6-2022

The Role and Immunogenicity of CBFA2T3-GLIS2 in Pediatric Acute Megakaryoblastic Leukemia

Elizabeth A. Garfinkle

The Role and Immunogenicity of CBFA2T3-GLIS2 in Pediatric Acute Megakaryoblastic Leukemia

Abstract

CBFA2T3-GLIS2 is the most prevalent fusion oncogene in pediatric acute megakaryoblastic leukemia in patients without Down syndrome (non-DS-AMKL) and is associated with an event free survival of only 8% even with high intensity chemotherapy and stem cell transplant in first remission. A cryptic inversion event on chromosome 16 joins the three nervy homology regions (NHR) of CBFA2T3 to the five zinc fingers of GLIS2. This configuration enables the encoded chimeric transcription factor to bind GLIS consensus sequences throughout the genome and recruit transcriptional activators and repressors to alter gene expression and enhance self-renewal capability. Few cooperating mutations have been identified in patients harboring this fusion which suggests it is the sole oncogenic driver. The molecular mechanism by which CBFA2T3-GLIS2 drives leukemogenesis is not well understood. Further, there are currently no studies exploring the immunogenicity of the fusion positive AMKL cells and immunotherapeutic approaches are lacking. To address these knowledge gaps, I developed two mouse models to study transcriptional regulation by the CBFA2T3-GLIS2 fusion oncogene and the ability of the immune system to recognize this malignancy with a goal of identifying therapeutic vulnerabilities for future translational studies. To understand transcriptional regulation by the fusion, I first turned to the literature. Studies on the wild type CBFA2T3 and GLIS2 proteins have demonstrated interactions with the transcriptional regulators ETO and CtBP1, respectively. Further, p300 has been shown to play a role in transcriptional regulation imparted by both transcription factors. I therefore hypothesized the fusion promotes transcriptional activation when the histone acetyltransferase p300 and the transcription factor ETO are recruited through NHR1 and NHR2, respectively. When the co-repressor CtBP1 is recruited through the PXDLS motif, located in the GLIS2 portion of the fusion, transcriptional repression predominates. Association of these co-factors with the fusion was confirmed through co-immunoprecipitations. Site-directed mutagenesis was then used to systematically delete NHR1 and NHR2 and mutate the PXDLS motif to evaluate the resultant effects on leukemogenesis imparted by the fusion. Murine models harboring the CBFA2T3-GLIS2 fusion without cooperating mutations have been unsuccessful and patient-derived xenograft (PDX) models are limited and difficult to manipulate. Therefore, I developed a novel humanized model of CBFA2T3-GLIS2 driven leukemia to understand the functional consequence of the mutant constructs. CD34+ stem cells were isolated from human cord blood and transduced with a lentivirus construct encoding the fusion and a GFP reporter. The cells were then differentiated to megakaryoblasts using human TPO and IL1β and sorted for purity prior to transplantation into immunodeficient NSG-SGM3 recipient mice. The fusion positive human primary megakaryoblasts induced a serially transplantable leukemia within 180 days that recapitulates CBFA2T3-GLIS2 positive patient samples on a transcriptional and translational level. Using the model, I found the fusion co-occupies sites throughout the genome with ETO and CtBP1 and induces upregulation of oncogenic signaling pathways including Hedgehog, JAK/STAT, Notch, and RUNX. Loss of NHR2 disrupted ETO association with the fusion, decreased fusion homodimerization, and abrogated leukemogenesis in vivo. I observed downregulation of JAK/STAT, Notch, and Hedgehog signaling pathways in cells with a fusion construct that lacks NHR2, suggesting these pathways may play a critical role in transformation. Future studies using CRISPR editing technology will be used to further elucidate the role of these pathways in AMKL transformation. An alternative to identifying proteins and biologic pathways necessary for transformation to target therapeutically is the use of immunotherapeutic approaches. In addition to a lack of knowledge about the CBFA2T3-GLIS2 transcriptional complex, there are currently no studies exploring the immunogenicity of the fusion positive AMKL cells. Since chemotherapy and stem cell transplantation fail to cure this disease, innovative treatment approaches are needed. While monoclonal antibody and chimeric antigen receptor T-cell (CAR-T) therapies have proven successful in acute lymphoblastic leukemia (ALL), similar successes haven't been realized in acute

myeloid leukemia (AML) due to the lack of targetable antigens that eradicate leukemia cells while minimizing off target toxicities such as depletion of normal myeloid cells which can lead to prolonged neutropenia. An alternative is T cell receptor engineered T cell (TCR-T) immunotherapy which uses heterodimers consisting of alpha and beta peptide chains to recognize polypeptide fragments presented by HLA molecules on the tumor cells. An advantage of this approach is the ability to recognize intracellular tumor specific and tumor associated antigen fragments in addition to extracellular proteins which results in a wider range of targets. To investigate the immunogenicity of CBFA2T3-GLIS2 positive AMKL cells, I first interrogated potential neoantigens spanning the fusion junction using the algorithm from NetMHCcons which identified two significant peptides. I then utilized immunopeptidomics to identify HLA class I presented peptides on the surface of AMKL PDX cells. One of the top hits was a peptide that corresponds to the 5' untranslated region (UTR) of bone morphogenetic protein 2 (BMP2). Although its role in leukemia is not well described, BMP2 is aberrantly upregulated in CBFA2T3-GLIS2 positive AMKL and required for serial replating in colony forming unit assays; therefore, this peptide is of great interest as a candidate leukemia-associated antigen. To identify potential AMKL directed effector CD8+ T cells, I developed a novel HLA class I exact matched PBMC-humanized CBFA2T3-GLIS2 positive AMKL PDX murine model. I observed a variable response in leukemia burden reduction, with the greatest responses correlating with an increase in granzyme A production and a decrease in TCR diversity; consistent with a potential clonal expansion of leukemia-specific effector CD8+ T cells. Single TCR expressing stable cell lines will be generated for the top expanded clones and used to assess leukemia clearance in the model and to measure specificity and functionality for the putative candidate antigens through tetramer staining and peptide stimulation assays, respectively. These novel studies have shed light on the immunogenicity of CBFA2T3-GLIS2 mediated AMKL and have the potential to contribute to the development of a TCR-T immunotherapy in the setting of minimal residual disease to provide an alternative, effective treatment for this chemotherapy resistant malignancy.

Document Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Program
Biomedical Sciences

Research Advisor
Tanja Gruber, PhD

Keywords
Immunotherapy, Leukemia, Transcription Factor

Subject Categories
Analytical, Diagnostic and Therapeutic Techniques and Equipment | Congenital, Hereditary, and Neonatal Diseases and Abnormalities | Diseases | Medical Molecular Biology | Medicine and Health Sciences | Neoplasms

UNIVERSITY OF TENNESSEE HEALTH SCIENCE CENTER

DOCTORAL DISSERTATION

The Role and Immunogenicity of CBFA2T3-GLIS2 in Pediatric Acute Megakaryoblastic Leukemia

Author:
Elizabeth A. Garfinkle

Advisor:
Tanja. A. Gruber, PhD

A Dissertation Presented for The Graduate Studies Council of
The University of Tennessee Health Science Center
in Partial Fulfillment of the Requirements for the Doctor of Philosophy degree from
The University of Tennessee

in

Biomedical Sciences: Cancer and Developmental Biology
College of Graduate Health Sciences

June 2022

DEDICATION

This dissertation is dedicated in memory of my aunt, Jacqueline G. McNatt,
for her unconditional love and fervent support.

ACKNOWLEDGEMENTS

Thank you to my mentor, Dr. Tanja Gruber, for your support, guidance, and all the opportunities you afforded me that has helped me grow as a scientist. And thank you for being a wonderful role model. I will apply everything I have learned from you throughout the rest of my scientific career.

I appreciate the time and support from my committee members: Drs. Paul Thomas, Gerard Zambetti, Stacey Ogden, and Ronald (Nick) Laribee. Their suggestions and guidance played an integral role in the studies presented here and in my graduate career as a whole. Thank you especially to Dr. Paul Thomas and members of his lab, specifically Jeremy Crawford, for all the time and expert advice you provided in collaboration for the immunotherapy studies. I would also like to thank the members of the Gruber Lab, past and present, for their help throughout my graduate studies. A special thanks to Pratima Nallagatla for her bioinformatics expertise, Anitria Cotton for my training in murine work, and Jennifer Kamens and Jinjun Dang for their unwavering encouragement.

Thank you to J Racquel Collins-Milinkovich for over 10 years of invaluable mentorship. I would not be where I am today without your support. Finally, to my parents, Charlene and Jeffrey Garfinkle. Thank you for everything you have done and continue to do to help me achieve my childhood dream. I am forever appreciative and grateful.

ABSTRACT

CBFA2T3-GLIS2 is the most prevalent fusion oncogene in pediatric acute megakaryoblastic leukemia in patients without Down syndrome (non-DS-AMKL) and is associated with an event free survival of only 8% even with high intensity chemotherapy and stem cell transplant in first remission. A cryptic inversion event on chromosome 16 joins the three nervy homology regions (NHR) of CBFA2T3 to the five zinc fingers of GLIS2. This configuration enables the encoded chimeric transcription factor to bind GLIS consensus sequences throughout the genome and recruit transcriptional activators and repressors to alter gene expression and enhance self-renewal capability. Few cooperating mutations have been identified in patients harboring this fusion which suggests it is the sole oncogenic driver. The molecular mechanism by which CBFA2T3-GLIS2 drives leukemogenesis is not well understood. Further, there are currently no studies exploring the immunogenicity of the fusion positive AMKL cells and immunotherapeutic approaches are lacking. To address these knowledge gaps, I developed two mouse models to study transcriptional regulation by the *CBFA2T3-GLIS2* fusion oncogene and the ability of the immune system to recognize this malignancy with a goal of identifying therapeutic vulnerabilities for future translational studies.

To understand transcriptional regulation by the fusion, I first turned to the literature. Studies on the wild type CBFA2T3 and GLIS2 proteins have demonstrated interactions with the transcriptional regulators ETO and CtBP1, respectively. Further, p300 has been shown to play a role in transcriptional regulation imparted by both transcription factors. I therefore hypothesized the fusion promotes transcriptional activation when the histone acetyltransferase p300 and the transcription factor ETO are recruited through NHR1 and NHR2, respectively. When the co-repressor CtBP1 is recruited through the PXDLS motif, located in the GLIS2 portion of the fusion, transcriptional repression predominates. Association of these co-factors with the fusion was confirmed through co-immunoprecipitations. Site-directed mutagenesis was then used to systematically delete NHR1 and NHR2 and mutate the PXDLS motif to evaluate the resultant effects on leukemogenesis imparted by the fusion. Murine models harboring the CBFA2T3-GLIS2 fusion without cooperating mutations have been unsuccessful and patient-derived xenograft (PDX) models are limited and difficult to manipulate. Therefore, I developed a novel humanized model of CBFA2T3-GLIS2 driven leukemia to understand the functional consequence of the mutant constructs. CD34+ stem cells were isolated from human cord blood and transduced with a lentivirus construct encoding the fusion and a GFP reporter. The cells were then differentiated to megakaryoblasts using human TPO and IL1β and sorted for purity prior to transplantation into immunodeficient NSG-SGM3 recipient mice. The fusion positive human primary megakaryoblasts induced a serially transplantable leukemia within 180 days that recapitulates CBFA2T3-GLIS2 positive patient samples on a transcriptional and translational level. Using the model, I found the fusion co-occupies sites throughout the genome with ETO and CtBP1 and induces upregulation of oncogenic signaling pathways including Hedgehog, JAK/STAT, Notch, and RUNX. Loss of NHR2 disrupted ETO association with the fusion, decreased fusion homodimerization, and abrogated

leukemogenesis *in vivo*. I observed downregulation of JAK/STAT, Notch, and Hedgehog signaling pathways in cells with a fusion construct that lacks NHR2, suggesting these pathways may play a critical role in transformation. Future studies using CRISPR editing technology will be used to further elucidate the role of these pathways in AMKL transformation.

An alternative to identifying proteins and biologic pathways necessary for transformation to target therapeutically is the use of immunotherapeutic approaches. In addition to a lack of knowledge about the CBFA2T3-GLIS2 transcriptional complex, there are currently no studies exploring the immunogenicity of the fusion positive AMKL cells. Since chemotherapy and stem cell transplantation fail to cure this disease, innovative treatment approaches are needed. While monoclonal antibody and chimeric antigen receptor T-cell (CAR-T) therapies have proven successful in acute lymphoblastic leukemia (ALL), similar successes haven't been realized in acute myeloid leukemia (AML) due to the lack of targetable antigens that eradicate leukemia cells while minimizing off target toxicities such as depletion of normal myeloid cells which can lead to prolonged neutropenia. An alternative is T cell receptor engineered T cell (TCR-T) immunotherapy which uses heterodimers consisting of alpha and beta peptide chains to recognize polypeptide fragments presented by HLA molecules on the tumor cells. An advantage of this approach is the ability to recognize intracellular tumor specific and tumor associated antigen fragments in addition to extracellular proteins which results in a wider range of targets.

To investigate the immunogenicity of CBFA2T3-GLIS2 positive AMKL cells, I first interrogated potential neoantigens spanning the fusion junction using the algorithm from NetMHCcons which identified two significant peptides. I then utilized immunopeptidomics to identify HLA class I presented peptides on the surface of AMKL PDX cells. One of the top hits was a peptide that corresponds to the 5' untranslated region (UTR) of bone morphogenetic protein 2 (BMP2). Although its role in leukemia is not well described, *BMP2* is aberrantly upregulated in CBFA2T3-GLIS2 positive AMKL and required for serial replating in colony forming unit assays; therefore, this peptide is of great interest as a candidate leukemia-associated antigen. To identify potential AMKL directed effector CD8+ T cells, I developed a novel HLA class I exact matched PBMC-humanized CBFA2T3-GLIS2 positive AMKL PDX murine model. I observed a variable response in leukemia burden reduction, with the greatest responses correlating with an increase in granzyme A production and a decrease in TCR diversity; consistent with a potential clonal expansion of leukemia-specific effector CD8+ T cells. Single TCR expressing stable cell lines will be generated for the top expanded clones and used to assess leukemia clearance in the model and to measure specificity and functionality for the putative candidate antigens through tetramer staining and peptide stimulation assays, respectively. These novel studies have shed light on the immunogenicity of CBFA2T3-GLIS2 mediated AMKL and have the potential to contribute to the development of a TCR-T immunotherapy in the setting of minimal residual disease to provide an alternative, effective treatment for this chemotherapy resistant malignancy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ALL	Acute lymphoblastic leukemia
AMKL	Acute megakaryoblastic leukemia
AML	Acute myeloid leukemia
CAR-T	Chimeric antigen receptor T cell
C-G+ pMks	CBFA2T3-GLIS2 positive primary megakaryoblasts
C-G- pMks	CBFA2T3-GLIS2 negative primary megakaryoblasts
HLA	Human leukocyte antigen
LAA	Leukemia-associated antigen
LSA	Leukemia specific antigen
NHR	Nervy homology region
ORA	Over-representation analysis
PBMC	Peripheral blood mononuclear cell
PDX	Patient-derived xenograft
TCR	T cell receptor
TCR-T	T cell receptor-engineered T cells
ZF	Zinc finger

CHAPTER 1. INTRODUCTION AND BACKGROUND

Background on Pediatric Acute Megakaryoblastic Leukemia

Leukemia, a cancer of blood-forming tissues, develops from acquired genetic and epigenetic changes in hematopoietic stem cells (HSCs) and downstream progenitor cells that confer enhanced proliferative capability (Reya et al., 2001; Radtke et al., 2009). The World Health Organization (WHO) classifies subtypes of leukemia based on the proliferation rate (acute or chronic), the cell of origin (myeloid or lymphoid branch), and the genetic aberrations. One subtype, acute myeloid leukemia (AML), comprises 15-20% of all pediatric acute leukemias. AML is a heterogenous group of acute leukemia and includes several additional subtypes arising from more differentiated progenitor cells (Arber et al., 2016; de Rooij et al., 2015). One of these subtypes is acute megakaryoblastic leukemia (AMKL), a cancer of megakaryoblasts which are the precursors to platelets.

Pediatric AMKL comprises 4-15% of pediatric AML patients and is found in patients with Down syndrome (DS-AMKL) and in patients without Down syndrome (non-DS-AMKL) (Gruber et al., 2015). Children less than 4 years of age with Down syndrome have a 500-fold greater incidence of AMKL than the general population and have a three-year overall survival greater than 80% with standard chemotherapy (Khan et al., 2011). In contrast, non-DS-AMKL patients only have a survival rate of 14-34% (Gruber et al., 2012). Malignant transformation of megakaryoblasts in DS-AMKL is due to trisomy 21 and *GATA1* gene somatic mutations which are rarely found in non-DS-AMKL patients (Khan et al., 2011; de Rooij et al., 2017). Initially, the genomic alterations driving non-DS-AMKL were not well-defined.

In 2012, Gruber et al. led a next generation sequencing effort through the St. Jude Children's Research Hospital-Washington University Pediatric Cancer Genome Project to perform transcriptome and exome sequencing on 14 diagnostic bone marrow samples from non-DS-AMKL patients. In half of the patient samples, a cryptic inversion event on chromosome 16 was detected that produced the coding fusion *CBFA2T3-GLIS2*. A follow up sequencing study in 2017 by de Rooij et al. on a cohort of 87 pediatric and 24 adult non-DS-AMKL samples found that there are seven subtypes, five of which contain recurrent chromosomal translocations that code for chimeric transcription factors exclusive to the pediatric subset: *HOX* rearrangements (*HOX*r), *RBM15-MKL1*, *NUP98-KDM5a*, *KMT2A* rearrangements (*KMT2A*r), and *CBFA2T3-GLIS2*. Few cooperating mutations were found in these patients, suggesting the oncogenic fusions are the drivers of leukemogenesis. *CBFA2T3-GLIS2* was found to be the most prevalent and the most aggressive, with an event free survival of only 8% and an overall survival of only 14% despite maximum tolerated high-intensity chemotherapy and stem cell transplant in first remission. The molecular mechanism by which the CBFA2T3-GLIS2 chimeric transcription factor drives non-DS-AMKL in not well-defined. Identification of critical components to the transcriptional complex will contribute toward the development of novel, targeted therapies to improve patient outcomes for this deadly cancer.

The Role of CBFA2T3-GLIS2 in Non-DS-AMKL

The Chimeric Transcription Factor CBFA2T3-GLIS2

The CBFA2T3-GLIS2 chimeric transcription factor is encoded by a fusion oncogene produced from a cryptic inversion event on chromosome 16 that joins the N-terminus of *CBFA2T3* to the C-terminus of *GLIS2*. In the majority of cases, exon 10 of *CBFA2T3* is joined to exon 3 of *GLIS2* (Gruber et al., 2012). Full-length CBFA2T3 (also referred to as MTG16 or ETO2) is a member of the myeloid translocation gene (MTG) family of transcriptional corepressors that also include the closely related ETO (also referred to as CBFA2T1) protein (Lutterbach et al., 1998). CBFA2T3 acts in the hematopoietic compartment to repress myeloid differentiation, thus maintaining hematopoietic stem cell quiescence (Gelmetti et al., 1998; Steinauer et al., 2019). In AML, CBFA2T3 plays a role in driving a gene expression profile conducive to leukemia stem cell transformation (Steinauer et al., 2020). All members of the MTG family, including CBFA2T3, share four highly conserved Nervy homology regions (NHR1-4) that can recruit transcriptional co-factors to alter gene expression (Hug et al., 2004; Steinauer et al., 2017). NHR1-3 are retained in the CBFA2T3-GLIS2 fusion (Gruber et al., 2012).

Full-length GLIS2 is a member of the Gli-similar family of transcription factors. It can act as a transcriptional activator and repressor and contains five highly conserved zinc fingers, which are retained in the CBFA2T3-GLIS2 fusion. These zinc fingers (ZF) bind DNA through the GLIS2 consensus sequence. GLIS2 is not normally expressed in the hematopoietic compartment, its canonical tissue is the kidneys; however, the fusion of GLIS2 to CBFA2T3 leads to high levels of C-terminal GLIS2 expression in megakaryoblasts (Gill et al., 2006; Vasanth et al., 2011, Gruber et al., 2012).

Although full-length CBFA2T3 and GLIS2 proteins are well characterized, their contribution in the CBFA2T3-GLIS2 fusion is not well described. Based on the reported functions of full-length CBFA2T3 and GLIS2, and the configuration imparted by the cryptic inversion event, I hypothesize the CBFA2T3-GLIS2 chimeric transcription factor binds GLIS2 consensus sequences throughout the genome via the five ZFs in the GLIS2 portion and recruits transcriptional co-factors through the three CBFA2T3 NHRs to alter gene expression at bound sites to drive leukemic transformation.

CBFA2T3

As a member of the MTG family of transcriptional co-repressors, full-length CBFA2T3 contains four highly conserved NHRs at the N-terminus. NHR1 is homologous with several TATA binding proteins (TBP)-associated factors, such as hTAF 105 and 130, which are transcriptional co-activators (Hug et al., 2004; Steinauer et al., 2017). NHR1 is a binding site for p300, a transcriptional co-activator and histone acetyltransferase (Ortega et al., 2018). ETO, a closely related member of the MTG

family, is well known for its involvement in the *AML1-ETO* fusion oncogene found in AML. When NHR1 is deleted in the ETO portion of the fusion, repressor activity is unchanged but binding of p300 is abrogated and upregulation of key leukemic transformation genes is disrupted (Hildebrand et al., 2001; Wang et al., 2011; Chen et al., 2019). AML1-ETO cells treated with C646, a p300 inhibitor, arrested cell cycle progression, became apoptotic, and had a decrease in fusion protein expression (Chen et al., 2019).

NHR2 contains a hydrophobic heptad repeat region that plays a key role in homo-oligomerization and hetero-oligomerization (Hug et al., 2004). Deletion of NHR2 in the ETO portion of the AML1-ETO fusion abrogated leukemic transformation, highlighting the importance of oligomerization in leukemogenesis (Steinauer et al., 2017; Liu et al., 2006). NHR2 can bind ETO, a co-repressor. However, ETO can also bind p300, a transcriptional activator. This suggests both transcriptional repression and activation can be altered through the oligomerization of ETO and the recruitment of p300 to the transcriptional complex via NHR1 and ETO (Wang et al., 2011).

NHR3 does not contain any highly conserved domains and is not well characterized and is thought to support oligomerization and work in conjunction with NHR4, which contains two non-DNA binding zinc fingers that interact with nuclear co-repressors (Hug et al., 2004; Steinauer et al., 2017).

GLIS2

Full-length GLIS2, the other fusion partner of *CBFA2T3-GLIS2*, is one of three Gli-Similar (GLIS1-3) transcription factors in the Kruppel-like zinc finger family. GLIS transcription factors are closely related to other subfamilies of Kruppel-like zinc finger proteins including GLI and ZIC protein family. GLIS2 can act as a transcriptional activator and repressor and plays a critical role in embryonic kidney development, structure, and function. Mutations in GLIS2 in humans are associated with kidney disease and renal failure (Vasanth et al., 2011). GLIS1-3 contain five highly conserved zinc finger binding motifs that recognize and bind DNA at promoters of target genes, preferentially those that contain a guanine-rich GLIS consensus sequence. ZF 3 is required for nuclear localization of GLIS2 (Vasanth et al., 2011). ZF 1 interacts with β-catenin, a regulator of WNT signaling pathway which is important in cell proliferation, differentiation, migration, and stem cell renewal (Kim et al., 2007). In relation, GLI signaling is a downstream target of the Hedgehog signaling pathway, which is essential for cell differentiation including in kidney development (Attanasio et al., 2007).

