and stem cell factor (20 ng/mL; all cytokines were obtained from Stem Cell Technologies, Vancouver, Canada).

FLC transduction and sorting: Approximately 5×10^6 FLCs were resuspended in 2 mL of high-titer retroviral supernatant supplemented with 12 $\mu\text{g}/\text{mL}$ polybrene and fresh cytokines (IL-3, IL-6, and SCF). Plates were then spun at 600xg for 90 minutes at room temperature and incubated at 32°C for 48-72 hours. After transduction, cells with GFP were sorted for GFP positivity using the MoFlo Astrios cell sorter (Beckman Coulter, Brea, CA, USA) at the MD Anderson Cancer Center North Campus Flow Cytometry Core Facility.

RNA-Sequencing: RNA was extracted and purified from cells using Zymo Quick-RNA columns as per manufacturer's instructions (Zymo Research, Irvine, CA, USA). Barcoded, Illumina stranded total RNA libraries were prepared using the TruSeq Stranded Total RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). Briefly 250ng of DNase I treated total RNA was depleted of cytoplasmic and mitochondrial ribosomal RNA (rRNA) using Ribo-Zero Gold (Illumina, San Diego, CA, USA). After purification, the RNA was fragmented using divalent cations and double stranded cDNA was synthesized using random primers. The ends of the resulting double stranded cDNA fragments were repaired, 5'-phosphorylated, 3'-A tailed and Illumina-specific indexed adapters are were ligated. The products were purified and enriched by 12 cycles of PCR to create the final cDNA library. The libraries were quantified by qPCR and assessed for size distribution using the 4200 TapeStation High Sensitivity D1000 ScreenTape (Agilent Technologies, Santa Clara, CA, USA) then multiplexed, 3 libraries per lane and sequenced on the Illumina HiSeq4000 sequencer (Illumina, San Diego, CA, USA) using the 75 bp paired end format.

RNA-Seq analysis: Fastq files were pseudoaligned using Kallisto v0.44.0²³⁹ with 30 bootstrap samples to a transcriptome index based on the *Mus musculus* GRCm38 release (Ensembl). The resulting abundance data was further analyzed with Sleuth v0.30.0²⁴⁰ using models with covariates for both batch and condition. Gene-level

abundance estimates were calculated as the sum of transcripts per million (TPM) estimates of all transcripts mapped to a given gene. Wald tests were performed at a gene level for the “condition” covariate with a significance threshold of FDR < 10%.

qRT-PCR: RNA was extracted and purified from cell lines using Zymo Quick-RNA columns as per manufacturer’s instructions (Zymo Research, Irvine, CA, USA). Samples were subjected to in-column DNase treatment as per manufacturer’s instructions. Purified RNA was quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA). 1µg of RNA was reverse transcribed using iScript (BioRad, Hercules, CA, USA). *qRT-PCR* was performed using iTaq Univerval SYBR Green Supermix as per manufacturer instructions (BioRad, Hercules, CA, USA) using an ABI StepOnePlus Real Time PCR System. Primers used are listed in Table 4. Individual samples were assayed in triplicate. Calculations were performed using the Pfaffl method comparing expression changes between target genes and housekeeping control.¹⁵⁸

Primer	Sequence (5' to 3')
Human PPIA Forward	CCCACCGTGTCTTCGACATT
Human PPIA Reverse	GGACCCGTATGCTTTAGGATGA
Human RPLP0 Forward	CCTTCTCCTTTGGGCTGGTCATCCA
Human RPLP0 Reverse	CAGACACTGGCAACATTGCGGACAC
Human RUNX1 Ex5 Reverse	CCATCAAATCACAGTGGAT
Human RUNX1 Ghanem forward	GAAGTGGAAGAGGGAAAAGCTTCA
Human RUNX1 Ghanem reverse	GCACGTCCAGGTGAAATGCG
Human RUNX1 Ex3 forward (for total RUNX1)	CTGCTCCGTGCTGCCTAC
Human RUNX1 Ex4 reverse (for total RUNX1)	AGCCATCACAGTGACCAGAGT
Mouse Runx1 Ghanem Forward	CACTCTGACCATCACCGTCTT
Mouse Runx1 Ghanem Reverse	GGATCCCAGGTACTGGTAGGA
Mouse Runx1 Ghanem Total Forward	TGGGATCCATCACCTCTTCTT
Mouse Runx1 Ghanem Total reverse	ACGGCAGAGTAGGGAAGTGG

Table 4. Primers used in PCR. List of primers used in *qRT-PCR* and *RT-PCR*. All sequences are listed from 5' to 3'.

RT-PCR for splicing analysis: One μg of RNA was converted to cDNA using iScript (BioRad, Hercules, CA, USA) as per manufacturer's instructions. PCR reactions were performed using Apex buffer, 1.5mM MgCl_2 , 0.2mM of each dNTP, 0.5 units Taq polymerase, 0.5 μM forward primer, 0.5 μM reverse primer, and 100ng cDNA. Primers used are included in Table 4. PCR was run on a BioRad Thermocycler at 95°C x 3 minutes, followed by 35 cycles of 95°C for 1 minute, 60°C for 1 minute, and 72°C for 3 minutes, and finally 72°C for 5 minutes. After amplification, equal amounts of PCR products were run on a 2% agarose gel with ethidium bromide at 180V x 30 minutes, then visualized using a Syngene G:Box EF2 gel doc system with GENESys image capture software.

Sanger sequencing: DNA was purified from agarose gels using a gel purification kit (Invitrogen, Carlsbad, CA, USA) as per manufacturer's instructions. 200ng of purified DNA and 1pmol of each primer used in the PCR reaction were provided to the MDACC Sequencing and Microarray Core Facility. Sequencing was performed on an ABI 3730XL sequencer using BigDye terminator cycle sequencing chemistry with the forward primer used in the RT-PCR reaction. For validation, another sequencing run was performed with the reverse primer used in the RT-PCR reaction. Data analysis was provided as text files and chromatograms.

Plasmids and cloning:

hnRNP K mammalian cell expression plasmids: Full-length human hnRNP K cDNA was PCR amplified using cDNA isolated from 293T cells as a template. The PCR product was subsequently digested with restriction enzymes XhoI and EcoRI and ligated into the corresponding sites in the pcDNA c-flag vector (Addgene plasmid #20011).²⁴¹ hnRNP K domain deletions were amplified using the nested PCR technique²⁴², and subsequently ligated into the

XhoI/EcoRI restriction enzyme sites in the pcDNA c-flag vector. All of the above plasmids had an in-frame C-terminal flag tag to verify expression and for downstream applications.

hnRNP K viral expression plasmids: Full-length hnRNP K and its domain deletions were amplified using standard PCR techniques from the pcDNA c-flag plasmids, described above. These were then cloned into the XhoI/NotI sites in the all-in-one tetracycline inducible lentiviral vector TRE3G-ORF-P2A-eGFP-PGK-Tet3G-bsd (TLO2026, transOMIC Technologies, Alabama, USA). All of the viral plasmids had an in-frame C-terminal flag tag, and a P2A linker followed by eGFP. The viral selection marker in the plasmid was blasticidin. These viral plasmids were used to generate the tetracycline-inducible stable cell lines for overexpression of hnRNP K.

For fetal liver cell infection, retroviruses were made in an MSCV backbone and were modified from those described previously.¹⁸⁰ MSCV-AML1/ETO-IRES-GFP plasmid was obtained from Addgene (plasmid #60832).¹⁸⁰ The green fluorescent protein (GFP) coding sequence was replaced with firefly luciferase ORF obtained from a luciferase-pcDNA plasmid (Addgene plasmid #18964).¹⁸² For generation of empty vectors, AML1/ETO was excised using EcoRI and BamH1 and plasmid re-ligated. To generate MSCV-HNRNPK-IRES-GFP and MSCV-HNRNPK-IRES-Luciferase, AML1/ETO was replaced with HNRNPK that was PCR amplified from HEK 293T cells.

RUNX1 mammalian expression plasmids: Full-length human RUNX1(b) cDNA was PCR amplified using a RUNX1 expression plasmid as a template. The PCR product was subsequently digested with restriction enzymes XhoI and EcoRI and ligated into the corresponding sites in the pcDNA c-flag vector (Addgene plasmid #20011).²⁴¹ *RUNX1ΔEx6* was amplified using the nested PCR technique and subsequently ligated into the XhoI/EcoRI restriction enzyme sites in the pcDNA c-flag vector. Both plasmids had an in-frame C-terminal flag tag to verify expression and for downstream applications.

RUNX1 viral expression plasmids: Full-length *RUNX1(b)* and *RUNX1ΔEx6* were amplified using standard PCR techniques from the pcDNA c-flag plasmids, described above. These were then cloned into the XhoI/NotI sites in the all-in-one tetracycline inducible lentiviral vector TRE3G-ORF-P2A-eGFP-PGK-Tet3G-bsd (TLO2026, transOMIC Technologies, Alabama, USA). All of the viral plasmids had an in-frame C-terminal flag tag, a P2A linker followed by eGFP. The viral selection marker in the plasmid was blasticidin. These viral plasmids were used to make tetracycline-inducible stable cell lines for RUNX1 overexpression.

For infecting fetal liver cells, full-length *RUNX1(b)* and *RUNX1ΔEx6* were subcloned from the pcDNA c-flag plasmids into the MSCV-IRES-GFP vector (Addgene plasmid #20672).

RUNX1 luciferase reporter: An artificial oligo containing 13 consensus RUNX1 DNA binding sites (TGTGG) flanked by HindIII restriction enzyme sites, and its corresponding reverse complement pair, were synthesized. The two oligos were annealed and placed into HindIII sites in the PG13-Luc vector (Addgene plasmid #16442).²⁴³ p53 binding sites from the PG13-Luc plasmid were excised out during the HindIII digestion. The resulting RUNX1 reporter was referred to as RBG-Luc.

All plasmids were verified by sequencing. The primers used for cloning are listed in Table 5.

Transient transfections: 293T cells were plated at identical confluencies and transfected using equal amounts of DNA with JetPrime reagents as per manufacturer's instructions (Polyplus Transfection, New York, USA). All transfections were for 48 hours unless otherwise noted.

Generation of stable cell lines: Virus was produced as described above. Approximately 0.5×10^6 cells were resuspended in 2 mL of high-titer viral supernatant supplemented with 12 $\mu\text{g/mL}$ polybrene. Plates were then spun at 600xg for 90 minutes at room temperature and incubated at 32°C for 72 hours. After 72 hours, cells were spun down, viral supernatant removed, and cells were resuspended in RPMI 1640 with 10% tetracycline-free FBS. Appropriate selection antibiotics were added to media (2-6 $\mu\text{g/mL}$ of blasticidin or 1-2 $\mu\text{g/mL}$ of puromycin) and cells were placed at 37°C. When cells began to grow through antibiotic selection (3-15 days), an aliquot of cells was induced with doxycycline (0.2 $\mu\text{g/mL}$ for knockdown cell lines or 0.4 $\mu\text{g/mL}$ for overexpression cell lines) for 24-72 hours before evaluation of appropriate fluorescent protein expression by fluorescence microscopy was performed. If the appropriate fluorescent protein expression (GFP and/or RFP) was observed in >90% of cells in response to doxycycline, cells were utilized in downstream experiments. All cells were maintained in constant antibiotic selection and kept in stringently tetracycline-free conditions to avoid unintended gene induction. Prior to cell harvest for each experiment, GFP and/or RFP expression were visualized by fluorescence microscopy, and downstream assays only conducted if >90% of cells were appropriately positive.

Western blotting: Cells were homogenized in NP40 lysis buffer containing protease and phosphatase inhibitors (Millipore Sigma, Burlington, MA, USA). Soluble proteins were boiled in Laemmli buffer, resolved on an SDS-PAGE gel, and transferred to a PVDF membrane. Membranes were blocked with 5% milk for one hour at room temperature and incubated with primary antibody at 4°C overnight while rocking. Primary antibodies were hnRNP K (3C2,

1:1000, Abcam, Cambridge, MA, USA), RUNX1 (EPR3099, 1:1000, ab92336, Abcam, Cambridge, MA, USA), β -actin (AC-15, 1:2000, Santa Cruz Biotechnology, Dallas, TX, USA). Membranes were incubated with secondary antibody and antibody-protein interactions were visualized using enhanced chemiluminescence (GE Healthcare, Chicago, IL, USA) or BCIP/NBT color development substrate (VWR International, Radnor, PA, USA).

Luciferase assay: 293T cells were transiently transfected with luciferase-based reporter plasmids and expression plasmids using jetPRIME (Polyplus, New York, NY, USA). Human V5-tagged SIN3A pLX304 plasmid was purchased from Arizona State University Plasmid repository (clone HsCD00445676; Tempe, AZ, USA). Total DNA quantity was constant across all wells. 48 hours post-transfection, luciferase assay reagent was mixed in a 1:1 ratio with cell lysate in accordance with manufacturer's protocol (Luciferase Assay System kit, Promega, Madison, WI, USA;). Luciferase activity was measured with Synergy H4 Hybrid Reader (BioTek, Winooski, VT, USA). Transfection efficiency for each well was normalized using 62.5ng of a pCMV β -galactosidase plasmid, which was co-expressed in each experiment. All experiments were performed in at least triplicate.

Protein stability assays: 293T cells stably transduced with tetracycline-inducible constructs to overexpress *RUNX1*, either full-length or lacking exon 6. 400ng/mL doxycycline was added to cells prior to the addition of cycloheximide (10 μ M, Sigma-Aldrich, St. Louis, MO, USA) with or without MG-132 (10 μ M, SelleckChem, Houston, TX, USA) were then added to cells for 1-8 hours. Cells were collected and lysed in NP40 lysis buffer with protease and phosphatase inhibitors prior to western blot.

Colony formation assay: GFP sorted FLCs were cultured in quadruplicate wells of a 12-well plate in methylcellulose medium with cytokines IL-3, IL-6, erythropoietin, and stem cell factor (Methocult GF M3434, StemCell Technologies, Vancouver, Canada). 50,000 cells were plated

per well and colonies were counted after 7 days. Colonies were gently disrupted in PBS and cells counted manually with a hemacytometer using trypan blue dye exclusion, or subjected to cytospin or flow cytometry. Cytospins were stained with Wright-Giemsa.

Polysome fractionation assay: This assay was performed using a previously described protocol.²⁴⁴ 293T cells were transfected with a non-targeting control (siScramble; 5'-UAAGGCUAUGAAGAGAUAC-3') or a pool of siHNRNPK (5'-3': GAGCGCAUUAUUGAGUAUCA, GAUCUUGGUGGACCUAUUA, GGUCAGCGGAUUAACAAA, GUCGGGAGCUUCGAUCAA) (Dharmacon, Lafayette, CO, USA) using JetPrime (Polyplus, New York, NY, USA) as per manufacturer's instructions. 24 hours after transfection, cells were split and harvested 72 hours later. 10 minutes prior to harvest, 100mg/mL cycloheximide was added. Cells were collected using trypsin, pelleted, and washed with cold PBS supplemented with cycloheximide. Lysis was performed with polysome extraction buffer (20mM Tris-HCl pH 7.5, 100mM KCl, 5mM MgCl₂, and 0.5% Nonidet P-40) with added cycloheximide, protease inhibitors, and RNA inhibitors. Lysate was centrifuged to clear insoluble material, then loaded onto a 0-50% sucrose gradient and spun in an ultracentrifuge at 35000 rpm at 4°C for 3.75 hours. The gradient was then fractionated into 28 fractions with a piston gradient fractionator (BioComp, Fredericton, NB, Canada). Polysome profiles were created using absorbance readings at 260nm and aligned by manually identifying peaks, and scaling positions with linear transformation. The minimum absorbance value from each sample was subtracted out in order to account for background absorbance. RNA from each fraction was isolated using Trizol, reverse transcribed, and utilized for downstream qPCR analysis as described above. Further details are available elsewhere.⁸³

5.3 Results

hnRNP K globally affects splicing

Given the stringent binding of hnRNP K to RNA, and the increasing emphasis on aberrant RNA splicing in myeloid malignancies, we sought to evaluate whether hnRNP K could influence splicing patterns. To this end, we performed RNA-Seq on FLCs overexpressing hnRNP K compared to empty vector controls. In addition to notable gene expression differences between groups (Figure 30), hnRNP K influenced splicing of numerous transcripts (Figure 31).

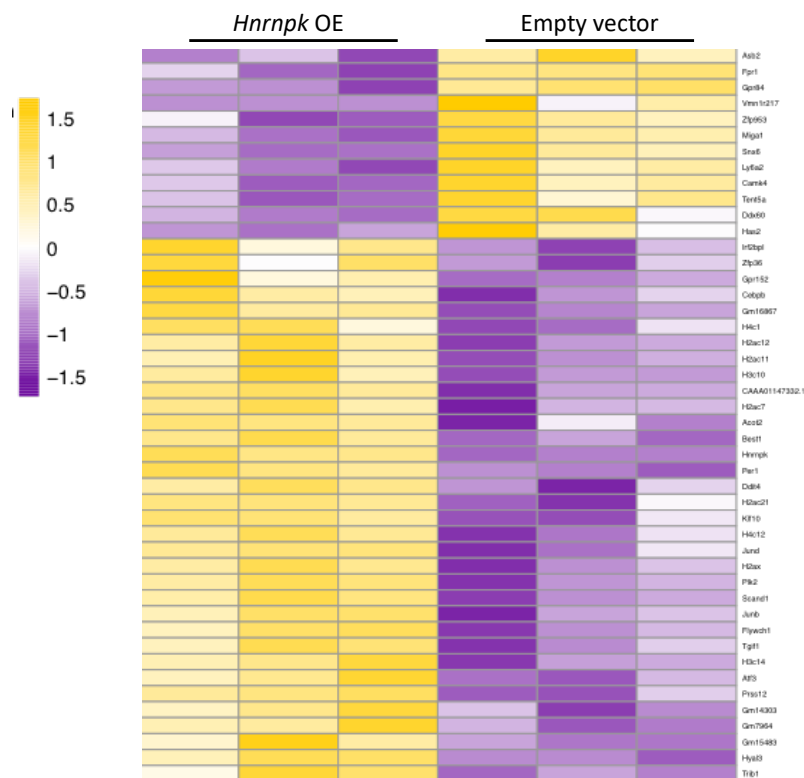


Figure 30. *Hnrrpk* overexpression results in differential gene expression in murine FLCs.

Heatmap of the most differentially expressed genes in FLCs overexpressing hnRNP K (left) compared to empty vector controls (right). Each column represents a biologic replicate.

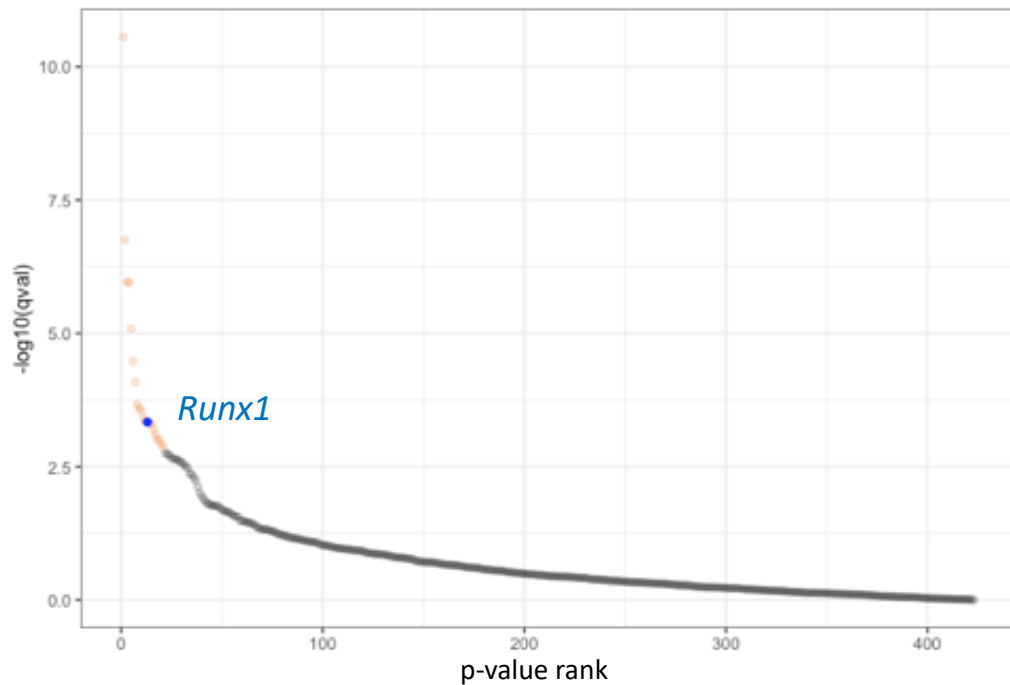


Figure 31. *Hnrnpk* overexpression alters splicing in FLCs. Plot of differentially spliced transcripts in *Hnrnpk* overexpressing FLCs compared with empty vector controls. Orange dots represent transcripts that were differentially spliced in a statistically significant manner. Blue dot indicates *Runx1*.

hnRNP K alters splicing of Runx1

One of the most compelling differentially spliced transcripts from our RNA-Seq data was *Runx1* (Figure 31). Since the *Runx1* transcript has an hnRNP K binding site at the junction of intron 5 and exon 6, near the 3' splice site (Figures 26, 28, 29), we checked for *Runx1* splice variants with alterations in exon 6. Samples overexpressing hnRNP K had a statistically significant increase in the amount of *Runx1* exon 6 skipping compared to empty vector controls (Figures 32A, B). We refer to this isoform as *Runx1* Δ Ex6. While both isoforms of *Runx1* were present in all samples, *Runx1* Δ Ex6 was more abundant in samples overexpressing *Hnrnpk*. Therefore, we sought to more specifically evaluate the impact of hnRNP K in this area. To validate these RNA-Seq findings, we performed *Runx1* RT-PCR with primers spanning exons 5-

7. In FLCs that were infected with *Hnrnpk* overexpressing virus, a PCR product corresponding to the predicted size of *Runx1* Δ Ex6 was dramatically enriched compared to controls (Figure 32C).

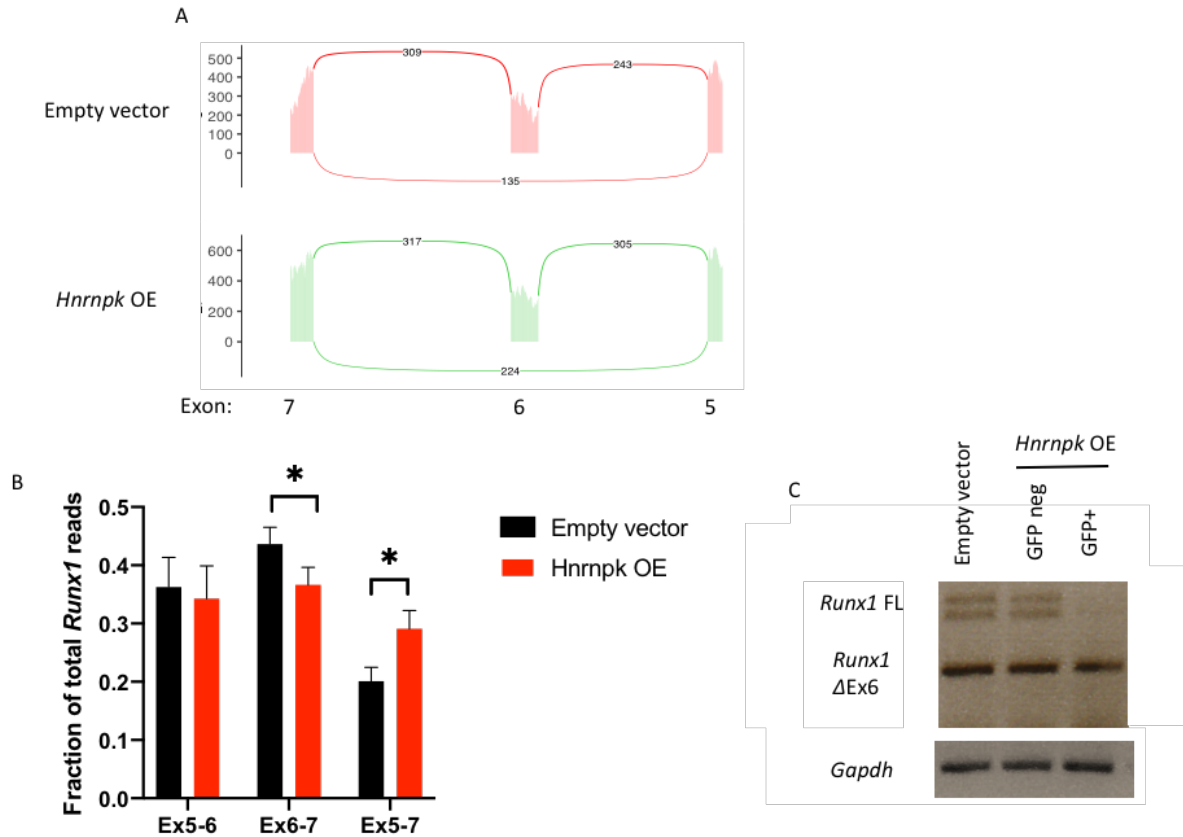


Figure 32. *Runx1* is differentially spliced in hnRNP K overexpressing FLCs. A. Sashimi plot of *Runx1* exons 5-7 from a representative empty vector sample (top) and an *Hnrnpk* overexpressing sample (bottom). Y-axis shows the number of reads. Numbers above lines between exons represent the number of reads spanning those exon junctions. B. Data from four empty vector samples and four *Hnrnpk* overexpressing samples represented as a fraction of total *Runx1* reads. C. *Runx1* RT-PCR from FLCs infected with empty vector or *Hnrnpk*. As an additional control, FLCs exposed to *Hnrnpk* OE virus but that were not infected (GFP neg) were included. GFP+ indicates FLCs that successfully integrated the *Hnrnpk* construct and were FACS sorted for GFP positivity. Empty vector controls were also sorted for GFP positivity. *Gapdh* is shown as a loading control.