In 2005, Kim, S et al. described the interaction of the co-transcriptional repressor protein CtBP1 (C-terminal binding protein 1) exclusively with GLIS2 out of the three GLI family transcription factors. CtBP1 is expressed in many tissues and embryonic development and recruits histone deacetylases to repress transcription. CtBP1 can interact with other proteins through a PXDLS or similar consensus sequence; however, it is not required to repress transcription through the recruitment of histone deacetylases. Full-

length GLIS2 contains one PXDLS sequence and three degenerate sequences. Kim, S et al. mutated each site using site-directed mutagenesis to change DL to AS to assess the effects of CtBP1 binding to GLIS2. Each of the four mutated sites individually caused a reduction in transcriptional activation in a luciferase reporter assay, but did not completely disrupt CtBP1 binding. Only when all four sites were mutated simultaneously was transcriptional activation and CtBP1 binding abolished (Kim, S et al., 2005). The CBFA2T3-GLIS2 fusion contains one PXDLS consensus sequence and one degenerate sequence at the C-terminus of the GLIS2 portion of the fusion (Gruber et al., 2012).

p300, like many histone acetyltransferases, contains a highly conserved PXDLS sequence in the center of its bromodomain and has been shown to bind directly to CtBP1. CtBP1 represses the transactivation activity of p300, thereby regulating the accessibility of co-factors with histone acetyltransferase activity to regions of chromatin (Kim, J et al., 2005).

CBFA2T3-GLIS2 in Non-DS-AMKL

Despite the poor prognosis associated with CBFA2T3-GLIS2 positive non-DS-AMKL, there are relatively few published studies on the role of the fusion in leukemogenesis (Gruber et al., 2012; de Rooij et al., 2017; Dang et al., 2017; Thirant et al., 2017, Lopez et al., 2019; Bertuccio et al., 2020; Benbarche et al., 2022). The Gruber research group has shown that the gene expression profile of pediatric non-DS-AMKL patients cluster according to the fusion oncogene they harbor; however, CBFA2T3-GLIS2 positive patients have a gene expression profile that is highly distinct from the other fusions and carry the lowest burden of somatic cooperating mutations (de Rooij et al., 2017; Gruber at al., 2012). Contributing to this distinct gene expression profile is the role of GLIS2 in sonic hedgehog (SHH), WNT/β-catenin, and bone morphogenic protein (BMP2) pathways which leads to an overexpression of downstream targets including *PTCH1*, *HHIP*, *BMP2*, and *BMP4*.

CBFA2T3-GLIS2 transduced mouse bone marrow and CBFA2T3-GLIS2 positive AMKL human cell lines were more sensitive in a dose dependent manner to dorsomorphin, a small molecule inhibitor of BMP type 1 receptors, than un-transduced mouse bone marrow and CBFA2T3-GLIS2 negative AMKL cells lines. Interestingly, *CBFA2T3-GLIS2* and full-length *GLIS2* transduced mouse bone marrow conferred enhanced self-renewal capability in an *in vitro* colony forming units (CFU) assay, with colonies displaying morphological characteristics and cell surface marker expression associated with megakaryocyte lineage differentiation (CD41/CD61). However, when *CBFA2T3-GLIS2* transduced mouse bone marrow is transplanted into lethally irradiated syngeneic recipient mice, they consistently fail to induce an *in vivo* leukemia (Gruber et al., 2012; Dang et al., 2017).

The failure to induce leukemia *in vivo* has been reported for other fusion driven leukemias, specifically AML1-ETO and MLL-AF4 (Higuchi et al., 2002; Yan et al., 2006; Milne, 2017). *AML1-ETO* transduced mouse bone marrow does lead to altered

gene expression and increased hematopoietic cell proliferation, but does not induce leukemia in mice without the inclusion of an extra exon 9a, resulting in the truncated AML1-ETO9a isoform that lacks the C-terminal NHR3 and NHR4 domains. *AML1-ETO* and *AML1-ETO9a* co-expression reduced the latency of disease on-set, which the authors suggest indicates fusion proteins and their alternatively spliced isoforms may work in collaboration to induce oncogenic transformation (Yan et al., 2006).

AF4 is the most common MLL fusion partner in MLL-fusion driven acute lymphoblastic leukemia (ALL) and is associated with a very aggressive phenotype. Although murine models of this leukemia have been challenging to develop and the reasoning for this is not well-defined, success has been realized in *MLL-AF4* transduced CD34+ primary human stem cells when using the mouse *AF4* cDNA sequence. This results in a pro-B-ALL when transplanted into immunodeficient mice (Milne, 2017; Lin et al., 2016). These findings in other fusion driven leukemia subtypes suggest that CBFA2T3-GLIS2 may require collaboration with a spliced isoform, expression in primary human stem cells, and/or additional cooperating mutations, although few are found in patient specimens, to induce leukemia in a model system for subsequent studies.

To this extent, Drenberg et al. reported that transduction of lineage depleted murine bone marrow with a retroviral construct encoding the *CBFA2T3-GLIS2* sequence plus the *JAK2* V617F mutation, a clinically relevant JAK mutation found in a CBFA2T3-GLIS2 positive patient, induced a serially transplantable megakaryoblastic leukemia in mice. Transduction with a construct encoding only the *CBFA2T3-GLIS2* fusion or only the *JAK2* V617F mutation did not induce leukemia (Drenberg et al., 2019). This shows that in a murine setting, CBFA2T3-GLIS2 requires a cooperating mutation to drive leukemic transformation.

Although not readily explored and reported in the literature, there is a possibility that *CBFA2T3-GLIS2* transduced mouse bone marrow cells may be immunogenic and cleared by the competent mouse immune system of the transplanted recipients. This can be overcome by transplantation of human primary cells overexpressing the fusion into immunodeficient mice.

The Mercher research group has overcome the absence of a mouse model lacking cooperating mutations for CBFA2T3-GLIS2 non-DS-AMKL by utilizing CRISPR/Cas9 technology to add a C-terminal GFP tag to the endogenous *CBFA2T3-GLIS2* fusion expressed in the human MO7e cell line (derived from a CBFA2T3-GLIS2 positive AMKL patient). Then they assessed genomic occupancy of the fusion using chromatin immunoprecipitation sequencing (ChIP-seq) in which they immunoprecipitated for the GFP tag and CBFA2T3 and then took the overlapping genes to differentiate fusion bound genes from endogenous, full-length, wild type CBFA2T3 bound sites. They found genome wide binding of the fusion at sites containing motifs associated with *ETS, GATA, RUNX, GLIS2*, and *ZIC* gene families. They also found co-occupancy of the fusion with the FLI-1 related ETS factor *ERG* at super enhancer sites (defined by the presence of H3K27ac, a transcriptional activation epigenetic mark), a known megakaryoblastic oncogene that is upregulated in CBFA2T3-GLIS2 positive AMKL patients. Patient-

derived xenograft AMKL cells harboring the CBFA2T3-GLIS2 fusion, and a luciferase reporter for tracking leukemia burden, expanded in transplanted immunodeficient mice. This proliferation was halted when CRISPR/Cas9 was used to knock-out *ERG* and also when NHR2 of the fusion was blocked by NC128, a competitive peptide consisting of 128 amino acids of NHR2. NC128 also successfully disrupts AML-ETO positive AML cells proliferative ability and reverses their differentiation block (Wichmann et al., 2007). This suggests ERG and oligomerization of the fusion, as NHR2 plays a crucial role in homo- and hetero-oligomerization of MTG proteins, are essential components of the CBFA2T3-GLIS2 transcriptional complex (Thirant et al., 2017, Lopez et al., 2019, Benbarche et al., 2022).

The lack of an easily manipulatable murine model, syngeneic or humanized, is a barrier to studies needed to understand and define essential components of the CBFA2T3-GLIS2 transcriptional complex which would yield new druggable targets for the development of small molecule inhibitors to improve patient outcomes for this devastating pediatric leukemia. Development of a small molecule inhibitor targeting the fusion itself is extremely difficult and frequently not feasible for chimeric transcription factors from a structural biology level; however, targeting a transcriptional co-factor essential to the complex may be more viable. The first half of the studies presented in this dissertation focus on the development of a novel humanized murine model of CBFA2T3-GLIS2 AMKL used to characterize the role of three key transcriptional co-factors: ETO, CtBP1, and p300. These were selected for subsequent studies based on published research describing the interaction of these co-factors with full-length CBFA2T3 and GLIS2.

The Immunogenicity of CBFA2T3-GLIS2 in Non-DS-AMKL

Immuno-Oncology

Cancer treatments originally fell into one of three groups: surgery, radiotherapy, or chemotherapy (Debela et al., 2021). The most common treatment for leukemia is chemotherapy and stem cell transplant in first remission; however, this has proved unsuccessful for pediatric patients with CBFA2T3-GLIS2 driven non-DS-AMKL (Gurber et al., 2012, de Rooij et al., 2017). Identification of critical co-factors in the CBFA2T3-GLIS2 transcriptional complex that can be targeted by small molecule inhibitors, part of the emerging targeted therapy field in cancer treatment, is one approach to develop a novel therapy for non-DS-AMKL patients to improve patient outcomes (Debela et al., 2021). Another approach is to exploit components of the patient's own immune system to direct targeted responses against their cancer cells. This constitutes the basis of the field of immune-oncology research (Allard et al., 2018).

The immune system has two main branches: innate immunity and adaptive immunity. Innate immunity is independent of antigen and is a rapid, non-specific response to a pathogen. Due to the lack of antigen specificity, no immunological memory

is generated. In contrast, adaptive immunity is antigen specific and therefore can generate immunological memory (Marshall et al., 2018). In the context of immunotherapy, cells belonging to both branches have been utilized. Natural killer (NK) cells, a member of the innate immune system, can recognize cancer cells independent of antigen presentation and can mediate tumor lysis through degranulation and have been harnessed for therapeutic treatments; however, the exact mechanism by which NK cells target cancer cells is not well-defined and cancer cells can confer resistance (Shimasaki et al., 2020). A more popular immunotherapy is the use of cytotoxic CD8+ T cells, a member of the adaptive immune system (Raskov et al., 2021). CD8+ T cells are activated upon engagement of their T cell receptor (TCR) with peptide presented in the binding groove of major histocompatibility complex (MHC) class I molecules (referred to as human leukocyte antigen, or HLA, in human cells) expressed on the surface of antigen-presenting cells (APCs), such as cancer cells. Activation of CD8+ T cells upon engagement with antigen can promote clonal expansion of effector cells that can release cytotoxic granules to induce apoptosis of target cells (ex: cancer cells) (Marshall et al., 2018).

Cancer exome data is a valuable tool for the identification of potential neoantigens, HLA class I presented peptides that are absent from the normal human genome and arise from mutations that often contribute to oncogenesis. Due to their novelty, infiltrating CD8+ T cells have the potential to mount a functional immune response. It was originally thought that for a cancer to be immunogenic, it had to have a high neoantigen burden to distinguish the cancer cells from normal, healthy tissue. For example, melanoma, lung, and stomach cancers have high mutational burdens and have had success with immunotherapy-based treatments (Schumacher et al., 2015). Even though pediatric cancers, especially acute leukemias, notoriously have a low mutational burden, studies have shown that an effective, strong immune response can be stimulated by high-quality, albeit few, neoantigens (Wang, P et al., 2021).

Immunotherapy for Leukemia

Monoclonal Antibodies

There are three main classes of monoclonal antibodies. The first class is called naked antibodies. These bind directly to their intended cell surface marker and induce apoptosis of the target cell through the activation of the antibody-dependent and complement-dependent cytotoxicity immune system pathways. The second class of monoclonal antibodies are antibody-drug conjugates (ADC). These are similar to naked antibodies in that they target a specific cell surface maker, but instead of relying on apoptosis activation through the immune system pathways, they bring a linked cytotoxin or radioisotope to the target cell which, once internalized, induces cell death. The third class is T-cell re-directing antibodies which include BiTEs (bispecific T cell engager) (Li et al., 2020).

Daratumumab is a naked anti-CD38 antibody treatment for AML. CD38 is a transmembrane glycoprotein on the surface of immune cells that plays a role in the regulation of cytokine release. Inflammatory cytokines are elevated in AML filled human bone marrow and promotes leukemic growth. Anti-CD38 antibody treatment in an AML xenograft transplantation model did demonstrate antibody-dependent phagocytosis and a block in AML cell trafficking, but was ultimately unable to successful control AML progression (Farber et al., 2021).

Gemtuzumab-ozogamicin (GO) is an FDA approved ADC against CD33 in pediatric AML. CD33 is an attractive target because it is expressed on the AML cell surface in more than 80% of AML patients and absent on hematopoietic stem cells, thereby minimizing off-target effects. The Children's Oncology Group (COG) trial AAML0531 evaluated GO efficacy in children to young adults with primary AML and found addition of GO to standard chemotherapy regiments significantly improved event free survival (EFS) due to a significant reduction in relapse risk (RR), however it did not improve overall survival (OS) (Gamis et al., 2014).

BiTEs link a cell surface marker monoclonal antibody to an anti-CD3 monoclonal antibody which results in engagement of an effector T cell with the target cell. Blinatumomab is an FDA approved BiTE that has been successful for relapsed/refractory B cell ALL in pediatric and adult patients that targets CD19 on the B-ALL target cells. Early studies in relapsed refractory adult AML patients with a CD123 BiTE demonstrate a benefit in a subset of patients that have evidence of an immune-infiltrated microenvironment (Uy et al., 2021).

CAR-T Cells

Chimeric antigen receptors (CAR) are genetically engineered T cell receptors that consist of an extracellular antigen-binding domain, a linker and spacer region, a transmembrane domain, and intracellular T cell signaling and co-stimulatory domains. The CAR can be custom engineered in a laboratory and then overexpressed in a patients' autologous or donor T cells, which are then adoptively transferred to the patient as a CAR-T immunotherapy. A unique advantage of CAR-T immunotherapy is that it is HLA class I independent; therefore, a single CAR-T construct can be used for groups of patients with a shared cancer antigen regardless of their haplotype. Like Blinatumomab, CAR-T cells have been successful in targeting CD19 on B-ALL target cells. Kymirah was the first CD19-targeting CAR-T therapy to be awarded FDA approval for children and young adults with relapsed or refractory B-ALL (Salmikangas et al., 2018).

The successful use of CAR-T based immunotherapy to treat AML patients has lagged behind ALL. The main targets of interest, CD33 and CD123, are often highly expressed on AML blasts but, unfortunately, are also expressed on myeloid progenitors and hematopoietic stem cells, respectively; thereby leading to severe myeloablation that can be lethal (Marvin-Peek et al., 2022). Advancements in CAR-T constructs to target new antigens, enhance efficacy, and modulate harmful off-target effects are in progress (Sheykhhasan et al., 2022). For example, Richards et al. reported successful targeting of

CD93 expressed on AML cells while sparing CD93+ endothelial cells in an *in vitro* model system using a NOT-gated CD93 CAR-T cell. NOT-gated CAR-T cells express two CAR constructs; one for the target antigen and an additional inhibitory CAR for an antigen present exclusively on the healthy cells that also happen express the target antigen. When the CAR-T cells bind on the target cells they will induce a functional response; however, when the CAR-T cells bind the off-target, healthy cells the secondary inhibitor CAR-T cell will also be engaged and mitigate the effector response (Richards et al., 2021).

TCR-T

T cell receptor engineered T cell (TCR-T) immunotherapy uses heterodimers consisting of an alpha and beta peptide chain to recognize polypeptide fragments presented by HLA class I molecules on the cancer cells; therefore, unlike CAR-T cell therapy, TCR-T cell therapy is HLA class I dependent. An advantage of this approach is the ability to recognize intracellular tumor-specific and tumor-associated antigen fragments in addition to extracellular proteins which results in a wider range of targets (Liu et al., 2022). TCR-T immunotherapy has been limited in the AML setting, although progress is being made.

In 2013, Chapuis et al. identified a high affinity HLA-A*02:01 WT1 specific TCR. WT1 is Wilm's Tumor Antigen 1, an intracellular protein that plays a role in AML proliferation and leukemogenesis. WT1 is overexpressed in AML blasts and leukemia stem cells compared to CD34+ stem cells. In 2019, Chapuis et al. reported persistent T cell responses against WT1 positive AML in high-risk patients after stem cell transplant that helped prevent relapse upon TCR-T adoptive cell transfer. They overexpressed the identified WT1-specific TCR in donor-derived CD8+ T cells that were EBV (Epstein-Barr virus) specific. They propose the use of EBV+ cells contribute to TCR-T survival and proliferation through the stimulation of a pro-inflammatory microenvironment triggered by a low-level reactive response to reservoirs of dormant EBV infected cells.

A recent 2021 study by van der Lee et al. reported identification of neoantigens from mutations in nucleophosmin 1 (NPM1) in primary AML. They isolated T cells that recognized one of the HLA-A*11:01 presented neoantigens, AVEEVSLRK, and cloned the responsive TCR which was effective against primary AML cells *in vitro* but was not able to display any anti-leukemia activity when transplanted into AML engrafted NSG mice. This highlights the need for more pre-clinical models to study the mechanisms of TCR:antigen engagement.

Limitations

As with any cancer therapy, there are limitations to the effectiveness of immunotherapy. All three immunotherapies described here rely on cell surface presentation of a marker or antigen, preferably one restricted only to the cancer cells to minimize harmful off-target effects. Cancer cells can evade immune targeting by

downregulating expression of the targeted marker or antigen, thereby conferring therapeutic resistance (Zhou et al., 2021). Cancer cells can also create an immunosuppressive microenvironment to hinder T cell functional activity. Upregulation of PD-L1 on cancer cells, the ligand for PD1 on T cells, inhibits T cell proliferation which is essential for a successful effector response. Anti-PD1/PD-L1 checkpoint inhibitors can successfully reverse this repressive environment in some patients (Pan et al., 2020).

Targeting CBFA2T3-GLIS2 AMKL With Immunotherapy

As described here, monoclonal antibody and CAR-T therapies have proven successful in ALL, but similar successes have not been realized in AML due to a lack of targetable antigens that can lead to eradication of leukemia cells without causing prolonged neutropenia. Even though CBFA2T3-GLIS2 positive AMKL has a low mutation burden, with the malignancy driven by a single fusion gene event, I hypothesize the fusion junction and peptides derived from the dysregulated gene expression imparted by the fusion can act as leukemia-specific and -associated antigens that can be recognized by specific T cell clones. Such clones would have the potential to be translated into effective adoptive TCR-T therapies that will limit off-target toxicities through the use of modified endogenous T cells. This hypothesis is supported Zamora et al., who showed that tumor infiltrating CD8+ T cells in diagnostic bone marrow from pediatric ALL patients were able to abundantly and functionally respond to ALL neoantigens, despite the relative paucity of mutations in these patients. Notably, they found a robust response directed against a neoantigen from the ETV6-RUNX1 fusion in seven of the nine patients that were screened, supporting the immunogenicity of fusion gene driven acute leukemias (Zamora et al., 2019).

The second half of the studies presented in this dissertation focus on the characterization of AMKL cell surface antigens and the development of a novel PBMC-humanized PDX murine model to study the interaction between AMKL cells and CD8+ T cells *in vivo*.

Specific Aims

Pediatric non-DS-AMKL is a genomically heterogeneous malignancy that, unlike the more prevalent DS-AMKL, carries the worst prognosis reported to date with five-year event survival as low as 8% (Gruber et al., 2012; de Rooij et al., 2017). Non-DS-AMKL includes seven subtypes with distinct driver mutations; the chimeric oncogenic transcription factor CBFA2T3-GLIS2 is the most prevalent and confers the worst prognosis even with current maximum intensity chemotherapy and stem cell transplant in first remission (Gruber et al., 2012). Therefore, new therapeutic approaches are needed for this aggressive leukemia.

Patient leukemia cells harboring the CBFA2T3-GLIS2 fusion have a distinct gene expression profile that distinguishes them from CBFA2T3-GLIS2 negative non-DS-AMKL cases (Gruber et al., 2012; de Rooij et al., 2017), indicating a unique mechanism of leukemogenesis. The exact molecular mechanism by which the fusion imparts differential gene expression is not well-defined. Identification of components critical to the transcriptional complex and their role in gene regulation may reveal novel therapeutic targets.

Immunotherapeutic approaches to treating chemotherapy resistant acute myeloid malignancies have lagged behind their lymphoid counterparts due to a lack of targetable antigens. Well established CBFA2T3-GLIS2 positive patient-derived xenograft (PDX) models have the potential to be PBMC-humanized for antigen discovery and the study of cellular immune responses to non-DS-AMKL, which will aid in the development of immunotherapy approaches, such as TCR-T adoptive cell transfer, for which there are currently no options.

Aim 1: Delineate the Role of CBFA2T3-GLIS2 in Pediatric Non-DS-AMKL

In this aim, a novel humanized murine model of non-DS-AMKL was used to demonstrate that the *CBFA2T3-GLIS2* fusion can act as a driver of leukemogenesis by binding GLIS2 consensus sequences throughout the genome and imparting differential gene expression through dimerization and the recruitment of the transcriptional co-factors ETO and CtBP1. This aim also identifies components of the CBFA2T3-GLIS2 transcriptional complex, which will aid in the development of targeted therapies to improve patient outcomes.

Aim 2: Evaluate the Immunogenicity of CBFA2T3-GLIS2 Positive Pediatric Non-DS-AMKL

An alternative therapeutic approach to improve patient outcomes for this aggressive leukemia is immunotherapy, for which there are currently no options. In this aim, the surfaceome of a CBFA2T3-GLIS2 positive patient sample was probed for potential HLA class I presented immunogenic antigens. A novel PBMC-humanized PDX model was used to evaluate the potential of CD8+ T cells to recognize and respond to these antigens. These studies present the first evaluation of non-DS-AMKL immunogenic antigens and will contribute to the advance in adoptive cell therapy-based immunotherapies for low mutational burden leukemias.

CHAPTER 2. THE ROLE OF CBFA2T3-GLIS2 IN PEDIATRIC NON-DOWN SYNDROME ACUTE MEGAKARYOBLASTIC LEUKEMIA

Introduction

Models of CBFA2T3-GLIS2 Non-DS-AMKL

Current models to study the role of CBFA2T3-GLIS2 in pediatric non-DS-AMKL are limited. Immortalized cell lines derived from patients harboring the CBFA2T3-GLIS2 fusion, such as M07e, WSU-AML, and RS1 (Gruber et al., 2012), have additional acquired mutations that are not represented in the patient population such as *TP53*. In contrast to other fusions found in pediatric non-DS-AMKL, CBFA2T3-GLIS2 is unable to transform mouse bone marrow cells in a transplant model (Dang et al., 2017). Several CBFA2T3-GLIS2 positive patient-derived xenograft models are available; however, patient cells are difficult to manipulate for experiments, cell quantity is limited, and they do not grow well *in vitro*. A reliable model of CBFA2T3-GLIS2 leukemia that recapitulates primary patient samples and can be experimentally manipulated is an essential step toward the development of novel targeted therapies to improve patient outcomes.

ETO, CtBP1, p300, and CBFA2T3-GLIS2

Identification of components critical to the CBFA2T3-GLIS2 transcriptional complex and their role in gene regulation may reveal novel therapeutic targets. I hypothesize the fusion's five zinc fingers are directing binding of genomic regions that contain a GLIS consensus sequence and imparting altered gene expression at these sites through the recruitment of transcriptional co-factors, which ultimately leads to leukemic transformation (**Figure 2-1**). As described in the introduction, studies on the wild type CBFA2T3 and GLIS2 proteins have demonstrated interactions with the transcriptional regulators ETO and CtBP1, respectively. Further, p300 has been shown to play a role in transcriptional regulation imparted by both transcription factors. Therefore, as summarized in **Figure 2-1**, I predict ETO is binding to NHR2 and inducing transcriptional repression or transcriptional activation through the recruitment of p300 depending on the target genes bound. NHR2 is an oligomerization domain for MTG proteins and can also recruit the fusion and endogenous full-length CBFA2T3 which, in turn, can also contribute to bringing ETO and p300 to the fusion transcriptional complex. p300 can also be recruited through NHR1. The histone acetyltransferase activity of p300 can be inhibited by CtBP1, a transcriptional co-repressor, which can bind the PXDLS domain in the GLIS2 C-terminal portion of the fusion. Taken together, I propose the CBFA2T3-GLIS2 fusion recruits ETO, CtBP1, and p300 to the transcriptional complex at key regulatory genes throughout the genome to impart changes in gene expression conducive to enhanced proliferation and inhibited differentiation.

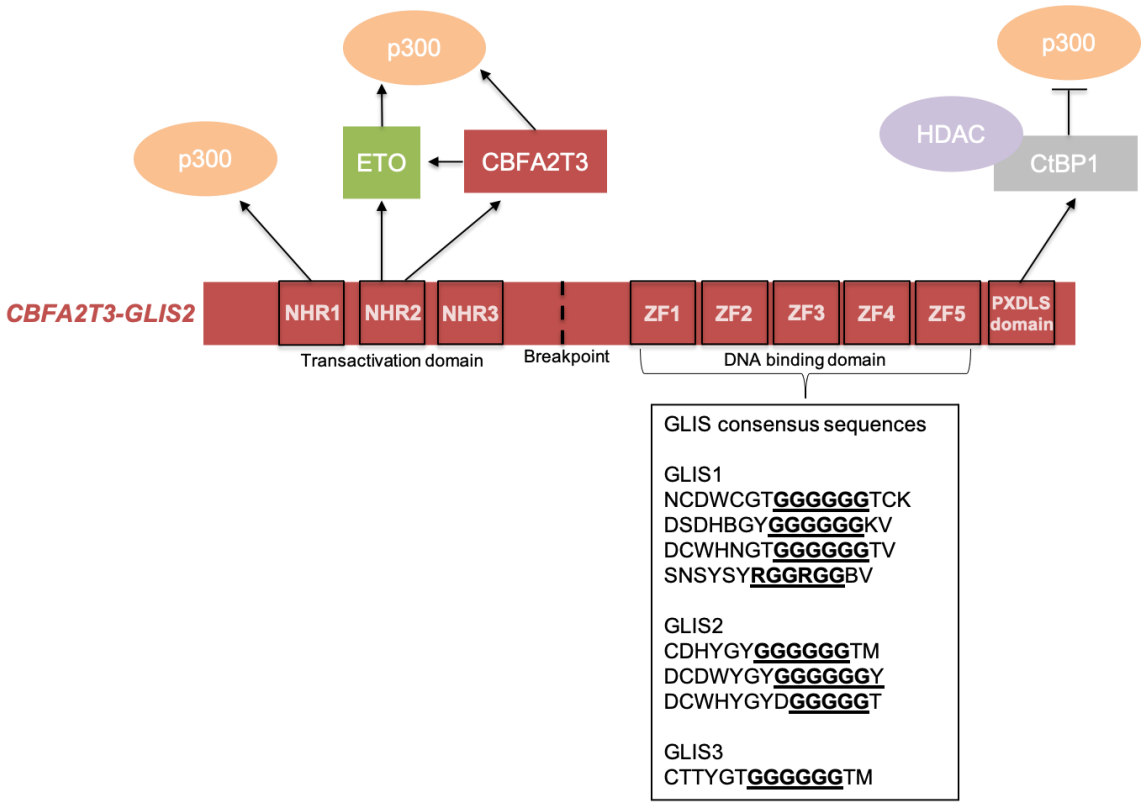


Figure 2-1. Schematic of the CBFA2T3-GLIS2 Transcriptional Complex.
Protein structure of the fusion is show in red and is not drawn to scale. GLIS2 consensuses sequences are listed below the DNA binding domain. Breakpoint indicates junction where the N-terminal portion of CBFA2T3 is joined to the C-terminal portion of GLIS2. NHR, nervy homolog region; ZF, zinc finger.

Materials and Methods

Generation of Primary Human Megakaryoblasts

As outlined in **Figure 2-2**, CD34+ stem cells were isolated from commercially available (St. Louis Cord Blood Bank, Cleveland Cord Blood Bank) human cord blood units. Cord blood was layered onto Ficoll (Fisher) and CD34+ stem cells were positively selected from the peripheral blood mononuclear cell (PBMC) layer using the human CD34 MicroBead kit (Miltenyi Biotec). CD34+ stem cells were cultured in GMP SCGM serum-free medium supplemented with 1% penicillin-streptomycin (10,000 U/mL Thermo Fisher) and 100 ng/ml each of human TPO, FLT3 Ligand, and SCF (Tonbo Bioscience) for 48 hours. Cultured CD34+ stem cells were then transduced using RetroNectin reagent (Takara Bio) and the third generation pCDH.MSCV.MCS.EF1acopGFP_CD711B-1 lentivirus encoding the *CBFA2T3-GLIS2* cDNA with an N-terminus 3xTY1 or 3xFLAG tag.

48 hours after transduction, the cells were removed from the RetroNectin and differentiated to megakaryoblasts in GMP SCGM serum-free medium supplemented with 1% penicillin-streptomycin (10,000 U/mL Thermo Fisher) and 100 ng/ml human TPO and 10 ng/ml human IL1β (Tonbo Bioscience) for 7-9 days which allows for about 30-fold expansion (Tijssen et al., 2011). On day 7-9, megakaryoblast differentiation status was assessed by flow cytometry for CD41 and CD34 and then sorted for GFP using fluorescence-activated cell sorting to obtain a pure GFP, CBFA2T3-GLIS2 positive population. Transduction efficiencies averaged 15 - 30%.

Animals

I chose to use the NSG-SGM3 (NOD.Cg-*Prkdc*$_{scid}$ *Il2rg*$_{tm1Wjl}$ ITg(CMV-IL3, CSF2, KITLG) 1Eav/MloySzJ) mouse strain for the murine models described in this dissertation because they are immunodeficient and contain three human transgenes: human interleukein-2 (IL2), human granulocyte/macrophage-stimulating factor (CM-CSF), and human stem cell factor (SCF), each driven by a human cytomegalovirus promoter/enhancer sequence. Therefore, these mice are ideal for engrafting the primary human cells because they lack mouse immune cells that may invoke rejection and they harbor human transgenes that promote myeloid engraftment. NSG-SGM3 breeder mice were purchased from the Jackson Laboratory (013062) and bred in the Animal Resource Center at St. Jude Children's Research Hospital and the Research Animal Facility at Stanford University. Mice are maintained on the NOD.Cg-Prkdscid Il2rgtm1Wjl/SzJ background.

All animal work was conducted under Dr. Tanja Gruber's IACUC Protocols: St. Jude Children's Research Hospital "Establishing Murine Models of Leukemia" 519-100568 (11/16/2018 – 11/16/2021); Stanford University "Murine Modeling of High Risk Pediatric Acute Leukemias" 33830 (08/25/2020 – 08/25/2023).

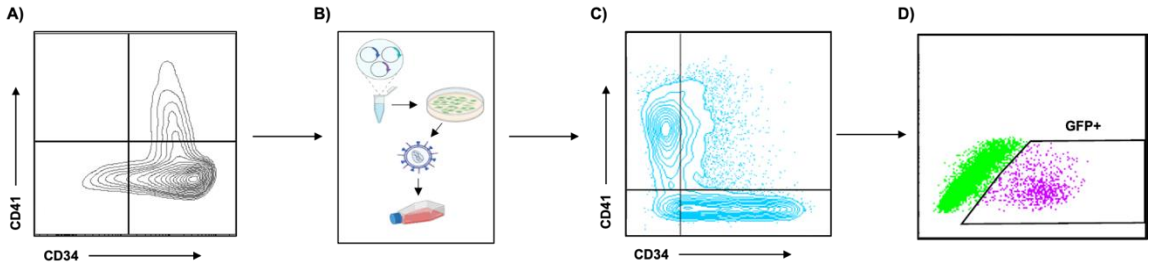


Figure 2-2. Generation of CBFA2T3-GLIS2 Transduced Primary Human Megakaryoblasts.
A) CD34+ stem cells are isolated from commercially available human cord blood. **B)** CD34+ stem cells are cultured for 48 hours and then transduced with a third-generation lentiviral vector encoding the *CBFA2T3-GLIS2* cDNA sequence with an N-terminal 3xTY1 or 3xFLAG tag for downstream applications. **C)** 48 hours after transduction, the CD34+ stem cells are differentiated to megakaryoblasts with human TPO and IL1β for 7-9 days. **D)** A pure population of CBFA2T3-GLIS2+ primary megakaryoblasts are obtained using fluorescence-activated cell sorting for GFP+. Cells are then used for downstream experiments such as transplantation and next generation sequencing.

Mice were observed for any signs of pain or distress, including the inability to eat or drink, hunched body position, scruffed fur, labored or otherwise abnormal breathing, and/or presentation of hind leg paralysis, by myself daily in addition to observation by the St. Jude and Stanford Husbandry staff twice daily, including weekends and holidays.

Irradiation

Mice were pre-conditioned with 200 rad. At St. Jude Children's Research Hospital, a cesium irradiator was used. At Stanford University, an x-ray irradiator was used. To minimize pain and distress that may be caused by the radiation, mice were irradiated twenty-four hours prior to injection to allow for recovery time. Mice were treated with prophylaxis Baytril and Amoxicilin to minimize contraction of bacterial infections. Mush food was provided if mice seemed in distress or had trouble eating the solid food provided.

Tail Vein Injections

Tail veins were dilated using a lamp or hand warmers for several minutes. Mice were then placed in a sterile, commercially available standard restraint. The tail was then grasped at mid-length, cleaned with 70% ethanol, and 50,000 – 400,000 human megakaryoblasts in 100 ul of sterile PBS were injected into the tail vein with a 1 ml insulin syringe. The injection site was cleaned again after the injection with 70% ethanol and pressure was applied with a piece of gauze to reduce any blood flow from the site of the injection. Mice were then returned to their cage and monitored for about ten minutes to ensure they recovered well from the procedure.

Intrafemoral Injections

Mice were anesthetized with 1-5% isoflurane administered through a small tube attached to their nose throughout the entire procedure. The left leg was cleaned with 70% ethanol and a 1 ml insulin syringe was used to make a small puncture in the femoral bone. 10 ul of sterile PBS was injected into the femur to ensure cell injection is smooth and does not have any leakage. Once confirmed, 50,000 – 400,000 human megakaryoblasts in 20 ul of sterile PBS was injected into the femur. The leg was cleaned again after the injection with 70% ethanol and the mouse was given a sub-cutaneous injection of Meloxicam (Metacam) pain reliever. The mouse was returned to its cage which was placed on a heating pad to aid in recovery. Mice were monitored for 30 minutes post-injection to ensure proper recovery.

Moribund Mice

Moribund mice were identified by the presentation of hind leg paralysis, hunched body position, labored breathing, and/or inability to eat or drink. They were euthanized

with carbon dioxide inhalation as defined by the American Veterinary Medical Association (AVMA) Guidelines for Euthanasia of Animals.

Necropsy and Histology

Mice were preserved in 10% formalin. Full body necropsies and histology (hematoxylin and eosin stain) were performed by American College of Veterinary Pathologists (ACVP) board-certified veterinary pathologists in the Veterinary Pathology Laboratory Core at St. Jude Children's Research Hospital and in the Animal Histology Services at Stanford University.

Cell Collection and Processing

Hind legs and spleen were harvested from euthanized mice. Bone marrow cells were flushed from the hind legs using PBS (Invitrogen) and a 23-gauge needle. Spleen cells were collected by crushing the spleen with the plunger of a 5 ml disposable syringe. Red blood cells were lysed with 1X RBC lysis buffer (Invitrogen 501129743) followed by mouse CD45 depletion using mouse CD45 microbeads (Miltenyi 130052301) according to the manufacturer's instructions. Unbound cells, the recovered and expanded primary human megakaryoblasts, were analyzed for purity by flow cytometry and used for downstream applications including Western blot, RNA-sequencing, and CUT&RUN-sequencing.

Flow Cytometry

Analysis

Fresh or thawed cells from liquid nitrogen storage were washed twice with PBS and re-suspended in 50 ul of PBS with 1 ul of mouse and/or human Fc block (mouse: Tonbo 700161; human: BD 564220) and incubated at room temperature for 10 minutes. Cells were then stained in a total of 50 ul of PBS for 20 minutes in the dark at room temperature with the following antibodies: live/dead ghost dye violet 510 (Tonbo 130870), APC-human CD34 (Miltenyi 130113176, Clone AC136), PE-human CD41 (Life Technologies MHCD4104, Clone VIPL3), PerCP-Cy5.5-mouse CD45 (BioLegend 103132, Clone 30-F11). After surface staining, cells were washed twice with staining buffer (1X PBS, 2% FBS, 1 mM EDTA), resuspended in 250 ul of staining buffer, filtered and transferred to 5 ml polystyrene flow cytometry tubes and analyzed on a BD (Becton-Dickinson) Fortessa (St. Jude Children's Research Hospital) and Cytek DxP10 (Stanford University) flow cytometers. Data was analyzed using FlowJo software.

Sorting

Fresh or thawed cells from liquid nitrogen storage were washed twice with PBS and re-suspended in 750 ul of staining buffer (1X PBS, 2% FBS, 1 mM EDTA), filtered, and transferred to 15 ml polypropylene tubes. 10 ul of DAPI staining solution (Miltenyi Biotec 130111570) was added and cells were sorted on BD Biosciences high-speed cell sorters at the St. Jude Children's Research Hospital Flow Cytometry and Cell Sorting Shared Resource and the Stanford Shared FACS Facility. Sorted cells were collected in 1 ml of staining buffer in 15 ml polypropylene tubes and washed twice with PBS before downstream applications including injection and RNA extraction. Data was analyzed using FlowJo software.

RT-PCR

The iScript cDNA synthesis kit (Bio-Rad 1708890) was used to generate cDNA from primary human megakaryoblasts according to the manufacturer's instructions. The Advantage 2 PCR Enzyme System (Takara Bio USA 639207) and the following primer pair were used to amplify the fusion junction DNA sequence:
FW: 5'CATCTGGAGGAAGGCTGA3'
RV: 5'CGGTTGTGGATCTTCAGGT3'

Lentiviral Vectors

Two CBFA2T3-GLIS2 encoding lentiviral vectors were cloned for these studies. Overlap extension PCR was used to clone the N-terminal 3xFLAG tag (Sigma) *CBFA2T3-GLIS2* coding sequence into the pCDH.MSCV.MCS.EF1acopGFP_CD711B-1 lentivirus backbone. Overlap extension PCR was also used to clone the N-terminal 3xTY1 tag *CBFA2T3-GLIS2* coding sequence into the CSI.pEF1a.IRES.EmGFP (CEI) lentivirus backbone.

Site-Directed Mutagenesis

The Pfu Ultra High-Fidelity DNA Polymerase and 10X Buffer kit (Agilent 600380) was used according to the manufacturer's instructions to delete NHR1, NHR2, and make the PXDLS mutation DL380,487AS using the following primers:
NHR1Δ FW:
5'CAGCACCTGCCCCCAGCCTGCGGGCTGCTGGACGCCAGCGCCTCCTCCCCC3'

NHR1Δ RV:
5'GGGGGAGGAGGCGCTGGCGTCCAGCAGCCCGCAGGCTGGGGGCAGGTGCTG3'

NHR2Δ FW:
5’CGGCAGGAAGAAGTGATCGACCACAAGGACACAAAGAAGGGCCCCGCTCCC3’

NHR2Δ RV:
5’GGGAGCGGGGCCCTTCTTTGTGTCCTTGTGGTCGATCACTTCTTCCTGCCG3’

PXDLS DL380AS FW:
5’CCTGCTGCCAGGCACCGTGCTGGCCTCGTCCACGGGCGTCAACTCAGCTG3’

PXDLS DL380AS RV:
5’CAGCTGAGTTGACGCCCGTGGACGAGGCCAGCACGGTGCCTGGCAGCAGG3’

PXDLS DL487AS FW:
5’ACCCCTGGCCCCCGGCCCCTTGCCTCCAGTGCCCTGGCCTGTGGCAACG3’

PXDLS DL487AS RV:
5’CGTTGCCACAGGCCAGGGCACTGGAGGCAAGGGGGCCGGGGGCCAGGGGT3’

Co-Immunoprecipitation and Western Blotting

Transiently transfected 293T cells or stably transduced primary human megakaryoblasts with the pCDH-3xFLAG-CBFA2T3-GLIS2 or CEI-3xTY-CBFA2T3-GLIS2 lentivirus vectors were washed twice with PBS. 293T cells were lysed in RIPA buffer (Pierce 89901) supplemented with 1 complete mini protease inhibitor tablet (Sigma 4693124001) according to the manufacturer’s instructions on ice for 25 minutes. Debris was pelleted by centrifugation at 14,000xg for 15 min at 4 degrees C and the supernatant was collected. Primary human megakaryoblasts were lysed in NP40 lysis buffer (Thermo Fisher FNN0021) supplemented with 1 complete mini protease inhibitor tablet (Sigma 44693124001) according to the manufacturer’s instructions on ice for 30 minutes. Debris was pelleted by centrifugation at 13,000 rpm for 10 min at 4 degrees C and the supernatant was collected. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher 23225) according to manufacturer’s instructions. 25 ug of protein was saved for input and 500 ug of protein was incubated overnight at 4 degrees C with 5 ug of primary antibody: ETO (polyclonal rabbit Bethyl A303-509A); CtBP1 (polyclonal rabbit Thermo 10972-1-AP); p300 (polyclonal rabbit Thermo PA1-848), TY1 (polyclonal rabbit GenScript A01004-40), FLAG (polyclonal rabbit Sigma F7425).

The next morning, 67 ul of Protein A Dynabeads (Thermo 1002D) was added to the protein lysate and incubated for two hours at 4 degrees C. After the incubation, beads were collected on a magnet and washed twice with ChIP low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH8.0, 150 mM NaCl) followed by two washes with PBS supplemented with 0.02% Tween 20. Bound proteins were eluted off

the beads with 50 ul of 2x Laemmli Sample Buffer (Bio-Rad 1610737) with β-mercaptoethanol. Samples were incubated at 95 degrees C at 500 rpm for 45 minutes. Eluted proteins were collected from beads on magnet stand and separated on a 4-20% Mini Protean TGX Stain Free Gel (Bio-Rad 4568094) at 100V for 90 minutes. The gel was transferred using Trans-Blot Turbo Mini 0.2 μm PVDF Transfer Packs (Bio-Rad 1704156) and the mixed molecular weight setting on the Bio-Rad Turbo Transfer System.

The PVDF membrane was blocked in 5% milk for one hour and incubated with 1:1000 dilution of primary antibody in 1% milk overnight at 4 degrees with rocking: ETO (monoclonal mouse Santa Cruz 134335); CtBP1 (monoclonal mouse Santa Cruz 17759); p300 (monoclonal mouse Santa Cruz 48343); TY1 (monoclonal mouse Thermo MA523513); FLAG (monoclonal mouse Sigma F3165). The next morning the membranes were washed five times with TBS-T (Fisher 28358) and then incubated with secondary antibody (HRP-Goat-Anti-Mouse IgG Light Chain Specific Secondary, Jackson 115035174) at 1:2000 dilution in 1% milk for one hour at room temperature with rocking. After the incubation, membranes were washed five times with TBS-T and developed with Clarity Western ECL Substrate (Bio-Rad 1705061) on a LI-COR imager. Membranes were stripped of bound antibody using Restore Western Blot Stripping Buffer (Thermo Fisher 21059) for 30 minutes at room temperature with rocking. Membranes were re-blocked with 5% milk for one hour at room temperature with rocking and then incubated overnight at 4 degrees with rocking with the primary antibody used for the immunoprecipitation to confirm enrichment. The next day, goat anti-rabbit IgG (H+L) superclonal HRP secondary (Thermo Fisher A27036) was applied and blots were developed with ECL Substrate.

RNA-Sequencing

RNA Extraction

50,000 – 500,000 flow sorted primary megakaryoblasts were washed twice with PBS and re-suspended in 500 ul of TRIzol Reagent (ThermoFisher 15596026) for five minutes at room temperature. RNA was extracted according to the manufacturer's instructions with all volumes cut in half due to the low input number of cells. 100 ul of chloroform was added and the top aqueous phase was collected. 1.2 ul of glycogen and 250 ul of 100% 2-propanol was added and the pellet was washed once with 500 ul of 75% ethanol. The pellets were air dried and re-suspended in 10 ul of warmed, RNAse-free water. RNA concentration was measured on a Nanodrop.

Sequencing

Low input library prep (Ovation/NEB) and sequencing on the Illumina NovaSeq6000 (S4) PE150 sequencer was performed at the Hartwell Center for Bioinformatics and Biotechnology at St. Jude Children's Research Hospital and

Vanderbilt Technologies for Advanced Genomics at Vanderbilt University Medical Center.

Differential Gene Expression Analysis

Raw paired-end FASTQ files were filtered to remove low quality reads and reads with adaptor contamination. STAR v2.7.la was used to align clean reads to the human genome (hg38) and to generate gene counts. STAR output for each sample was combined for differential expression testing using DESeq2_1.32.0 R package installed in R v4.1 using default settings. PCA was performed on rlog-transformed data (for sample size <20) or vst-transformed (for sample size $\geq$ 20) using DESeq2. Gene set enrichment analysis was performed using GSEA v4.2.2 (Mootha et al., 2003; Subramanian et al., 2005) parameters recommended for sample size.

Pearson Correlation

Intensity raw data was screened for missing values for exclusion before log2 transformation in R. Data was subset into Transplant and Patient groups, then the rowMean was taken by ID in each group. Cross-correlation plot was generated and correlation was measured using Pearson method.

Mass Spectrometry-Based Proteomics

100 ug of protein from 10 million normal primary megakaryoblasts, fusion transduced primary megakaryoblasts, and CBFA2T3-GLIS2 patient samples were analyzed in technical replicate by tandem mass tag (TMT) mass spectrometry at the Center for Proteomics and Metabolomics at St. Jude Children's Research Hospital. Protein lysates were digested and desalted and then underwent 16-plex TMT labeling. Samples were analyzed for total proteome by low pH reverse-phase liquid chromatography tandem mass spectrometry (LC-MS/MS). Retrieved data was searched against the UniProt human database to identify peptides. This was determined by assigning related mass addition to all possible amino acid residues followed by an algorithm that evaluates the accuracy of modification localization. Peptide quantification data was corrected for mixing errors and summarized to derive protein quantification results. Using this approach, the proteomics core routinely analyzes about 10,000 proteins which provides a comprehensive proteome analysis.

CUT&RUN-Sequencing

Fresh or thawed primary megakaryoblasts were cross-linked with 0.1% formaldehyde for one minute at room temperature. The reaction was quenched with 2.5M glycine. Cross-linked cells were then used with the CUTANA ChIC/CUT&RUN Kit (EpiCypher 14-1048) per the manufacturer's instructions. 500,000 live cells per

immunoprecipitation were washed twice with wash buffer (pre-wash buffer supplemented with complete mini protease inhibitor (Sigma 4693124001), 1M spermidine, 1% Triton X-100, and 0.05% SDS) and then bound to ConA Beads at room temperature for 10 minutes.