Exclusion of RUNX1 exon 6 occurs in an hnRNP K-dependent manner

Given that exclusion of exon 6 (appearance of *Runx1ΔEx6*) occurred in FLCs in an hnRNP K-dependent manner, and since the hnRNP K binding site in intron 5-6 is largely conserved between mouse and human, we hypothesized that *RUNX1* splicing surrounding exon 6 would be similarly affected in human cells. To this end, we generated several human AML cell lines with doxycycline-inducible hnRNP K overexpression (Figure 33A). We then utilized PCR primers to amplify *RUNX1* with a forward primer in exon 5, and a reverse primer in exon 7 (Figure 33B).¹⁴⁹ With this strategy, we were able to distinguish clear alterations in *RUNX1* splicing patterns surrounding this exon that were dependent on hnRNP K expression levels (Figure 33C). This is consistent with the findings from our RNA sequencing of murine fetal liver cells, where hnRNP K overexpression leads to an enrichment of *Runx1ΔEx6* (Figure 32).

To confirm that the smaller band (160bp) visualized in RT-PCR was indeed the expected product lacking exon 6, we performed Sanger sequencing on the PCR products. From these results, it was evident that the larger band in these gels included exon 6, whereas the smaller band completely lacked exon 6, leaving a clean junction between exons 5 and 7 (Figure 33D-H). Therefore, cells overexpressing hnRNP K have an increase in *RUNX1ΔEx6*.

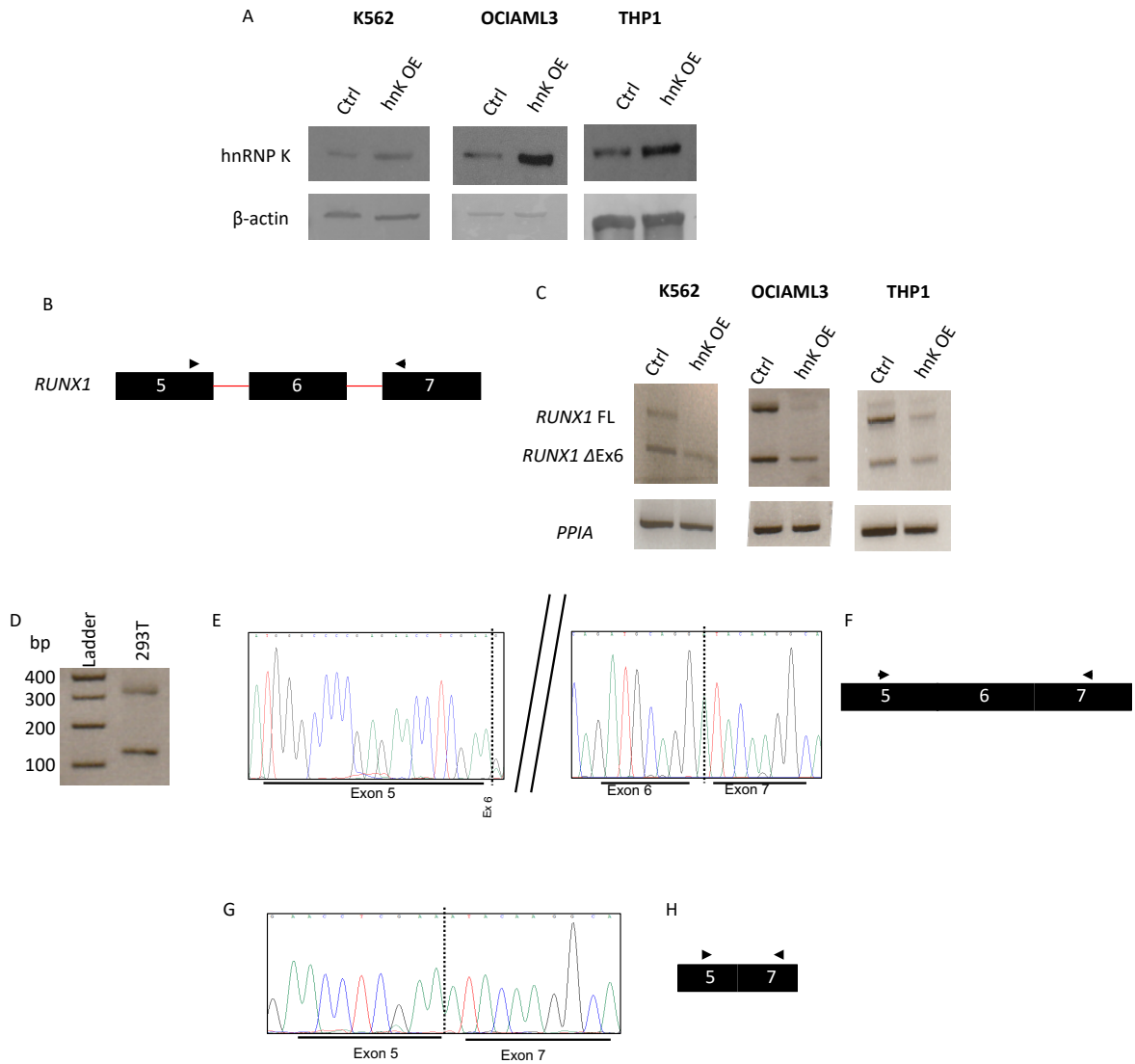


Figure 33. hnRNP K overexpression leads to exclusion of *RUNX1* exon 6. A. Western blot of stable human AML cell lines with tetracycline-inducible empty vector control (ctrl) or hnRNP K overexpression (hnK OE). Cells were treated with doxycycline for 24 hours prior to harvest. B. PCR strategy for *RUNX1* amplification. C. RT-PCR from the indicated human AML cell lines treated with doxycycline for 24 hours to overexpress an empty vector (Ctrl) or hnRNP K (hnK OE). *PPIA* serves as a loading control. D. *RUNX1* RT-PCR from 293T cells. Ladder with associated number of base pairs (bp) is clearly labeled. E. Sanger sequencing of 352 bp band from gel in panel A. Double lines are placed in lieu of the entirety of exon 6 sequencing results. F. Schematic of the exons present in the PCR product containing exons 5-7, herein referred to

as *RUNX1 FL* (full length). Arrows indicate the location of RT-PCR primers. G. Sanger sequencing of 160 bp band from gel in panel A. H. Schematic of the exons present in the PCR product lacking exon 6. Arrows indicate the location of RT-PCR primers.

hnRNP K knockdown favors inclusion of RUNX1 exon 6

In both human and mouse, hnRNP K overexpression led to an accumulation of *RUNX1 Δ Ex6* (Figures 32, 33). To further assess to the functional relationship between *RUNX1* splicing and hnRNP K expression, we performed a series of knockdown and rescue experiments in human cell lines. In K562s, three *shHNRNPK* constructs were used to create stable tetracycline-inducible lines. Indeed, knockdown of hnRNP K was evident at the RNA and protein level 24-hours after doxycycline induction (Figures 34A, B). Importantly, RT-PCR revealed a decrease in *RUNX1 Δ Ex6* in the context of hnRNP K knockdown (Figure 34C). A concomitant increase in full length *RUNX1* was also evident. This is consistent with recently published data in this cell line.¹⁴⁹

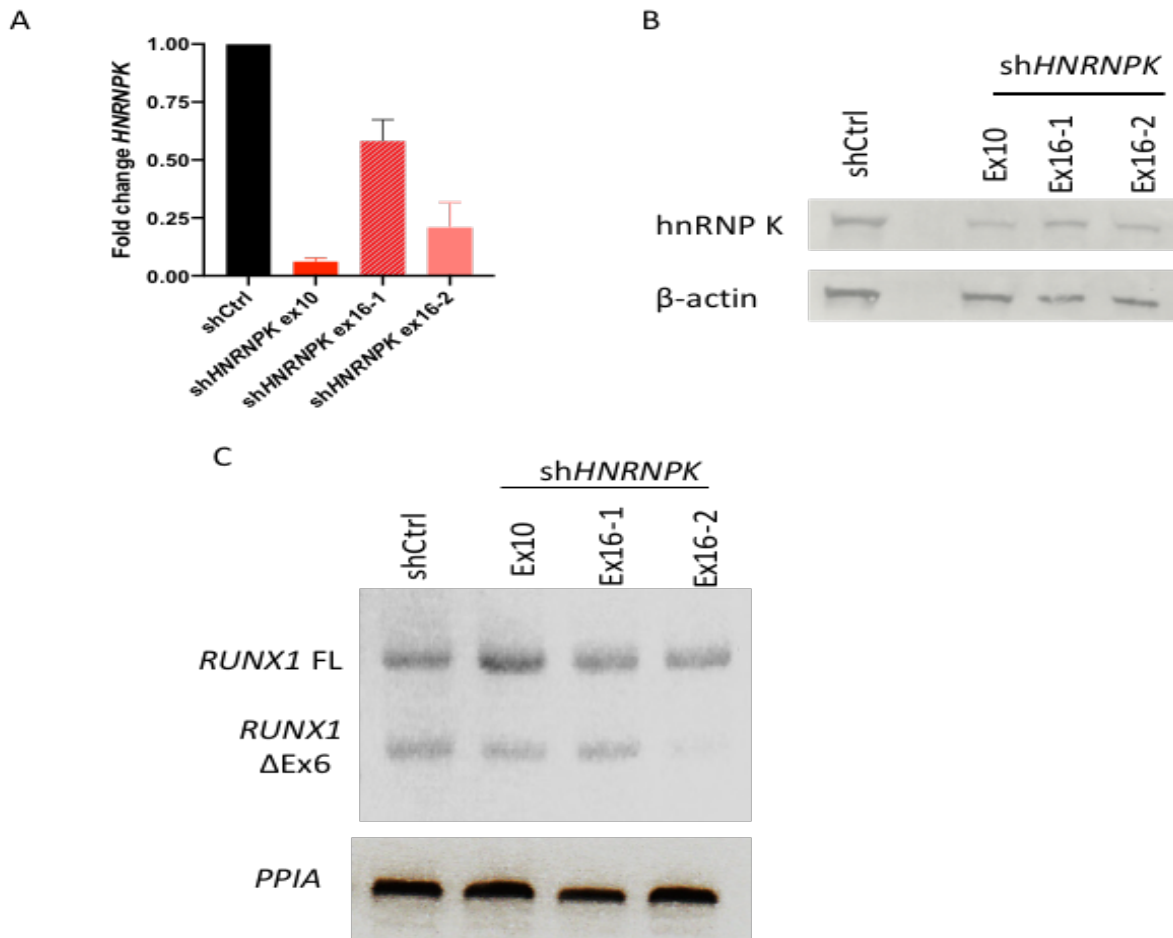


Figure 34. Knockdown of hnRNP K in K562 increases *RUNX1* exon 6 inclusion. A. *HNRNPK* qRT-PCR in K562 stably infected with the indicated constructs treated with doxycycline for 24 hours. Ex10 construct targets *HNRNPK* in exon 10, while ex16-1 and ex16-2 constructs target the 3' UTR of *HNRNPK*. B. Western blot of the cell lines used in A induced with doxycycline for 24 hours. C. *RUNX1* RT-PCR of the same cell lines. *PPIA* is used as a loading control.

Add-back of hnRNP K to an shHNRNPK cell line rescues RUNX1 splicing alterations

Since the shHNRNPK Ex16-2 construct induced significant knockdown by targeting the 3' UTR, we used this line to create another stable cell line with exogenous hnRNP K lacking this targeted region. Therefore, we could use these cells as a tool to 'rescue' phenotypes associated with a decrease in hnRNP K expression. As expected, in K562 cells, this resulted in hnRNP K

protein expression that approximated that of wildtype controls (Figure 35A). Remarkably, *RUNX1* RT-PCR in cells with rescued hnRNP K expression also mimicked wildtype (Figure 35B). This finding was echoed in OCIAML3 cells (Figure 35C). These findings indicate that alterations in hnRNP K expression are sufficient to alter splicing of *RUNX1* exon 6.

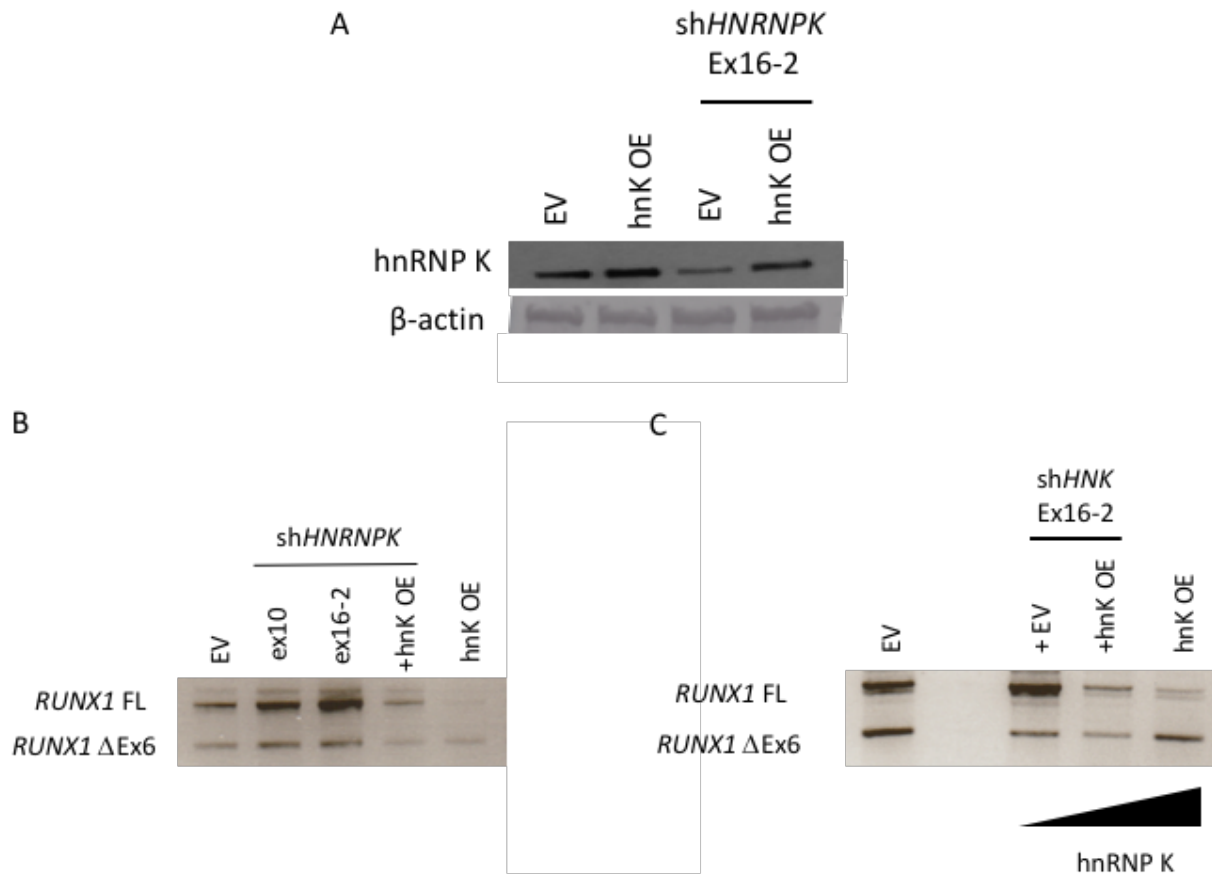


Figure 35. Addback of hnRNP K to shHNRNPK cells rescues *RUNX1* splicing. A. Western blot of K562 cells infected with empty vector (EV), hnRNP K overexpression (hnK OE), or either of those in addition to shHNRNPK ex16-2. Cells were treated with doxycycline for 24 hours prior to harvest. B. *RUNX1* RT-PCR in K562 cells. C. *RUNX1* RT-PCR in OCIAML3 cells. Samples are ordered from lowest to highest hnRNP K expression, indicated by the black triangle.

RUNX1 Δ Ex6 is translated

Since the known function of RUNX1 is as transcription factor, we sought to evaluate whether *RUNX1 Δ Ex6* had a corresponding protein product with similar function. To assess whether *RUNX1 Δ Ex6* was translated, we performed western blots of cell lines overexpressing hnRNP K. These revealed the presence of a smaller protein product (Figures 36A, B). Compared to empty vector controls, this smaller protein product (~45 kDa), was more abundant in cell lines overexpressing hnRNP K. Since the *RUNX1 Δ Ex6* isoform is 192bp (encoding 64 amino acids), shorter than the full-length form (Figure 33), this size difference lends credence to the conclusion that this smaller protein product is *RUNX1 Δ Ex6*. Strikingly, in cells with decreased hnRNP K, this isoform was undetectable (Figure 36C).

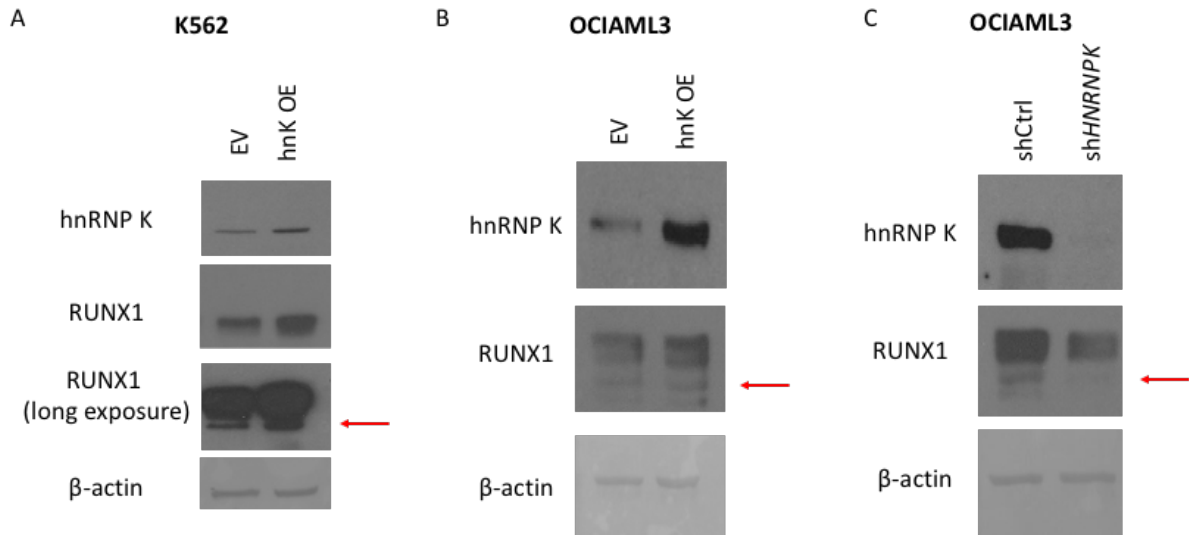


Figure 36. A smaller RUNX1 isoform corresponds with hnRNP K expression at the protein level. Western blots of A. K562 or B. OCIAML3 cells stably infected with empty vector (EV) or hnRNP K overexpression (hnK OE) treated with doxycycline for 24 hours. C. Western blot of OCIAML3 cells stably infected with shCtrl or sh*HNRNPK* ex10 and treated with doxycycline for 24 hours. Red arrows indicate the smaller isoform of RUNX1 that appears at ~45 kDa.

RUNX1 Δ Ex6 is present in AML patient samples

In Chapter 2, we identified that hnRNP K is overexpressed in AML. Given the data in this chapter that overexpression of hnRNP K leads to an increase in *RUNX1 Δ Ex6*, we examined AML patient samples for the presence of this isoform. While full-length *RUNX1* was abundant in healthy bone marrow and peripheral blood samples, expression of this isoform was drastically reduced in nearly every case of AML (Figure 37). Instead, *RUNX1 Δ Ex6* was the dominant isoform expressed. This is a crucial finding, as the association of *RUNX1 Δ Ex6* with AML supports the clinical relevance of this entity.

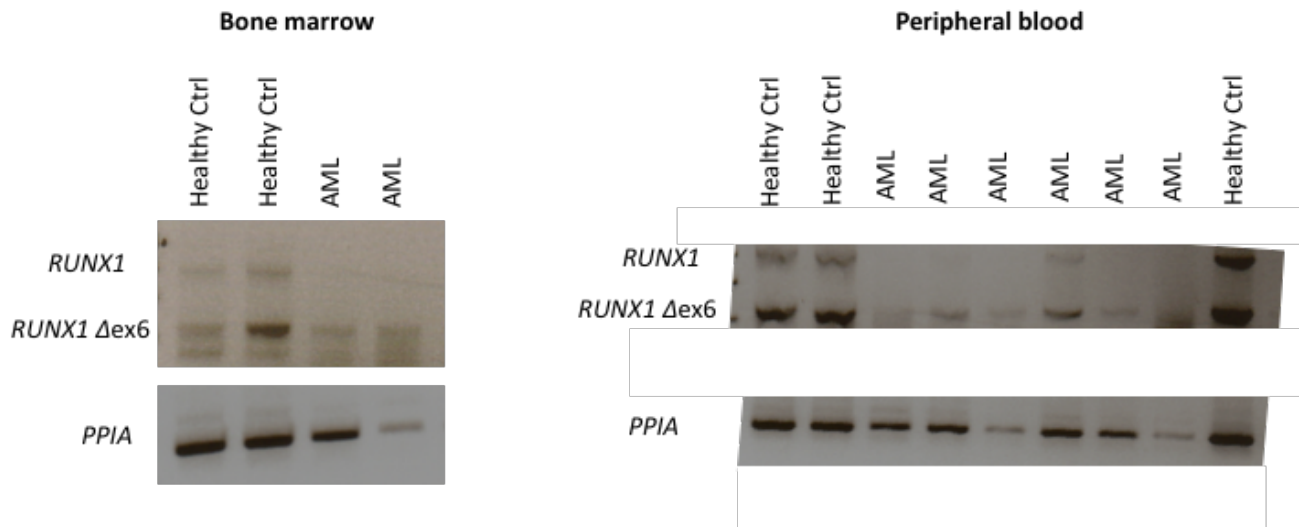


Figure 37. *RUNX1 Δ Ex6* is present in AML patient samples. *RUNX1* RT-PCR from bone marrow (left) and peripheral blood (right) of healthy human control donors or AML patient samples. *PPIA* is used as a loading control.

Runx1 Δ Ex6 is present in mice transplanted with hnRNP K-overexpressing FLCs

Thus far, our data have shown that hnRNP K overexpression leads to an enrichment of *Runx1 Δ Ex6* in murine FLCs (RNA-Seq and RT-PCR; Figures 32, 33) and *RUNX1 Δ Ex6* in human AML cell lines (RT-PCR, western blot; Figures 34, 36). In addition, patients with AML also

express *RUNX1ΔEx6* (RT-PCR; Figure 37). To further evaluate how hnRNP K overexpression influences this splicing *in vivo*, we evaluated Runx1 expression in the mouse model described in Chapter 3. In spleen and bone marrow of mice transplanted with FLCs overexpressing hnRNP K, there was a substantial enrichment of *Runx1ΔEx6* at the expense of full-length *Runx1* (Figure 38A). At the protein level, before transplantation (Figure 38B) and in the spleen after transplantation (Figure 38C), higher protein expression of Runx1ΔEx6 was evident. As expected, these findings corresponded nicely with hnRNP K overexpression.