Cell-bound beads were collected and re-suspended in Antibody Buffer (Wash Buffer supplemented with 5% digitonin and 0.5M EDTA) and 0.5 ug polyclonal rabbit antibody of interest was added. Polyclonal rabbit antibodies used in the studies described here are as follows: ETO (Bethyl A303-509A); CtBP1 (Thermo 10972-1-AP); TY1 (GenScript A01004-40). Beads were incubated at 4 degrees C overnight on a nutator. The next day, beads were collected and washed twice with cold Cell Permeabilization Buffer (Wash Buffer supplemented with 5% digitonin) prior to incubation with pAG-MNase for ten minutes at room temperature. Beads were collected and washed twice again with Cell Permeabilization Buffer and then 100 mM of Calcium Chloride was added on ice to activate the enzymatic reaction. Beads were incubated at 4 degrees C for two hours on a nutator. After the incubation, Stop Buffer was added and immunoprecipitated protein-DNA complexes were eluted off the beads via incubation at 37 degrees C for 10 min.

The supernatant was collected and de-crosslinked was performed by adding SDS and Proteinase K and incubating at 55 degrees C overnight. The next day, DNA was purified using the provided DNA Cleanup Columns and DNA Binding and Wash Buffer reagents. DNA concentration was measured using the Qubit dsDNA HS kit (Fisher Q32851). Library prep (NEBNext/ChIPSeq) and sequencing on the Illumina NovaSeq600 (S4) PE150 sequencer was performed at Vanderbilt Technologies for Advanced Genomics at Vanderbilt University Medical Center. Samples were sequenced in technical (ETO) or biological replicate (CtBP1, fusion).

Fastq files were trimmed and aligned to the human Hg19 genome via Basepair (https://app.basepairtech.com). Bam files were extracted and MACS v2.2.71 was used to call peaks in each sample relative to IgG control background (using parameters -f BAMPE, -g 2700000000, --pvalue 1e-3, --call-summits, -B and -SPMR). To build the single tracks for each sample, macs2 bdgcmp was used to generate fold-enrichment track (using paramters – M FE and -p 0.00001) and resulting bedGraphs were fixed for conversion to bigwig files (using bdg2bw.sh and bedGraphToBigWig method by bedtools v2.27.1). To calculate correlation between replicates before combining replicates, multiBigWigSummary (paramters bins and -chromosomesToSkip, “chrX”, “chrY”, “chrM”, and “random”) and plotCorrelation (paramters – corMethod pearson – skipZeros, -- whatToPlot scatterplot and -removeOutliers) from deeptools v3.3.0 were used. Annotation of narrowPeak files was performed using Homer v4.11 and hg19 as reference. Genes associated with the promoter regions (in chr1-22) were intersected between replicates for each sample set to generate a gene list for each sample set. For pathway enrichment analysis on subsets of genes, Over Representation Analysis was performed using biomaRt_2.50.3, ReactomePA1_1.38.0, and clusterProfiler_4.2.2 R packages (Boyle et al., 2014).

Homer Motif Analysis

Motif discovery was performed using Homers' findMotifsGenome.pl with genomics regions as input and hg19 as reference (using parameters -size given and –mask). For motif scanning analysis using custom motif files, the same method and parameters were used along with the -find function.

Schematics

Schematics were made with Biorender (https://biorender.com).

Results

CBFA2T3-GLIS2 Can Act as a Driver of Non-DS-AMKL in a Humanized Murine Model

CBFA2T3-GLIS2 positive primary human megakaryoblasts, herein referred to as C-G+ pMks, induce a serially transplantable leukemia in immunodeficient mice within 180 days compared to empty vector transduced pMKs, herein referred to as C-G- pMks, which do not proliferate in vivo (**Figure 2-3** and **Supplemental Mouse Data for Chapter 2**). The lack of decreased disease onset in the secondary and tertiary transplants suggests the model is genomically stable. Moribund mice presented with hind leg paralysis and enlarged spleens. Necropsy and histology analysis showed the presence of neoplastic infiltrates in the bone marrow, spleen, and central nervous system, the latter of which caused the hind leg paralysis due to compression of the spinal cord (**Figure 2-4A**). Flow cytometry analysis of mouse CD45 depleted cells from the bone marrow and spleen show that the infiltrating leukemic cells are GFP positive, confirming they are expanded transplanted cells (**Figure 2-4B**). Expression of the CBFA2T3-GLIS2 fusion in transformed cells was confirmed through RT-PCR (**Figure 2-4C**) and Western blot (**Figure 2-4D**). The humanized murine model confirms the hypothesis that the CBFA2T3-GLIS2 chimeric transcription factor can act as a driver of leukemia when expressed in pMks.

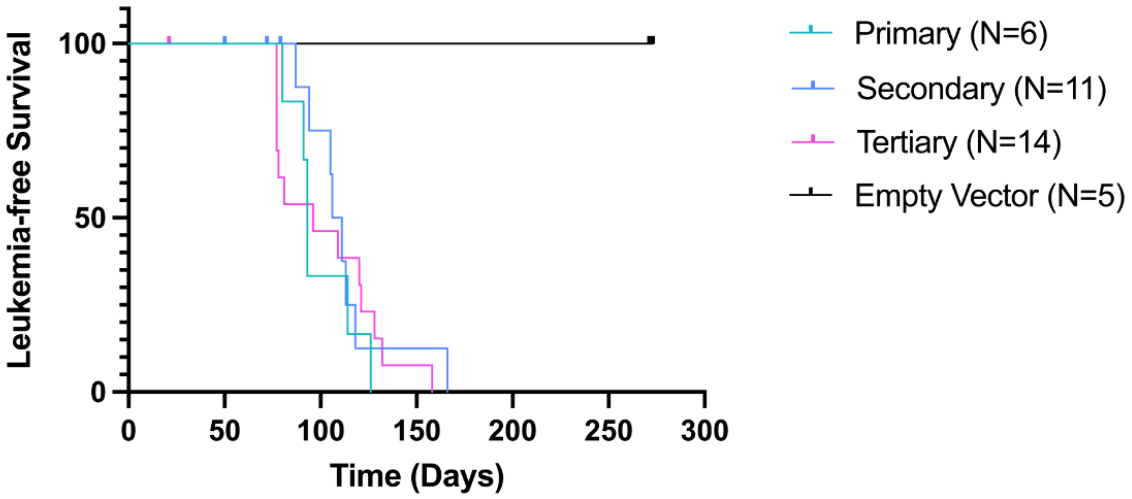


Figure 2-3. CBFA2T3-GLIS2 Transduced Primary Human Megakaryoblasts Induce a Serially Transplantable AMKL.
Primary human megakaryoblasts were transduced with an empty lentiviral vector or a vector encoding the *CBFA2T3-GLIS2* cDNA sequence and transplanted into immunodeficient mice. Transformed cells were isolated from moribund mice and injected into secondary and subsequent tertiary recipients. Leukemia-free survival is shown.

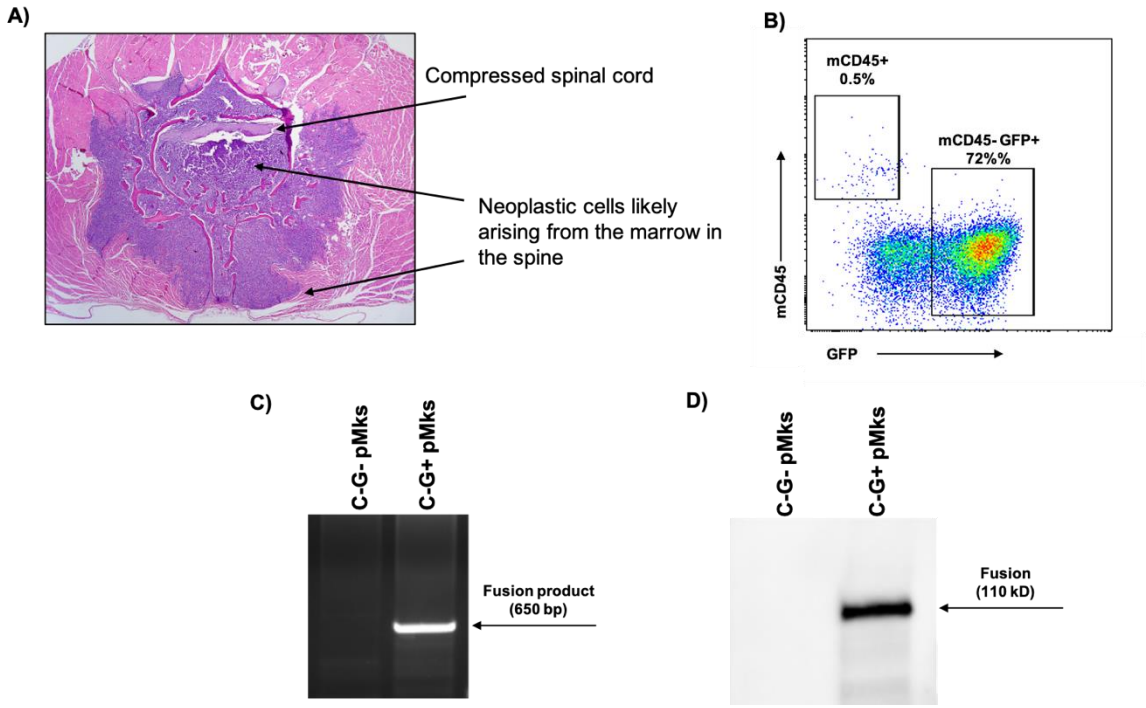


Figure 2-4. Characterization of Transplanted CBFA2T3-GLIS2 Positive Primary Human Megakaryoblasts.

A) Cross-section from a representative spinal column of a moribund primary transplant that presented with hind leg paralysis. Arrows point to compression of the spinal cord from the infiltrating neoplastic cells that caused the hind leg paralysis. **B)** Representative flow cytometry analysis plot of mouse CD45 depleted bone marrow from a moribund tertiary transplant. Cells were stained with a fluorescent antibody against mouse CD45. **C)** Reverse-transcriptase PCR (RT-PCR) and **D)** Western blot using an anti-FLAG antibody for detection of the CBFA2T3-GLIS2 fusion was performed on bone marrow cells from a moribund primary transplant (C-G+ pMks) and in normal megakaryoblasts lacking the fusion (C-G- pMks). The arrows point to the fusion band.

CBFA2T3-GLIS2 Transformed Primary Human Megakaryoblasts Recapitulate Patient Samples on a Transcriptional and Translational Level

RNA-sequencing of C-G+ pMks cultured *in vitro* and mouse CD45 depleted spleen cells from serially transplanted immunodeficient mice show a gene expression pattern similar to that of CBFA2T3-GLIS2 positive patient diagnostic bone marrow samples. As previously reported (Gruber et al., 2012; de Rooij et al., 2017), CBFA2T3-GLIS2 positive AMKL displays a distinct gene expression pattern compared to all other AMKL genomic subsets (**Figure 2-5A**). There is a strong correlation ($R = 0.91$, $p < 2.2e-16$) between transplanted C-G+ pMks and fusion positive patient samples, confirming the transplanted samples recapitulate the patient samples on a transcriptional level (**Figure 2-5B**).

To evaluate the correlation on a translational level, mass spectrometry protein quantification was performed on mouse CD45 depleted spleen cells from one secondary transplant and one tertiary transplant, mouse CD45 depleted spleen cells from four CBFA2T3-GLIS2 positive patient-derived xenografts, and C-G- pMks cultured *in vitro* to serve as a normal control. Each sample was analyzed in technical replicate. Tandem mass tagging was utilized to accurately and sensitively identify proteins and measure their relative abundance in each sample (Thompson et al., 2003). Proteomics analysis also found a strong correlation ($R = 0.98$, $p < 2.2e-16$) between transplant and patient samples on a translational level (**Figure 2-5C).**

Collectively, this data validates that the transformed primary human megakaryoblasts recapitulate the expression profile of CBFA2T3-GLIS2 positive patient AMKL and highlights the novelty of the humanized murine model, which can easily be manipulated for downstream experiments to understand components critical to the CBFA2T3-GLIS2 transcriptional complex that will contribute to the development of novel targeted therapies.

CBFA2T3-GLIS2 Binds Sites Throughout the Genome and Induces Differential Gene Expression

To assess the genomic binding sites of CBFA2T3-GLIS2, cleavage under targets and release under nuclease sequencing (CUT&RUN-seq) was used in which the fusion was immunoprecipitated in secondary transplant C-G+ pMks using an antibody against the 3xTY1 tag at the N-terminus of the fusion cDNA sequence followed by sequencing of the associated DNA and comparison to an IgG control immunoprecipitation. A total of 117,357 gene regions were identified which corresponded to 6,289 unique genes bound by the fusion at promoters. 88% of the genes were expressed and 54% were differentially expressed when compared to C-G- pMks (**Figure 2-6A** and **Supplemental Sequencing Data for Chapter 2**). Although a significant proportion of peaks were at gene promoters (22% of total gene regions), 40% and 28% were bound at intron and intergenic regions, respectively,

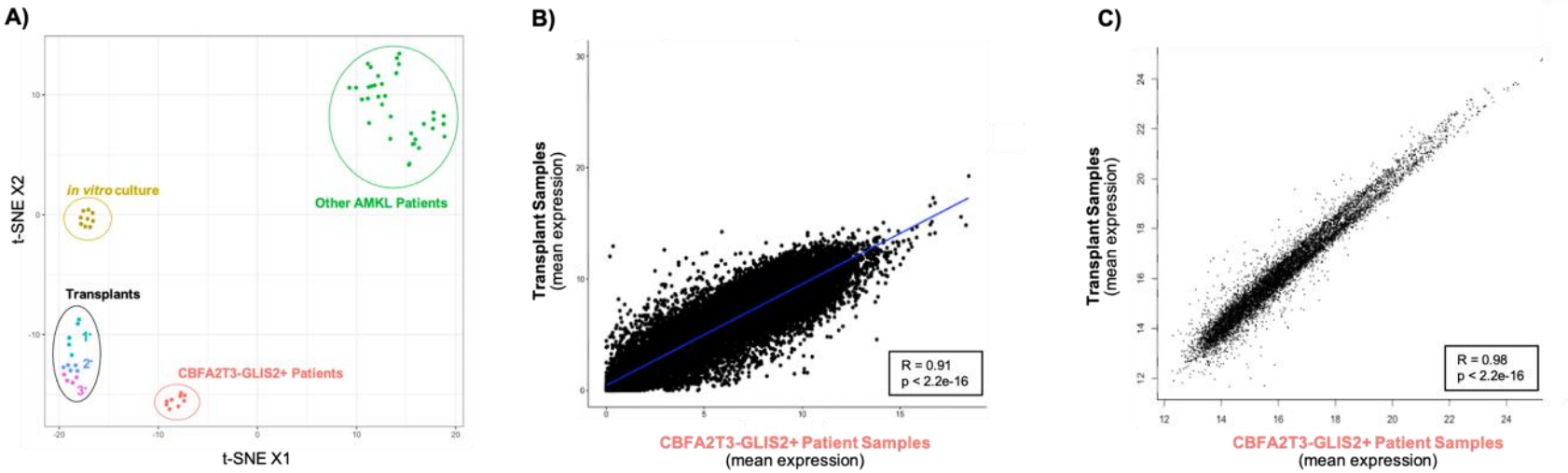


Figure 2-5. CBFA2T3-GLIS2 Positive Primary Human Megakaryoblasts Recapitulate Patient Samples on a Transcriptional and Translational Level.

A) t-SNE visualization of the top 1% differentially expressed genes. **B)** Cross-correlation plot of the top 1% differentially expressed genes between CBFA2T3-GLIS2 positive patient samples (n = 9) and primary, secondary, and tertiary transplant samples (n = 14). The correlation coefficient and probability value are displayed on the plot. **C)** Cross-correlation plot of the top 1% differentially expressed proteins between CBFA2T3-GLIS2 positive patient samples (n = 4 in technical replicate) and secondary and tertiary transplant samples (n = 2 in technical replicate). The correlation coefficient and probability value are displayed on the plot.

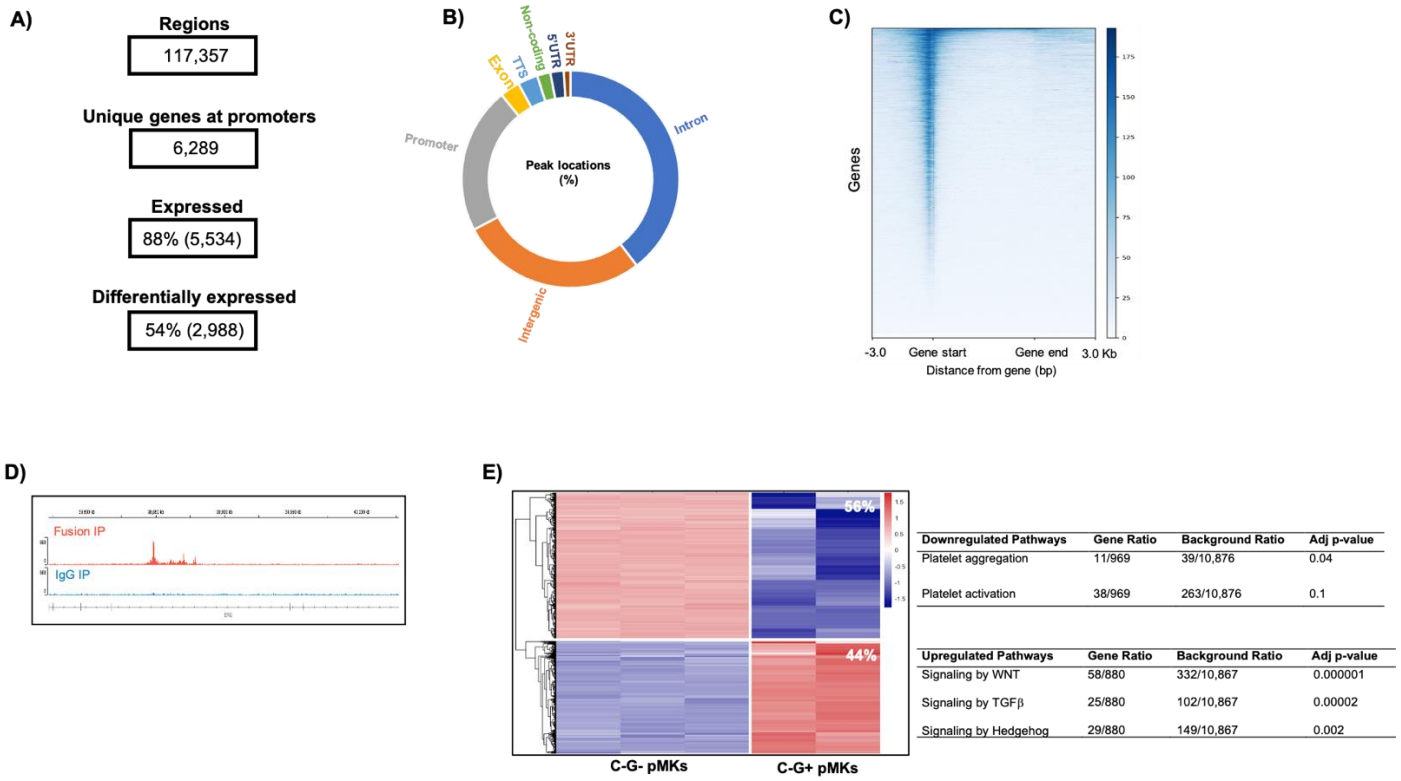


Figure 2-6. CBFA2T3-GLIS2 Binds Promoter Sites Throughout the Genome and Induces Differential Gene Expression. **A)** Summary of called CUT&RUN-seq peaks from 3xTY1 immunoprecipitation for the CBFA2T3-GLIS2 fusion in secondary transplant C-G+ pMks. **B)** Peak locations distribution of the 117,357 bound regions. 40% intron; 27% intergenic; 22% promoter; 3% exon, 3% TSS (transcription termination site); 2% non-coding, 2% 5'UTR (untranslated region); 1% 3'UTR. **C)** Heat map displaying the distance of the peaks to the nearest gene start site. **D)** Example called peak for *ERG* using Integrative Genomics Viewer (IGV) browser tool. **E)** Heatmap of differentially expressed, fusion bound genes in C-G+ pMks (n = 2) compared to C-G- pMks (n = 3). A total of 2,988 genes are displayed with a fold change > 1 and an adjusted p-value $\leq$ 0.01. Pathway analysis was performed using over-representation analysis (ORA) and the Reactome Pathway Database.

suggesting the fusion is also regulating genes within introns and binding at enhancer regions located upstream of promoter sites (**Figure 2-6B**). Overall, there is a strong enrichment at gene start sites (**Figure 2-6C**). One of the genes bound by the fusion was *ERG*, which was also reported by Thirant et al. in 2017 (**Figure 2-6D**).

The most striking result was that 54% of the expressed, fusion bound genes were differentially expressed when compared to C-G- pMks (**Figure 2-6A**). 56% of the genes were downregulated and 44% were upregulated, which confirms the hypothesis that the fusion alters gene expression in both directions (**Figure 2-6E**). Over-representation analysis (ORA), a statistical analysis to measure enrichment of certain pathways in a given dataset (Wieder et al., 2021), using the Reactome Pathway Database revealed enrichment of pathways associated with platelet aggregation and activation in the downregulated fusion-bound genes. This suggests the fusion is suppressing genes involved in differentiated, functional megakaryoblasts, consistent with AMKL cells that have a block in differentiation and are non-functional platelet pre-cursors. ORA on the subset of fusion bound genes that are upregulated revealed enrichment of the WNT, TGFβ, and Hedgehog signaling pathways (**Figure 2-6E** and **Supplemental Sequencing Data for Chapter 2**). These are particularly of interest because downstream targets of these pathways are upregulated in CBFA2T3-GLIS2 positive patient samples (Gruber et al., 2012), further supporting that the gene modified, fully transformed C-G+ pMks closely mimic patient AMKL cells.

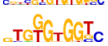
Given that the fusion retains the five highly conserved DNA binding zinc fingers of full-length GLIS2, I hypothesize a majority of genes bound by the fusion will contain a GLIS, or the closely related GLI, consensus sequence. First, the Homer known motif search was used to identify the most represented consensus sequences from the peak regions associated with the 6,289 fusion bound genes at promoters (**Figure 2-6A**) which, after merging peaks within 3 kb of each other and filtering for peaks with a minimum width of 51 and a max width of 10,533, resulted in 6,175 total peak regions. Consistent with Mercher's group, ETS, GATA, and RUNX motifs were found.

In contrast to Mercher's group, however, the search did not identify any Homer known motif matches belonging to the ZIC or GLIS2 family (Thirant et al., 2017) (**Table 2-1**). This is likely due to motif redundancy and lack of representation of GLI and GLIS consensus sequences in the Homer database. Homer known motif analysis relies on a curated database based on the analysis of published ChIP-seq data. Only high-quality ChIP-seq experiments where the top Homer motif resembled the consensus site for factors with the given DNA binding domain are used. To focus the search on the GLI superfamily of transcription factors that share a high degree of homology in the DNA binding zinc finger domains, a custom Homer motif search was used to determine the abundance of the GLIS and GLI family motifs. This search also included the ETS, GATA, and RUNX motifs as controls to verify consistent results between the Homer known motif analysis and Homer motif scanning using the custom motif files. The custom search revealed GLIS1-4 (SNSYSYRGGRGGBV) is the preferential sequence represented in fusion bound peak regions, as 28.6% of the input genomic regions contained this motif (**Table 2-2**).

Table 2-1. Known Homer Discovery Motif Analysis of Fusion Bound Genes.

Motif	Name	% of Targets	p-Value
	ETS	45.7	1e-76
	GATA	29.1	1e-9
	RUNX	22.7	1e-9

Table 2-2. Homer Custom Motif Search of Fusion Bound Genes.

Motif	Name	% Abundance of Motif	Threshold
	ETS	72.5 (4476/6175)	6.83
	GATA	40.0 (2448/6175)	5.94
	RUNX	27.9 (1721/6175)	6.65
	GLIS1-4	28.6 (1766/6175)	10.94
	GLI2-3	2.8 (173/6175)	11.93
	GLI2-2	1.7 (105/6175)	11.76
	GLI3-4	1.5 (90/6175)	12.33
	GLIS1-2	1.3 (78/6175)	13.19
	GLI3-1	0.4 (27/6175)	13.14
	GLIS2-3	0.2 (13/6175)	14.00
	GLI2-4	0.1 (8/6175)	14.12
	GLIS2-2	0.1 (7/6175)	14.40
	GLI2-1	0.1 (6/6175)	14.12
	GLIS1-3	0.08 (5/6175)	15.38
	GLI1-1	0.06 (4/6175)	14.52
	GLIS2-1	0.06 (4/6175)	15.10
	GLI3-2	0.03 (2/6175)	16.08
	GLIS3-1	0.02 (1/6175)	17.98

This is consistent with Vasanth et al. who reported full-length GLIS2 binds GLIS binding sequences with stronger affinity than GLI binding sequences (Vasanth et al., 2011).

ETO, CtBP1, and p300 Associate with the CBFA2T3-GLIS2 Transcriptional Complex

To test the theory that CBFA2T3-GLIS2 recruits the co-factors ETO, CtBP1, and p300 to the transcriptional complex (**Figure 2-1**), a series of co-immunoprecipitations were performed in transfected 293T cells, a well-established embryonic kidney cell line that lacks expression of the fusion. 293T cells were selected because they are easily expandable in culture and have high transfection efficiency. Cells were transiently transfected for forty-eight hours with a lentiviral construct encoding 3xFLAG-CBFA2T3-GLIS2 or one of several mutant constructs outlined in **Table 2-3**. I hypothesize deletion of NHR1 will lead to a partial loss of p300, deletion of NHR2 will lead to a complete loss of ETO, mutation of the PXDLS motifs will lead to a complete loss of CtPB1 as was shown previously in full-length GLIS2 (Kim, S et al., 2005), and a combination of all three alterations will lead to complete disruption of p300, ETO, and CtBP1 with the transcriptional complex (**Table 2-3**).

Co-immunoprecipitation for the fusion and ETO confirmed that the fusion directly binds ETO through NHR2 as the interaction is lost when NHR2 is deleted (**Figure 2-7A, B**). The association of ETO with the fusion was confirmed in C-G+ pMks from a tertiary transplant (**Figure 2-7C, D**). NHR2 also promotes oligomerization of MTG family members (**Figure 2-1**). Therefore, I hypothesize that loss of NHR2 will not only disrupt ETO binding to the fusion, but also dimerization of the fusion. To test this, 293T cells were co-transfected for forty-eight hours with a lentiviral construct encoding 3xTY1-CBFA2T3-GLIS2 and a lentiviral construct encoding either unmanipulated 3xFLAG-CBFA2T3-GLIS2 or 3xFLAG-CBFA2T3-GLIS2 with NHR1 or NHR2 deleted. Consistent with the hypothesis, co-immunoprecipitation for the TY and FLAG tagged fusion constructs showed a reduction of dimerization when NHR2 is deleted but not NHR1 (**Figure 2-8A, B**). Unlike ETO binding, dimerization is not completely abrogated when NHR2 is deleted. This may be due to the ability of NHR3 to support oligomerization (Wu et al., 2005). The ability of NHR2 deletion constructs to dimerize, albeit at a significantly reduced level, suggests the loss of leukemogenesis cannot be solely due to the loss of oligomerization, arguing in favor of ETO binding playing an important role.

In contrast to the Kim et al. 2005 paper, which showed mutation of all four PXDLS sequences in full-length GLIS2 abrogated binding of CtBP1, mutation of the only two PXDLS sequences (DL380, 487AS) present in the GLIS2 portion of the fusion did not disrupt CtBP1 binding. Deletion of NHR1 and 2 in addition to the mutations did abrogate CtBP1 direct binding to the fusion (**Figure 2-9A**); however, the fusion was still detected when CtBP1 was pulled down albeit at reduced levels compared to wild type,

Table 2-3. CBFA2T3-GLIS2 Mutants.

Motif	Predicted Effect
NHR1Δ	Partial loss of p300
NHR2Δ	Loss of ETO, disruption of dimerization
PXDLS DL380,487AS	Loss of CtBP1
NHR1-2Δ; PXDLS DL380,487AS	Loss of p300, ETO, and CtBP1

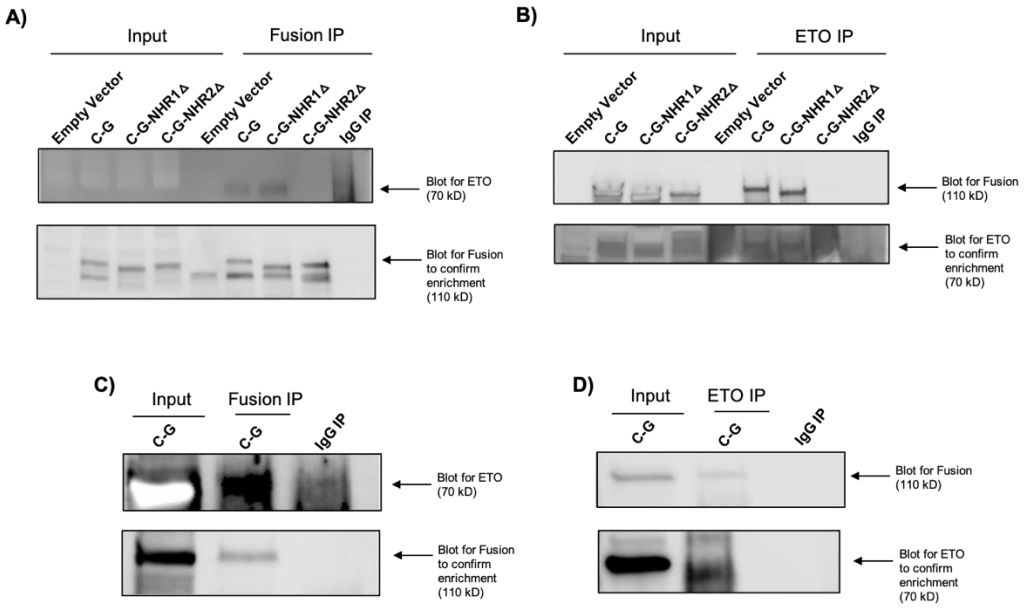


Figure 2-7. ETO Directly Binds CBFA2T3-GLIS2 Through NHR2.
A) Immunoprecipitation for the fusion via the N-terminal 3xFLAG tag and Western blot for ETO showed association in the unmanipulated fusion construct that is lost upon deletion of NHR2 in transiently transfected 293T cells. **B)** This was confirmed via the reverse co-immunoprecipitation. **C-D)** The association of ETO and the fusion was verified in C-G+ pMks from a tertiary transplant in both directions. Arrows point to fusion or ETO bands.

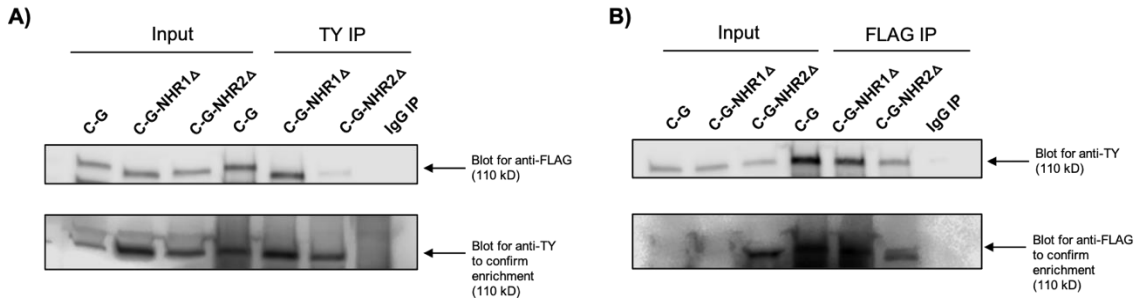


Figure 2-8. Dimerization of the Fusion Is Dependent on NHR2.
A) Immunoprecipitation for the TY-CBFA2T3-GLIS2 construct and Western blot for the unmanipulated or manipulated FLAG-CBFA2T3-GLIS2 construct shows that dimerization is mediated, in part, by NHR2. **B)** This was confirmed via the reverse co-immunoprecipitation.

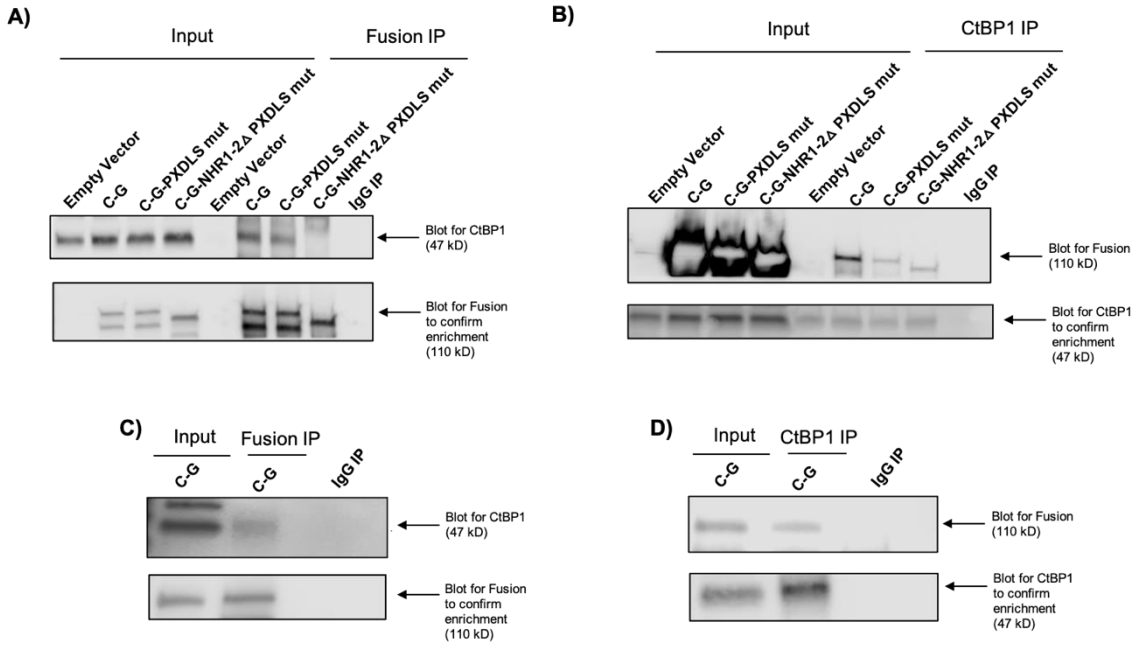


Figure 2-9. CtBP1 Directly Binds CBFA2T3-GLIS2 Through NHR1 and the PXDLS Domain.
A) Immunoprecipitation for the fusion via the N-terminal 3xFLAG tag and Western blot for CtBP1 showed abrogation of direct CtBP1 binding when NHR1-2 was deleted and the PXDLS domain was mutated. **B)** In contrast, CtBP1 still indirectly associates with the fusion even when NHR1-2 was deleted and the PXDLS domain was mutated. **C-D)** The association of CtBP1 and the fusion was verified in C-G+ pMks from a tertiary transplant in both directions. Arrows point to fusion or CtBP1 bands.

suggesting CtBP1 can still indirectly associates with the fusion transcriptional complex to some degree (**Figure 2-9B**). This discrepancy is potentially mediated by CtBP1 binding the bromodomain of p300 (Kim, J et al., 2005). The interaction between CtBP1 and the fusion was confirmed in C-G+ pMks from a tertiary transplant (**Figure 2-9C,D**). Because the mutant constructs did not completely block the association of CtBP1 with the fusion, it is not possible to determine the requirement of CtBP1 in leukemogenesis using this experimental design. CRISPR-Cas9 editing technology would need to be employed to knock-out CtBP1 from the transduced megakaryoblasts.

Surprisingly, a direct binding of p300 to CBFA2T3-GLIS2 was not confirmed, as there was no co-immunoprecipitation of p300 when the fusion was enriched (**Figure 2-10A**). However, p300 does indirectly associate with the fusion transcriptional complex as is evidenced by co-immunoprecipitation of the fusion when p300 is enriched (**Figure 2-10B**). Due to the lack of direct binding, deletion of NHR1 and 2 in combination with mutation of the PXDLS motifs did not abrogate p300 recruitment, although it is reduced (**Figure 2-10B**). The indirect interaction between p300 and the fusion was confirmed in C-G+ pMks from a tertiary transplant (**Figure 2-10C**). Collectively, this data supports the theory that the reduced, although still present, association of CtBP1 with the fusion transcriptional complex even in the presence of the NHR1-2 deletion and PXDLS motif mutation construct (**Figure 2-9B)** is mediated, in part, by the indirect recruitment of p300.

In summary, ETO directly binds to CBFA2T3-GLIS2 through NHR2. NHR2 plays a significant role in the dimerization of the fusion. CtBP1 directly binds to the fusion through NHR1 and the PXDLS sequences; however, due to the indirect association of p300 with the fusion transcriptional complex, this data supports a model whereby CtBP1 can also be indirectly be recruited to the complex even when those sites are altered. Because CtBP1 and p300 retain the ability to associate with the mutant constructs, I predict loss of NHR2, and thereby abrogation of ETO heterodimerization will be the most detrimental to leukemogenesis.

ETO and CtBP1 Co-Occupy CBFA2T3-GLIS2-Bound Genes

After confirming ETO and CtBP1 bind directly with the fusion, I next wanted to map the genomic co-occupancy of these two transcription factors and assess their effects on gene expression. CUT&RUN-seq was performed in C-G+ pMks from a secondary transplant and C-G- pMks as a normal control. ETO bound 698 genes in the normal, C-G- pMks whereas ETO bound 4,745 genes in the malignant C-G+ pMks; a 580% increase in bound sites. CtBP1 bound 2,246 genes in the normal, C-G- pMks and 1,689 in the C-G+ pMks; a 25% decrease in bound genes (**Figure 2-11** and **Supplemental Sequencing Data for Chapter 2**). For both ETO and CtBP1, a significant percentage of sites gained were also co-occupied by the fusion, 68% (3,220/4,745) and 54% (913/1,689) respectively, arguing in favor of recruitment to these sites by the fusion itself (**Figure 2-12**).

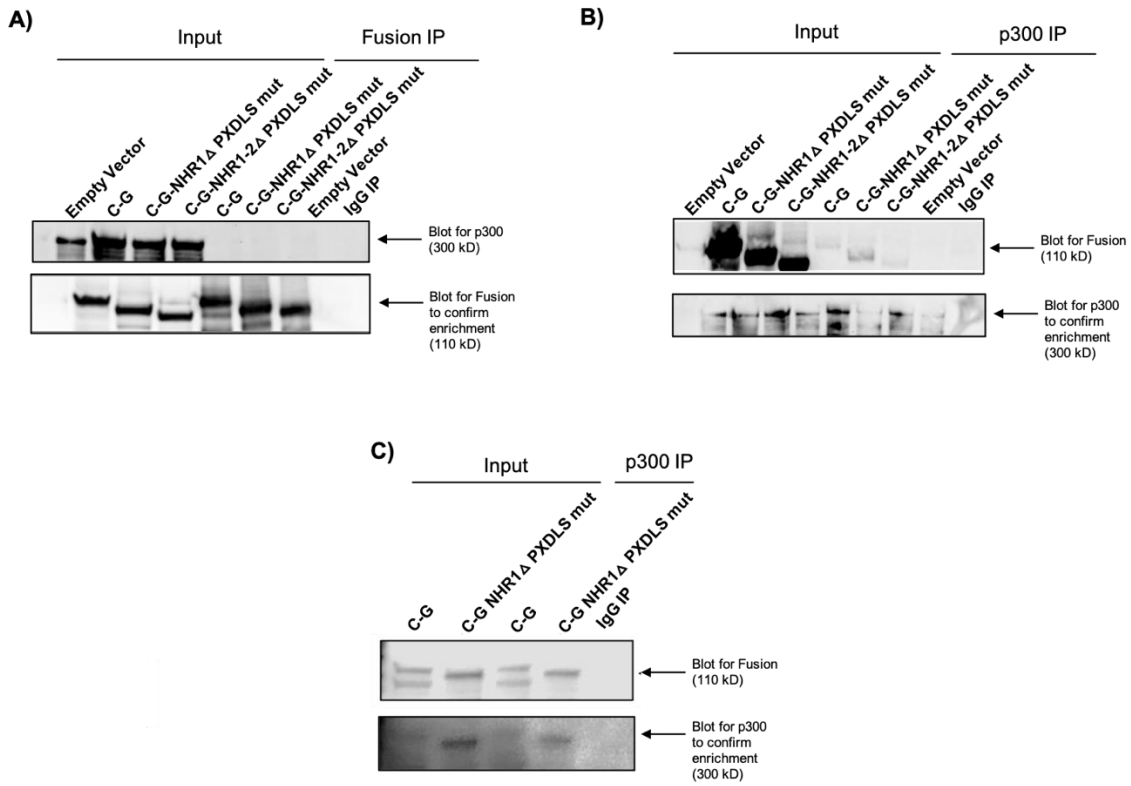


Figure 2-10. p300 Indirectly Associates with CBFA2T3-GLIS2.
A) Immunoprecipitation for the fusion via the N-terminal 3xFLAG tag and Western blot for p300 did not show a direct interaction. **B)** However, p300 does indirectly associate with the fusion transcriptional complex even with NHR1-2 deleted and the PXDLS domain mutated, likely due to the lack of requirement for direct fusion binding. **C)** The indirect association of p300 with the fusion was verified in C-G+ pMks from a tertiary transplant. Arrows point to fusion or p300 bands.

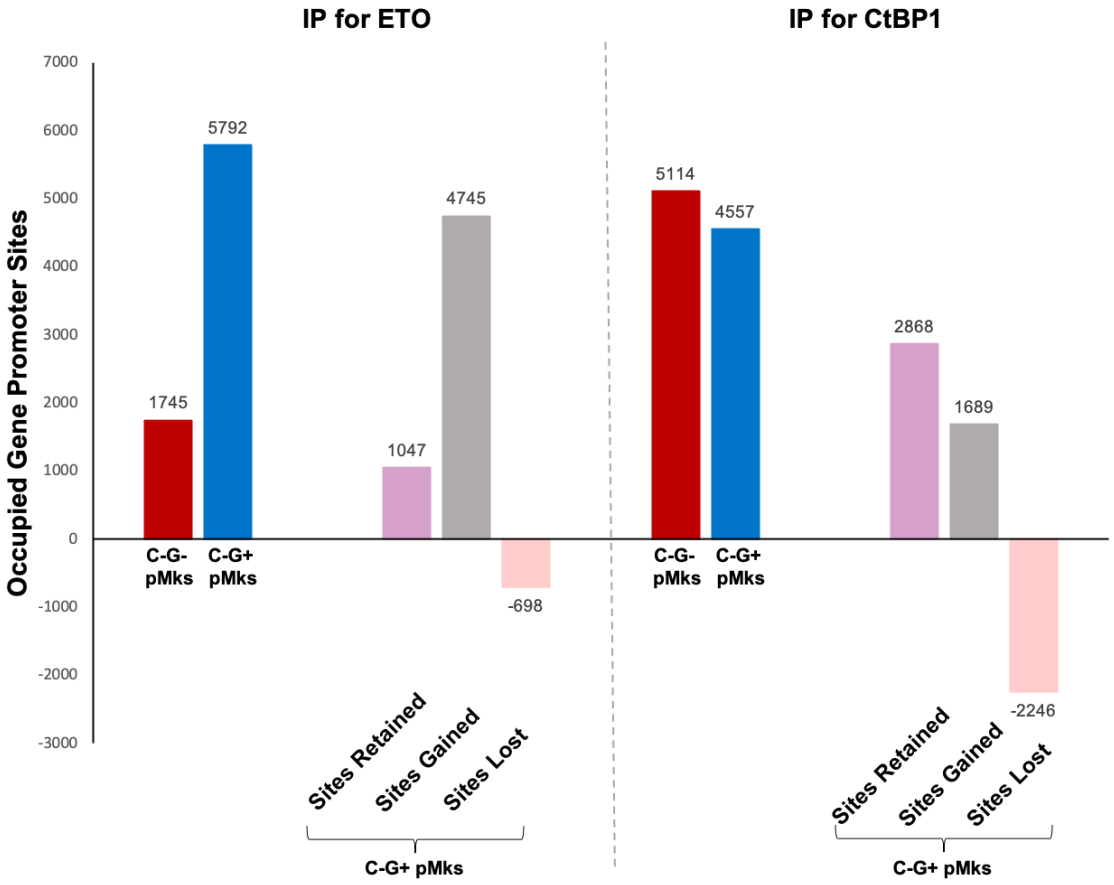


Figure 2-11. Summary of ETO- and CtBP1-Bound Genes.
CUT&RUN-sequencing was performed to immunoprecipitate (IP) for ETO or CtBP1 in C-G- pMks and C-G+ pMks. The number of genes occupied by ETO or CtBP1 at promoter sites and the number of sites retained, gained, and lost in C-G+ pMks compared to C-G- pMks are summarized.