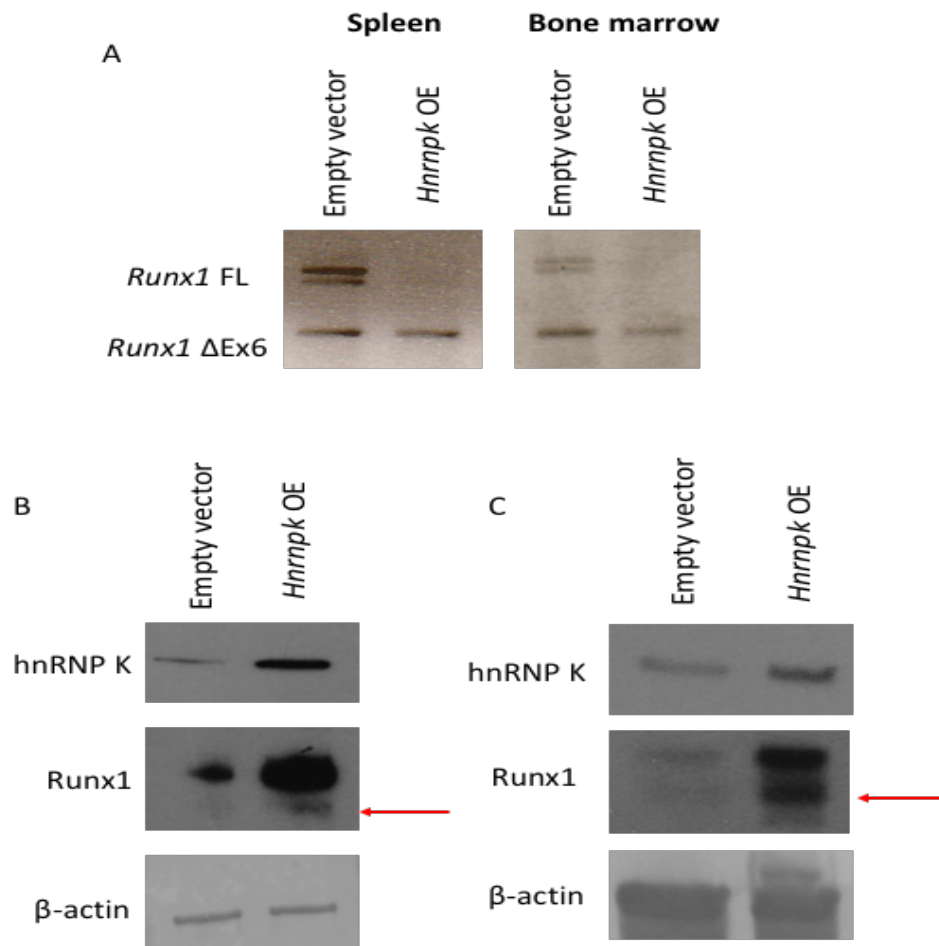


Figure 38. FLCs overexpressing hnRNP K have more Runx1ΔEx6. A. *Runx1* RT-PCR from spleen (left) or bone marrow (right) of mice transplanted with empty vector control FLCs or hnRNP K-overexpressing FLCs. B-C. Western blots of FLCs pre-transplant (B) or from spleen

of representative mice (C) transplanted with FLCs overexpressing empty vector or hnRNP K. Red arrow denotes Runx1 Δ Ex6.

RUNX1 Δ Ex6 is more stable than full-length RUNX1

To begin evaluating the functional consequences of RUNX1 Δ Ex6, we compared the stability of this protein isoform to that of full-length RUNX1. Adding cycloheximide to eukaryotic cells impairs ribosomal translocation, thereby inhibiting cytosolic translation.^{245, 246} Since *de novo* protein synthesis is dramatically hindered in the presence of cycloheximide, western blotting can be used to visualize protein degradation over time. Proteins that degrade more quickly (decrease in intensity on western blot) are said to be less stable than proteins that maintain expression for longer periods of time in the presence of cycloheximide.²⁴⁷ Cycloheximide chase assays showed that RUNX1 Δ Ex6 was substantially more stable than its full length counterpart (Figure 39A). Consistent with this observation, addition of MG132, an inhibitor of proteasomal degradation, stabilized full-length RUNX1, but had no effect on RUNX1 Δ Ex6 (Figure 39B), indicating that the proteasomal pathway is responsible for degradation of full-length RUNX1. While RUNX1 Δ Ex6 is present in relatively low quantities in cells, its long half-life may allow for biologic consequences despite low expression. These data are consistent with reports that exon 6 contains ubiquitination sites that target RUNX1 for proteasomal degradation.²⁴⁸

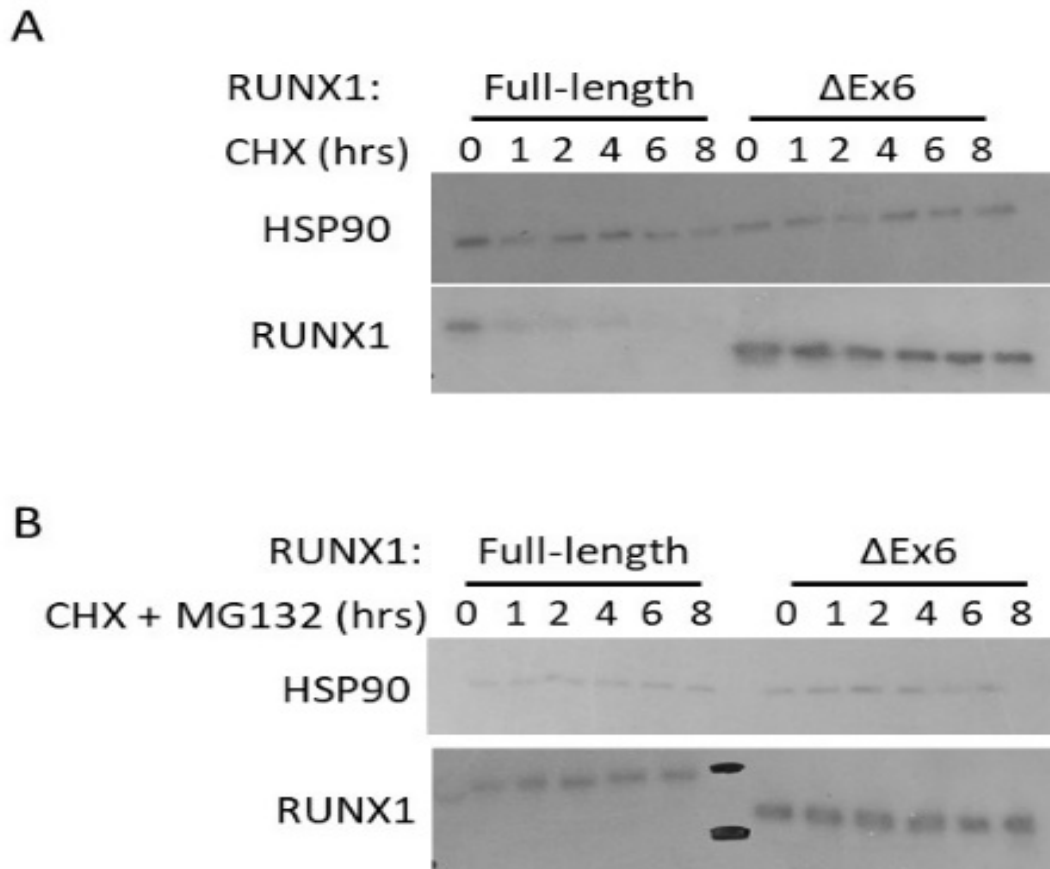


Figure 39. RUNX1ΔEx6 is more stable than full-length RUNX1. Western blot of 293T cells with stably integrated tetracycline-inducible *RUNX1* (full-length or ΔEx6) were treated with doxycycline for 24 hours prior to addition of A. cycloheximide (CHX) or B. CHX and MG-132 for up to 8 hours. HSP90 is used as a loading control. Molecular weight ladder is marked in black marker in panel B.

RUNX1ΔEx6 has differential transcriptional activity compared to full length RUNX1

The predominantly known function of RUNX1 is as a transcription factor. To further investigate whether RUNX1ΔEx6 has biologic function, we examined its transcriptional capabilities using luciferase reporter assays with an artificial *RUNX1* reporter containing 13 consensus RUNX1 binding sites. We observed that cells transfected with full-length RUNX1 dramatically repressed luciferase activity; however, cells with RUNX1ΔEx6 had diminished

capability to repress luciferase activity (Figure 40B). Addition of SIN3A, a transcriptional co-repressor that binds RUNX1 within a region encoded by exon 6, also demonstrated differential impact on full-length RUNX1 compared to RUNX1 Δ Ex6 (Figure 40C). This suggests that the lack of exon 6 interferes with the ability of SIN3A to interact with RUNX1, and supports the notion that RUNX1 Δ Ex6 has altered transcriptional activity compared to full-length RUNX1.

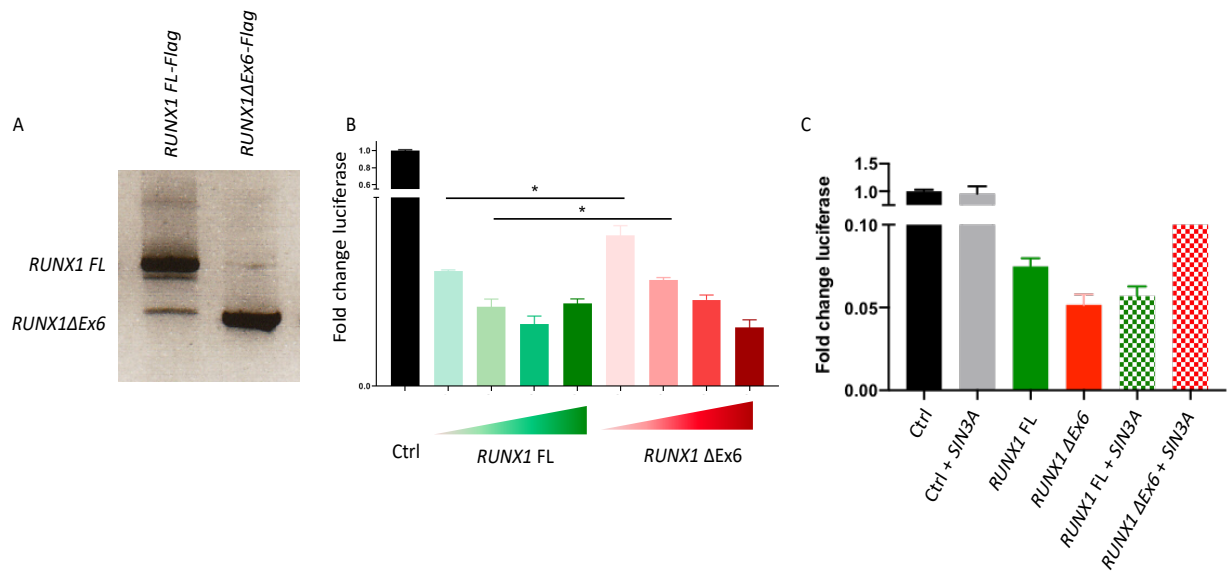


Figure 40. Full-length RUNX1 and RUNX1 Δ Ex6 have differential transcriptional activity. A. *RUNX1* RT-PCR in 293T cells transfected with *RUNX1* full length (FL) or Δ Ex6 overexpression plasmids B. Luciferase assays performed in 293T cells transfected with luciferase reporter construct and empty vector (Ctrl) or increasing concentrations of full length *RUNX1* or *RUNX1* Δ Ex6. *RUNX1* concentrations range from 0.0625 μ g to 0.5 μ g. C. Luciferase assays performed in 293T cells transfected with luciferase reporter construct and empty vector (Ctrl) and 0.0625 μ g of *RUNX1* with or without exon 6 as well as 0.5 μ g of *SIN3A*.

The KH3 domain of hnRNP K mediates *RUNX1* splicing

The previous data suggest that *RUNX1* Δ Ex6 has different biologic function than its full-length counterpart. Since the presence of *RUNX1* Δ Ex6 is hnRNP K-dependent, we sought to examine whether a particular domain of hnRNP K was responsible for this splicing. To this end, we transfected 293T cells with constructs lacking sequential domains of *HNRNPK*. When *RUNX1* splicing was evaluated in these lines, cells overexpressing any form of *HNRNPK* had a near complete elimination of the full-length *RUNX1* in favor of *RUNX1* Δ Ex6, except those lacking the KNS or KH3 domain (Figure 41A). In these cells transfected with *HNRNPK* Δ KNS or *HNRNPK* Δ KH3, splicing of *RUNX1* was nearly identical to cells transfected with an empty vector control.

To evaluate this phenomenon in another cell line, we made K562 cells stably transduced with tetracycline-inducible overexpression of hnRNP K with deletions of the KH1, KH2, KI, or KH3 domain. When overexpressing these constructs, only those lacking KH3 showed a reversion to splicing like that of the empty vector control (Figure 41B). The data from these two cell lines thus strongly suggest that KH3 is the domain of hnRNP K that mediates inclusion of *RUNX1* exon 6.

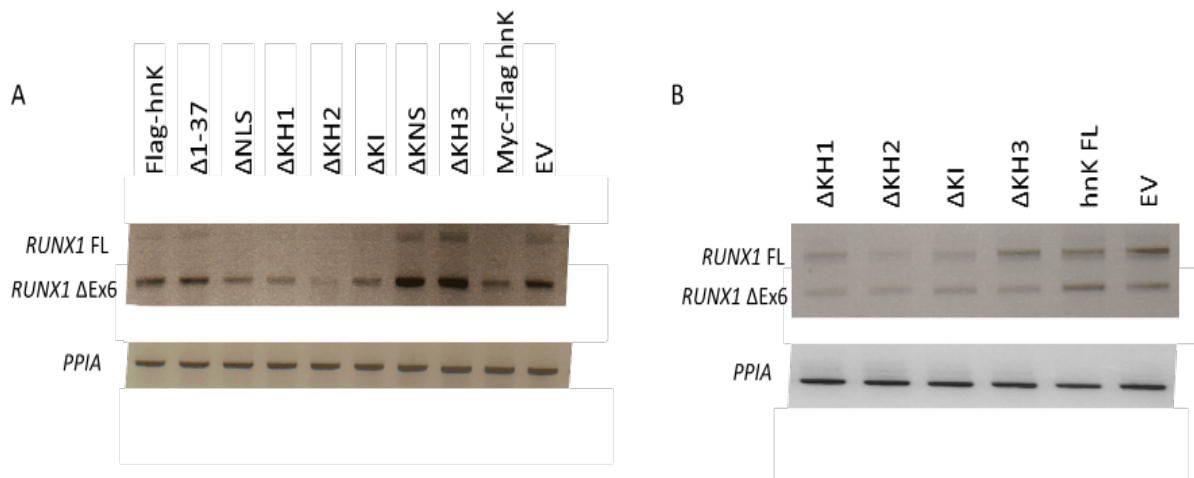


Figure 41. KH3 domain of hnRNP K mediates *RUNX1* exon 6 splicing. RT-PCR of *RUNX1* from A. 293T cells transfected for 24 hours with the indicated flag-tagged hnRNP K constructs

or B. K562 cells with tetracycline-inducible expression of the indicated hnRNP K constructs. K562 cells were treated with doxycycline for 24 hours before RNA was harvested. Deletion constructs are arranged in the order in which the functional domains occur within hnRNP K from amino to carboxy terminus of hnRNP K. EV: empty vector control. *PPIA* is used as a loading control.

The KH3 domain of hnRNP K is required for hnRNP K-overexpressing phenotypes in vitro

Since it appears that the KH3 domain of hnRNP K is largely responsible for mediating *RUNX1* splicing, we wanted to evaluate whether this domain contributed to the ability of hnRNP K-overexpressing FLCs to form colonies *in vitro*. Compared to those overexpressing full-length hnRNP K, FLCs overexpressing hnRNP K without KH3 had substantially decreased colony formation (Figure 42). Strikingly, cells with hnRNP K Δ KH3 formed even fewer colonies *in vitro* than empty vector controls. Together, these data indicate that hnRNP K-mediated increase in Runx1 Δ Ex6 is critical to the *in vitro* phenotypes associated with hnRNP K overexpression.

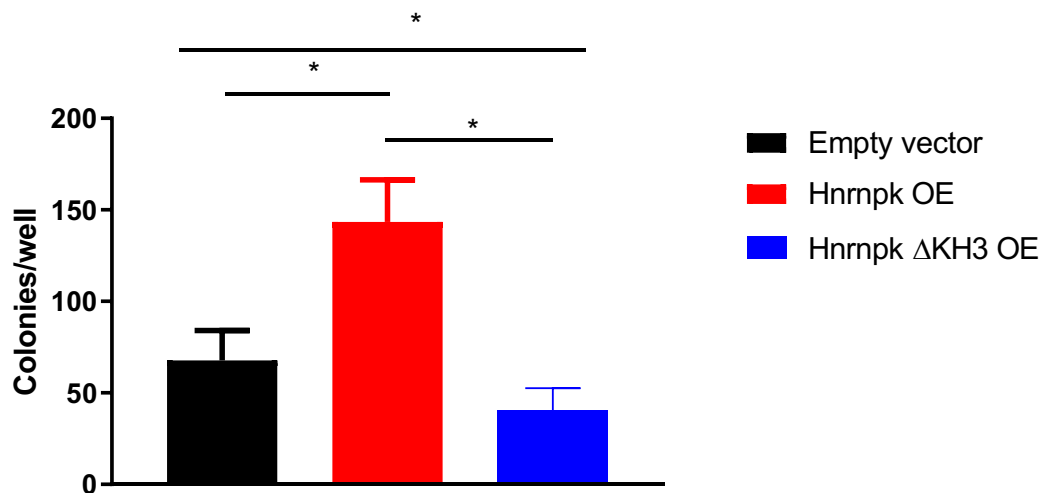


Figure 42. The KH3 domain of hnRNP K is required for an *in vitro* phenotype associated with hnRNP K overexpression. Bar graph of the number of hematopoietic colonies formed per well from FLCs stably transduced with empty vector (black), full-length hnRNP K (red), or hnRNP K lacking the KH3 domain (blue). Cells were sorted for GFP expression to ensure stable

integration of the viral construct. Note: This figure includes empty vector and hnRNP K OE data points that were included in figure 12.

hnRNP K influences RUNX1 and global translation

In addition to the increase in RUNX1 Δ Ex6 observed with hnRNP K overexpression, it was evident that total RUNX1 protein expression also increased in this setting (Figures 36, 38). Given the binding of hnRNP K to the 5' UTR of *RUNX1* identified in Chapter 4 (Figures 26, 28, 29), and the interaction between hnRNP K and ribosomal subunits identified in Chapter 4 (Figure 23), we sought to evaluate whether hnRNP K influenced translation of *RUNX1*. The notion that hnRNP K regulates RUNX1 expression at a post-transcriptional/translational level is supported by the observation that total *RUNX1* RNA is not increased in hnRNP K-overexpressing cells (Figure 43A).

To more directly address this question, we performed polysome fractionation assays. In these experiments, ribosomal subunits and their associated mRNAs are segregated based on density in a sucrose gradient. Fractions of this gradient are inspected by a spectrophotometer to assess absorbance at 260nm. By tracing these absorbances over each fraction, we can produce the beloved “polysome trace”, representing an overview of a global translational profile for a given population of cells. Furthermore, RNA from these individual fractions can be isolated and analyzed by qRT-PCR to measure translation of an individual RNA species.

Knockdown of hnRNP K in 293T cells resulted in a marked change in the global polysome trace (Figures 43B, C). Cells with decreased hnRNP K showed a substantial spike in the 80S (monosome) fraction compared to controls. Polysomes are traditionally deemed as the translationally active components, as this is where the majority of new peptide bonds are formed.²⁴⁹⁻²⁵¹ Thus, as monosomes accumulate in cells, translation generally becomes less active at a global scale, likely due to mRNA loading difficulties. Therefore, decreased hnRNP K results in less active translation.

To specifically evaluate the translational status of *RUNX1*, we performed qRT-PCR on the individual fractions comprising the polysome trace. Compared to cells transfected with a scrambled control, cells with decreased hnRNP K had less *RUNX1* in the 80S monosome fraction (Figure 43D). Given that there was not a compensatory increase of *RUNX1* in the low-molecular weight and high molecular weight polysome fractions, this suggests that *RUNX1* may fail to load into complete ribosomal subunits in the absence of hnRNP K.

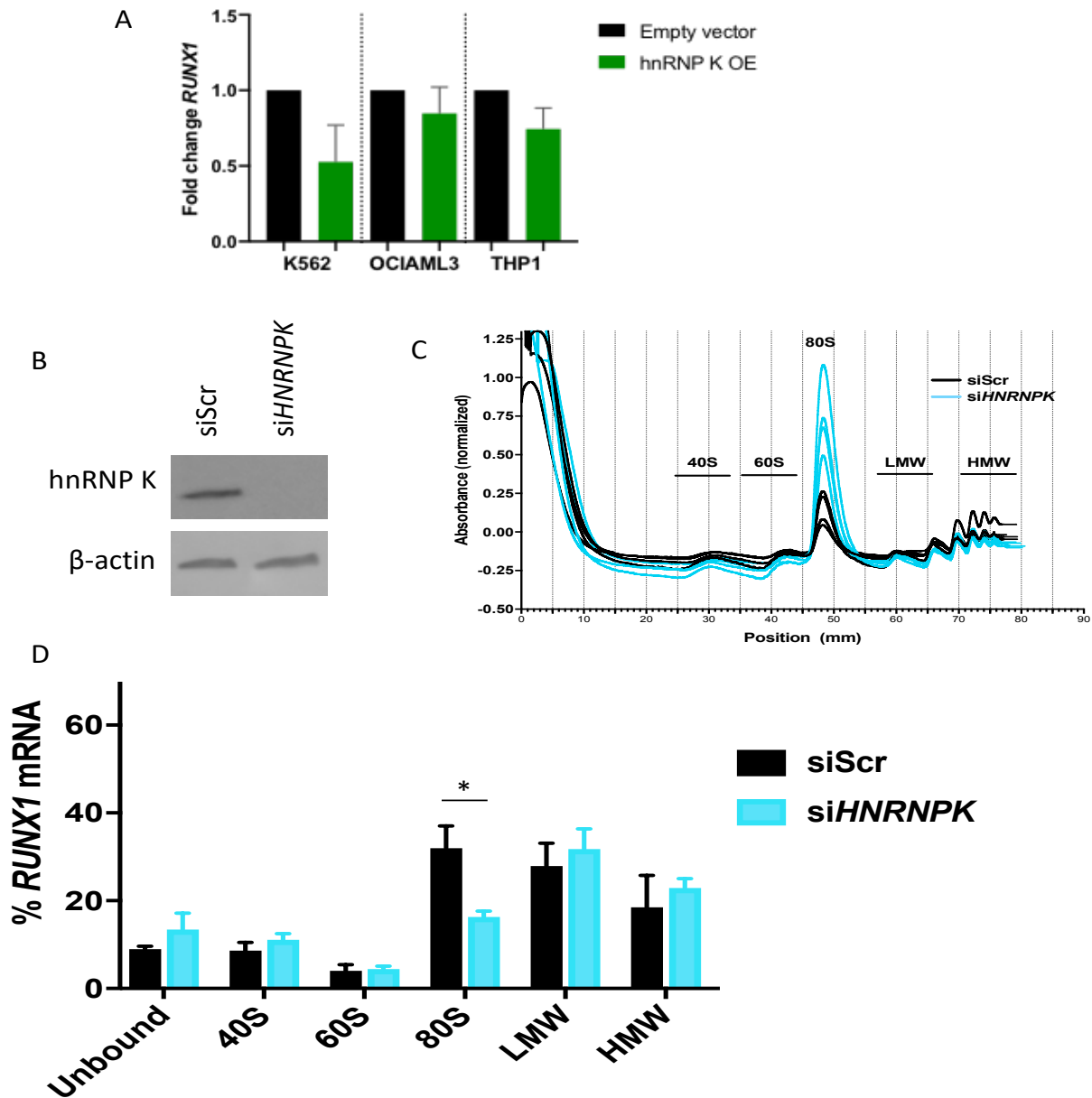


Figure 43. hnRNP K alterations affect global translation and that of *RUNX1*. A. qRT-PCR for total *RUNX1* in the indicated cell lines overexpressing hnRNP K or empty vector. Cells were treated with doxycycline for 24 hours prior to harvest. Primers in this reaction spanned exons 3-4 of *RUNX1*, thus measuring the total level of transcript independent of spliced products. B. Western blot of 293T cells transfected with siScramble control (siScr) or siHNRNPK for 72 hours. C. Polysome trace of 293T cells transfected as in panel A. X-axis represents the fraction of sucrose gradient collected as a position of the length of the physical gradient (nm). Y-axis

represents the absorbance of the fraction at 260nm normalized to a no protein control. Fractions are labeled according to the accepted location of peaks for 40S, 60S, and 80S ribosome fractions. LMW= low-molecular weight; HMW= high-molecular weight polysome fractions. Traces from four individual experiments are superimposed in this graph. D. *RUNX1* qRT-PCR from fractions collected in panel B. Portions of this figure were modified with permission from Gallardo & Malaney et al., JNCI, 2019.⁸³

5.4 Discussion

In this chapter, we identified that hnRNP K overexpression altered genome-wide RNA splicing. Strikingly, *Runx1* emerged as one of the most differentially spliced transcripts in this setting (Figure 31). This observation, along with the data in Chapter 4 that hnRNP K stringently and specifically bound human and mouse *RUNX1* at intron 5-6, and the role of RUNX1 as a key transcription factor in leukemia, we specifically honed in on the interaction between hnRNP K and *RUNX1*.