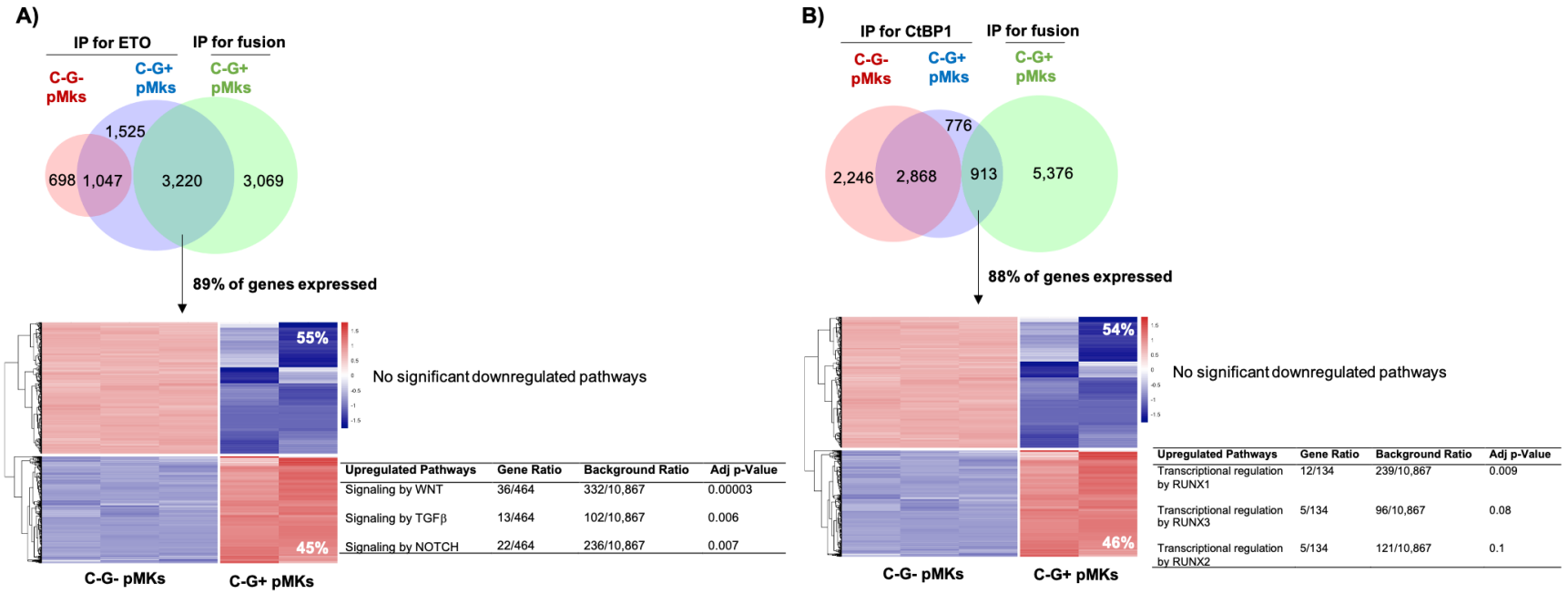


Figure 2-12. ETO and CtBP1 Co-Occupy CBFA2T3-GLIS2 Bound Genes and Alter Gene Expression.
A) Venn diagram comparing genes bound by ETO in C-G- pMks versus C-G+ pMks and comparing genes co-occupied by both ETO and the fusion in C-G+ pMks. Heatmap of differentially expressed genes co-occupied by ETO and the fusion. A total of 2,865 genes are displayed with a fold-change > 1 and an adjusted p-value < 0.01. Pathway analysis was performed using over-representation analysis (ORA) and the Reactome Pathway Database. **B)** Venn diagram comparing genes bound by CtBP1 in C-G- pMks versus C-G+ pMks and comparing genes co-occupied by both CtBP1 and the fusion in C-G+ pMks. Heatmap of differentially expressed genes co-occupied by CtBP1 and the fusion. A total of 803 genes are displayed with a fold change > 1 and an adjusted p-value < 0.01. Pathway analysis was performed using over-representation analysis (ORA) and the Reactome Pathway Database.

Specifically, when the C-G+ pMks ETO bound genes were compared with the fusion bound genes (**Figure 2-6**), 3,220 genes (51%) were co-occupied by both ETO and the fusion. 89% of the shared genes were expressed, with 55% downregulated and 45% upregulated. Over-representation analysis (ORA) using the Reactome Pathway Database did not yield any significantly enriched pathways for the downregulated gene set but did reveal enrichment of WNT, TGFβ, and Notch signaling pathways in the upregulated gene set (**Figure 2-12A** and **Supplemental Sequencing Data for Chapter 2**). This is consistent with the fusion bound genes in the CUT&RUN-seq analysis (**Figure 2-6**) and suggests ETO plays a role in upregulation of these key regulatory pathways.

When the C-G+ pMks CtBP1 bound genes were compared with the fusion bound genes (**Figure 2-6**), 913 genes (14%) were co-occupied by both CtBP1 and the fusion. 88% of the shared genes were expressed, with 54% downregulated and 46% upregulated much like ETO co-occupied genes. Also as seen with ETO co-occupied genes, but in contrast to the hypothesis that CtBP1 association with the fusion promotes transcriptional repression, ORA using the Reactome Pathway Database did not find any significantly enriched pathways for the downregulated gene set.

There was, however, enrichment of runt-related transcription factor RUNX1 (also referred to as AML1), RUNX2 (also referred to as CBFA1), and RUNX3 (also referred to as CBFA3) transcriptional regulation in co-occupied genes that were upregulated (**Figure 2-12B** and **Supplemental Sequencing Data for Chapter 2)**. RUNX1 is well known for its role in AML1-ETO AML. Recent studies reported that upregulated *RUNX1* cooperates with the *FLT3-ITD* fusion, present in about 25% of all AML cases (Daver et al., 2019), to induce leukemogenesis whereas downregulation of *RUNX1* drives differentiation of AML cells (Said et al., 2021). Matthijssens et al. demonstrated that RUNX2 was required for survival of T-ALL cells *in vitro* and *in vivo* (Matthjssens et al., 2021) and RUNX3 has been shown as a super-enhancer associated gene that is highly expressed in AML cells, which is associated with a poor prognosis (Zhang et al., 2021).

ETO and CtBP1 are well known for their repressor activity; but, in the context of co-occupancy with CBFA2T3-GLIS2, this data suggests they are contributing to transcriptional activation of key pathways associated with proliferation and development. Both ETO and CtBP1 can bind the transcriptional activator p300 which may be contributing to the upregulation of fusion bound genes although I have shown p300 does not directly bind to the fusion. CUT&RUN-seq to map the co-occupancy of p300 with ETO, CtBP1, and the fusion will be used to test this theory. Jin et al. described a novel role for CtBP1 as an activator of *MDR1* (multidrug resistance mutation 1) in multidrug resistant cancer cell lines, suggesting CtBP1 itself may act as a transcriptional activator in some environments (Jin et al., 2007).

Loss of NHR2 Abrogates Leukemogenesis *in vivo*

To determine which structural components of the fusion are essential for leukemogenesis, I transplanted immunodeficient mice with pMks overexpressing the

unmanipulated CBFA2T3-GLIS2 coding sequence or one of the mutated sequences as outlined in **Table 2-3**. Deletion of NHR1 or mutation of the PXDLS motif, individually or in combination, did not affect leukemogenesis compared to the unmanipulated fusion, even though these modifications cause major structural changes to the fusion (**Figure 2-13A** and **Supplemental Mouse Data for Chapter 2**). Deletion of NHR2 alone or in combination with NHR1 deletion and mutation of PXDLS motif abrogated leukemogenesis compared to the unmanipulated fusion. These mutant pMks did not expand *in vivo* out to almost 300 days and flow cytometry and necropsy analysis of the bone marrow and spleen did not show any evidence of engrafted pMks (**Figure 2-13B** and **Supplemental Mouse Data for Chapter 2**). This demonstrates NHR2 is critical for CBFA2T3-GLIS2 driven leukemic transformation.

I previously showed that deletion of NHR2 causes a loss of ETO binding to the fusion and reduces the ability of the fusion to dimerize (**Figures 2-7, 2-8**). Therefore, I hypothesize leukemogenesis is abrogated *in vivo* upon deletion of NHR2 because ETO and dimerization of the fusion play critical roles in the dysregulation of the WNT, TGFβ, and Hedgehog signaling pathways.

Due to the lack of expansion *in vivo* when NHR2 is deleted, unmanipulated C-G+ pMks and C-G+ with NHR2 deleted pMks were cultured *in vitro* for eight days and then RNA-sequencing in technical triplicate was performed to investigate gene expression changes upon deletion of NHR2. Global differential gene expression analysis revealed 51% of genes were downregulated and 49% of genes were upregulated in the NHR2Δ pMks compared to the unmanipulated C-G+ pMks. Gene set enrichment analysis (GSEA) using the Hallmark pathways revealed downregulation of IL6-JAK/STAT3, IL2-STAT5 signaling, Notch, Hedgehog, and TGFβ signaling in the NHR2Δ cells. Interestingly, WNT β-catenin signaling was upregulated in the NHR2Δ cells; however, several of the upregulated genes in the pathway are negative regulators of WNT signaling including DKK1, RBPJ, and LEF1 (**Figure 2-14** and **Supplemental Sequencing Data for Chapter 2**). This data suggests NHR2Δ halts leukemic transformation *in vivo* due to the impaired upregulation of JAK/STAT, Notch, Hedgehog, and TGFβ signaling pathways and negative regulation of the WNT signaling pathway, which may be mediated through the loss of ETO and reduction of fusion dimerization. These pathways may serve as new therapeutic targets, although systematic knock-out of key pathway components is needed to prove which are essential for transformation.

When focusing in on genes co-occupied by the fusion and ETO, there is a similar pattern with 60% of fusion and ETO co-occupied genes upregulated in NHR2Δ pMks and 40% downregulated in unmanipulated C-G+ pMks (**Figure 2-15** and **Supplemental Sequencing Data for Chapter 2**). In agreement with the global pathway analysis (**Figure 2-14**), signaling by NOTCH4 and Hedgehog ligand biogenesis were downregulated in NHR2Δ pMks (**Figure 2-15**). However, in contrast, signaling by WNT is downregulated in this sub-set of genes and includes decreased expression of several histone modifying genes including LEO1, RUVBL1, and XPO1 (**Figure 2-15**), suggesting ETO may play a role in the upregulation of this pathway.

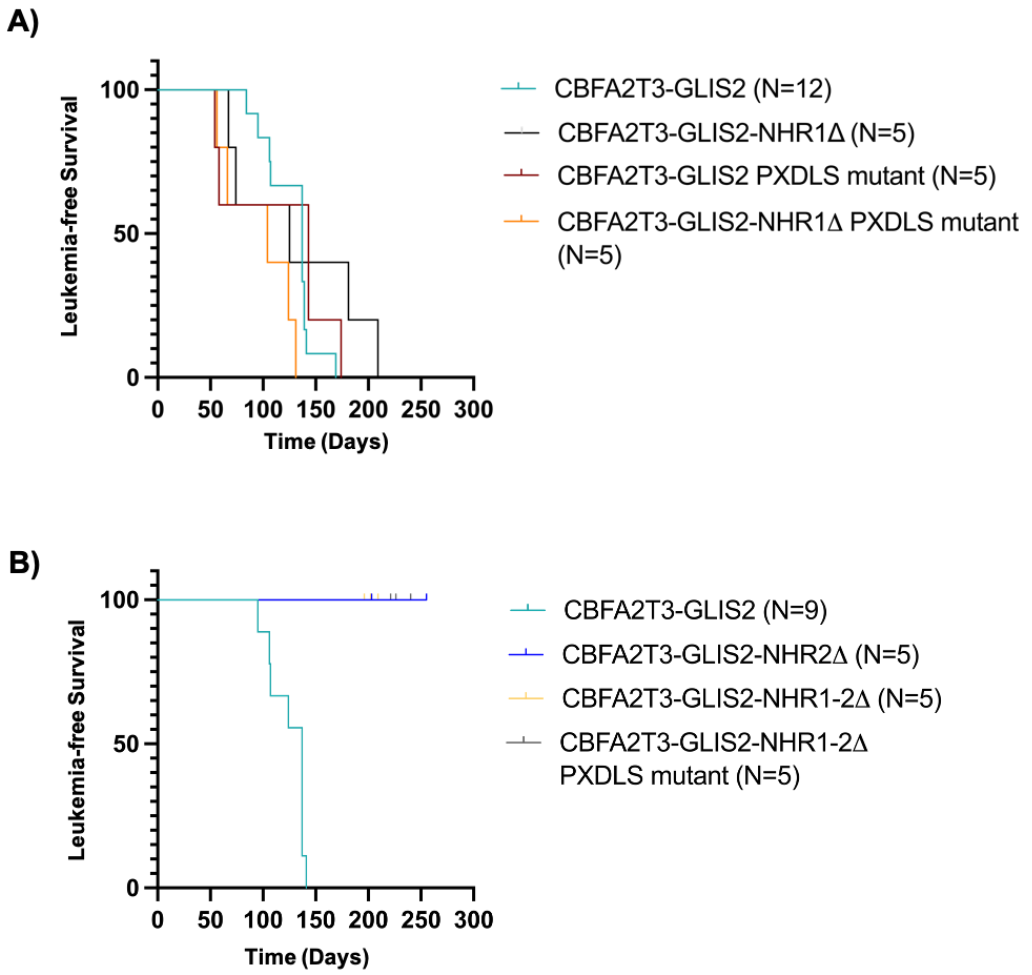
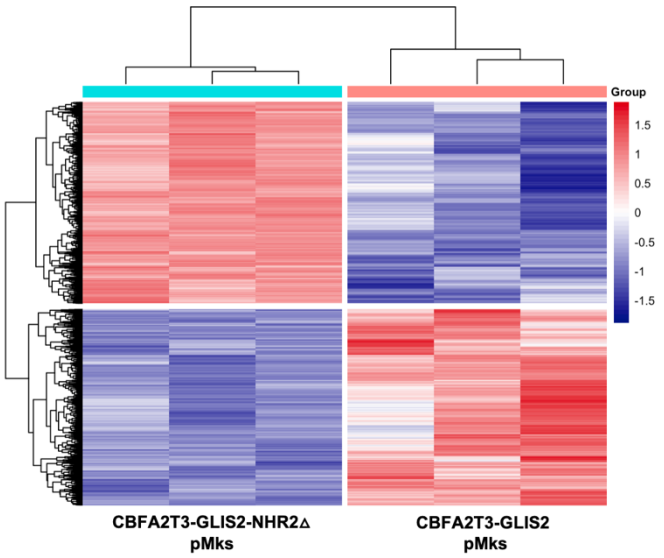


Figure 2-13. NHR2 Abrogates Leukemogenesis *in vivo*.
Primary human megakaryoblasts were transduced with a lentiviral vector encoding the *CBFA2T3-GLIS2* unmanipulated cDNA sequence or a mutant version. **A)** NHR1Δ and the PXDLS mutation do not affect leukemic potential of transduced pMks. **B)** Deletion of NHR2 abrogates leukemogenesis. Leukemia-free survival is shown.

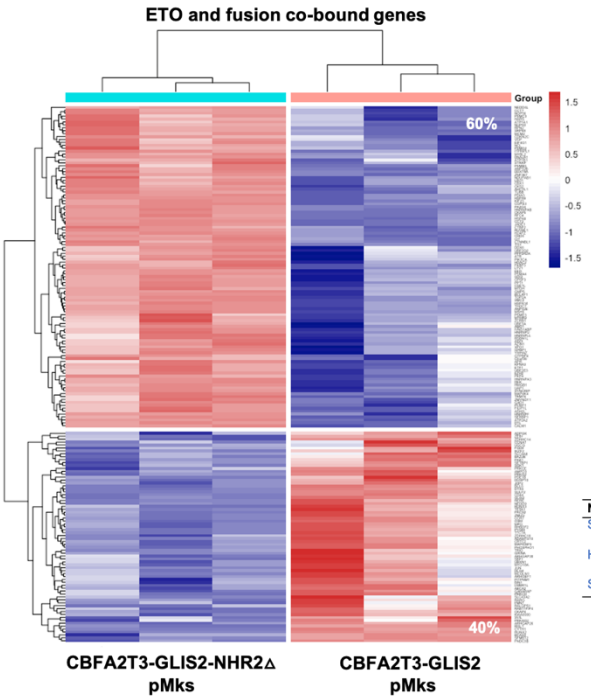


GSEA (Hallmark) CG-NHR2Δ vs CG

Pathway	Normalized Enrichment Score	FDR q-value
IL6 JAK STAT3 Signaling	-2.09	0.00
IL2 STAT5 Signaling	-1.7	0.00
Notch Signaling	-1.5	0.02
Hedgehog Signaling	-1.4	0.04
TGFβ Signaling	-0.9	0.8
WNT β-catenin	1.04	0.4

Figure 2-14. Deletion of NHR2 Alters Gene Expression of Oncogenic Pathways on a Global Scale.

Heatmap of global differentially expressed genes in CBFA2T3-GLIS2-NHR2Δ pMks (n = 3) compared to unmanipulated CBFA2T3-GLIS2 pMks (n = 3) cultured *in vitro* for eight days. A total of 3,010 genes are displayed with a fold change > 1 and an adjusted p-value ≤ 0.01. Pathway analysis was performed using Gene Set Enrichment Analysis and the Hallmark gene sets.



NHR2Δ Downregulated Pathways	Gene Ratio	Background Ratio	Adj p-value
Signaling by NOTCH4	6/82	82/10,867	0.0007
Hedgehog ligand biogenesis	6/82	65/10,867	0.0003
Signaling by WNT	2/82	332/10,867	0.0003

Figure 2-15. Deletion of NHR2 Alters Gene Expression of Oncogenic Pathways Bound by ETO and CBFA2T3-GLIS2.

Heatmap of ETO and fusion co-bound expressed genes in CBFA2T3-GLIS2-NHR2Δ pMks (n = 3) compared to unmanipulated CBFA2T3-GLIS2 pMks (n = 3) cultured *in vitro* for eight days. A total of 182 genes co-occupied by CBFA2T3-GLIS2 and ETO are displayed with a fold change > 1 and an adjusted p-value $\leq$ 0.01. Pathway analysis was performed using over-representation analysis (ORA) and the Reactome Pathway Database.

One downregulated gene of particular interest in NHR2Δ pMks is DTX4, a gene in the Notch signaling pathway. The Notch signaling pathway plays a role in cell development and fate decisions. Dysregulation of Notch signaling has been observed in many cancers including leukemia. DTX4 is one of five deltex (Dx)-related proteins that act as ligand independent activators of Notch signaling. All members of the DTX family have highly conserved C-terminal E3 ubiquitin ligase domains (Chastagner et al., 2017). DTX4 ubiquitinates Notch1 at the cell surface which promotes internalization of Notch1. Notch1 is then cleaved by a disintegrin and metalloproteinase 10 (ADAM10) which releases Notch-IC, the intracellular domain of Notch receptor. Notch-IC translocates to the nucleus and binds to the co-repressor CSL (CBF1, Su(H), Lag-1) and induces its conversion to a transcriptional activator of Notch target genes. DTX4 drives the development and metastasis of several cancers including colorectal and melanoma (Wang, L et al., 2021).

Although not widely described in AML, Huang et al. reported DTX4 may serve as a biomarker associated with poor outcome in AML cohorts, from The Cancer Genome Atlas and the TARGET AML cohort from UCSC Xena database, as part of a larger study to define a set of tumor microenvironment-related differentially expressed genes (Huang et al., 2019). Although the Notch pathway has not been previously implicated in CBFA2T3-GLIS2 mediated leukemia, I propose a model whereby ETO and the fusion contribute to upregulation of DTX4 and, in turn, upregulation of Notch signaling, which is disrupted when NHR2 is deleted and ETO is no longer associating with the fusion transcriptional complex and dimerization of the fusion is reduced. The requirement of Notch signaling in leukemogenesis, however, remains unknown and will require additional experiments to clarify.

Conclusion

I have developed a novel model of CBFA2T3-GLIS2 positive non-DS-AMKL that faithfully recapitulates CBFA2T3-GLIS2 positive patients on a transcriptional and translational level and can be easily manipulated for further studies towards defining the fusion transcriptional complex and testing new therapeutic treatments. The model shows that the *CBFA2T3-GLIS2* fusion can act as an oncogenic driver in primary human megakaryoblasts. The highly conserved GLIS2 zinc fingers bind sites throughout the genome, preferentially at GLIS1-4 consensus sequences, and gene expression is altered in both directions through the homodimerization of the fusion and recruitment of the transcriptional co-factors ETO, CtBP1, and p300. The fusion upregulates WNT, TGFβ, and Hedgehog signaling; key regulatory pathways in cell proliferation and development (**Figure 2-16A**).

Deletion of NHR1 and/or mutation of the PXDLS motif does not abrogate association of ETO, CtBP1, and p300 with the fusion and did not have an effect on leukemic potential *in vivo*. However, deletion of NHR2 does abrogate ETO heterodimerization with the fusion, reduces fusion homodimerization, and blocks leukemic transformation *in vivo*. Fusion positive primary megakaryoblasts harboring the

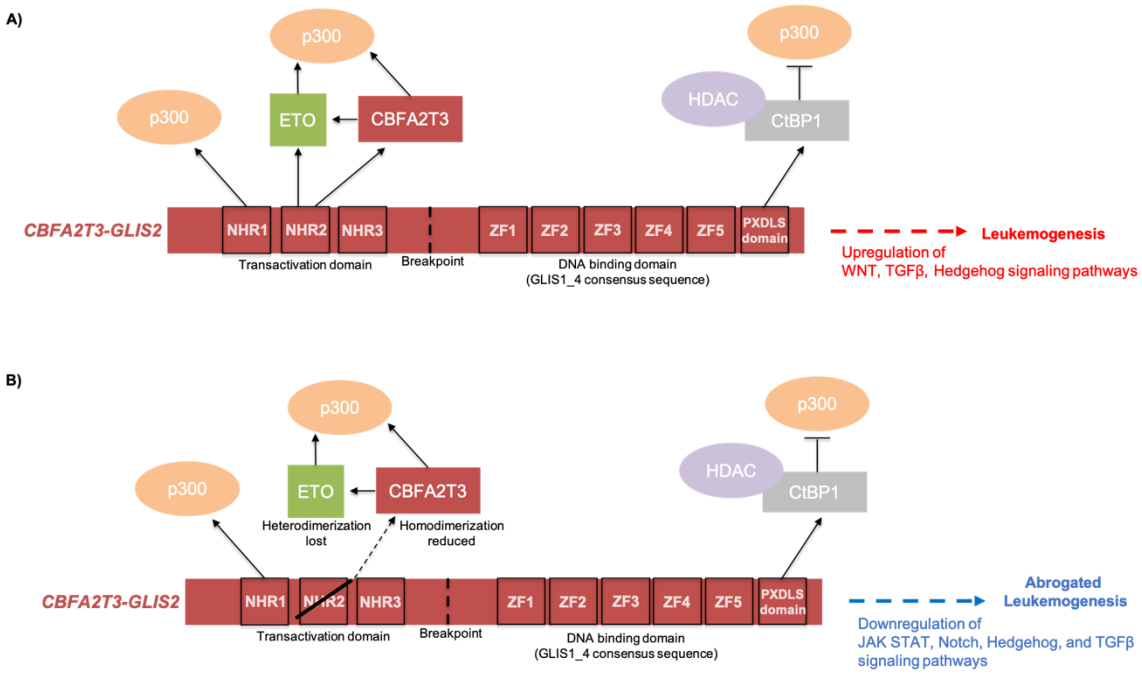


Figure 2-16. Schematic of CBFA2T3-GLIS2 Gene Dysregulation.
Protein structure of the fusion is shown in red and is not drawn to scale. Breakpoint indicates junction where the N-terminal portion of CBFA2T3 is joined to the C-terminal portion of GLIS2. NHR, nervy homology region; ZF, zinc finger. **A)** Upregulated signaling pathways and leukemogenesis occurs when NHR2 is intact and facilitates ETO and fusion hetero- and homo-dimerization, respectively. **B)** Downregulated signaling pathways and abrogated leukemogenesis occurs when NHR2 is deleted and ETO is unable to bind and there is a reduction in fusion homodimerization.

NHR2Δ have downregulation of JAK/STAT, Notch, Hedgehog, and TGFβ signaling at the transcriptional level compared to unmanipulated cells (**Figure 2-16B**). This demonstrates the importance of NHR2 in the fusion transcriptional complex and provides several candidate signaling pathways to therapeutically target in future studies.

Discussion

I have shown that ETO and CtBP1 are associated with upregulation of distinct oncogenic pathways; WNT/TGFβ/Notch and RUNX1/2/3, respectively (**Figure 2-12**). The ETO upregulated pathways are of great interest because deletion of NHR2 alone, which completely disrupts ETO binding to the fusion, abrogates leukemogenesis *in vivo*. RUNX1, 2, and 3 upregulation associated with CtBP1 and fusion co-bound genomic sites likely contribute to, but do not appear to be required for, malignant transformation as they remain unaltered when NHR2 is deleted.

Within the scope of these studies, I did not investigate the role of p300 and H3K27ac in the CBAF2T3-GLIS2 transcriptional complex. A recent study by Benbarche et al. described the results of a CRISPR interference (CRISPRi) screen of super enhancers (SE), clusters of enhancers with high levels of transcription factor binding (Pott et al., 2015), in CBFA2T3-GLIS2 positive AMKL. Their screen revealed two SEs required for AMKL cell proliferation; *KIT* and *PDGFRA*. KIT and PDGFRA are tyrosine-kinase receptors, therefore treatment with tyrosine kinase inhibitors disrupted growth of fusion positive PDX cells *in vivo* (Benbarche et al., 2022). I predict p300 association with the fusion complex is recruiting H3K27ac to increase gene expression conducive to leukemic transformation, in part through the regulation of SE sites.

CHAPTER 3. THE IMMUNOGENICITY OF CBFA2T3-GLIS2 IN NON-DS-AMKL

Introduction: A Novel PBMC-Humanized PDX Murine Model

There are currently no immunotherapy options for non-DS-AMKL patients. Models to study the interaction between AMKL cells and the immune system are limited, but essential for the development of immunotherapies for these patients. To study the *CBFA2T3-GLIS2* fusion oncogene, the Gruber lab has established CBFA2T3-GLIS2 positive AMKL patient-derived xenograft (PDX) murine models from four patients (**Figure 3-1A**). Each of these has the potential to be engrafted with human peripheral blood mononuclear cells (PBMCs) to investigate the immunogenicity of this leukemia sub-type. M7007 secondary transplant gives rise to leukemia in about 60 days (**Figure 3-1B)**.

Roth et al. reconstituted the immune system of immunodeficient mice engrafted with human prostate cancer cells, from the PC3 line, with human PBMCs and autologous human dendritic cells. Mice with both cancer cells and PBMCs had smaller tumors than those with only the cancer cells, although PBMCs were unable to completely clear tumor. They found their humanized model promoted accumulation of tumor infiltrating lymphocytes (TILs) that recapitulated the phenotype of TILs found in patients with prostate cancer, highlighting the value of their model to recapitulate patient's tumor microenvironment. Roth et al. did not match the haplotype of the PBMCs to that of the tumor cells and, therefore, did not perform any studies on antigen-specific T cell responses (Roth et al., 2015).

Engraftment of the well-established CBFA2T3-GLIS2 positive PDX murine models with PBMCs from the patient would create a valuable, autologous PBMC-humanized PDX model to identify and study leukemia-antigen specific CD8+ T cells that may be used as the basis for the development of a novel, effective TCR-T immunotherapy for patients. Unfortunately, banked remission PBMCs are not available from any of the four AMKL PDX models. To overcome this problem, I hypothesized that I could reconstitute the immune system of the PDX models with 6/6 class I exact matched PBMCs from a healthy, unrelated donor to study the potential expansion of AMKL specific CD8+ T cells. I predict that the CBFA2T3-GLIS2 positive AMKL cells act as antigen-presenting cells to present leukemia-specific antigens (ex: those derived from the fusion junction) and leukemia-associated antigens (ex: peptides derived from dysregulated proteins) on surface HLA class I molecules to infiltrating CD8+ T cells that can induce a functional effector response to clear the leukemia cells (**Figure 3-2).**

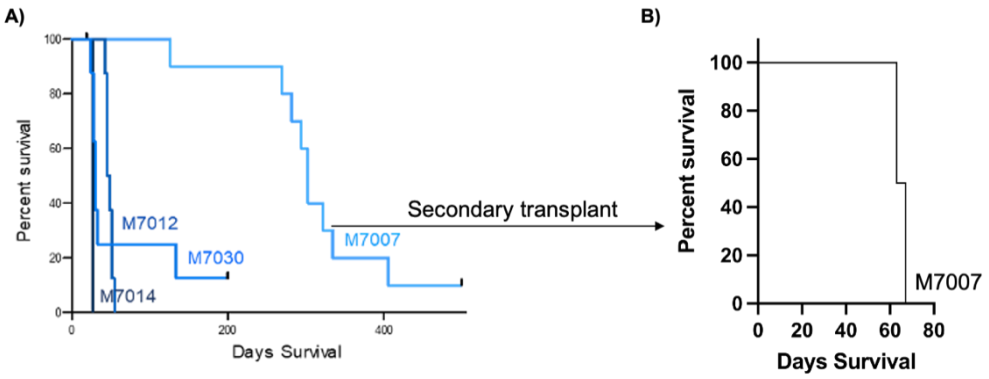


Figure 3-1. CBFA2T3-GLIS2 Positive PDX Murine Models.
A) 1-5 million primary leukemic blasts from four patients were transplanted into NSG-SGM3 immunodeficient mice conditioned with 200 rads on day -1. **B)** 7 million M7007 PDX cells were transplanted (secondary transplant) into NSG-SGM3 immunodeficient mice conditioned with 200 rads on day -1.

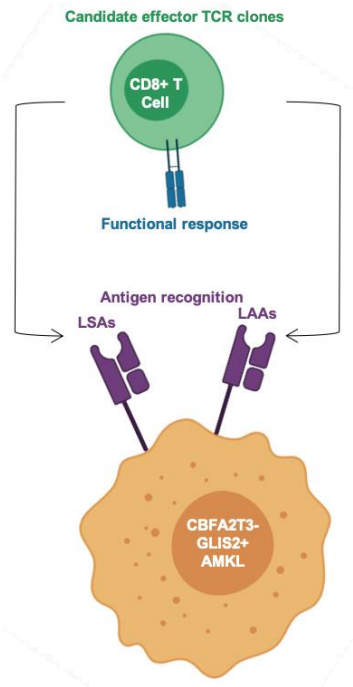


Figure 3-2. Summary of Theorized Effector CD8+ T Cell Recognition and Response to AMKL Antigens.
I hypothesize CBFA2T3-GLIS2 positive AMKL cells (orange) present LSAs (leukemia-specific antigens) and LAAs (leukemia-associated antigens) via HLA class I complexes (purple) that are recognized and targeted by specific CD8+ TCR clones (green).

Materials and Methods

M7007 Patient Cells

M7007 patient cells were obtained from a diagnostic bone marrow sample donated by a 1.67-year-old female at St. Jude Children's Research Hospital. RNA-sequencing, whole exome sequencing, and Cicero CBFA2T3-GLIS2 fusion detection was performed as part of the discovery cohort published by Gruber et al. in Cancer Cell in 2012. HLA typing was performed at St. Jude Children's Research Hospital.

PBMCs

PBMCs from an individual, healthy, unrelated, female donor with a 6/6 HLA class I match to the M7007 patient cells were purchased from AllCells.

M7007 Patient-Derived Xenograft Model

Previously established in the lab, 1 million M7007 patient cells were injected via tail vein into an NSG-SGM3 mouse (see Chapter 2 methods) pre-conditioned with 200 rads 24 hours prior. 322 days post-transplant the mouse was moribund and sacrificed (see Chapter 2 methods). The spleen weighed 1.54 g. Spleen and bone marrow cells were depleted of mCD45 cells (see Chapter 2 methods) and viably frozen in FBS supplemented with 10% DMSO and stored in liquid nitrogen. Secondary transplant via tail vein of 7 million cells thawed cells led to a survival of about 60 days in two biologic replicates. In my preliminary studies, secondary transplant of 3.5 million cells led to high leukemia burden in the spleen and bone marrow around 40 days.

NetMHCcons and NetMHCpan-4.1

The binding affinity between the fusion junction peptide sequence and mass spectrometry identified peptides and the HLA class I alleles of M7007 patient was determined using NetMHCcons (Karosiene et al., 2012) and NetMHC-pan4.1 (Reynisson et al., 2020). Based on the criteria defined in Zamora et al., peptides were classified as significant binders if the IC_{50} was between 5 and 500 nM. Strong binders were classified as having an IC_{50} between 5 and 150 nM and intermediate binders were classified as having an IC_{50} between 150 and 500 nM (Zamora et al., 2019).

Immunopeptidomics

Secondary transplant of 3.5 million M7007 PDX cells into an NSG-SGM3 mouse pre-conditioned with 200 rads 24 hours prior was treated with 2.5 ug of rhIFNγ on day 45

post-transplant for 48 hours prior to collection. Spleen cells were depleted of mCD45 cells. 144 million expanded PDX cells were recovered after depletion. 1 million cells were stained for flow cytometry analysis with mouse CD45, human CD45, human CD34, and pan HLA-A,B,C to assess surface HLA expression. Gating was performed as follows: lymphocytes, singlets, live cells, mCD45 negative, CD45loCD34+, pan-HLA.

The remaining 143 million cells were viably frozen and shipped on dry ice to Monash University in Melbourne, Australia to the lab of Prof. Anthony Purcell. Dr. Nathan Croft, a Senior Lecturer in Dr. Purcell's research group, performed enrichment of pan-HLA class I, using clone w6/32 purified from the supernatant of in-house grown hybridoma cells, complexes from the cell surface followed by peptide elution (Purcell et al., 2019). The peptides were then analyzed via on-line separation using a Dionex UltiMate 3000 RSLCnano system acquisition on a Thermo Q Exactive™ Plus mass spectrometer, operating in data dependent acquisition mode and acquiring the top 20 most intense MS1 peaks per cycle for subsequent MS2 fragmentation. The mass spectrometer used a resolution of 70,000 in MS1 mode and 17,500 in MS2, acquiring across a scan range of 375-1575 m/z and 200-2000 m/z, respectively. Acquired spectra were then analyzed using PEAKS Studio (v10) (Ma et al., 2003) with settings of Orbitrap, HCD fragmentation, no digestion enzyme, precursor tolerance of 10 ppm and product tolerance of 0.02 Da (Jappe et al., 2020).

From the results, data were screened at a local FDR of 1% with confidence values expressed as -10logPvalues. Separate searchers were conducted so that peptide results could be compared against the standard human proteome (Uniprot) as well as the translated RNA-sequencing from the M7007 patient (Gruber et al., 2012). Significant peptides (FDR 1% and $\leq$ 12 amino acids in length) found in the patient RNA-sequencing data were compared against the HLA ligand atlas database, a comprehensive collection of naturally presented HLA allele specific antigens on human tissue, including bone marrow (Marcu et al, 2021), to identify peptides not present in normal human tissues. The binding affinity (IC_{50} of 5 – 500 nM) of peptides absent from the HLA ligand atlas database for the given HLA alleles of M7007 were determined using NetMHCpan-4.1 (Reynisson et al., 2020).

Colony Forming Unit Assay

B6;129S4-*Bmp2*$^{tm1Jfm/}$J mice were purchased from the Jackson Laboratory (016230). These mice have *loxP* sites flanking exon 3 of the *Bmp2* gene. In the presence of Cre recombinase, BMP2 is knocked out. Bone marrow was harvested from 4–6-week-old mice and cultured overnight in bone marrow media (RPMI, 10% FBS, 1% Pen/Strep, 10 ng/ml mouse IL-3, 30 ng/ml mouse IL6, and 50 ng mouse SCF (Tonbo Bioscience). The next day, bone marrow cells were co-transduced via spinfection with MSCV retroviral constructs encoding mCherry-CBFA2T3-GLIS2 and GFP-Cre, mCherry-CBFA2T3-GLIS2 and GFP-empty vector, GFP-Cre alone, or empty vector mCherry or GFP. After 48-hour transduction, mouse bone marrow cells were sorted and 20,000 cells were plated into MethoCult GF M3434 (StemCell Technologies) in technical replicates.

Colonies were counted and re-plated weekly for a total of six weeks. The assay was performed in biological replicate.

PBMC-Humanized M7007 PDX Murine Model

15 NSG-SGM3 mice were pre-conditioned with 200 rads on day -1. On day 0, M7007 only (N=3) and Experimental (N=9) groups were injected via tail vein with 3.5 million M7007 PDX cells. On day 19, retro-orbital bleeds were performed to confirm PDX engraftment by flow cytometry (CD45lo CD34+). On day 20, PBMC only (N=3) and Experimental (N=9) groups were injected with 2.5 million 6/6 HLA class I matched PBMCs. On day 40, all mice were sacrificed and bone marrow and spleen cells were collected and depleted of mouse CD45 cells. 1 million bone marrow cells were used for immunophenotyping and the rest were used for CD8+ T cell enrichment and 10X single cell sequencing. Mouse serum was collected from retro-orbital bleeds for cytometric bead array analysis.

Retro-Orbital Bleeding

Mice were anesthetized in an isoflurane chamber. They were removed from the chamber and 100 ul of blood was immediately collected from the venous sinus using a non-heparinized glass capillary tube. Collected blood was dispensed into a 1.7 ml Eppendorf tube containing 10 ul of 0.5M EDTA to prevent clotting. Pressure was applied to the eye of the mouse with a piece of gauze until the bleeding stopped. The mouse was returned to its cage and monitored for 10 minutes to ensure proper recovery. The mouse was checked again after 24 hours to ensure proper recovery.

Flow Cytometry

Mouse CD45 depleted bone marrow cells were stained for flow cytometry as described in Chapter 2 methods. Two panels were used for staining. Panel 1 was to immunophenotype the PDX cells as follows: leukemia burden (CD45lo CD34+), HLA class I expression (CD45lo, CD34+, HLA-A*11:01, HLA-Bw6), PDL1 expression (CD45lo, CD34+, PDL1+). The following human monoclonal antibodies were used: APC/Cy7-hCD45 (BioLegend 368516, Clone 2D1), APC-hCD34 (Miltenyi 13-113-176, Clone AC136), BV605-hHLA-A*11:01 (OneLambda BIH0084), FITC-hBw6 (OneLambda FH0038), PE-hPDL1 (BioLegend 393608, Clone MIH2). Panel 2 was to immunophenotype the T cells as follows: PBMCs (CD45high, CD34-), CD3 T cells (CD45high, CD34-, CD3+), CD8 T cells (CD45high, CD34-, CD3+, CD4-, CD8+). The following human monoclonal antibodies were used: APC/Cy7-hCD45 (BioLegend 368516, Clone 2D1), APC-hCD34 (Miltenyi 13-113-176, Clone AC136), FITC-hCD3 (BioLegend 300406 Clone UCHT1), PE-Cy7-hCD4 (BioLegend 357410, Clone A161A1), PB-hCD8 (BioLegend 980906, Clone SK1).

Cytometric Bead Array

Human cytokine levels were measured in serum collected from mouse blood using the LEGENDplex Human CD8/NK Panel (13-plex) multiplex bead-based assay from BioLegend (740267). The kit uses fluorescence-encoded beads to simultaneously quantify 13 human proteins by flow cytometry. These include IL-2, IL-4, IL-10, IL-6, IL-17A, TNF-α, sFas, sFasL, IFNγ, granzyme A, granzyme B, perforin, and granulysin. Mouse serum from eligible mice was tested according to the kit's instructions. Samples were analyzed in technical replicate.

10X Single Cell Sequencing

Live CD8+ T cells were isolated from mouse depleted bone marrow cells from the PBMC-humanized PDX mice using first the EasySep Dead Cell Removal (Annexin V) Kit (StemCell Technologies 17899) to isolate live cells followed by the EasySep Human CD8+ T Cell Isolation Kit (StemCell Technologies 17953) to negatively select for CD8+ T cells. Single-cell expression and TCR data were generated using version 1 of the 5' Single Cell Immune Profiling kit (10X Genomics) with corresponding human V(D)J amplification. Single-index libraries were sequenced on the Illumina NovaSeq platform following the configuration recommended by 10X Genomics. Sequencing reads were processed using CellRanger v6.1.1 (10X Genomics) with the corresponding GRCh38-2020-A reference for gene expression data and the GRCh38-alts-ensembl-5.0.0 reference for TCR data. Across libraries, the mean reads obtained per cell were in excess of 120,000, and expression data were aggregated using CellRanger with default parameters to normalize sequencing depth across cells. Downstream analyses were performed in R with Seurat (Hao et al., 2021), and TCR clonotypes were integrated with the expression data using scripts from TCRdist (Dash et al., 2017). Data were filtered to exclude cells above the 98th percentile for number of unique molecular identifiers or genes, with fewer than 300 genes, and with greater than 10% of expression owed to mitochondrial expression. Expression counts were log-normalized, and variable features were identified using the VST method after excluding genes associated with TCR and IG genes. Cell cycle phase was inferred as in (Tirosh et al., 2016), and data were scaled to regress out the effects of the number of RNA molecules, the percentage of expression owed to mitochondrial genes, and cell cycle scores. PCA, UMAP, and clustering were performed using default parameters, and SignacX (Chamberlain et al., 2021) was used to annotate cell subsets.

Results

CBFA2T3-GLIS2 Patient Cells Present Leukemia-Associated Antigens

To screen for potential immunogenic antigens in CBFA2T3-GLIS2 positive non-DS-AMKL, the HLA class I surface presentation of neoantigen(s) spanning the fusion

junction was interrogated using the publicly available algorithms NetMHCcons and NetMHCpan-4.1 (Karosiene et al., 2012; Reynisson et al., 2020). Both algorithms predict peptide binding affinity for a given HLA allele which aids in the prediction of T cell response. NetMHCcons is trained solely on binding affinity data whereas NetMHCpan-4.1 is an integration of binding affinity data and mass spectrometry-eluted ligands.

The M7007 secondary PDX was chosen for subsequent studies due to the short onset of disease (**Figure 3-1B**). Using the HLA haplotype of the M7007 cells (**Table 3-1**) and fusion junction-containing 8-11 amino acid peptide sequences (**Figure 3-3**), HLA binding affinity was predicted using NetMHCcons, which identified two peptides capable of binding HLA class I with intermediate affinity, defined by a predicated binding affinity IC_{50} of 150 – 500 nM, that span the fusion junction (**Table 3-2**). These peptides were confirmed in NetMMHCpan-4.1.

Although algorithms are useful for making predictions about peptide-HLA binding affinity, they are unable to truly determine surface presentation. To validate the putative neoantigens identified using NetMHCcons, and explore the surface immunopeptidome of the M7007 cells, I collaborated with Prof. Anthony Purcell at Monash University in Melbourne, Australia to perform pan-HLA class I elution mass spectrometry analysis on the cells after expansion in immunodeficient mice. M7007 PDX cells do not inherently express high levels of surface HLA class I (**Figure 3-4A**); therefore, the transplanted mice were treated with recombinant human IFNγ (rhIFNγ) for 48 hours which increased the expression to 79% (**Figure 3-4B**). The immunopeptidomics analysis identified 242 peptides (false discovery rate (FDR) 1% and $\leq$ 12 amino acids in length, the typical length to bind HLA class I complexes (Wieczorek et al., 2017). Importantly, 141 of these peptides were not found in the HLA ligand atlas database, a comprehensive collection of naturally presented HLA allele specific antigens on human tissue, including bone marrow (Marcu et al, 2021). This suggests these 141 peptides are 1) leukemia-specific and 2) represent leukemia-associated antigens that are not presented on normal cells, which is essential for minimizing harmful off-target effects of TCR-T immunotherapy. When these peptides were analyzed for binding affinity through NetMHCpan-4.1, 28 of the 141 peptides had a significant binding affinity (IC_{50} of 5 – 500 nM) for one of the HLA alleles present in the PDX cells. (**Figure 3-5** and **Supplemental Immunopeptidomics Data for Chapter 3**). 19 peptides had a strong binding affinity (IC_{50} of 5 – 150 nM) and 9 had an intermediate binding affinity (IC_{50} of 150 – 500 nM). Most of the peptides were predicted to bind HLA-A*02:01 (**Table 3-3**).

One of the top hits from this analysis was a peptide derived from the coding exon of the DDX3X protein, a member of the DEAD-box helicase family (**Table 3-3, in bold**). DDX3X plays a role in cancer progression and has been reported as an immunogenic protein in melanoma (Mo et al., 2021). Although outside the significant binding affinity range (IC_{50} of 5 – 500 nM), a notable peptide was identified that corresponded to the 5'UTR (untranslated region) of BMP2 protein (bone morphogenetic protein 2) (**Table 3-3, in bold**). The derivation from the 5'UTR suggests non-canonical peptide expression of BMP2 in the AMKL cells (Prakash et al, 2021). BMP2 is a key regulator in the

Table 3-1. Haplotype of M7007 PDX.

HLA-A	HLA-B	HLA-C
A*02:01	B*07:02	C*04:01
A*11:01	B*35:01	C*07:02

Figure 3-3. CBFA2T3-GLIS2 Fusion Junction Peptide Sequence.
The CBFA2T3-GLIS2 fusion junction is an attractive neoantigen due to its exclusive expression in the AMKL cells.

Table 3-2. NetMHCcons Fusion Junction Peptide Predictions.

Peptide Sequence	HLA-Allele	IC_{50} (nM)
SSEDFQPLR	A*11:01	304.37
DSSEDFQPLRY	B*35:01	384.09

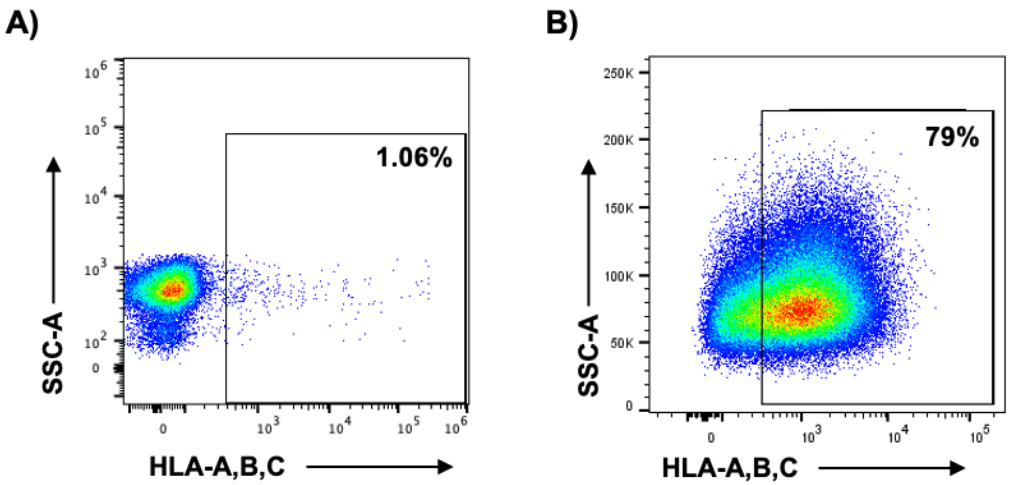


Figure 3-4. rhIFNγ Increases Surface HLA Expression on PDX Cells.
Flow cytometry gating as follows: lymphocytes, singlets, live, hCD34+, pan-HLA-A,B,C. **A)** HLA expression on mouse CD45 depleted spleen cells from a mouse not treated with rhIFNγ was only 1.06%. **B)** HLA expression on mouse CD45 depleted spleen cells from a mouse treated with rhIFNγ for 48 hours prior to cell collection increased to 79%.

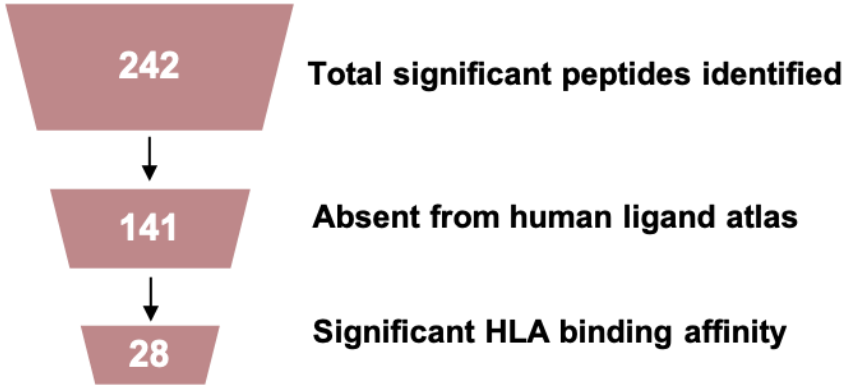


Figure 3-5. Immunopeptidomics Analysis.
242 significant peptides (FDR 1% and ≤ 12 amino acids in length). 141/242 were not present in the human ligand atlas. 28/141 had a significant binding affinity (IC_{50} 5 – 500 nM) as calculated by NetMHCpan-4.1.