The impact of hnRNP K on splicing of individual transcripts has been described.^{56, 58, 59, 150, 151} However, to our knowledge, genome-wide analyses of splicing alterations occurring in the context of differential hnRNP K expression has not been described. While we focused on *RUNX1*, these global analyses certainly provide insights for future inspection of the role of hnRNP K in hematopoietic splicing.

Our observations that hnRNP K reduction leads to inclusion of exon 6 is in line with what others have briefly identified.^{56, 149} However, these prior observations have been made in the context of hnRNP K knockdown. To our knowledge, we are the first to show that overexpression of hnRNP K results in *RUNX1* exon 6 exclusion. This was a critical finding given our observation that this isoform is readily identifiable in AML patient samples (Figure 37). Unlike other alternatively spliced isoforms of *RUNX1*, *RUNX1A* and *RUNX1B*, the isoforms we studied are expected to be identical in their C terminal domains. In addition, we are confident that what we refer to as *RUNX1* Δ Ex6 is not merely *RUNX1A*, as *RUNX1A*: 1) lacks exon 7, where our reverse primer for splicing PCR is placed and 2) has a predicted molecular weight of ~27kDa—substantially less than the observed ~45kDa for *RUNX1* Δ Ex6.

Functional consequences due to loss of a single exon are biologically plausible. The amino acid residues encoded in exon 6 of *RUNX1* comprise a negative regulatory region, where SIN3A, a corepressor, binds.²⁵² This region also limits *RUNX1*-DNA interactions²⁵³. We showed that *RUNX1* Δ Ex6 had differential transcriptional activity compared to its full-length counterpart

(Figure 40). Our data suggest that this is due in part to altered cooperativity with SIN3A (Figure 40B). Though we did not formally test the DNA-binding affinity of RUNX1 Δ Ex6 compared to full-length RUNX1, one group identified murine Runx1 Δ Ex6 as having increased DNA binding affinity.⁹³ Consistent with that observation, another group described RUNX1 Δ Ex6 as having dominant negative activity over RUNX1 when evaluated in a human ovarian cancer cell line.¹⁵⁵ More complete evaluation of the transcriptional differences between full-length RUNX1 and RUNX1 Δ Ex6 in AML is ongoing in our laboratory.

In addition to altered transcriptional activity, we also observed increased stability of RUNX1 Δ Ex6 compared to full-length RUNX1 (Figure 39). This is consistent with observations of murine Runx1 lacking this exon.⁸⁸ Since exon 6 contains two ubiquitination sites²⁴⁸, the absence of these sites in RUNX1 Δ Ex6 may explain this apparent increase in protein stability.

To gather more insight into the mechanism of this hnRNP K-dependent *RUNX1* exon 6 splicing, we narrowed down domains of hnRNP K that appeared to be critical for this phenomenon. Absence of the KH3 domain had a striking effect on the inclusion of exon 6, and reverted splicing to near wildtype levels (Figure 41). In support of this observation, early descriptions of hnRNP K described poly(C) binding as mediated largely by KH3.¹⁵³ However, other groups have concluded that all KH domains play a role in poly(C) binding.⁷⁴

Deletion of the KH3 domain in hnRNP K largely rescued *RUNX1* splicing (Figure 41) and also abrogates an *in vitro* phenotype associated with hnRNP K overexpression (Figure 42). Of note, the KH3 domain of hnRNP K can bind ssDNA acid as a single domain²⁵⁴, and is the only portion of the protein in which a crystal structure has been determined (in complex with ssDNA).²⁵⁵ This could be of particular interest from a therapeutic standpoint for several reasons. First, given the myriad of biologic functions of hnRNP K, completely eliminating this protein would likely be extremely toxic. Indeed, *Hnmpk* knockout mice are not viable.⁵³ Second, hnRNP K has multiple functional domains, so understanding which domain to target could expedite drug discovery while also mitigating potential toxicity. Finally, the existing crystal structure of the KH3

domain may expedite drug development, as this can be used for computer-based predictive drug discovery models and can also be used to structurally optimize compounds identified in cell-based assays.

Our studies did not evaluate other splicing factors in the context of hnRNP K overexpression. It has been shown that hnRNP K competes with U2AF2 (also known as U2AF65) for binding at poly(C) tracts near 3' splice sites, like that near the *Runx1* intron 5-6 junction.⁵⁶ Little is known about how hnRNP K itself regulates splicing factors, though this is certainly a topic of great interest given the impact of hnRNP K on splicing alterations.

Given the vast array of hnRNP K-involved biologic processes, it seems unlikely that hnRNP K overexpression is oncogenic by means of a single mechanism. To begin evaluating these other mechanisms, we briefly evaluated hnRNP K's translational properties. Indeed, our polysome studies indicated that hnRNP K alterations globally affect translation (Figure 43). Specifically, this results in disrupted translation of *RUNX1*. We were not able to evaluate translation of full-length *RUNX1* compared to *RUNX1* Δ Ex6, but this would certainly be of interest. Given the hnRNP K binding to the 5' UTR of *RUNX1* (Figures 26, 28, 29), we anticipate that there would be no difference between these isoforms. It is also notable that while hnRNP K clearly bound the 5' UTR of human *RUNX1* (Figure 28, 29), this hnRNP K binding site was absent in murine *Runx1*. However, hnRNP K overexpression still resulted in increased protein expression of total Runx1 in mouse (Figure 38). The mechanisms underlying this observation are unclear, but may be due to hnRNP K's global modulation of translation.

Together, our studies indicate that hnRNP K mediates splicing and translation of *RUNX1*, and overexpression of hnRNP K results in an enrichment of *RUNX1* Δ Ex6 in mouse and (wo)man.

Chapter 6

Discussion

Treatment of acute myeloid leukemia (AML) is evolving at an unprecedented pace due to development of drugs targeting actionable genetic lesions. This therapeutic renaissance has dramatically improved outcomes for many patients with AML. However, many individuals do not harbor these targetable alterations, or relapse by various mechanisms. Therefore, it is imperative to identify novel mechanisms controlling the fundamental biologic processes underlying leukemogenesis and to develop therapies that target these vital nodes of dysregulated proliferation, survival, and/or differentiation. To this end, our laboratory has adopted a multistep approach to identify novel alterations driving hematologic malignancies. Accordingly, we have identified hnRNP K, an RNA-binding protein, as a novel driver of these diseases through its regulation of critical oncogenes required for cell growth, survival, and differentiation. This body of work focuses on the oncogenic role of hnRNP K in AML.

In Chapter 2, we identified overexpression of hnRNP K in acute myeloid leukemia and observed that this overexpression was associated with inferior outcomes, including decreased remission durations and overall survival (Figure 5). These data allude to an oncogenic role for hnRNP K when overexpressed.

Consequently, we identified overexpression of hnRNP K as a driving event underlying myeloproliferation in an *in vivo* mouse model in Chapter 3. Our data indicate that elevated hnRNP K expression is pathologic; however, the mechanisms underlying the observed overexpression of hnRNP K are still unclear. Many cases of AML harbored an additional copy of the *HNRNPK* locus identified by FISH (Figure 10). These additional copies were not part of any known chromosomes, indicating that the extra copies may exist as sSMCs. Beyond the *HNRNPK* locus, the contents of these so-called sSMCs were not identified. While these cases also had increased hnRNP K protein expression, it remains unclear whether the *HNRNPK* locus on the sSMC is transcriptionally active. Identification of the regulatory elements in these sSMCs may shed light on mechanisms of endogenous *HNRNPK* regulation in diseased states.

It is also possible that increased hnRNP K expression may be due to *trans* effects—where alterations at another genomic location (e.g., mutated or deleted transcription factors, altered epigenetic regulators, etc.) influence expression of a gene at a distant locus.²⁵⁶ One group recently reported that in AML cases with del(9q) (usually associated with *HNRNPK* haploinsufficiency), those with co-occurring mutations in the transcription factor *CEBPA* had near normal *HNRNPK* RNA expression; del(9q) cases with unmutated *CEBPA* had low *HNRNPK* expression.²⁵⁷ Given the necessity of hnRNP K at an organismal level⁵³, as well as the array of tumor types associated with its overexpression⁷⁵⁻⁸², identifying the regulation of this oncogene would be worthy of future pursuit. Indeed, several high-throughput methods now exist to evaluate possible *trans* effects on a more global scale.^{258, 259}

To evaluate the consequences of hnRNP K overexpression *in vivo*, we developed a mouse model in Chapter 3. These mice had reduced survival and developed myeloproliferative phenotypes. This model utilized retroviral transduction of fetal liver cells transplanted into recipient mice. During the course of the current studies, we also developed transgenic models of hnRNP K overexpression using both constitutive and conditional expression in the hematopoietic system. While these mice exhibited gross hematopoietic abnormalities, they failed to maintain reliable overexpression of the hnRNP K protein, thereby confounding the role of hnRNP K in disease pathogenesis in these systems. In contrast, overexpression of hnRNP K in fetal liver cells was substantially more consistent. Whether this is due to a technical aspect of stable viral integration or is an effect secondary to altering cells during a different stage of organismal development remains unclear. This once again highlights a need to understand regulation of hnRNP K expression *in vivo*.

In our mouse model of hnRNP K overexpression, the disease phenotypes were exclusively myeloid in nature (Chapter 3). While these mice developed phenotypes substantially similar to those seen in *Hnrnpk*^{+/-} mice, a percentage of *Hnrnpk*^{+/-} animals developed lymphoid malignancies.⁵³ Further supporting a role for hnRNP K in lymphoid development and

lymphomagenesis, mice overexpressing hnRNP K in B-cells develop overt lymphomas.⁸³ These observations demonstrate that altered hnRNP K expression, in either direction, can influence hematopoiesis. As *Hnrnpk* null mice die *in utero*⁵³, evaluation of hematopoiesis in these embryos may elucidate the most fundamental roles of hnRNP K in blood development. Interestingly, children with germline *HNRNPK* mutations, presumed to be inactivating insults, do not have grossly observable hematopoietic defects.^{38, 167, 260}

Several large-scale global analyses were performed to elucidate the oncogenic basis of hnRNP K overexpression in myeloid disease. Consequently, we identified that the oncogenicity of hnRNP K may be attributed to its RNA-binding activities. hnRNP K stringently and specifically binds RNA, leading to altered splicing of genes such as *RUNX1*. We further showed that the KH3 domain of hnRNP K was largely responsible for mediating this splicing alteration (Figure 41), and that deletion of KH3 abrogated the effects of hnRNP K overexpression (Figure 42). Previous reports have concluded that the RNA-binding function of hnRNP K is mediated through cooperative activity of its three KH domains, and binding is abrogated when any domain is disrupted.^{74, 217} Therefore, we anticipated that loss of any KH domain would alter these splicing effects. We instead identified KH3 as the key domain mediating *RUNX1* splicing. This finding was particularly compelling, as it suggests that these KH domains have varied roles in nucleic-acid directed processes. Furthermore, these findings place the KH3 domain as being critical in the phenotypes described in this work.

The observation that hnRNP K overexpression results in altered splicing of *RUNX1*, enriching for an isoform with altered transcriptional activity (*RUNX1*ΔEx6), is a captivating finding, as disrupted *RUNX1* transcriptional programs have been shown to contribute to leukemogenesis.^{261, 262} In *RUNX1*ΔEx6, the DNA-binding domain of *RUNX1* is predicted to remain intact, since the Runt-homology domain (RHD) is upstream of exon 6.²⁶³ However, the interaction with transcriptional co-factors and the regulation of this isoform could feasibly contribute to leukemogenesis. Indeed, expression of only the *RUNX1* RHD, which lacks many

of the regulatory elements of *RUNX1*, is sufficient to cause myelodysplastic syndrome in mice.²⁶¹ The lack of exon 6 described in our studies may result in broadly altered *RUNX1*-mediated transcription. Evaluation of the transcriptome landscape of cells enriched for *RUNX1*ΔEx6 will be useful in clarifying this point. To this end, our laboratory is pursuing RNA-sequencing as well as chromatin immunoprecipitation followed by sequencing (ChIP-Seq) of *RUNX1*ΔEx6-overexpressing cells.

Well-characterized alterations in *RUNX1* transcriptional programs are perhaps best exemplified by those observed in core-binding factor acute myeloid leukemia (CBF-AML). Many cases of CBF-AML harbor t(8;21)¹⁵, which encodes a fusion protein known as *RUNX1-RUNX1T1* (also called *AML1-ETO* or *RUNX1-CBFA2T1*)¹⁶. While globally, *RUNX1-RUNX1T1* binds to similar genomic locations as *RUNX1*, vast transcriptional alterations are observed.^{264, 265}

The breakpoint within *RUNX1* involved in t(8;21) most commonly occurs in the intron spanning exons 5 and 6.^{266, 267} However, these breakpoints are almost invariably located 5' to the hnRNP K binding site, though rare instances of *RUNX1-RUNX1T1* containing *RUNX1* exon 6 have been reported.²⁶⁸ Thus, hnRNP K is unlikely to be directly involved in alternative splicing of the *RUNX1* portion of this gene product. However, we have not evaluated the possibility that hnRNP K alters splicing of *RUNX1T1*, though preliminary analysis did not identify putative hnRNP K binding sites in this gene. Given that several splice variants of *RUNX1-RUNX1T1* have been reported²⁶⁹⁻²⁷¹, future studies in this area may be worthwhile.

Our initial evaluation has determined that the hnRNP K binding site in the 5' UTR of *RUNX1* is, however, maintained in the most common forms of t(8;21), suggesting that hnRNP K may be involved in *RUNX1-RUNX1T1*-associated leukemogenesis. Interestingly, in cases of t(8;21) AML, deletions of chromosome 9q have been observed as a common secondary abnormality²⁷¹⁻²⁷³, further alluding to a role for hnRNP K in this disease. To begin evaluating this scenario, we have added overexpression of *RUNX1-RUNX1T1* to our hnRNP K-overexpressing

fetal liver cells prior to transplant in recipient mice. These mice develop frank transplantable leukemia, though the relative impact of hnRNP K in this setting is not yet clear. Continuation of these studies to evaluate the interaction between hnRNP K and RUNX1-RUNX1T1 may provide valuable insight into this disease.

Given the data presented in Chapter 5 that hnRNP K alters splicing of not only *RUNX1*, but also other genes involved in leukemia, it would be of great interest to assess the mutational status of splicing factor genes in these cases and others that overexpress hnRNP K. While mutations in *SF3B1*, *SRSF2*, and *U2AF1* are mutually exclusive events in cancers, including AML^{143, 274, 275}, identification of hnRNP K overexpression as yet another mutually exclusive event would strongly indicate that a major mechanism of hnRNP K's oncogenicity stems from its deleterious effects on alternative splicing. Furthermore, comparing splicing alterations in AML patients with hnRNP K overexpression to splicing alterations in patients harboring splicing factor mutations could also shed light on recurrently altered splicing events that may underlie leukemogenic transformation.

In light of the vast array of cellular processes in which hnRNP K appears to be involved, the possibility that hnRNP K exerts oncogenic functions through mechanisms beyond altered splicing of *RUNX1* must be acknowledged. Consistent with this sentiment, mice exclusively expressing *Runx1* Δ Ex6 exhibit aberrant hematopoiesis, but do not develop leukemia.¹⁴⁹ In our previous work evaluating hnRNP K-overexpression in the setting of B-cell lymphomas, disease was driven largely via hnRNP K's interactions with *MYC*, where it stabilized the *MYC* transcript and promoted c-MYC translation.⁸³ While we did not extensively evaluate the hnRNP K-*MYC* interaction in the current studies, it remains possible that *MYC* alterations could contribute to the observed phenotypes. Indeed, our collaborators have described c-MYC overexpression as an important prognostic factor in AML.²⁷⁶ Furthermore, hnRNP K overexpression resulted in global gene expression and splicing alterations—underscoring the notion that altered *RUNX1* splicing may only represent a portion of the oncogenic facets of hnRNP K overexpression.

To provide a therapeutic output for our findings, work is currently underway in our laboratory to screen compounds that kill leukemia cells in an hnRNP K-dependent manner. As splicing inhibitors have been used in early clinical trials^{277, 278}, understanding a specific context for using these drugs is beneficial. These experiments will evaluate whether hnRNP K-overexpressing AML cells—harboring vastly altered splicing—are more sensitive to splicing inhibitors than cells with average hnRNP K expression. Given that hnRNP K is itself a splicing inhibitor, effects seen with these drugs may represent a synthetically lethal relationship²⁷⁹.

Given the oncogenic properties of hnRNP K when overexpressed (Chapter 3, Gallardo & Malaney et al., 2019⁸³), it is logical that targeting this protein directly may have therapeutic benefit. While no specific inhibitors of hnRNP K are currently available, few compounds have been reported to modify hnRNP K-dependent processes.²⁸⁰ Since the RNA-binding properties of hnRNP K appear to be most central to its oncogenic functions—whether mediating splicing (*RUNX1*; Chapter 5), or mRNA stability and translation (*RUNX1*, Chapter 5; *MYC*⁸³)—disrupting this interaction is a viable approach for drug development. To this end, our laboratory is currently utilizing fluorescence anisotropy assays (Figure 28) as screening tools to evaluate the capability of small molecule compounds to disrupt hnRNP K-RNA interactions. Presently, several lead compounds have emerged that appear to effectively interrupt these interfaces. Ongoing work in developing a drug to inhibit hnRNP K-RNA interactions is an exciting venture, as patients with AML (Chapter 2) and DLBCL⁸³ overexpressing hnRNP K have shortened overall survival.

Compounds that inhibit hnRNP K-RNA interactions may also find substantial utility outside of oncology. For example, replication of Influenza A virus requires the splicing activity of hnRNP K; depletion of hnRNP K reduces the pathogenicity of this ubiquitous virus.^{59, 154} In human immunodeficiency virus-1 (HIV-1), hnRNP K acts as an inhibitory splicing factor.²⁸¹ hnRNP K has been implicated in viral pathogenesis of numerous other RNA viruses, including enterovirus 71 (EV71)²⁸², dengue²⁸³, chikungunya²⁸⁴, sindbis²⁸⁵, hepatitis B (HBV)²⁸⁶, and hepatitis C (HCV)²⁸⁷. Interestingly, some of these viruses, such as HBV and HCV, have known

roles in oncogenic transformation.^{288, 289} Therefore, inhibiting hnRNP K's RNA-directed processes may also be beneficial in virally-induced cancers. Interestingly, hnRNP K has been implicated in the life cycles of several DNA viruses, as well. These include herpes simplex virus-1 (HSV1)²⁹⁰, Epstein-Barr virus (EBV)²⁹¹, cytomegalovirus (CMV)²⁹², and African swine fever virus²⁹³. Indeed, EBV can cause lymphoma²⁹⁴, thereby indirectly linking hnRNP K to hematopoietic oncogenesis. However, in the case of human papillomavirus (HPV), an oncogenic DNA virus, hnRNP K appears to inhibit translation of viral proteins, suggesting it may have anti-viral activity in this context.⁶⁵ This emphasizes that hnRNP K's activities can vary according to the particular cellular context.

In support of the notion that hnRNP K can directly promote cellular growth programs in tumor cells as well as promote viral infection, hnRNP K was the first described human pan-granzyme substrate.²⁹⁵ This apparent dire desire to cleave and destroy this protein supports claims that hnRNP K is often pathogenic—particularly in high amounts. Indeed, knockdown of hnRNP K sensitizes tumor cells to death mediated by cytotoxic lymphocytes.²⁹⁵

Altered hnRNP K may directly or indirectly result in immunologic consequences that may predispose organisms to infection and/or inflammatory milieus. As described above, hnRNP K may influence a cell's propensity to viral infection—which can itself be oncogenic. Of note, hnRNP K plays a role in several immunological processes. For example, hnRNP K is required for the DNA cleavage involved in B-cell somatic hypermutation and class switch recombination.²⁹⁶ While these processes are inhibited with depletion of hnRNP K, the impact of hnRNP K overexpression has not been defined. In T-cells, hnRNP K has been identified as an important target of ERK signaling and subsequent IL-2 production.²⁹⁷ Again, to our knowledge, hnRNP K overexpression has not been studied in this context. In uninduced macrophages, hnRNP K has been shown to translationally repress central downstream components of TLR4 signaling, acting to basally repress secretion of inflammatory cytokines.²⁹⁸ In the context of hnRNP K overexpression, it is feasible that such a repressive effect may be exaggerated, and

thus greater stimuli required to achieve an appropriate inflammatory response. We did not evaluate the role of hnRNP K overexpression on the functions of these immune cells in our mouse models. Thus, it is possible that the myeloproliferative phenotypes we observed were secondary to a vastly dysregulated immune system. Evaluating the functional capabilities of component immune cells in these mice may help clarify the sincerity of this possibility. Additionally, future studies involving infection challenges or in-depth analysis of the cytokine levels in these mice would be of interest.

Understanding the biology underlying hematopoietic disease, including AML, is what fundamentally allows for development of effective therapeutic modalities. Findings from our studies have substantially advanced our knowledge regarding the clinical relevance and *in vivo* functions of hnRNP K and its interactions with *RUNX1*. Additionally, these studies have led to the identification and understanding of mechanisms that drive hnRNP K-mediated leukemogenesis and have determined how hnRNP K overexpression cooperates with *RUNX1*, one of the most critical transcription factors in hematopoiesis. Given the overexpression of hnRNP K in AML (Chapter 2), its ability to cause myeloid disease in mice (Chapter 3), and alter *RUNX1* splicing (Chapters 4, 5), we conclude that hnRNP K is an oncogene in AML that functions in part via its interaction with and splicing regulation of *RUNX1*.

References

1. Yamamoto JF and Goodman MT. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. *Cancer Causes Control* 2008; 19: 379-390. 2007/12/08. DOI: 10.1007/s10552-007-9097-2.
2. Siegel RL, Miller KD and Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015; 65: 5-29. 2015/01/07. DOI: 10.3322/caac.21254.
3. Juliusson G, Antunovic P, Derolf A, Lehmann S, Mollgard L, Stockelberg D, Tidefelt U, Wahlin A and Hoglund M. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood* 2009; 113: 4179-4187. DOI: 10.1182/blood-2008-07-172007.
4. Creutzig U, Buchner T, Sauerland MC, Zimmermann M, Reinhardt D, Dohner H and Schlenk RF. Significance of age in acute myeloid leukemia patients younger than 30 years: a common analysis of the pediatric trials AML-BFM 93/98 and the adult trials AMLCG 92/99 and AMLSG HD93/98A. *Cancer* 2008; 112: 562-571. 2007/12/14. DOI: 10.1002/cncr.23220.
5. SEER Stat Fact Sheets: Acute Myeloid Leukemia (AML). *Surveillance, Epidemiology, and End Results (SEER) Program (<http://www.seercancer.gov>) Research Data* 2019; National Cancer Institute, DCCPS, surveillance Research Program, Surveillance Systems Branch
6. Seita J and Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med* 2010; 2: 640-653. 2010/10/05. DOI: 10.1002/wsbm.86.
7. Chang HY, Rodriguez V, Narboni G, Bodey GP, Luna MA and Freireich EJ. Causes of death in adults with acute leukemia. *Medicine (Baltimore)* 1976; 55: 259-268. 1976/05/01. DOI: 10.1097/00005792-197605000-00005.

8. Hillestad LK. Acute promyelocytic leukemia. *Acta Med Scand* 1957; 159: 189-194. 1957/11/29.
9. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, Bloomfield CD, Cazzola M and Vardiman JW. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016; 127: 2391-2405. 2016/04/14. DOI: 10.1182/blood-2016-03-643544.
10. Welch JS, Yuan W and Ley TJ. PML-RARA can increase hematopoietic self-renewal without causing a myeloproliferative disease in mice. *J Clin Invest* 2011; 121: 1636-1645. 2011/03/03. DOI: 10.1172/JCI42953.
11. Brown D, Kogan S, Lagasse E, Weissman I, Alcalay M, Pelicci PG, Atwater S and Bishop JM. A PMLRARAalpha transgene initiates murine acute promyelocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 1997; 94: 2551-2556. 1997/03/18. DOI: 10.1073/pnas.94.6.2551.
12. Kantarjian H. Acute myeloid leukemia--major progress over four decades and glimpses into the future. *Am J Hematol* 2016; 91: 131-145. 2015/11/26. DOI: 10.1002/ajh.24246.
13. Zhang TD, Chen GQ, Wang ZG, Wang ZY, Chen SJ and Chen Z. Arsenic trioxide, a therapeutic agent for APL. *Oncogene* 2001; 20: 7146-7153. 2001/11/13. DOI: 10.1038/sj.onc.1204762.
14. Platzbecker U, Avvisati G, Cicconi L, Thiede C, Paoloni F, Vignetti M, Ferrara F, Divona M, Albano F, Efficace F, Fazi P, Sborgia M, Di Bona E, Breccia M, Borlenghi E, Cairoli R, Rambaldi A, Melillo L, La Nasa G, Fiedler W, Brossart P, Hertenstein B, Salih HR, Wattad M, Lubbert M, Brandts CH, Hanel M, Rollig C, Schmitz N, Link H, Frairia C, Pogliani EM, Fozza C, D'Arco AM, Di Renzo N, Cortelezzi A, Fabbiano F, Dohner K, Ganser A, Dohner H, Amadori S, Mandelli F, Ehninger G, Schlenk RF and Lo-Coco F. Improved Outcomes With Retinoic Acid and Arsenic Trioxide Compared With Retinoic Acid and Chemotherapy in Non-High-Risk Acute Promyelocytic Leukemia: Final Results of the Randomized Italian-German APL0406 Trial.

Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2017; 35: 605-612. 2016/07/13. DOI: 10.1200/JCO.2016.67.1982.

15. Rowley JD. Identification of a translocation with quinacrine fluorescence in a patient with acute leukemia. *Ann Genet* 1973; 16: 109-112. 1973/06/01.

16. Erickson P, Gao J, Chang KS, Look T, Whisenant E, Raimondi S, Lasher R, Trujillo J, Rowley J and Drabkin H. Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* 1992; 80: 1825-1831. 1992/10/01.

17. Lutterbach B, Hou Y, Durst KL and Hiebert SW. The inv(16) encodes an acute myeloid leukemia 1 transcriptional corepressor. *Proceedings of the National Academy of Sciences of the United States of America* 1999; 96: 12822-12827. 1999/10/27. DOI: 10.1073/pnas.96.22.12822.

18. Liu P, Tarle SA, Hajra A, Claxton DF, Marlton P, Freedman M, Siciliano MJ and Collins FS. Fusion between transcription factor CBF beta/PEBP2 beta and a myosin heavy chain in acute myeloid leukemia. *Science* 1993; 261: 1041-1044. 1993/08/20. DOI: 10.1126/science.8351518.

19. Schlenk RF, Benner A, Krauter J, Buchner T, Sauerland C, Ehninger G, Schaich M, Mohr B, Niederwieser D, Krahl R, Pasold R, Dohner K, Ganser A, Dohner H and Heil G. Individual patient data-based meta-analysis of patients aged 16 to 60 years with core binding factor acute myeloid leukemia: a survey of the German Acute Myeloid Leukemia Intergroup. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2004; 22: 3741-3750. 2004/08/04. DOI: 10.1200/JCO.2004.03.012.

20. Marcucci G, Mrozek K, Ruppert AS, Maharry K, Kolitz JE, Moore JO, Mayer RJ, Pettenati MJ, Powell BL, Edwards CG, Sterling LJ, Vardiman JW, Schiffer CA, Carroll AJ, Larson RA and Bloomfield CD. Prognostic factors and outcome of core binding factor acute myeloid leukemia patients with t(8;21) differ from those of patients with inv(16): a Cancer and Leukemia Group B

study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005; 23: 5705-5717. 2005/08/20. DOI: 10.1200/JCO.2005.15.610.

21. Yates J, Glidewell O, Wiernik P, Cooper MR, Steinberg D, Dosik H, Levy R, Hoagland C, Henry P, Gottlieb A, Cornell C, Berenberg J, Hutchison JL, Raich P, Nissen N, Ellison RR, Frelick R, James GW, Falkson G, Silver RT, Haurani F, Green M, Henderson E, Leone L and Holland JF. Cytosine arabinoside with daunorubicin or adriamycin for therapy of acute myelocytic leukemia: a CALGB study. *Blood* 1982; 60: 454-462.

22. Yates JW, Wallace HJ, Jr., Ellison RR and Holland JF. Cytosine arabinoside (NSC-63878) and daunorubicin (NSC-83142) therapy in acute nonlymphocytic leukemia. *Cancer chemotherapy reports* 1973; 57: 485-488.

23. Prebet T, Boissel N, Reutenauer S, Thomas X, Delaunay J, Cahn JY, Pigneux A, Quesnel B, Witz F, Thepot S, Ugo V, Terre C, Recher C, Tavernier E, Hunault M, Esterni B, Castaigne S, Guilhot F, Dombret H, Vey N, Acute Leukemia French A, Groupe Ouest-Est des leucemies et autres maladies du s and Core Binding Factor Acute Myeloid Leukemia i. Acute myeloid leukemia with translocation (8;21) or inversion (16) in elderly patients treated with conventional chemotherapy: a collaborative study of the French CBF-AML intergroup. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009; 27: 4747-4753. 2009/09/02. DOI: 10.1200/JCO.2008.21.0674.

24. Slovak ML, Kopecky KJ, Cassileth PA, Harrington DH, Theil KS, Mohamed A, Paietta E, Willman CL, Head DR, Rowe JM, Forman SJ and Appelbaum FR. Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group Study. *Blood* 2000; 96: 4075-4083. 2000/12/09.

25. Dohner H, Weisdorf DJ and Bloomfield CD. Acute Myeloid Leukemia. *N Engl J Med* 2015; 373: 1136-1152. 2015/09/17. DOI: 10.1056/NEJMra1406184.

26. Li HY, Deng DH, Huang Y, Ye FH, Huang LL, Xiao Q, Zhang B, Ye BB, Lai YR, Mo ZN and Liu ZF. Favorable prognosis of biallelic CEBPA gene mutations in acute myeloid leukemia patients: a meta-analysis. *Eur J Haematol* 2015; 94: 439-448. 2014/09/18. DOI: 10.1111/ejh.12450.
27. Dohner K, Schlenk RF, Habdank M, Scholl C, Rucker FG, Corbacioglu A, Bullinger L, Frohling S and Dohner H. Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood* 2005; 106: 3740-3746. DOI: 10.1182/blood-2005-05-2164.
28. Stone RM, Mandrekar SJ, Sanford BL, Laumann K, Geyer S, Bloomfield CD, Thiede C, Prior TW, Dohner K, Marcucci G, Lo-Coco F, Klisovic RB, Wei A, Sierra J, Sanz MA, Brandwein JM, de Witte T, Niederwieser D, Appelbaum FR, Medeiros BC, Tallman MS, Krauter J, Schlenk RF, Ganser A, Serve H, Ehninger G, Amadori S, Larson RA and Dohner H. Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation. *N Engl J Med* 2017; 377: 454-464. DOI: 10.1056/NEJMoa1614359.
29. Perl AE, Martinelli G, Cortes JE, Neubauer A, Berman E, Paolini S, Montesinos P, Baer MR, Larson RA, Ustun C, Fabbiano F, Erba HP, Di Stasi A, Stuart R, Olin R, Kasner M, Ciceri F, Chou WC, Podoltsev N, Recher C, Yokoyama H, Hosono N, Yoon SS, Lee JH, Pardee T, Fathi AT, Liu C, Hasabou N, Liu X, Bahceci E and Levis MJ. Gilteritinib or Chemotherapy for Relapsed or Refractory FLT3-Mutated AML. *N Engl J Med* 2019; 381: 1728-1740. 2019/10/31. DOI: 10.1056/NEJMoa1902688.
30. Stein EM, DiNardo CD, Fathi AT, Pollyea DA, Stone RM, Altman JK, Roboz GJ, Patel MR, Collins R, Flinn IW, Sekeres MA, Stein AS, Kantarjian HM, Levine RL, Vyas P, MacBeth KJ, Tosolini A, VanOostendorp J, Xu Q, Gupta I, Lila T, Risueno A, Yen KE, Wu B, Attar EC, Tallman MS and de Botton S. Molecular remission and response patterns in patients with mutant-IDH2 acute myeloid leukemia treated with enasidenib. *Blood* 2019; 133: 676-687. 2018/12/05. DOI: 10.1182/blood-2018-08-869008.

31. DiNardo CD, Stein EM, de Botton S, Roboz GJ, Altman JK, Mims AS, Swords R, Collins RH, Mannis GN, Pollyea DA, Donnellan W, Fathi AT, Pigneux A, Erba HP, Prince GT, Stein AS, Uy GL, Foran JM, Traer E, Stuart RK, Arellano ML, Slack JL, Sekeres MA, Willekens C, Choe S, Wang H, Zhang V, Yen KE, Kapsalis SM, Yang H, Dai D, Fan B, Goldwasser M, Liu H, Agresta S, Wu B, Attar EC, Tallman MS, Stone RM and Kantarjian HM. Durable Remissions with Ivosidenib in IDH1-Mutated Relapsed or Refractory AML. *N Engl J Med* 2018; 378: 2386-2398. 2018/06/05. DOI: 10.1056/NEJMoa1716984.
32. Cancer Genome Atlas Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, Robertson A, Hoadley K, Triche TJ, Jr., Laird PW, Baty JD, Fulton LL, Fulton R, Heath SE, Kalicki-Veizer J, Kandoth C, Klco JM, Koboldt DC, Kanchi KL, Kulkarni S, Lamprecht TL, Larson DE, Lin L, Lu C, McLellan MD, McMichael JF, Payton J, Schmidt H, Spencer DH, Tomasson MH, Wallis JW, Wartman LD, Watson MA, Welch J, Wendl MC, Ally A, Balasundaram M, Birol I, Butterfield Y, Chiu R, Chu A, Chuah E, Chun HJ, Corbett R, Dhalla N, Guin R, He A, Hirst C, Hirst M, Holt RA, Jones S, Karsan A, Lee D, Li HI, Marra MA, Mayo M, Moore RA, Mungall K, Parker J, Pleasance E, Plettner P, Schein J, Stoll D, Swanson L, Tam A, Thiessen N, Varhol R, Wye N, Zhao Y, Gabriel S, Getz G, Sougnez C, Zou L, Leiserson MD, Vandin F, Wu HT, Applebaum F, Baylin SB, Akbani R, Broom BM, Chen K, Motter TC, Nguyen K, Weinstein JN, Zhang N, Ferguson ML, Adams C, Black A, Bowen J, Gastier-Foster J, Grossman T, Lichtenberg T, Wise L, Davidsen T, Demchok JA, Shaw KR, Sheth M, Sofia HJ, Yang L, Downing JR and Eley G. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013; 368: 2059-2074. DOI: 10.1056/NEJMoa1301689.
33. Metzeler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Gorlich D, Schneider S, Konstandin NP, Dufour A, Braundl K, Ksienzyk B, Zellmeier E, Hartmann L, Greif PA, Fiegl M, Subklewe M, Bohlander SK, Krug U, Faldum A, Berdel WE, Wormann B, Buchner T, Hiddemann W, Braess J, Spiekermann K and Group AS. Spectrum and prognostic relevance

of driver gene mutations in acute myeloid leukemia. *Blood* 2016; 128: 686-698. 2016/06/12. DOI: 10.1182/blood-2016-01-693879.

34. Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, Long N, Schultz AR, Traer E, Abel M, Agarwal A, Blucher A, Borate U, Bryant J, Burke R, Carlos A, Carpenter R, Carroll J, Chang BH, Coblentz C, d'Almeida A, Cook R, Danilov A, Dao KT, Degnin M, Devine D, Dibb J, Edwards DKt, Eide CA, English I, Glover J, Henson R, Ho H, Jemal A, Johnson K, Johnson R, Junio B, Kaempf A, Leonard J, Lin C, Liu SQ, Lo P, Loriaux MM, Luty S, Macey T, MacManiman J, Martinez J, Mori M, Nelson D, Nichols C, Peters J, Ramsdill J, Rofelty A, Schuff R, Searles R, Segerdell E, Smith RL, Spurgeon SE, Sweeney T, Thapa A, Visser C, Wagner J, Watanabe-Smith K, Werth K, Wolf J, White L, Yates A, Zhang H, Cogle CR, Collins RH, Connolly DC, Deininger MW, Drusbosky L, Hourigan CS, Jordan CT, Kropf P, Lin TL, Martinez ME, Medeiros BC, Pallapati RR, Pollyea DA, Swords RT, Watts JM, Weir SJ, Wiest DL, Winters RM, McWeeney SK and Druker BJ. Functional genomic landscape of acute myeloid leukaemia. *Nature* 2018; 562: 526-531. 2018/10/20. DOI: 10.1038/s41586-018-0623-z.

35. Sweetser DA, Peniket AJ, Haaland C, Blomberg AA, Zhang Y, Zaidi ST, Dayyani F, Zhao Z, Heerema NA, Boulwood J, Dewald GW, Paietta E, Slovak ML, Willman CL, Wainscoat JS, Bernstein ID and Daly SB. Delineation of the minimal commonly deleted segment and identification of candidate tumor-suppressor genes in del(9q) acute myeloid leukemia. *Genes Chromosomes Cancer* 2005; 44: 279-291. 2005/07/15. DOI: 10.1002/gcc.20236.

36. Dayyani F, Wang J, Yeh JR, Ahn EY, Tobey E, Zhang DE, Bernstein ID, Peterson RT and Sweetser DA. Loss of TLE1 and TLE4 from the del(9q) commonly deleted region in AML cooperates with AML1-ETO to affect myeloid cell proliferation and survival. *Blood* 2008; 111: 4338-4347. 2008/02/09. DOI: 10.1182/blood-2007-07-103291.

37. Kronke J, Bullinger L, Teleanu V, Tschurtz F, Gaidzik VI, Kuhn MW, Rucker FG, Holzmann K, Paschka P, Kapp-Schworer S, Spath D, Kindler T, Schittenhelm M, Krauter J, Ganser A, Gohring G, Schlegelberger B, Schlenk RF, Dohner H and Dohner K. Clonal evolution

in relapsed NPM1-mutated acute myeloid leukemia. *Blood* 2013; 122: 100-108. 2013/05/25. DOI: 10.1182/blood-2013-01-479188.

38. Dentici ML, Barresi S, Niceta M, Pantaleoni F, Pizzi S, Dallapiccola B, Tartaglia M and Digilio MC. Clinical spectrum of Kabuki-like syndrome caused by HNRNPK haploinsufficiency. *Clin Genet* 2018; 93: 401-407. 2017/04/05. DOI: 10.1111/cge.13029.

39. Bustelo XR, Suen KL, Michael WM, Dreyfuss G and Barbacid M. Association of the vav proto-oncogene product with poly(rC)-specific RNA-binding proteins. *Molecular and cellular biology* 1995; 15: 1324-1332. 1995/03/01. DOI: 10.1128/mcb.15.3.1324.

40. Hobert O, Jallal B, Schlessinger J and Ullrich A. Novel signaling pathway suggested by SH3 domain-mediated p95vav/heterogeneous ribonucleoprotein K interaction. *The Journal of biological chemistry* 1994; 269: 20225-20228. 1994/08/12.

41. Ostrowski J, Schullery DS, Denisenko ON, Higaki Y, Watts J, Aebersold R, Stempka L, Gschwendt M and Bomsztyk K. Role of tyrosine phosphorylation in the regulation of the interaction of heterogeneous nuclear ribonucleoprotein K protein with its protein and RNA partners. *The Journal of biological chemistry* 2000; 275: 3619-3628. 2000/02/01. DOI: 10.1074/jbc.275.5.3619.

42. Schullery DS, Ostrowski J, Denisenko ON, Stempka L, Shnyreva M, Suzuki H, Gschwendt M and Bomsztyk K. Regulated interaction of protein kinase Cdelta with the heterogeneous nuclear ribonucleoprotein K protein. *The Journal of biological chemistry* 1999; 274: 15101-15109. 1999/05/18. DOI: 10.1074/jbc.274.21.15101.

43. Mandal M, Vadlamudi R, Nguyen D, Wang R-A, Costa L, Bagheri-Yarmand R, Mendelsohn J and Kumar R. Growth factors regulate heterogeneous nuclear ribonucleoprotein K expression and function. *Journal of Biological Chemistry* 2001; 276: 9699-9704.

44. Ostrowski J, Kawata Y, Schullery DS, Denisenko ON and Bomsztyk K. Transient recruitment of the hnRNP K protein to inducibly transcribed gene loci. *Nucleic acids research* 2003; 31: 3954-3962. 2003/07/11. DOI: 10.1093/nar/gkg452.

45. Denisenko ON and Bomsztyk K. The product of the murine homolog of the *Drosophila* extra sex combs gene displays transcriptional repressor activity. *Molecular and cellular biology* 1997; 17: 4707-4717. 1997/08/01. DOI: 10.1128/mcb.17.8.4707.
46. Shnyreva M, Schullery DS, Suzuki H, Higaki Y and Bomsztyk K. Interaction of two multifunctional proteins. Heterogeneous nuclear ribonucleoprotein K and Y-box-binding protein. *The Journal of biological chemistry* 2000; 275: 15498-15503. 2000/05/16. DOI: 10.1074/jbc.275.20.15498.
47. Thompson PJ, Dulberg V, Moon KM, Foster LJ, Chen C, Karimi MM and Lorincz MC. hnRNP K coordinates transcriptional silencing by SETDB1 in embryonic stem cells. *PLoS Genet* 2015; 11: e1004933. 2015/01/23. DOI: 10.1371/journal.pgen.1004933.
48. Michelotti EF, Michelotti GA, Aronsohn AI and Levens D. Heterogeneous nuclear ribonucleoprotein K is a transcription factor. *Molecular and cellular biology* 1996; 16: 2350-2360. 1996/05/01. DOI: 10.1128/mcb.16.5.2350.
49. Denisenko ON, O'Neill B, Ostrowski J, Van Seuning I and Bomsztyk K. Zik1, a transcriptional repressor that interacts with the heterogeneous nuclear ribonucleoprotein particle K protein. *The Journal of biological chemistry* 1996; 271: 27701-27706. 1996/11/01. DOI: 10.1074/jbc.271.44.27701.
50. Choi HS, Song KY, Hwang CK, Kim CS, Law PY, Wei LN and Loh HH. A proteomics approach for identification of single strand DNA-binding proteins involved in transcriptional regulation of mouse mu opioid receptor gene. *Mol Cell Proteomics* 2008; 7: 1517-1529. 2008/05/06. DOI: 10.1074/mcp.M800052-MCP200.
51. Lau JS, Baumeister P, Kim E, Roy B, Hsieh TY, Lai M and Lee AS. Heterogeneous nuclear ribonucleoproteins as regulators of gene expression through interactions with the human thymidine kinase promoter. *J Cell Biochem* 2000; 79: 395-406. 2000/09/06. DOI: 10.1002/1097-4644(20001201)79:3<395::aid-jcb50>3.0.co;2-m.

52. Da Silva N, Bharti A and Shelley CS. hnRNP-K and Pur(alpha) act together to repress the transcriptional activity of the CD43 gene promoter. *Blood* 2002; 100: 3536-3544. 2002/11/02. DOI: 10.1182/blood.V100.10.3536.
53. Gallardo M, Lee HJ, Zhang X, Bueso-Ramos C, Pagoon LR, McArthur M, Multani A, Nazha A, Manshouri T, Parker-Thornburg J, Rapado I, Quintas-Cardama A, Kornblau SM, Martinez-Lopez J and Post SM. hnRNP K Is a Haploinsufficient Tumor Suppressor that Regulates Proliferation and Differentiation Programs in Hematologic Malignancies. *Cancer cell* 2015; 28: 486-499. DOI: 10.1016/j.ccell.2015.09.001.
54. Lynch M, Chen L, Ravitz MJ, Mehtani S, Korenblat K, Pazin MJ and Schmidt EV. hnRNP K binds a core polypyrimidine element in the eukaryotic translation initiation factor 4E (eIF4E) promoter, and its regulation of eIF4E contributes to neoplastic transformation. *Molecular and cellular biology* 2005; 25: 6436-6453. DOI: 10.1128/MCB.25.15.6436-6453.2005.
55. Expert-Bezancon A, Le Caer JP and Marie J. Heterogeneous nuclear ribonucleoprotein (hnRNP) K is a component of an intronic splicing enhancer complex that activates the splicing of the alternative exon 6A from chicken beta-tropomyosin pre-mRNA. *The Journal of biological chemistry* 2002; 277: 16614-16623. 2002/02/28. DOI: 10.1074/jbc.M201083200.
56. Cao W, Razanau A, Feng D, Lobo VG and Xie J. Control of alternative splicing by forskolin through hnRNP K during neuronal differentiation. *Nucleic acids research* 2012; 40: 8059-8071. DOI: 10.1093/nar/gks504.
57. Venables JP, Koh CS, Froehlich U, Lapointe E, Couture S, Inkel L, Bramard A, Paquet ER, Watier V, Durand M, Lucier JF, Gervais-Bird J, Tremblay K, Prinos P, Klinck R, Elela SA and Chabot B. Multiple and specific mRNA processing targets for the major human hnRNP proteins. *Molecular and cellular biology* 2008; 28: 6033-6043. 2008/07/23. DOI: 10.1128/MCB.00726-08.
58. Revil T, Pelletier J, Toutant J, Cloutier A and Chabot B. Heterogeneous nuclear ribonucleoprotein K represses the production of pro-apoptotic Bcl-xS splice isoform. *The Journal of biological chemistry* 2009; 284: 21458-21467. 2009/06/13. DOI: 10.1074/jbc.M109.019711.

59. Tsai P-L, Chiou N-T, Kuss S, García-Sastre A, Lynch KW and Fontoura BM. Cellular RNA binding proteins NS1-BP and hnRNP K regulate influenza A virus RNA splicing. *PLoS pathogens* 2013; 9.
60. Skalweit A, Doller A, Huth A, Kahne T, Persson PB and Thiele BJ. Posttranscriptional control of renin synthesis: identification of proteins interacting with renin mRNA 3'-untranslated region. *Circ Res* 2003; 92: 419-427. 2003/02/26. DOI: 10.1161/01.RES.0000059300.67152.4E.
61. Shanmugam N, Reddy MA and Natarajan R. Distinct roles of heterogeneous nuclear ribonuclear protein K and microRNA-16 in cyclooxygenase-2 RNA stability induced by S100b, a ligand of the receptor for advanced glycation end products. *The Journal of biological chemistry* 2008; 283: 36221-36233. 2008/10/16. DOI: 10.1074/jbc.M806322200.
62. Yano M, Okano HJ and Okano H. Involvement of Hu and heterogeneous nuclear ribonucleoprotein K in neuronal differentiation through p21 mRNA post-transcriptional regulation. *The Journal of biological chemistry* 2005; 280: 12690-12699. 2005/01/27. DOI: 10.1074/jbc.M411119200.
63. Bomsztyk K, Van Seuningem I, Suzuki H, Denisenko O and Ostrowski J. Diverse molecular interactions of the hnRNP K protein. *FEBS Lett* 1997; 403: 113-115. 1997/02/17. DOI: 10.1016/s0014-5793(97)00041-0.
64. Evans JR, Mitchell SA, Spriggs KA, Ostrowski J, Bomsztyk K, Ostarek D and Willis AE. Members of the poly (rC) binding protein family stimulate the activity of the c-myc internal ribosome entry segment in vitro and in vivo. *Oncogene* 2003; 22: 8012-8020. 2003/09/13. DOI: 10.1038/sj.onc.1206645.
65. Collier B, Goobar-Larsson L, Sokolowski M and Schwartz S. Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. *The Journal of biological chemistry* 1998; 273: 22648-22656. 1998/08/26. DOI: 10.1074/jbc.273.35.22648.

66. Ostareck-Lederer A and Ostareck DH. Control of mRNA translation and stability in haematopoietic cells: the function of hnRNPs K and E1/E2. *Biol Cell* 2004; 96: 407-411. 2004/09/24. DOI: 10.1016/j.biocel.2004.03.010.
67. Ostareck DH, Ostareck-Lederer A, Wilm M, Thiele BJ, Mann M and Hentze MW. mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-lipoxygenase translation from the 3' end. *Cell* 1997; 89: 597-606.
68. Naarmann IS, Harnisch C, Flach N, Kremmer E, Kuhn H, Ostareck DH and Ostareck-Lederer A. mRNA silencing in human erythroid cell maturation: heterogeneous nuclear ribonucleoprotein K controls the expression of its regulator c-Src. *The Journal of biological chemistry* 2008; 283: 18461-18472. 2008/04/29. DOI: 10.1074/jbc.M710328200.
69. Mukhopadhyay NK, Kim J, Cinar B, Ramachandran A, Hager MH, Di Vizio D, Adam RM, Rubin MA, Raychaudhuri P, De Benedetti A and Freeman MR. Heterogeneous nuclear ribonucleoprotein K is a novel regulator of androgen receptor translation. *Cancer Res* 2009; 69: 2210-2218. 2009/03/05. DOI: 10.1158/0008-5472.CAN-08-2308.
70. Chkheidze AN and Liebhaber SA. A novel set of nuclear localization signals determine distributions of the alphaCP RNA-binding proteins. *Molecular and cellular biology* 2003; 23: 8405-8415. 2003/11/13. DOI: 10.1128/mcb.23.23.8405-8415.2003.
71. Michael WM, Eder PS and Dreyfuss G. The K nuclear shuttling domain: a novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J* 1997; 16: 3587-3598. 1997/06/01. DOI: 10.1093/emboj/16.12.3587.
72. Ostareck-Lederer A, Ostareck DH, Rucknagel KP, Schierhorn A, Moritz B, Huttelmaier S, Flach N, Handoko L and Wahle E. Asymmetric arginine dimethylation of heterogeneous nuclear ribonucleoprotein K by protein-arginine methyltransferase 1 inhibits its interaction with c-Src. *The Journal of biological chemistry* 2006; 281: 11115-11125. 2006/02/24. DOI: 10.1074/jbc.M513053200.

73. Van Seuning I, Ostrowski J, Bustelo XR, Sleath PR and Bomsztyk K. The K protein domain that recruits the interleukin 1-responsive K protein kinase lies adjacent to a cluster of c-Src and Vav SH3-binding sites. Implications that K protein acts as a docking platform. *The Journal of biological chemistry* 1995; 270: 26976-26985. 1995/11/10. DOI: 10.1074/jbc.270.45.26976.
74. Siomi H, Matunis MJ, Michael WM and Dreyfuss G. The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic acids research* 1993; 21: 1193-1198. 1993/03/11. DOI: 10.1093/nar/21.5.1193.
75. Carpenter B, McKay M, Dundas SR, Lawrie LC, Telfer C and Murray GI. Heterogeneous nuclear ribonucleoprotein K is over expressed, aberrantly localised and is associated with poor prognosis in colorectal cancer. *British journal of cancer* 2006; 95: 921-927. DOI: 10.1038/sj.bjc.6603349.
76. Chen LC, Chung IC, Hsueh C, Tsang NM, Chi LM, Liang Y, Chen CC, Wang LJ and Chang YS. The antiapoptotic protein, FLIP, is regulated by heterogeneous nuclear ribonucleoprotein K and correlates with poor overall survival of nasopharyngeal carcinoma patients. *Cell Death Differ* 2010; 17: 1463-1473. 2010/03/13. DOI: 10.1038/cdd.2010.24.
77. Ciarlo M, Benelli R, Barbieri O, Minghelli S, Barboro P, Balbi C and Ferrari N. Regulation of neuroendocrine differentiation by AKT/hnRNP/AR/beta-catenin signaling in prostate cancer cells. *International journal of cancer* 2012; 131: 582-590. 2011/10/22. DOI: 10.1002/ijc.26402.
78. Wen F, Shen A, Shanas R, Bhattacharyya A, Lian F, Hostetter G and Shi J. Higher expression of the heterogeneous nuclear ribonucleoprotein k in melanoma. *Ann Surg Oncol* 2010; 17: 2619-2627. 2010/05/26. DOI: 10.1245/s10434-010-1121-1.
79. Mandal M, Vadlamudi R, Nguyen D, Wang RA, Costa L, Bagheri-Yarmand R, Mendelsohn J and Kumar R. Growth factors regulate heterogeneous nuclear ribonucleoprotein K expression and function. *The Journal of biological chemistry* 2001; 276: 9699-9704. 2000/12/31. DOI: 10.1074/jbc.M008514200.

80. Barboro P, Repaci E, Rubagotti A, Salvi S, Boccardo S, Spina B, Truini M, Introini C, Puppo P, Ferrari N, Carmignani G, Boccardo F and Balbi C. Heterogeneous nuclear ribonucleoprotein K: altered pattern of expression associated with diagnosis and prognosis of prostate cancer. *British journal of cancer* 2009; 100: 1608-1616. 2009/04/30. DOI: 10.1038/sj.bjc.6605057.
81. Zhou R, Shanas R, Nelson MA, Bhattacharyya A and Shi J. Increased expression of the heterogeneous nuclear ribonucleoprotein K in pancreatic cancer and its association with the mutant p53. *International journal of cancer* 2010; 126: 395-404. DOI: 10.1002/ijc.24744.
82. Chen X, Gu P, Xie R, Han J, Liu H, Wang B, Xie W, Xie W, Zhong G, Chen C, Xie S, Jiang N, Lin T and Huang J. Heterogeneous nuclear ribonucleoprotein K is associated with poor prognosis and regulates proliferation and apoptosis in bladder cancer. *J Cell Mol Med* 2017; 21: 1266-1279. 2016/11/20. DOI: 10.1111/jcmm.12999.
83. Gallardo M, Malaney P, Aitken MJL, Zhang X, Link TM, Shah V, Alybayev S, Wu MH, Pagon LR, Ma H, Jacamo R, Yu L, Xu-Monette ZY, Steinman H, Lee HJ, Sarbassov D, Rapado I, Barton MC, Martinez-Lopez J, Bueso-Ramos C, Young KH and Post SM. Uncovering the role of hnRNP K, an RNA-binding protein, in B-cell lymphomas. *J Natl Cancer Inst* 2019 2019/05/12. DOI: 10.1093/jnci/djz078.
84. Okuda T, Nishimura M, Nakao M and Fujita Y. RUNX1/AML1: a central player in hematopoiesis. *Int J Hematol* 2001; 74: 252-257. 2001/11/28. DOI: 10.1007/bf02982057.
85. Wang Q, Stacy T, Binder M, Marin-Padilla M, Sharpe AH and Speck NA. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proceedings of the National Academy of Sciences of the United States of America* 1996; 93: 3444-3449. 1996/04/16. DOI: 10.1073/pnas.93.8.3444.
86. Gowney JD, Shigematsu H, Li Z, Lee BH, Adelsperger J, Rowan R, Curley DP, Kutok JL, Akashi K, Williams IR, Speck NA and Gilliland DG. Loss of Runx1 perturbs adult

hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* 2005; 106: 494-504. 2005/03/24. DOI: 10.1182/blood-2004-08-3280.

87. Ghози MC, Bernstein Y, Negreanu V, Levanon D and Groner Y. Expression of the human acute myeloid leukemia gene AML1 is regulated by two promoter regions. *Proceedings of the National Academy of Sciences of the United States of America* 1996; 93: 1935-1940. 1996/03/05. DOI: 10.1073/pnas.93.5.1935.

88. Komeno Y, Yan M, Matsuura S, Lam K, Lo MC, Huang YJ, Tenen DG, Downing JR and Zhang DE. Runx1 exon 6-related alternative splicing isoforms differentially regulate hematopoiesis in mice. *Blood* 2014; 123: 3760-3769. DOI: 10.1182/blood-2013-08-521252.

89. Miyoshi H, Ohira M, Shimizu K, Mitani K, Hirai H, Imai T, Yokoyama K, Soeda E and Ohki M. Alternative splicing and genomic structure of the AML1 gene involved in acute myeloid leukemia. *Nucleic acids research* 1995; 23: 2762-2769. 1995/07/25. DOI: 10.1093/nar/23.14.2762.

90. Tanaka T, Tanaka K, Ogawa S, Kurokawa M, Mitani K, Nishida J, Shibata Y, Yazaki Y and Hirai H. An acute myeloid leukemia gene, AML1, regulates hemopoietic myeloid cell differentiation and transcriptional activation antagonistically by two alternative spliced forms. *EMBO J* 1995; 14: 341-350. 1995/01/16.

91. Kagoshima H, Shigesada K, Satake M, Ito Y, Miyoshi H, Ohki M, Pepling M and Gergen P. The Runt domain identifies a new family of heteromeric transcriptional regulators. *Trends Genet* 1993; 9: 338-341. 1993/10/01. DOI: 10.1016/0168-9525(93)90026-e.

92. Meyers S, Downing JR and Hiebert SW. Identification of AML-1 and the (8;21) translocation protein (AML-1/ETO) as sequence-specific DNA-binding proteins: the runt homology domain is required for DNA binding and protein-protein interactions. *Molecular and cellular biology* 1993; 13: 6336-6345. 1993/10/01. DOI: 10.1128/mcb.13.10.6336.

93. Bae SC, Ogawa E, Maruyama M, Oka H, Satake M, Shigesada K, Jenkins NA, Gilbert DJ, Copeland NG and Ito Y. PEBP2 alpha B/mouse AML1 consists of multiple isoforms that

- possess differential transactivation potentials. *Molecular and cellular biology* 1994; 14: 3242-3252. 1994/05/01. DOI: 10.1128/mcb.14.5.3242.
94. Zhang YW, Bae SC, Huang G, Fu YX, Lu J, Ahn MY, Kanno Y, Kanno T and Ito Y. A novel transcript encoding an N-terminally truncated AML1/PEBP2 alphaB protein interferes with transactivation and blocks granulocytic differentiation of 32Dcl3 myeloid cells. *Molecular and cellular biology* 1997; 17: 4133-4145. 1997/07/01. DOI: 10.1128/mcb.17.7.4133.
95. Challen GA and Goodell MA. Runx1 isoforms show differential expression patterns during hematopoietic development but have similar functional effects in adult hematopoietic stem cells. *Exp Hematol* 2010; 38: 403-416. 2010/03/09. DOI: 10.1016/j.exphem.2010.02.011.
96. Brady G, Elgueta Karstegl C and Farrell PJ. Novel function of the unique N-terminal region of RUNX1c in B cell growth regulation. *Nucleic acids research* 2013; 41: 1555-1568. 2012/12/21. DOI: 10.1093/nar/gks1273.
97. Sroczynska P, Lancrin C, Kouskoff V and Lacaud G. The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. *Blood* 2009; 114: 5279-5289. 2009/10/28. DOI: 10.1182/blood-2009-05-222307.
98. Liu X, Zhang Q, Zhang DE, Zhou C, Xing H, Tian Z, Rao Q, Wang M and Wang J. Overexpression of an isoform of AML1 in acute leukemia and its potential role in leukemogenesis. *Leukemia* 2009; 23: 739-745. 2009/01/20. DOI: 10.1038/leu.2008.350.
99. Tsuzuki S, Hong D, Gupta R, Matsuo K, Seto M and Enver T. Isoform-specific potentiation of stem and progenitor cell engraftment by AML1/RUNX1. *PLoS Med* 2007; 4: e172. 2007/05/17. DOI: 10.1371/journal.pmed.0040172.
100. Tsuzuki S and Seto M. Expansion of functionally defined mouse hematopoietic stem and progenitor cells by a short isoform of RUNX1/AML1. *Blood* 2012; 119: 727-735. 2011/12/02. DOI: 10.1182/blood-2011-06-362277.

101. Guo H, Ma O, Speck NA and Friedman AD. Runx1 deletion or dominant inhibition reduces Cebpa transcription via conserved promoter and distal enhancer sites to favor monoopoiesis over granulopoiesis. *Blood* 2012; 119: 4408-4418. 2012/03/28. DOI: 10.1182/blood-2011-12-397091.
102. Wotton SF, Blyth K, Kilbey A, Jenkins A, Terry A, Bernardin-Fried F, Friedman AD, Baxter EW, Neil JC and Cameron ER. RUNX1 transformation of primary embryonic fibroblasts is revealed in the absence of p53. *Oncogene* 2004; 23: 5476-5486. 2004/05/11. DOI: 10.1038/sj.onc.1207729.
103. Wolyniec K, Wotton S, Kilbey A, Jenkins A, Terry A, Peters G, Stocking C, Cameron E and Neil JC. RUNX1 and its fusion oncoprotein derivative, RUNX1-ETO, induce senescence-like growth arrest independently of replicative stress. *Oncogene* 2009; 28: 2502-2512. 2009/05/19. DOI: 10.1038/onc.2009.101.
104. Miyoshi H, Shimizu K, Kozu T, Maseki N, Kaneko Y and Ohki M. t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proceedings of the National Academy of Sciences of the United States of America* 1991; 88: 10431-10434.
105. Nucifora G, Begy CR, Kobayashi H, Roulston D, Claxton D, Pedersen-Bjergaard J, Parganas E, Ihle JN and Rowley JD. Consistent intergenic splicing and production of multiple transcripts between AML1 at 21q22 and unrelated genes at 3q26 in (3;21)(q26;q22) translocations. *Proceedings of the National Academy of Sciences of the United States of America* 1994; 91: 4004-4008. 1994/04/26. DOI: 10.1073/pnas.91.9.4004.
106. Golub TR, Barker GF, Bohlander SK, Hiebert SW, Ward DC, Bray-Ward P, Morgan E, Raimondi SC, Rowley JD and Gilliland DG. Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 1995; 92: 4917-4921.

107. De Braekeleer E, Douet-Guilbert N, Morel F, Le Bris MJ, Ferec C and De Braekeleer M. RUNX1 translocations and fusion genes in malignant hemopathies. *Future Oncol* 2011; 7: 77-91. 2010/12/23. DOI: 10.2217/fon.10.158.
108. Osato M, Asou N, Abdalla E, Hoshino K, Yamasaki H, Okubo T, Suzushima H, Takatsuki K, Kanno T, Shigesada K and Ito Y. Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2alphaB gene associated with myeloblastic leukemias. *Blood* 1999; 93: 1817-1824. 1999/03/09.
109. Osato M. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene* 2004; 23: 4284-4296. 2004/05/25. DOI: 10.1038/sj.onc.1207779.
110. Tsai SC, Shih LY, Liang ST, Huang YJ, Kuo MC, Huang CF, Shih YS, Lin TH, Chiu MC and Liang DC. Biological Activities of RUNX1 Mutants Predict Secondary Acute Leukemia Transformation from Chronic Myelomonocytic Leukemia and Myelodysplastic Syndromes. *Clin Cancer Res* 2015; 21: 3541-3551. 2015/04/05. DOI: 10.1158/1078-0432.CCR-14-2203.
111. Tang JL, Hou HA, Chen CY, Liu CY, Chou WC, Tseng MH, Huang CF, Lee FY, Liu MC, Yao M, Huang SY, Ko BS, Hsu SC, Wu SJ, Tsay W, Chen YC, Lin LI and Tien HF. AML1/RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia: prognostic implication and interaction with other gene alterations. *Blood* 2009; 114: 5352-5361. 2009/10/08. DOI: 10.1182/blood-2009-05-223784.
112. Schnittger S, Dicker F, Kern W, Wendland N, Sundermann J, Alpermann T, Haferlach C and Haferlach T. RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood* 2011; 117: 2348-2357. 2010/12/15. DOI: 10.1182/blood-2009-11-255976.
113. Greif PA, Konstandin NP, Metzeler KH, Herold T, Pasalic Z, Ksienzyk B, Dufour A, Schneider F, Schneider S, Kakadia PM, Braess J, Sauerland MC, Berdel WE, Buchner T, Woermann BJ, Hiddemann W, Spiekermann K and Bohlander SK. RUNX1 mutations in cytogenetically normal acute myeloid leukemia are associated with a poor prognosis and up-

regulation of lymphoid genes. *Haematologica* 2012; 97: 1909-1915. 2012/06/13. DOI: 10.3324/haematol.2012.064667.

114. Mender JH, Maharry K, Radmacher MD, Mrozek K, Becker H, Metzeler KH, Schwind S, Whitman SP, Khalife J, Kohlschmidt J, Nicolet D, Powell BL, Carter TH, Wetzler M, Moore JO, Kolitz JE, Baer MR, Carroll AJ, Larson RA, Caligiuri MA, Marcucci G and Bloomfield CD. RUNX1 mutations are associated with poor outcome in younger and older patients with cytogenetically normal acute myeloid leukemia and with distinct gene and MicroRNA expression signatures. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012; 30: 3109-3118. 2012/07/04. DOI: 10.1200/JCO.2011.40.6652.

115. Grossmann V, Kern W, Harbich S, Alpermann T, Jeromin S, Schnittger S, Haferlach C, Haferlach T and Kohlmann A. Prognostic relevance of RUNX1 mutations in T-cell acute lymphoblastic leukemia. *Haematologica* 2011; 96: 1874-1877. 2011/08/11. DOI: 10.3324/haematol.2011.043919.

116. Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, Ratajczak J, Resende IC, Haworth C, Hock R, Loh M, Felix C, Roy DC, Busque L, Kurnit D, Willman C, Gewirtz AM, Speck NA, Bushweller JH, Li FP, Gardiner K, Poncz M, Maris JM and Gilliland DG. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* 1999; 23: 166-175. 1999/10/03. DOI: 10.1038/13793.

117. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R and Hubbard TJ. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* 2012; 22: 1760-1774. 2012/09/08. DOI: 10.1101/gr.135350.111.

118. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; 489: 57-74. 2012/09/08. DOI: 10.1038/nature11247.
119. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabudde NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LD, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K, Chatterjee A, Huang TC, Zhong J, Wu X, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TS, Hruban RH, Kerr CL, Bader GD, Iacobuzio-Donahue CA, Gowda H and Pandey A. A draft map of the human proteome. *Nature* 2014; 509: 575-581. 2014/05/30. DOI: 10.1038/nature13302.
120. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber F and Kuster B. Mass-spectrometry-based draft of the human proteome. *Nature* 2014; 509: 582-587. 2014/05/30. DOI: 10.1038/nature13319.
121. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP and Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008; 456: 470-476. 2008/11/04. DOI: 10.1038/nature07509.
122. Pan Q, Shai O, Lee LJ, Frey BJ and Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 2008; 40: 1413-1415. 2008/11/04. DOI: 10.1038/ng.259.
123. Gilbert W. Why genes in pieces? *Nature* 1978; 271: 501. 1978/02/09. DOI: 10.1038/271501a0.

124. Crick F. Split genes and RNA splicing. *Science* 1979; 204: 264-271. 1979/04/20. DOI: 10.1126/science.373120.
125. Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 2003; 72: 291-336. 2003/03/11. DOI: 10.1146/annurev.biochem.72.121801.161720.
126. Nilsen TW. The spliceosome: no assembly required? *Mol Cell* 2002; 9: 8-9. 2002/01/24. DOI: 10.1016/s1097-2765(02)00430-6.
127. Brow DA. Allosteric cascade of spliceosome activation. *Annu Rev Genet* 2002; 36: 333-360. 2002/11/14. DOI: 10.1146/annurev.genet.36.043002.091635.
128. Turunen JJ, Niemela EH, Verma B and Frilander MJ. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* 2013; 4: 61-76. 2012/10/18. DOI: 10.1002/wrna.1141.
129. Tilgner H, Knowles DG, Johnson R, Davis CA, Chakraborty S, Djebali S, Curado J, Snyder M, Gingeras TR and Guigo R. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res* 2012; 22: 1616-1625. 2012/09/08. DOI: 10.1101/gr.134445.111.
130. Singh R and Valcarcel J. Building specificity with nonspecific RNA-binding proteins. *Nat Struct Mol Biol* 2005; 12: 645-653. 2005/08/04. DOI: 10.1038/nsmb961.
131. Long JC and Cáceres JF. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* 2009; 417: 15-27. 2008/12/09. DOI: 10.1042/BJ20081501.
132. Krecic AM and Swanson MS. hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol* 1999; 11: 363-371. 1999/07/08. DOI: 10.1016/S0955-0674(99)80051-9.
133. Zahler AM, Lane WS, Stolk JA and Roth MB. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes & development* 1992; 6: 837-847. 1992/05/01. DOI: 10.1101/gad.6.5.837.

134. Kohtz JD, Jamison SF, Will CL, Zuo P, Luhrmann R, Garcia-Blanco MA and Manley JL. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* 1994; 368: 119-124. 1994/03/10. DOI: 10.1038/368119a0.
135. Pandit S, Zhou Y, Shiue L, Coutinho-Mansfield G, Li H, Qiu J, Huang J, Yeo GW, Ares M, Jr. and Fu XD. Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol Cell* 2013; 50: 223-235. 2013/04/09. DOI: 10.1016/j.molcel.2013.03.001.
136. Han J, Ding JH, Byeon CW, Kim JH, Hertel KJ, Jeong S and Fu XD. SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Molecular and cellular biology* 2011; 31: 793-802. 2010/12/08. DOI: 10.1128/MCB.01117-10.
137. Xue Y, Zhou Y, Wu T, Zhu T, Ji X, Kwon YS, Zhang C, Yeo G, Black DL, Sun H, Fu XD and Zhang Y. Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol Cell* 2009; 36: 996-1006. 2010/01/13. DOI: 10.1016/j.molcel.2009.12.003.
138. Llorian M, Schwartz S, Clark TA, Hollander D, Tan LY, Spellman R, Gordon A, Schweitzer AC, de la Grange P, Ast G and Smith CW. Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat Struct Mol Biol* 2010; 17: 1114-1123. 2010/08/17. DOI: 10.1038/nsmb.1881.
139. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M, Chalkidis G, Suzuki Y, Shiosaka M, Kawahata R, Yamaguchi T, Otsu M, Obara N, Sakata-Yanagimoto M, Ishiyama K, Mori H, Nolte F, Hofmann WK, Miyawaki S, Sugano S, Haferlach C, Koeffler HP, Shih LY, Haferlach T, Chiba S, Nakauchi H, Miyano S and Ogawa S. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011; 478: 64-69. 2011/09/13. DOI: 10.1038/nature10496.
140. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, Krysiak K, Harris CC, Koboldt DC, Larson DE, McLellan MD, Dooling DJ, Abbott RM, Fulton RS, Schmidt H, Kalicki-Veizer J, O'Laughlin M, Grillot M, Baty J, Heath S, Frater JL, Nasim T, Link DC, Tomasson MH,

Westervelt P, DiPersio JF, Mardis ER, Ley TJ, Wilson RK and Walter MJ. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet* 2011; 44: 53-57. 2011/12/14. DOI: 10.1038/ng.1031.

141. Papaemmanuil E, Cazzola M, Boultonwood J, Malcovati L, Vyas P, Bowen D, Pellagatti A, Wainscoat JS, Hellstrom-Lindberg E, Gambacorti-Passerini C, Godfrey AL, Rapado I, Cvejic A, Rance R, McGee C, Ellis P, Mudie LJ, Stephens PJ, McLaren S, Massie CE, Tarpey PS, Varela I, Nik-Zainal S, Davies HR, Shlien A, Jones D, Raine K, Hinton J, Butler AP, Teague JW, Baxter EJ, Score J, Galli A, Della Porta MG, Travaglino E, Groves M, Tauro S, Munshi NC, Anderson KC, El-Naggar A, Fischer A, Mustonen V, Warren AJ, Cross NC, Green AR, Futreal PA, Stratton MR, Campbell PJ and Chronic Myeloid Disorders Working Group of the International Cancer Genome C. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med* 2011; 365: 1384-1395. 2011/10/15. DOI: 10.1056/NEJMoa1103283.

142. Dvinge H, Kim E, Abdel-Wahab O and Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. *Nat Rev Cancer* 2016; 16: 413-430. 2016/06/11. DOI: 10.1038/nrc.2016.51.

143. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, Yoon CJ, Ellis P, Wedge DC, Pellagatti A, Shlien A, Groves MJ, Forbes SA, Raine K, Hinton J, Mudie LJ, McLaren S, Hardy C, Latimer C, Della Porta MG, O'Meara S, Ambaglio I, Galli A, Butler AP, Walldin G, Teague JW, Quek L, Sternberg A, Gambacorti-Passerini C, Cross NC, Green AR, Boultonwood J, Vyas P, Hellstrom-Lindberg E, Bowen D, Cazzola M, Stratton MR, Campbell PJ and Chronic Myeloid Disorders Working Group of the International Cancer Genome C. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* 2013; 122: 3616-3627; quiz 3699. 2013/09/14. DOI: 10.1182/blood-2013-08-518886.

144. Kim E, Ilagan JO, Liang Y, Daubner GM, Lee SC, Ramakrishnan A, Li Y, Chung YR, Micol JB, Murphy ME, Cho H, Kim MK, Zebari AS, Aumann S, Park CY, Buonamici S, Smith PG, Deeg HJ, Lobry C, Aifantis I, Modis Y, Allain FH, Halene S, Bradley RK and Abdel-Wahab O.

- SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer cell* 2015; 27: 617-630. 2015/05/13. DOI: 10.1016/j.ccell.2015.04.006.
145. Shirai CL, Ley JN, White BS, Kim S, Tibbitts J, Shao J, Ndonwi M, Wadugu B, Duncavage EJ, Okeyo-Owuor T, Liu T, Griffith M, McGrath S, Magrini V, Fulton RS, Fronick C, O'Laughlin M, Graubert TA and Walter MJ. Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. *Cancer cell* 2015; 27: 631-643. 2015/05/13. DOI: 10.1016/j.ccell.2015.04.008.
146. Dvinge H and Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. *Genome Med* 2015; 7: 45. 2015/06/27. DOI: 10.1186/s13073-015-0168-9.
147. Cha JY, Lambert QT, Reuther GW and Der CJ. Involvement of fibroblast growth factor receptor 2 isoform switching in mammary oncogenesis. *Mol Cancer Res* 2008; 6: 435-445. 2008/03/14. DOI: 10.1158/1541-7786.MCR-07-0187.
148. Surget S, Khoury MP and Bourdon JC. Uncovering the role of p53 splice variants in human malignancy: a clinical perspective. *Onco Targets Ther* 2013; 7: 57-68. 2014/01/01. DOI: 10.2147/OTT.S53876.
149. Ghanem LR, Kromer A, Silverman IM, Ji X, Gazzara M, Nguyen N, Aguilar G, Martinelli M, Barash Y and Liebhaber SA. Poly(C)-Binding Protein Pcbp2 Enables Differentiation of Definitive Erythropoiesis by Directing Functional Splicing of the Runx1 Transcript. *Molecular and cellular biology* 2018; 38 2018/06/06. DOI: 10.1128/MCB.00175-18.
150. Liu L, Luo C, Luo Y, Chen L, Liu Y, Wang Y, Han J, Zhang Y, Wei N, Xie Z, Wu W, Wu G and Feng Y. MRPL33 and its splicing regulator hnRNPK are required for mitochondria function and implicated in tumor progression. *Oncogene* 2018; 37: 86-94. 2017/09/05. DOI: 10.1038/onc.2017.314.
151. Peng WZ, Liu JX, Li CF, Ma R and Jie JZ. hnRNPK promotes gastric tumorigenesis through regulating CD44E alternative splicing. *Cancer Cell Int* 2019; 19: 335. 2019/12/21. DOI: 10.1186/s12935-019-1020-x.

152. Hong X, Song R, Song H, Zheng T, Wang J, Liang Y, Qi S, Lu Z, Song X, Jiang H, Liu L and Zhang Z. PTEN antagonises Tc1/hnRNP-mediated G6PD pre-mRNA splicing which contributes to hepatocarcinogenesis. *Gut* 2014; 63: 1635-1647. 2013/12/20. DOI: 10.1136/gutjnl-2013-305302.
153. Dejgaard K and Leffers H. Characterisation of the nucleic-acid-binding activity of KH domains. Different properties of different domains. *Eur J Biochem* 1996; 241: 425-431. 1996/10/15. DOI: 10.1111/j.1432-1033.1996.00425.x.
154. Thompson MG, Munoz-Moreno R, Bhat P, Roytenberg R, Lindberg J, Gazzara MR, Mallory MJ, Zhang K, Garcia-Sastre A, Fontoura BMA and Lynch KW. Co-regulatory activity of hnRNP K and NS1-BP in influenza and human mRNA splicing. *Nat Commun* 2018; 9: 2407. DOI: 10.1038/s41467-018-04779-4.
155. Nanjundan M, Zhang F, Schmandt R, Smith-McCune K and Mills GB. Identification of a novel splice variant of AML1b in ovarian cancer patients conferring loss of wild-type tumor suppressive functions. *Oncogene* 2007; 26: 2574-2584. 2006/10/31. DOI: 10.1038/sj.onc.1210067.
156. Ok CY, Patel KP, Garcia-Manero G, Routbort MJ, Peng J, Tang G, Goswami M, Young KH, Singh R, Medeiros LJ, Kantarjian HM, Luthra R and Wang SA. TP53 mutation characteristics in therapy-related myelodysplastic syndromes and acute myeloid leukemia is similar to de novo diseases. *J Hematol Oncol* 2015; 8: 45. 2015/05/09. DOI: 10.1186/s13045-015-0139-z.
157. Sambrook J and Russell DW. Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protoc* 2006; 2006 2006/01/01. DOI: 10.1101/pdb.prot4455.
158. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* 2001; 29: e45. 2001/05/09. DOI: 10.1093/nar/29.9.e45.
159. Tibes R, Qiu Y, Lu Y, Hennessy B, Andreeff M, Mills GB and Kornblau SM. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary

- leukemia specimens and hematopoietic stem cells. *Mol Cancer Ther* 2006; 5: 2512-2521. 2006/10/17. DOI: 10.1158/1535-7163.MCT-06-0334.
160. Kornblau SM, Tibes R, Qiu YH, Chen W, Kantarjian HM, Andreeff M, Coombes KR and Mills GB. Functional proteomic profiling of AML predicts response and survival. *Blood* 2009; 113: 154-164. 2008/10/09. DOI: 10.1182/blood-2007-10-119438.
161. Hoff FW, Hu CW, Qutub AA, Qiu Y, Hornbaker MJ, Bueso-Ramos C, Abbas HA, Post SM, de Bont E and Kornblau SM. Proteomic Profiling of Acute Promyelocytic Leukemia Identifies Two Protein Signatures Associated with Relapse. *Proteomics Clin Appl* 2019: e1800133. 2019/01/17. DOI: 10.1002/prca.201800133.
162. Hu CW, Qiu Y, Ligeralde A, Raybon AY, Yoo SY, Coombes KR, Qutub AA and Kornblau SM. A quantitative analysis of heterogeneities and hallmarks in acute myelogenous leukaemia. *Nat Biomed Eng* 2019; 3: 889-901. 2019/04/17. DOI: 10.1038/s41551-019-0387-2.
163. Hu J, He X, Baggerly KA, Coombes KR, Hennessy BT and Mills GB. Non-parametric quantification of protein lysate arrays. *Bioinformatics* 2007; 23: 1986-1994. 2007/06/30. DOI: 10.1093/bioinformatics/btm283.
164. Neeley ES, Baggerly KA and Kornblau SM. Surface Adjustment of Reverse Phase Protein Arrays using Positive Control Spots. *Cancer Inform* 2012; 11: 77-86. 2012/05/03. DOI: 10.4137/CIN.S9055.
165. Neeley ES, Kornblau SM, Coombes KR and Baggerly KA. Variable slope normalization of reverse phase protein arrays. *Bioinformatics* 2009; 25: 1384-1389. 2009/04/02. DOI: 10.1093/bioinformatics/btp174.
166. Liehr T, Claussen U and Starke H. Small supernumerary marker chromosomes (sSMC) in humans. *Cytogenet Genome Res* 2004; 107: 55-67. 2004/08/12. DOI: 10.1159/000079572.
167. Au PYB, You J, Caluseriu O, Schwartzentruber J, Majewski J, Bernier FP, Ferguson M, Care for Rare Canada C, Valle D, Parboosingh JS, Sobreira N, Innes AM and Kline AD. GeneMatcher aids in the identification of a new malformation syndrome with intellectual disability,

- unique facial dysmorphisms, and skeletal and connective tissue abnormalities caused by de novo variants in HNRNPK. *Hum Mutat* 2015; 36: 1009-1014. 2015/07/16. DOI: 10.1002/humu.22837.
168. Hahn CN, Venugopal P, Scott HS and Hiwase DK. Splice factor mutations and alternative splicing as drivers of hematopoietic malignancy. *Immunol Rev* 2015; 263: 257-278. 2014/12/17. DOI: 10.1111/imr.12241.
169. Tsai PL, Chiou NT, Kuss S, Garcia-Sastre A, Lynch KW and Fontoura BM. Cellular RNA binding proteins NS1-BP and hnRNP K regulate influenza A virus RNA splicing. *PLoS Pathog* 2013; 9: e1003460. 2013/07/05. DOI: 10.1371/journal.ppat.1003460.
170. Prelich G. Gene overexpression: uses, mechanisms, and interpretation. *Genetics* 2012; 190: 841-854. 2012/03/16. DOI: 10.1534/genetics.111.136911.
171. Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, McConkey ME, Habib N, Yosef N, Chang CY, Shay T, Frampton GM, Drake AC, Leskov I, Nilsson B, Preffer F, Dombkowski D, Evans JW, Liefeld T, Smutko JS, Chen J, Friedman N, Young RA, Golub TR, Regev A and Ebert BL. Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 2011; 144: 296-309. 2011/01/19. DOI: 10.1016/j.cell.2011.01.004.
172. Bochtler T, Granzow M, Stolzel F, Kunz C, Mohr B, Kartal-Kaess M, Hinderhofer K, Heilig CE, Kramer M, Thiede C, Endris V, Kirchner M, Stenzinger A, Benner A, Bornhauser M, Ehninger G, Ho AD, Jauch A, Kramer A and Study Alliance Leukemia I. Marker chromosomes can arise from chromothripsis and predict adverse prognosis in acute myeloid leukemia. *Blood* 2017; 129: 1333-1342. 2017/01/26. DOI: 10.1182/blood-2016-09-738161.
173. Kwong YL, Wong KF, Chan V and Chan CH. Persistence of AML1 rearrangement in peripheral blood cells in t(8;21). *Cancer Genet Cytogenet* 1996; 88: 151-154. 1996/06/01. DOI: 10.1016/0165-4608(95)00282-0.
174. Miyamoto T, Nagafuji K, Akashi K, Harada M, Kyo T, Akashi T, Takenaka K, Mizuno S, Gondo H, Okamura T, Dohy H and Niho Y. Persistence of multipotent progenitors expressing

AML1/ETO transcripts in long-term remission patients with t(8;21) acute myelogenous leukemia. *Blood* 1996; 87: 4789-4796. 1996/06/01.

175. Zuna J, Burjanivova T, Mejstrikova E, Zemanova Z, Muzikova K, Meyer C, Horsley SW, Kearney L, Colman S, Ptoszkova H, Marschalek R, Hrusak O, Stary J, Greaves M and Trka J. Covert preleukemia driven by MLL gene fusion. *Genes Chromosomes Cancer* 2009; 48: 98-107. 2008/10/22. DOI: 10.1002/gcc.20622.

176. Ema H and Nakauchi H. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood* 2000; 95: 2284-2288. 2000/03/25.

177. Morrison SJ, Hemmati HD, Wandycz AM and Weissman IL. The purification and characterization of fetal liver hematopoietic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 1995; 92: 10302-10306.

178. Gekas C, Dieterlen-Lievre F, Orkin SH and Mikkola HK. The placenta is a niche for hematopoietic stem cells. *Dev Cell* 2005; 8: 365-375. 2005/03/02. DOI: 10.1016/j.devcel.2004.12.016.

179. Mikkola HK and Orkin SH. The journey of developing hematopoietic stem cells. *Development* 2006; 133: 3733-3744. 2006/09/14. DOI: 10.1242/dev.02568.

180. Zuber J, Radtke I, Pardee TS, Zhao Z, Rappaport AR, Luo W, McCurrach ME, Yang MM, Dolan ME, Kogan SC, Downing JR and Lowe SW. Mouse models of human AML accurately predict chemotherapy response. *Genes & development* 2009; 23: 877-889. DOI: 10.1101/gad.1771409.

181. Schmitt CA, Fridman JS, Yang M, Baranov E, Hoffman RM and Lowe SW. Dissecting p53 tumor suppressor functions in vivo. *Cancer cell* 2002; 1: 289-298. 2002/06/28. DOI: 10.1016/s1535-6108(02)00047-8.

182. Safran M, Kim WY, O'Connell F, Flippin L, Gunzler V, Horner JW, Depinho RA and Kaelin WG, Jr. Mouse model for noninvasive imaging of HIF prolyl hydroxylase activity: assessment of an oral agent that stimulates erythropoietin production. *Proceedings of the National Academy of*

- Sciences of the United States of America* 2006; 103: 105-110. 2005/12/24. DOI: 10.1073/pnas.0509459103.
183. Naviaux RK, Costanzi E, Haas M and Verma IM. The pCL vector system: rapid production of helper-free, high-titer, recombinant retroviruses. *J Virol* 1996; 70: 5701-5705. 1996/08/01.
184. Bock TA. Assay systems for hematopoietic stem and progenitor cells. *Stem Cells* 1997; 15 Suppl 1: 185-195. 1997/01/01. DOI: 10.1002/stem.5530150824.
185. Okada S, Nakauchi H, Nagayoshi K, Nishikawa S, Nishikawa S, Miura Y and Suda T. Enrichment and characterization of murine hematopoietic stem cells that express c-kit molecule. *Blood* 1991; 78: 1706-1712. 1991/10/01.
186. Spangrude GJ, Heimfeld S and Weissman IL. Purification and characterization of mouse hematopoietic stem cells. *Science* 1988; 241: 58-62. 1988/07/01. DOI: 10.1126/science.2898810.
187. Akashi K, Traver D, Miyamoto T and Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 2000; 404: 193-197. 2000/03/21. DOI: 10.1038/35004599.
188. Kondo M, Weissman IL and Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* 1997; 91: 661-672. 1997/12/11. DOI: 10.1016/s0092-8674(00)80453-5.
189. Brenner AK, Aasebo E, Hernandez-Valladares M, Selheim F, Berven F, Gronningsaeter IS, Bartaula-Brevik S and Bruserud O. The Capacity of Long-Term in Vitro Proliferation of Acute Myeloid Leukemia Cells Supported Only by Exogenous Cytokines Is Associated with a Patient Subset with Adverse Outcome. *Cancers (Basel)* 2019; 11 2019/01/13. DOI: 10.3390/cancers11010073.
190. Chiu SC, Liu HH, Chen CL, Chen PR, Liu MC, Lin SZ and Chang KT. Extramedullary hematopoiesis (EMH) in laboratory animals: offering an insight into stem cell research. *Cell Transplant* 2015; 24: 349-366. 2015/02/04. DOI: 10.3727/096368915X686850.

191. Lessard-Beaudoin M, Laroche M, Demers MJ, Grenier G and Graham RK. Characterization of age-associated changes in peripheral organ and brain region weights in C57BL/6 mice. *Exp Gerontol* 2015; 63: 27-34. 2015/01/20. DOI: 10.1016/j.exger.2015.01.003.
192. Tobler A, Miller CW, Johnson KR, Selsted ME, Rovera G and Koeffler HP. Regulation of gene expression of myeloperoxidase during myeloid differentiation. *J Cell Physiol* 1988; 136: 215-225. 1988/08/01. DOI: 10.1002/jcp.1041360203.
193. Shultz LD, Lyons BL, Burzenski LM, Gott B, Chen X, Chaleff S, Kotb M, Gillies SD, King M, Mangada J, Greiner DL and Handgretinger R. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol* 2005; 174: 6477-6489. 2005/05/10. DOI: 10.4049/jimmunol.174.10.6477.
194. Shultz LD, Schweitzer PA, Christianson SW, Gott B, Schweitzer IB, Tennent B, McKenna S, Mobraaten L, Rajan TV, Greiner DL and et al. Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. *J Immunol* 1995; 154: 180-191. 1995/01/01.
195. Song HK and Hwang DY. Use of C57BL/6N mice on the variety of immunological researches. *Lab Anim Res* 2017; 33: 119-123. 2017/07/28. DOI: 10.5625/lar.2017.33.2.119.
196. Klevorn LE and Teague RM. Adapting Cancer Immunotherapy Models for the Real World. *Trends Immunol* 2016; 37: 354-363. 2016/04/24. DOI: 10.1016/j.it.2016.03.010.
197. Verhaak RG, Wouters BJ, Erpelinck CA, Abbas S, Beverloo HB, Lugthart S, Lowenberg B, Delwel R and Valk PJ. Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. *Haematologica* 2009; 94: 131-134. 2008/10/08. DOI: 10.3324/haematol.13299.
198. Gotlib J. World Health Organization-defined eosinophilic disorders: 2017 update on diagnosis, risk stratification, and management. *Am J Hematol* 2017; 92: 1243-1259. 2017/10/19. DOI: 10.1002/ajh.24880.

199. Simon HU, Plotz SG, Dummer R and Blaser K. Abnormal clones of T cells producing interleukin-5 in idiopathic eosinophilia. *N Engl J Med* 1999; 341: 1112-1120. 1999/10/08. DOI: 10.1056/NEJM1999100734111503.
200. Larson RA, Williams SF, Le Beau MM, Bitter MA, Vardiman JW and Rowley JD. Acute myelomonocytic leukemia with abnormal eosinophils and inv(16) or t(16;16) has a favorable prognosis. *Blood* 1986; 68: 1242-1249. 1986/12/01.
201. Loffler H, Gassmann W and Haferlach T. AML M1 and M2 with eosinophilia and AML M4Eo: diagnostic and clinical aspects. *Leuk Lymphoma* 1995; 18 Suppl 1: 61-63. 1995/01/01. DOI: 10.3109/10428199509075305.
202. Frolova O, Samudio I, Benito JM, Jacamo R, Kornblau SM, Markovic A, Schober W, Lu H, Qiu YH, Buglio D, McQueen T, Pierce S, Shpall E, Konoplev S, Thomas D, Kantarjian H, Lock R, Andreeff M and Konopleva M. Regulation of HIF-1alpha signaling and chemoresistance in acute lymphocytic leukemia under hypoxic conditions of the bone marrow microenvironment. *Cancer Biol Ther* 2012; 13: 858-870. 2012/07/13. DOI: 10.4161/cbt.20838.
203. Cox B and Emili A. Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nat Protoc* 2006; 1: 1872-1878. 2007/05/10. DOI: 10.1038/nprot.2006.273.
204. Bairoch A and Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic acids research* 2000; 28: 45-48. 1999/12/11. DOI: 10.1093/nar/28.1.45.
205. Perkins DN, Pappin DJ, Creasy DM and Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999; 20: 3551-3567. 1999/12/28. DOI: 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2.

206. Eng JK, McCormack AL and Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 1994; 5: 976-989. 1994/11/01. DOI: 10.1016/1044-0305(94)80016-2.
207. D GH, Kelley DR, Tenen D, Bernstein B and Rinn JL. Widespread RNA binding by chromatin-associated proteins. *Genome Biol* 2016; 17: 28. 2016/02/18. DOI: 10.1186/s13059-016-0878-3.
208. Dodt M, Roehr JT, Ahmed R and Dieterich C. FLEXBAR-Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology (Basel)* 2012; 1: 895-905. 2012/01/01. DOI: 10.3390/biology1030895.
209. Langmead B and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; 9: 357-359. 2012/03/06. DOI: 10.1038/nmeth.1923.
210. Schneider VA, Graves-Lindsay T, Howe K, Bouk N, Chen HC, Kitts PA, Murphy TD, Pruitt KD, Thibaud-Nissen F, Albracht D, Fulton RS, Kremitzki M, Magrini V, Markovic C, McGrath S, Steinberg KM, Auger K, Chow W, Collins J, Harden G, Hubbard T, Pelan S, Simpson JT, Threadgold G, Torrance J, Wood JM, Clarke L, Koren S, Boitano M, Peluso P, Li H, Chin CS, Phillippy AM, Durbin R, Wilson RK, Flicek P, Eichler EE and Church DM. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res* 2017; 27: 849-864. 2017/04/12. DOI: 10.1101/gr.213611.116.
211. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R and Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; 25: 2078-2079. 2009/06/10. DOI: 10.1093/bioinformatics/btp352.
212. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 2010; 28: 511-515. 2010/05/04. DOI: 10.1038/nbt.1621.

213. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL and Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 2013; 31: 46-53. 2012/12/12. DOI: 10.1038/nbt.2450.
214. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G and Mesirov JP. Integrative genomics viewer. *Nat Biotechnol* 2011; 29: 24-26. 2011/01/12. DOI: 10.1038/nbt.1754.
215. Sadelain M, Papapetrou EP and Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer* 2011; 12: 51-58. 2011/12/02. DOI: 10.1038/nrc3179.
216. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, Kok CY, Jia M, Jubb H, Sondka Z, Thompson S, De T and Campbell PJ. COSMIC: somatic cancer genetics at high-resolution. *Nucleic acids research* 2017; 45: D777-D783. 2016/12/03. DOI: 10.1093/nar/gkw1121.
217. Paziewska A, Wyrwicz LS, Bujnicki JM, Bomszyk K and Ostrowski J. Cooperative binding of the hnRNP K three KH domains to mRNA targets. *FEBS Lett* 2004; 577: 134-140. DOI: 10.1016/j.febslet.2004.08.086.
218. Aslanidis C and de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic acids research* 1990; 18: 6069-6074. 1990/10/25. DOI: 10.1093/nar/18.20.6069.
219. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM and Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; 25: 25-29. 2000/05/10. DOI: 10.1038/75556.
220. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498-2504. 2003/11/05. DOI: 10.1101/gr.1239303.

221. Sood R, Kamikubo Y and Liu P. Role of RUNX1 in hematological malignancies. *Blood* 2017; 129: 2070-2082. 2017/02/10. DOI: 10.1182/blood-2016-10-687830.
222. Swanson MS and Dreyfuss G. Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. *Molecular and cellular biology* 1988; 8: 2237-2241. 1988/05/01. DOI: 10.1128/mcb.8.5.2237.
223. Anderson BJ, Larkin C, Guja K and Schildbach JF. Using fluorophore-labeled oligonucleotides to measure affinities of protein-DNA interactions. *Methods Enzymol* 2008; 450: 253-272. 2009/01/21. DOI: 10.1016/S0076-6879(08)03412-5.
224. Pagano JM, Clingman CC and Ryder SP. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* 2011; 17: 14-20. 2010/11/26. DOI: 10.1261/rna.2428111.
225. Huynh K and Partch CL. Analysis of protein stability and ligand interactions by thermal shift assay. *Curr Protoc Protein Sci* 2015; 79: 28 29 21-28 29 14. 2015/02/03. DOI: 10.1002/0471140864.ps2809s79.
226. Bomsztyk K, Denisenko O and Ostrowski J. hnRNP K: one protein multiple processes. *Bioessays* 2004; 26: 629-638. DOI: 10.1002/bies.20048.
227. Mikula M, Dzwonek A, Karczmarski J, Rubel T, Dadlez M, Wyrwicz LS, Bomsztyk K and Ostrowski J. Landscape of the hnRNP K protein-protein interactome. *Proteomics* 2006; 6: 2395-2406. 2006/03/07. DOI: 10.1002/pmic.200500632.
228. Bellissimo DC and Speck NA. RUNX1 Mutations in Inherited and Sporadic Leukemia. *Front Cell Dev Biol* 2017; 5: 111. 2018/01/13. DOI: 10.3389/fcell.2017.00111.
229. Sakurai H, Harada Y, Ogata Y, Kagiya Y, Shingai N, Doki N, Ohashi K, Kitamura T, Komatsu N and Harada H. Overexpression of RUNX1 short isoform has an important role in the development of myelodysplastic/myeloproliferative neoplasms. *Blood Adv* 2017; 1: 1382-1386. 2018/01/04. DOI: 10.1182/bloodadvances.2016002725.

230. Fu L, Fu H, Tian L, Xu K, Hu K, Wang J, Wang J, Jing H, Shi J and Ke X. High expression of RUNX1 is associated with poorer outcomes in cytogenetically normal acute myeloid leukemia. *Oncotarget* 2016; 7: 15828-15839. DOI: 10.18632/oncotarget.7489.
231. Behrens K, Maul K, Tekin N, Kriebitzsch N, Indenbirken D, Prassolov V, Muller U, Serve H, Cammenga J and Stocking C. RUNX1 cooperates with FLT3-ITD to induce leukemia. *J Exp Med* 2017; 214: 737-752. 2017/02/19. DOI: 10.1084/jem.20160927.
232. Malik N, Yan H, Moshkovich N, Palangat M, Yang H, Sanchez V, Cai Z, Peat TJ, Jiang S, Liu C, Lee M, Mock BA, Yuspa SH, Larson D, Wakefield LM and Huang J. The transcription factor CFBFB suppresses breast cancer through orchestrating translation and transcription. *Nat Commun* 2019; 10: 2071. 2019/05/08. DOI: 10.1038/s41467-019-10102-6.
233. Dvinge H. Regulation of alternative mRNA splicing: old players and new perspectives. *FEBS Lett* 2018; 592: 2987-3006. 2018/06/02. DOI: 10.1002/1873-3468.13119.
234. Scotti MM and Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet* 2016; 17: 19-32. 2015/11/26. DOI: 10.1038/nrg.2015.3.
235. Clower CV, Chatterjee D, Wang Z, Cantley LC, Vander Heiden MG and Krainer AR. The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism. *Proceedings of the National Academy of Sciences of the United States of America* 2010; 107: 1894-1899. 2010/02/06. DOI: 10.1073/pnas.0914845107.
236. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, Potter NE, Heuser M, Thol F, Bolli N, Gundem G, Van Loo P, Martincorena I, Ganly P, Mudie L, McLaren S, O'Meara S, Raine K, Jones DR, Teague JW, Butler AP, Greaves MF, Ganser A, Dohner K, Schlenk RF, Dohner H and Campbell PJ. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* 2016; 374: 2209-2221. 2016/06/09. DOI: 10.1056/NEJMoa1516192.
237. Saez B, Walter MJ and Graubert TA. Splicing factor gene mutations in hematologic malignancies. *Blood* 2017; 129: 1260-1269. 2016/12/13. DOI: 10.1182/blood-2016-10-692400.

238. Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, Sabatini DM, Chen IS, Hahn WC, Sharp PA, Weinberg RA and Novina CD. Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* 2003; 9: 493-501. 2003/03/22. DOI: 10.1261/ma.2192803.
239. Bray NL, Pimentel H, Melsted P and Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016; 34: 525-527. 2016/04/05. DOI: 10.1038/nbt.3519.
240. Pimentel H, Bray NL, Puente S, Melsted P and Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods* 2017; 14: 687-690. 2017/06/06. DOI: 10.1038/nmeth.4324.
241. Sanjabi S, Williams KJ, Sacconi S, Zhou L, Hoffmann A, Ghosh G, Gerondakis S, Natoli G and Smale ST. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. *Genes & development* 2005; 19: 2138-2151. 2005/09/17. DOI: 10.1101/gad.1329805.
242. Ho SN, Hunt HD, Horton RM, Pullen JK and Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 1989; 77: 51-59. 1989/04/15. DOI: 10.1016/0378-1119(89)90358-2.
243. el-Deiry WS, Tokino T, Velculescu VE, Levy DB, Parsons R, Trent JM, Lin D, Mercer WE, Kinzler KW and Vogelstein B. WAF1, a potential mediator of p53 tumor suppression. *Cell* 1993; 75: 817-825. 1993/11/19. DOI: 10.1016/0092-8674(93)90500-p.
244. Panda AC, Martindale JL and Gorospe M. Polysome Fractionation to Analyze mRNA Distribution Profiles. *Bio Protoc* 2017; 7 2017/05/19. DOI: 10.21769/BioProtoc.2126.
245. Obrig TG, Culp WJ, McKeehan WL and Hardesty B. The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *The Journal of biological chemistry* 1971; 246: 174-181. 1971/01/10.
246. Ennis HL and Lubin M. Cycloheximide: Aspects of Inhibition of Protein Synthesis in Mammalian Cells. *Science* 1964; 146: 1474-1476. 1964/12/11. DOI: 10.1126/science.146.3650.1474.

247. Buchanan BW, Lloyd ME, Engle SM and Rubenstein EM. Cycloheximide Chase Analysis of Protein Degradation in *Saccharomyces cerevisiae*. *J Vis Exp* 2016 2016/05/12. DOI: 10.3791/53975.
248. Biggs JR, Peterson LF, Zhang Y, Kraft AS and Zhang DE. AML1/RUNX1 phosphorylation by cyclin-dependent kinases regulates the degradation of AML1/RUNX1 by the anaphase-promoting complex. *Molecular and cellular biology* 2006; 26: 7420-7429. 2006/10/04. DOI: 10.1128/MCB.00597-06.
249. Warner JR and Knopf PM. The discovery of polyribosomes. *Trends Biochem Sci* 2002; 27: 376-380. 2002/07/13. DOI: 10.1016/s0968-0004(02)02126-6.
250. Van Der Kelen K, Beyaert R, Inze D and De Veylder L. Translational control of eukaryotic gene expression. *Crit Rev Biochem Mol Biol* 2009; 44: 143-168. 2009/07/17. DOI: 10.1080/10409230902882090.
251. Aspden JL, Eyre-Walker YC, Phillips RJ, Amin U, Mumtaz MA, Brocard M and Couso JP. Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. *Elife* 2014; 3: e03528. 2014/08/22. DOI: 10.7554/eLife.03528.
252. Zhao X, Jankovic V, Gural A, Huang G, Pardanani A, Menendez S, Zhang J, Dunne R, Xiao A, Erdjument-Bromage H, Allis CD, Tempst P and Nimer SD. Methylation of RUNX1 by PRMT1 abrogates SIN3A binding and potentiates its transcriptional activity. *Genes & development* 2008; 22: 640-653. 2008/03/05. DOI: 10.1101/gad.1632608.
253. Huang H, Woo AJ, Waldon Z, Schindler Y, Moran TB, Zhu HH, Feng GS, Steen H and Cantor AB. A Src family kinase-Shp2 axis controls RUNX1 activity in megakaryocyte and T-lymphocyte differentiation. *Genes & development* 2012; 26: 1587-1601. 2012/07/05. DOI: 10.1101/gad.192054.112.
254. Baber JL, Levens D, Libutti D and Tjandra N. Chemical shift mapped DNA-binding sites and ¹⁵N relaxation analysis of the C-terminal KH domain of heterogeneous nuclear ribonucleoprotein K. *Biochemistry* 2000; 39: 6022-6032. 2000/05/23. DOI: 10.1021/bi000105e.

255. Backe PH, Messias AC, Ravelli RB, Sattler M and Cusack S. X-ray crystallographic and NMR studies of the third KH domain of hnRNP K in complex with single-stranded nucleic acids. *Structure* 2005; 13: 1055-1067. 2005/07/12. DOI: 10.1016/j.str.2005.04.008.
256. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, Bell JT, Yang TP, Meduri E, Barrett A, Nisbett J, Sekowska M, Wilk A, Shin SY, Glass D, Travers M, Min JL, Ring S, Ho K, Thorleifsson G, Kong A, Thorsteindottir U, Ainali C, Dimas AS, Hassanali N, Ingle C, Knowles D, Krestyaninova M, Lowe CE, Di Meglio P, Montgomery SB, Parts L, Potter S, Surdulescu G, Tsaprouni L, Tsoka S, Bataille V, Durbin R, Nestle FO, O'Rahilly S, Soranzo N, Lindgren CM, Zondervan KT, Ahmadi KR, Schadt EE, Stefansson K, Smith GD, McCarthy MI, Deloukas P, Dermitzakis ET, Spector TD and Multiple Tissue Human Expression Resource C. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* 2012; 44: 1084-1089. 2012/09/04. DOI: 10.1038/ng.2394.
257. Naarmann-de Vries IS, Sackmann Y, Klein F, Ostareck-Lederer A, Ostareck DH, Jost E, Ehninger G, Brummendorf TH, Marx G, Rollig C, Thiede C and Crysandt M. Characterization of acute myeloid leukemia with del(9q) - Impact of the genes in the minimally deleted region. *Leuk Res* 2019; 76: 15-23. 2018/11/27. DOI: 10.1016/j.leukres.2018.11.007.
258. Ha G, Roth A, Lai D, Bashashati A, Ding J, Goya R, Giuliany R, Rosner J, Oloumi A, Shumansky K, Chin SF, Turashvili G, Hirst M, Caldas C, Marra MA, Aparicio S and Shah SP. Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. *Genome Res* 2012; 22: 1995-2007. 2012/05/29. DOI: 10.1101/gr.137570.112.
259. Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, Dou Y, Zhang Y, Shi Z, Arshad OA, Gritsenko MA, Zimmerman LJ, McDermott JE, Clauss TR, Moore RJ, Zhao R, Monroe ME, Wang YT, Chambers MC, Slebos RJC, Lau KS, Mo Q, Ding L, Ellis M, Thiagarajan M, Kinsinger CR, Rodriguez H, Smith RD, Rodland KD, Liebler DC, Liu T, Zhang B and Clinical Proteomic Tumor Analysis C. Proteogenomic Analysis of Human Colon Cancer Reveals New

- Therapeutic Opportunities. *Cell* 2019; 177: 1035-1049 e1019. 2019/04/30. DOI: 10.1016/j.cell.2019.03.030.
260. Lange L, Pagnamenta AT, Lise S, Clasper S, Stewart H, Akha ES, Quaghebeur G, Knight SJ, Keays DA, Taylor JC and Kini U. A de novo frameshift in HNRNPK causing a Kabuki-like syndrome with nodular heterotopia. *Clin Genet* 2016; 90: 258-262. 2016/03/10. DOI: 10.1111/cge.12773.
261. Matsuura S, Komeno Y, Stevenson KE, Biggs JR, Lam K, Tang T, Lo MC, Cong X, Yan M, Neuberg DS and Zhang DE. Expression of the runt homology domain of RUNX1 disrupts homeostasis of hematopoietic stem cells and induces progression to myelodysplastic syndrome. *Blood* 2012; 120: 4028-4037. 2012/08/25. DOI: 10.1182/blood-2012-01-404533.
262. Rhoades KL, Hetherington CJ, Harakawa N, Yergeau DA, Zhou L, Liu LQ, Little MT, Tenen DG and Zhang DE. Analysis of the role of AML1-ETO in leukemogenesis, using an inducible transgenic mouse model. *Blood* 2000; 96: 2108-2115. 2000/09/09.
263. Backstrom S, Wolf-Watz M, Grundstrom C, Hard T, Grundstrom T and Sauer UH. The RUNX1 Runt domain at 1.25A resolution: a structural switch and specifically bound chloride ions modulate DNA binding. *J Mol Biol* 2002; 322: 259-272. 2002/09/10. DOI: 10.1016/s0022-2836(02)00702-7.
264. Gardini A, Cesaroni M, Luzi L, Okumura AJ, Biggs JR, Minardi SP, Venturini E, Zhang DE, Pelicci PG and Alcalay M. AML1/ETO oncoprotein is directed to AML1 binding regions and co-localizes with AML1 and HEB on its targets. *PLoS Genet* 2008; 4: e1000275. 2008/12/02. DOI: 10.1371/journal.pgen.1000275.
265. Ptasinska A, Assi SA, Mannari D, James SR, Williamson D, Dunne J, Hoogenkamp M, Wu M, Care M, McNeill H, Cauchy P, Cullen M, Tooze RM, Tenen DG, Young BD, Cockerill PN, Westhead DR, Heidenreich O and Bonifer C. Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. *Leukemia* 2012; 26: 1829-1841. 2012/02/22. DOI: 10.1038/leu.2012.49.

266. Zhang Y, Strissel P, Strick R, Chen J, Nucifora G, Le Beau MM, Larson RA and Rowley JD. Genomic DNA breakpoints in AML1/RUNX1 and ETO cluster with topoisomerase II DNA cleavage and DNase I hypersensitive sites in t(8;21) leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 2002; 99: 3070-3075. 2002/02/28. DOI: 10.1073/pnas.042702899.
267. Xiao Z, Greaves MF, Buffler P, Smith MT, Segal MR, Dicks BM, Wiencke JK and Wiemels JL. Molecular characterization of genomic AML1-ETO fusions in childhood leukemia. *Leukemia* 2001; 15: 1906-1913. 2001/12/26. DOI: 10.1038/sj.leu.2402318.
268. Solari L, Bauer T, Dicker F, Haferlach C, Griesshammer M, Schnittger S, Becker H and Lubbert M. A novel recurrent AML1-ETO fusion: tight in vivo association with BCR-ABL1. *Leukemia* 2013; 27: 1397-1400. 2013/02/22. DOI: 10.1038/leu.2013.53.
269. LaFiura KM, Edwards H, Taub JW, Matherly LH, Fontana JA, Mohamed AN, Ravindranath Y, Ge Y and Children's Oncology G. Identification and characterization of novel AML1-ETO fusion transcripts in pediatric t(8;21) acute myeloid leukemia: a report from the Children's Oncology Group. *Oncogene* 2008; 27: 4933-4942. 2008/05/13. DOI: 10.1038/onc.2008.134.
270. Faber ZJ, Chen X, Gedman AL, Boggs K, Cheng J, Ma J, Radtke I, Chao JR, Walsh MP, Song G, Andersson AK, Dang J, Dong L, Liu Y, Huether R, Cai Z, Mulder H, Wu G, Edmonson M, Rusch M, Qu C, Li Y, Vadodaria B, Wang J, Hedlund E, Cao X, Yergeau D, Nakitandwe J, Pounds SB, Shurtleff S, Fulton RS, Fulton LL, Easton J, Parganas E, Pui CH, Rubnitz JE, Ding L, Mardis ER, Wilson RK, Gruber TA, Mullighan CG, Schlenk RF, Paschka P, Dohner K, Dohner H, Bullinger L, Zhang J, Klco JM and Downing JR. The genomic landscape of core-binding factor acute myeloid leukemias. *Nat Genet* 2016; 48: 1551-1556. 2016/11/01. DOI: 10.1038/ng.3709.
271. Peterson LF, Boyapati A, Ahn EY, Biggs JR, Okumura AJ, Lo MC, Yan M and Zhang DE. Acute myeloid leukemia with the 8q22;21q22 translocation: secondary mutational events and

alternative t(8;21) transcripts. *Blood* 2007; 110: 799-805. 2007/04/07. DOI: 10.1182/blood-2006-11-019265.

272. Schoch C, Haase D, Haferlach T, Gudat H, Buchner T, Freund M, Link H, Lengfelder E, Wandt H, Sauerland MC, Löffler H and Fonatsch C. Fifty-one patients with acute myeloid leukemia and translocation t(8;21)(q22;q22): an additional deletion in 9q is an adverse prognostic factor. *Leukemia* 1996; 10: 1288-1295. 1996/08/01.

273. Peniket A, Wainscoat J, Side L, Daly S, Kusec R, Buck G, Wheatley K, Walker H, Chatters S, Harrison C, Boultonwood J, Goldstone A and Burnett A. Del (9q) AML: clinical and cytological characteristics and prognostic implications. *British journal of haematology* 2005; 129: 210-220. 2005/04/09. DOI: 10.1111/j.1365-2141.2005.05445.x.

274. Seiler M, Peng S, Agrawal AA, Palacino J, Teng T, Zhu P, Smith PG, Cancer Genome Atlas Research N, Buonamici S and Yu L. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep* 2018; 23: 282-296 e284. 2018/04/05. DOI: 10.1016/j.celrep.2018.01.088.

275. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, Schnittger S, Sanada M, Kon A, Alpermann T, Yoshida K, Roller A, Nadarajah N, Shiraishi Y, Shiozawa Y, Chiba K, Tanaka H, Koeffler HP, Klein HU, Dugas M, Aburatani H, Kohlmann A, Miyano S, Haferlach C, Kern W and Ogawa S. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* 2014; 28: 241-247. 2013/11/14. DOI: 10.1038/leu.2013.336.

276. Ohanian M, Rozovski U, Kanagal-Shamanna R, Abruzzo LV, Loghavi S, Kadia T, Futreal A, Bhalla K, Zuo Z, Huh YO, Post SM, Ruvolo P, Garcia-Manero G, Andreeff M, Kornblau S, Borthakur G, Hu P, Medeiros LJ, Takahashi K, Hornbaker MJ, Zhang J, Nogueras-Gonzalez GM, Huang X, Verstovsek S, Estrov Z, Pierce S, Ravandi F, Kantarjian HM, Bueso-Ramos CE and Cortes JE. MYC protein expression is an important prognostic factor in acute myeloid leukemia. *Leuk Lymphoma* 2019; 60: 37-48. 2018/05/10. DOI: 10.1080/10428194.2018.1464158.

277. Eskens FA, Ramos FJ, Burger H, O'Brien JP, Piera A, de Jonge MJ, Mizui Y, Wiemer EA, Carreras MJ, Baselga J and Tabernero J. Phase I pharmacokinetic and pharmacodynamic study of the first-in-class spliceosome inhibitor E7107 in patients with advanced solid tumors. *Clin Cancer Res* 2013; 19: 6296-6304. 2013/08/29. DOI: 10.1158/1078-0432.CCR-13-0485.
278. Steensma DP, Wermke M, Klimek VM, Greenberg PL, Font P, Komrokji RS, Yang J, Brunner AM, Carraway HE, Ades L, Al-Kali A, Alonso Dominguez JM, Alonso A, Coombs CC, Deeg HJ, Donnellan WB, Foran JM, Garcia-Manero G, Maris MB, McMasters M, Micol J-B, Perez De Oteyza J, Thol F, Wang ES, Watts JM, Buonamici S, Kim A, Gourineni V, Marino AJ, Rioux N, Schindler J, Smith S, Yao H, Yuan X, Yu K and Platzbecker U. Results of a Clinical Trial of H3B-8800, a Splicing Modulator, in Patients with Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML) or Chronic Myelomonocytic Leukemia (CMML). *Blood* 2019; 134: 673-673. DOI: 10.1182/blood-2019-123854.
279. Kaelin WG, Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 2005; 5: 689-698. 2005/08/20. DOI: 10.1038/nrc1691.
280. Boccitto M, Lee N, Sakamoto S, Spruce LA, Handa H, Clardy J, Seeholzer SH and Kalb RG. The Neuroprotective Marine Compound Psammaphysene A Binds the RNA-Binding Protein HNRNPK. *Mar Drugs* 2017; 15 2017/08/08. DOI: 10.3390/md15080246.
281. Marchand V, Santerre M, Aigueperse C, Fouillen L, Saliou JM, Van Dorselaer A, Sanglier-Cianferani S, Branlant C and Motorin Y. Identification of protein partners of the human immunodeficiency virus 1 tat/rev exon 3 leads to the discovery of a new HIV-1 splicing regulator, protein hnRNP K. *RNA Biol* 2011; 8: 325-342. 2011/03/04. DOI: 10.4161/ma.8.2.13984.
282. Lin JY, Li ML, Huang PN, Chien KY, Horng JT and Shih SR. Heterogeneous nuclear ribonuclear protein K interacts with the enterovirus 71 5' untranslated region and participates in virus replication. *J Gen Virol* 2008; 89: 2540-2549. 2008/09/18. DOI: 10.1099/vir.0.2008/003673-0.

283. Kanlaya R, Pattanakitsakul SN, Sinchaikul S, Chen ST and Thongboonkerd V. Vimentin interacts with heterogeneous nuclear ribonucleoproteins and dengue nonstructural protein 1 and is important for viral replication and release. *Mol Biosyst* 2010; 6: 795-806. 2010/06/23. DOI: 10.1039/b923864f.
284. Bourai M, Lucas-Hourani M, Gad HH, Drosten C, Jacob Y, Tafforeau L, Cassonnet P, Jones LM, Judith D, Couderc T, Lecuit M, Andre P, Kummerer BM, Lotteau V, Despres P, Tangy F and Vidalain PO. Mapping of Chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component. *J Virol* 2012; 86: 3121-3134. 2012/01/20. DOI: 10.1128/JVI.06390-11.
285. Burnham AJ, Gong L and Hardy RW. Heterogeneous nuclear ribonuclear protein K interacts with Sindbis virus nonstructural proteins and viral subgenomic mRNA. *Virology* 2007; 367: 212-221. 2007/06/15. DOI: 10.1016/j.virol.2007.05.008.
286. Ng LF, Chan M, Chan SH, Cheng PC, Leung EH, Chen WN and Ren EC. Host heterogeneous ribonucleoprotein K (hnRNP K) as a potential target to suppress hepatitis B virus replication. *PLoS Med* 2005; 2: e163. 2005/07/22. DOI: 10.1371/journal.pmed.0020163.
287. Poenisch M, Metz P, Blankenburg H, Ruggieri A, Lee JY, Rupp D, Rebhan I, Diederich K, Kaderali L, Domingues FS, Albrecht M, Lohmann V, Erfle H and Bartenschlager R. Identification of HNRNPK as regulator of hepatitis C virus particle production. *PLoS Pathog* 2015; 11: e1004573. 2015/01/09. DOI: 10.1371/journal.ppat.1004573.
288. Perz JF, Armstrong GL, Farrington LA, Hutin YJ and Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol* 2006; 45: 529-538. 2006/08/02. DOI: 10.1016/j.jhep.2006.05.013.
289. Jacobson IM, Davis GL, El-Serag H, Negro F and Trepo C. Prevalence and challenges of liver diseases in patients with chronic hepatitis C virus infection. *Clin Gastroenterol Hepatol* 2010; 8: 924-933; quiz e117. 2010/08/18. DOI: 10.1016/j.cgh.2010.06.032.

290. Schmidt T, Striebinger H, Haas J and Bailer SM. The heterogeneous nuclear ribonucleoprotein K is important for Herpes simplex virus-1 propagation. *FEBS Lett* 2010; 584: 4361-4365. 2010/10/05. DOI: 10.1016/j.febslet.2010.09.038.
291. Gross H, Hennard C, Masouris I, Cassel C, Barth S, Stober-Grasser U, Mamiani A, Moritz B, Ostareck D, Ostareck-Lederer A, Neuenkirchen N, Fischer U, Deng W, Leonhardt H, Noessner E, Kremmer E and Grasser FA. Binding of the heterogeneous ribonucleoprotein K (hnRNP K) to the Epstein-Barr virus nuclear antigen 2 (EBNA2) enhances viral LMP2A expression. *PloS one* 2012; 7: e42106. 2012/08/11. DOI: 10.1371/journal.pone.0042106.
292. van Domselaar R, de Poot SA, Remmerswaal EB, Lai KW, ten Berge IJ and Bovenschen N. Granzyme M targets host cell hnRNP K that is essential for human cytomegalovirus replication. *Cell Death Differ* 2013; 20: 419-429. 2012/10/27. DOI: 10.1038/cdd.2012.132.
293. Hernaez B, Escribano JM and Alonso C. African swine fever virus protein p30 interaction with heterogeneous nuclear ribonucleoprotein K (hnRNP-K) during infection. *FEBS Lett* 2008; 582: 3275-3280. 2008/09/09. DOI: 10.1016/j.febslet.2008.08.031.
294. Young LS, Yap LF and Murray PG. Epstein-Barr virus: more than 50 years old and still providing surprises. *Nat Rev Cancer* 2016; 16: 789-802. 2016/11/04. DOI: 10.1038/nrc.2016.92.
295. van Domselaar R, Quadir R, van der Made AM, Broekhuizen R and Bovenschen N. All human granzymes target hnRNP K that is essential for tumor cell viability. *The Journal of biological chemistry* 2012; 287: 22854-22864. 2012/05/15. DOI: 10.1074/jbc.M112.365692.
296. Hu W, Begum NA, Mondal S, Stanlie A and Honjo T. Identification of DNA cleavage- and recombination-specific hnRNP cofactors for activation-induced cytidine deaminase. *Proceedings of the National Academy of Sciences of the United States of America* 2015; 112: 5791-5796. 2015/04/23. DOI: 10.1073/pnas.1506167112.
297. Chang JW, Koike T and Iwashima M. hnRNP-K is a nuclear target of TCR-activated ERK and required for T-cell late activation. *Int Immunol* 2009; 21: 1351-1361. 2009/11/03. DOI: 10.1093/intimm/dxp106.

298. Liepelt A, Mossanen JC, Denecke B, Heymann F, De Santis R, Tacke F, Marx G, Ostareck DH and Ostareck-Lederer A. Translation control of TAK1 mRNA by hnRNP K modulates LPS-induced macrophage activation. *RNA* 2014; 20: 899-911. 2014/04/23. DOI: 10.1261/rna.042788.113.

Vita

Marisa Janelle Lynne Aitken was born Marisa Janelle Lynne Hornbaker in Colorado Springs, CO in January, 1990 to Charles and Connie Hornbaker. After graduating as valedictorian from Colorado Springs Christian School in 2007, she enrolled as a student-athlete (volleyball) at the University at Buffalo, The State University of New York. In May 2011, she graduated *summa cum laude* with a Bachelor's of Science in Pharmacology and Toxicology. She matriculated into the Medical Scientist Training Program (MSTP) at the McGovern Medical School/University of Texas MD Anderson Cancer UT Health Graduate School of Biomedical Sciences in Houston, TX, in July 2011. Prior to beginning full-time graduate school, she completed three years of medical education. Her initial appointment was in the Department of Pediatrics at MD Anderson Cancer Center under the supervision of Dr. Patrick Zweidler-McKay. She subsequently transitioned to Dr. Sean Post's laboratory in the Department of Leukemia at MD Anderson, where she completed her PhD studies.