Table 3-3. Top Peptides of Interest.

Peptide Sequence	Gene	HLA-Allele	IC_{50}	Annotation
SMMDVDHQI	CCT5	A*02:01	5.27	Strong
ALAAVVTEV	**DDX3X**	A*02:01	5.97	Strong
TPITQGVEM	IGLL1	B*35:01	16.22	Strong
MLQDFQIQA	ATP6AP1	A*02:01	17.90	Strong
SLLLHLPHVV	NUCB1	A*02:01	21.37	Strong
VLAFENPQV	NME9	A*02:01	22.05	Strong
AMEALVVEV	DHX9	A*02:01	31.50	Strong
TMINAIPVA	SUPT16H	A*02:01	39.97	Strong
NMLNPPAEV	XRCC5	A*02:01	48.05	Strong
FLVSVLSKVT	KPNA2	A*02:01	51.67	Strong
NLDNPIQTV	NOMO2	A*02:01	69.64	Strong
TVVSKKAL	LARP1	A*02:01	81.02	Strong
ALSSQDVMV	LDB1	A*02:01	84.20	Strong
VWAGLVPSI	ARID1B	A*02:01	85.75	Strong
ALLDQALSNA	MACF1	A*02:01	90.16	Strong
SLISNPPAM	NCKAP1L	A*02:01	103.75	Strong
GLDNQIQEI	PSMC1P1	A*02:01	123.17	Strong
KAISGVHTV	EIF2S3	A*02:01	126.19	Strong
FLFVSFIV	CDK16	A*02:01	148.78	Strong
SLQEEQQQL	RUFY1	A*02:01	155.36	Intermediate
AVLFLSVLLL	HUWE1	A*02:01	162.90	Intermediate
TIDPSNPMV	CTR9	A*02:01	163.10	Intermediate
VQDSLNPAV	ITGA2B	A*02:01	224.02	Intermediate
TMSSSIKTV	C5orf66	A*02:01	227.88	Intermediate
AVMAQVQAI	SCAF4	A*02:01	246.12	Intermediate
GLLGQPEATMV	SON	A*02:01	309.48	Intermediate
NVVCVVYDV	RHOT2	A*02:01	398.32	Intermediate
TLFLLVIMLI	OGT	A*02:01	480.15	Intermediate
AVPFPGPGV	**5'UTR BMP2**	A*02:01	1724.49	N.S.

Strong = IC_{50} (NetMHCpan-4.1) between 5 and 150 nM

Intermediate = IC_{50} (NetMHCpan-4.1) between 150 and 500 nM

N.S. (not significant) = IC_{50} (NetMHCpan-4.1) greater than 500 nM

development of bone and cartilage and, although its role in leukemia is not well-defined, aberrant upregulation of *BMP2* has been reported in CBFA2T3-GLIS2 positive AMKL (Gruber et al., 2012).

Further, the Gruber lab has preliminary data suggesting BMP2 is required for the self-renewal phenotype imparted by the fusion oncogene. Bone marrow from mice containing *loxP* sites flanking exon 3 of the *Bmp2* gene were transduced with retroviral constructs encoding the CBFA2T3-GLIS2 fusion and Cre recombinase or the fusion and empty vector, sorted for purity, and plated into methylcellulose containing mouse cytokines to promote bone marrow cell growth. Self-renewal potential was assessed by counting colony formation upon weekly re-plating. Cells transduced with the fusion alone re-plate out to week five. In contrast, cells co-transduced with the fusion and Cre recombinase, and thereby lack BMP2 expression, were unable to serially re-plate (**Figure 3-6**). BMP2 peptide is of great interest as a candidate leukemia-associated antigen because it is uniquely upregulated in this leukemia sub-type and may be required for leukemogenesis.

Interestingly, peptides derived from the CBFA2T3-GLIS2 fusion junction or full-length partners were not detected by the mass spectrometry analysis. Although the fusion is the driver of the leukemia, as proven in Chapter 2, and is detectable by RNA-sequencing (Gruber et al., 2012), the cell surface level expression of the junctional peptides may be below the level of detection of the methods used in this study.

PBMC-Humanization Variably Reduces Leukemic Burden in a PDX Murine Model

To further evaluate the immunogenicity of CBFA2T3-GLIS2 positive AMKL and explore whether the PDX cells can induce a functional effector CD8+ T cell response, potentially directed against one of the antigens discovered by the immunopeptidomics analysis, I established a PBMC-humanized version of the M7007 PDX murine model. Immunodeficient mice were engrafted with M7007 CBFA2T3-GLIS2 positive PDX cells for twenty days prior to the transplantation of exact HLA class I matched (**Table 3-1**) PBMCs from an unrelated, healthy donor for an additional twenty days. The goal was to assess if the leukemia cells could act as antigen-presenting cells and stimulate a leukemia-directed effector CD8+ T cell response.

Experimental mice (engrafted with both M7007 cells and PBMCs), M7007 only control mice, and PBMC only control mice were sacrificed at day 40 (**Figure 3-7** and **Supplemental Mouse Data for Chapter 3**). Mouse CD45 depleted cells from the bone marrow of each mouse was stained for flow cytometry to assess leukemia burden, surface HLA expression on the leukemia cells, and engraftment of CD8+ T cells (**Figure 3-8**). Flow cytometry analysis at day 40 showed a strong reduction of leukemia blasts in 3/9 experimental mice (mean = 30.1%, range = 1.7-56%) when compared to the M7007 only control mice (mean = 97.4%, range = 92.7 – 99.8%, p = 0.048) (**Figure 3-9A**, responders).

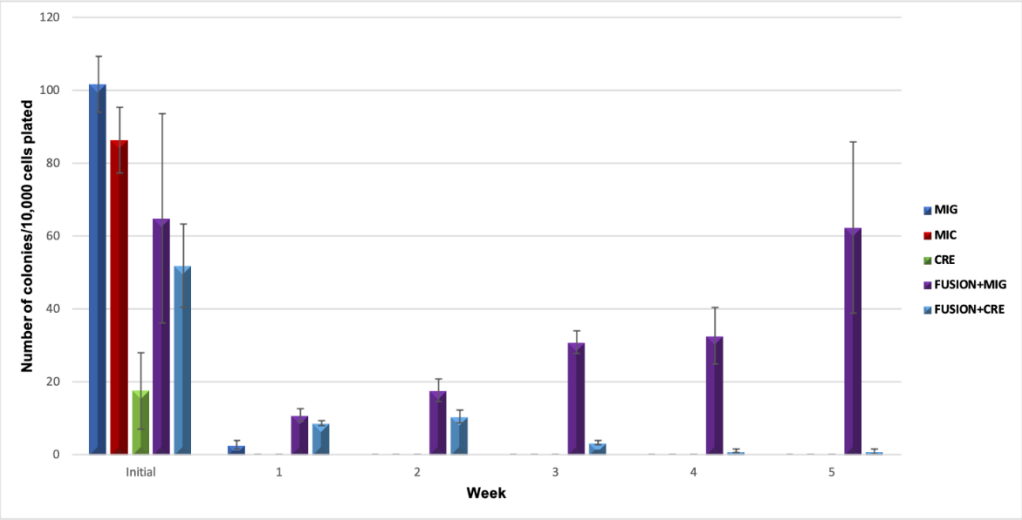


Figure 3-6. BMP2 Is Required for CBFA2T3-GLIS2 Self-Renewal.
Colony forming unit (CFU) assay over the course of 5 weeks. Bone marrow from 4–6-week-old B6;129S4-*Bmp2*$^{tm1Jfm/}$J mice were transduced with retroviral constructs and plated in methylcellulose media. Data is from two independent experiments and standard deviation is calculated. MIG, empty vector GFP; MIC, empty vector mCherry; Cre, vector encoding Cre recombinase; Fusion+MIG, co-transduction of vector encoding CBFA2T3-GLIS2 and empty vector GFP; Fusion+Cre, co-transduction of vector encoding CBFA2T3-GLIS2 and vector encoding Cre recombinase.

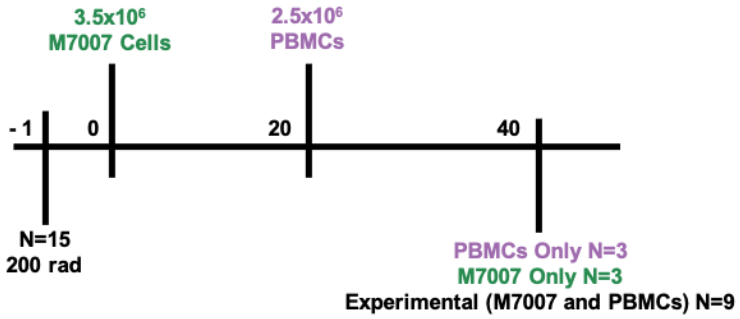


Figure 3-7. Establishment of a Novel PBMC-Humanized AMKL PDX Model.
Immunodeficient mice were pre-conditioned with 200 rads twenty-four hours prior to engraftment with M7007 PDX cells. After twenty days, HLA matched PBMCs from a healthy, unrelated donor were introduced. Mice were collected after an additional twenty days for downstream analyses.

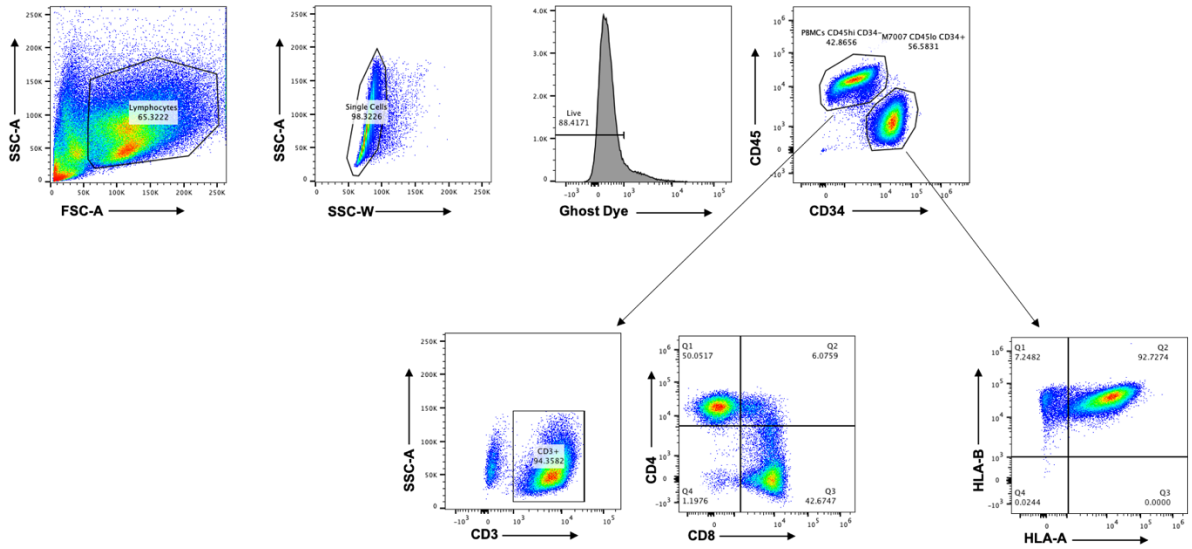


Figure 3-8. Flow Cytometry Gating Strategy.
Cells from the bone marrow of mice collected at day 40 were depleted of mCD45 cells and then stained for flow cytometry analysis with the following human antibodies: CD45, CD34, CD3, CD8, CD4, HLA-A, HLA-B.

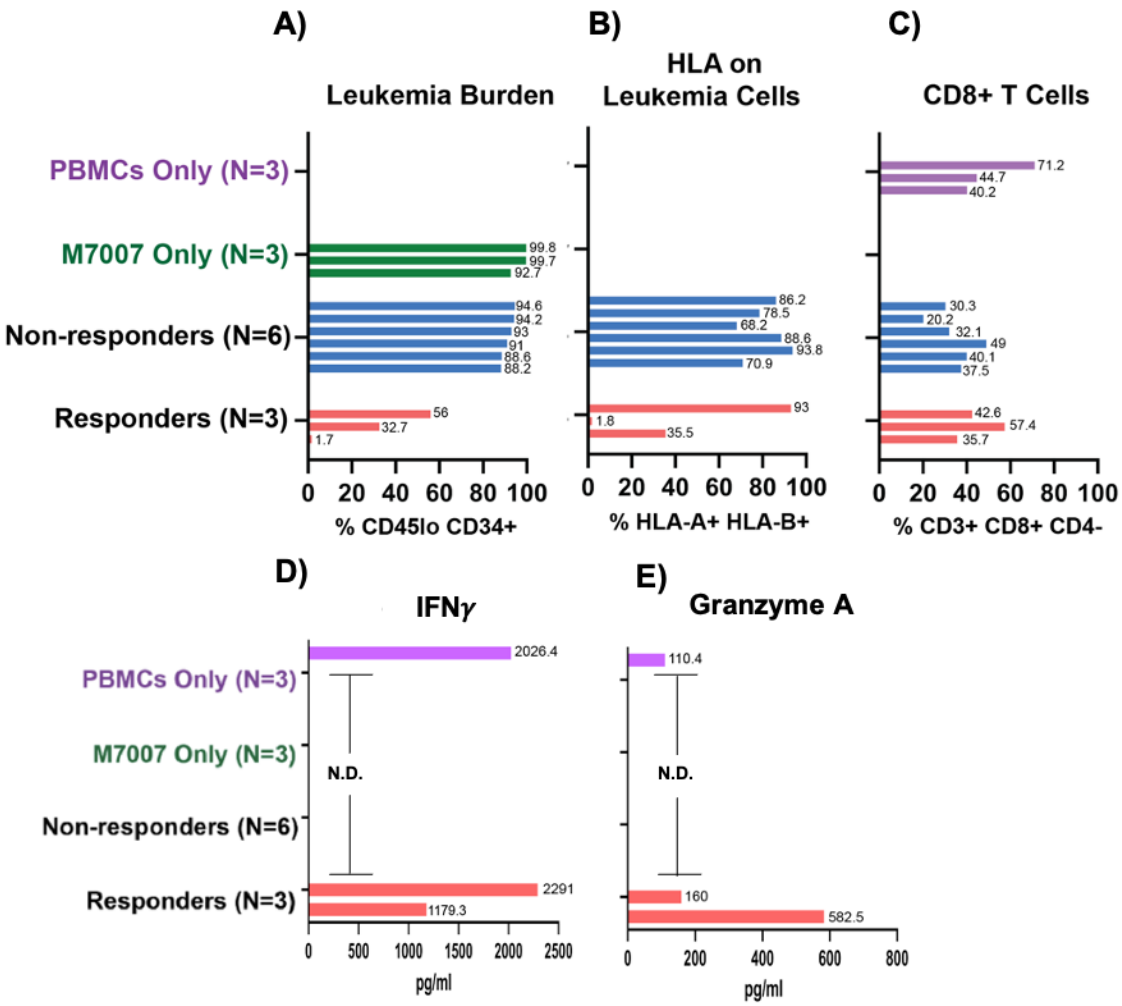


Figure 3-9. Characterization of the PBMC-Humanized PDX Model.
A – C) Flow cytometry immunophenotyping of mouse CD45 depleted cells from the bone marrow. Each bar represents one mouse. **A)** M7007 leukemia burden as defined by %CD45lo CD34+. **B)** Surface HLA-A and HLA-B double positive expression on M7007 PDX cells. **C)** CD8+ T cells as defined by CD3+ CD8+ CD4-. **D – E)** Cytometric bead array for a human CD8/NK cytokine panel was performed in serum collected from the mice. N.D.; no data available for those samples. **D)** IFNγ and **E)** Granzyme A measured in pg/ml.

In contrast, 6/9 experimental mice had a leukemia burden similar to that of the M7007 only control mice (mean = 91.6%, range = 88.2 – 94.6%, p = 0.109) (**Figure 3-9A**, non-responders).

As noted previously, M7007 cells do not express high levels of surface HLA class I. However, there was an increase in the expression of surface HLA on the M7007 cells when exposed to the PBMCs (**Figure 3-9B**). This is likely due to the IFNγ production by the CD8+ T cells, which is consistent with the findings that rhIFNγ greatly increased pan-HLA surface expression on the PDX cells prior to collection for mass spectrometry analysis (**Figure 3-4**). Notably, CD8+ T cell engraftment was similar between all groups (**Figure 3-9C**). Consistent with this equivalent engraftment, cytometric bead array for a panel of human CD8/NK cell associated cytokines in serum collected from the PBMC-humanized PDX model revealed similar levels of IFNγ between a mouse from the PBMC only group compared to the two greatest responders (no data is available for the M7007 only and non-responders groups) (**Figure 3-9D**). Strikingly, the responder mouse that showed the greatest M7007 PDX cell reduction also had high levels of granzyme A, a protease in cytotoxic granules released by cytotoxic killer cells (Lieberman, 2010), compared to the PBMC only group, further supporting the response is mediated by leukemia-directed effector T cells (**Figure 3-9E**). Collectively, the characterization of this novel murine model suggests the variability of response seen in this cohort is due to differences in the expansion of specific leukemia-directed CD8+ T cell clones between mice.

CBFA2T3-GLIS2 Positive PDX Cells Induce a Clonal Expansion of CD8+ T Cells

To characterize the TCR repertoire in the model, single cell 5’ V(D)J TCR sequencing with paired gene expression analysis was performed on positively enriched CD8+ T cells from the bone marrow. The gene expression pattern of the non-responders was distinct from the responders group. Interestingly, the responders group pattern is similar to that of the PBMC only group (**Figure 3-10**). There was a large reduction in TCRβ chain diversity in the non-responder and responder groups, with the greatest reduction of diversity observed in the responders when compared to the PBMC only and non-responders (**Figure 3-11** and **Supplemental Sequencing Data for Chapter 3**). This finding is consistent with a clonal expansion of potentially leukemia antigen-directed TCRs. 122 unique TCR clonotypes were identified in the responder’s group and compared against the PubMed TCRdatabase to determine if any were previously described against known antigens. This analysis revealed two prior descriptions against 1) Dengue and 2) Yellow fever (**Supplemental Sequencing Data for Chapter 3**); thus leaving 120 unique clonotypes not previously described in the literature. Access to information regarding the medical history of the PBMC donor is unavailable; therefore, it is possible the donor was vaccinated against these two viruses or these particular TCR clonotypes may be easily generated and similar to those associated with responses to those viruses notated in the database. The top nine clones were ranked based on expanded

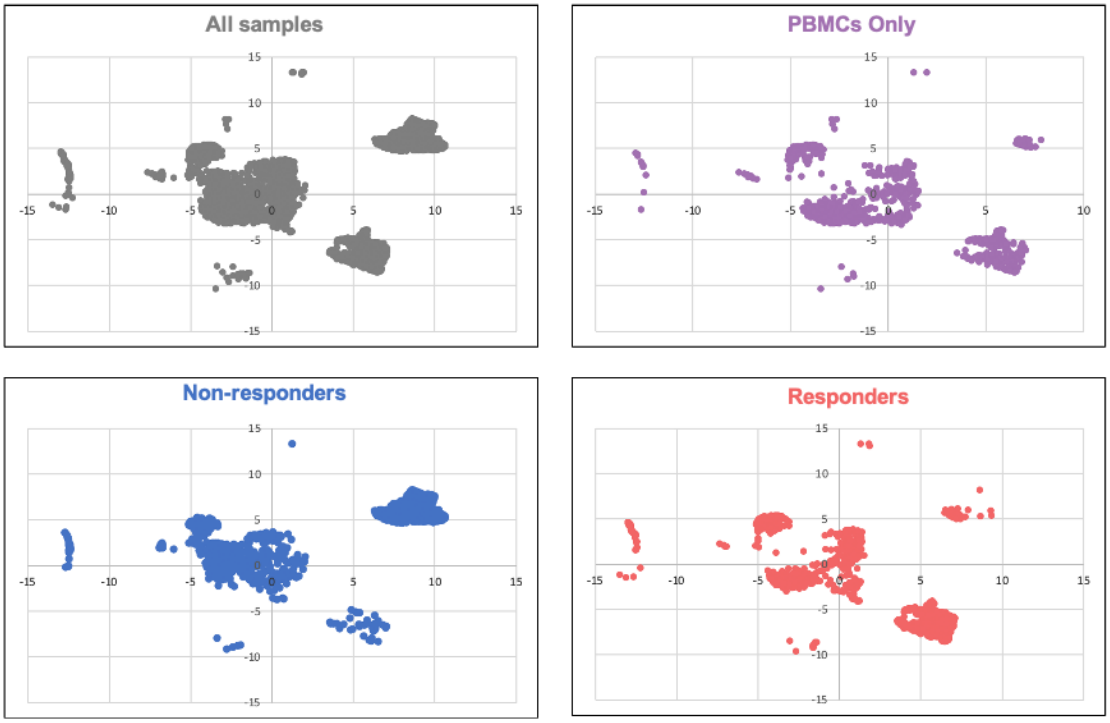


Figure 3-10. CD8+ T Cell Single Cell Gene Expression Analysis.
UMAP plots displaying single cell gene expression analysis of positively selected CD8+ T cells. Overlay of all cells is shown in gray. PBMCs only is in purple, non-responders only is in blue, and responders only is in red.

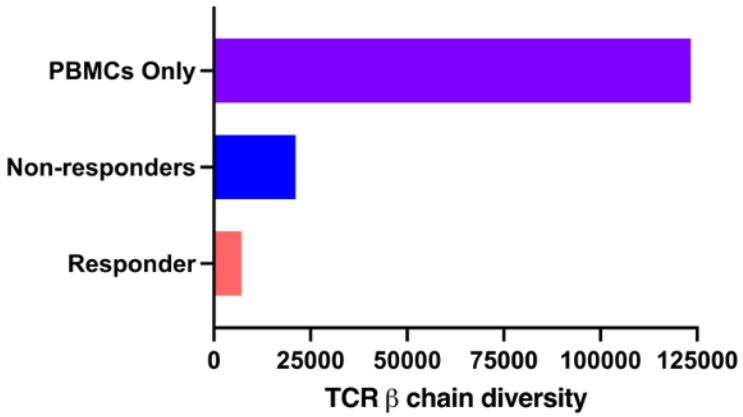


Figure 3-11. CD8+ T Cell Diversity.
TCRβ chain clonal diversity was analyzed from single cell TCR-sequencing. The CD8+ T cell repertoire was higher in the PBMCs only mice (purple) compared to those exposed to M7007 PDX cells with the responders group showing less diversity (red) than the non-responders group (blue), consistent with a leukemia directed clonal expansion.

clone size, as defined by the β chain complementarity-determining region 3 (CDR3), and are shown in **Table 3-4**.

I have shown, for the first time, successful engraftment of HLA class I matched PBMCs in an established CBFA2T3-GLIS2 positive AMKL PDX murine model and provided evidence of a leukemia-specific effector T cell response mediated by clonal expansion and granzyme A release (**Figure 3-12** and **Supplemental Sequencing Data for Chapter 3)**. This innovative model has the potential to identify nascent T cell clones directed at AMKL leukemia antigens. Studies to screen the top expanded clones from the responders group (**Table 3-4** and **Supplemental Sequencing Data for Chapter 3**) for recognition and response against the candidate fusion junction leukemia-specific (**Table 3-2**) and leukemia-associated antigens (**Table 3-3**) discovered by NetMHCcons and immunopeptidomics, respectively, are in progress.

Conclusion

These studies present the first evidence that CBFA2T3-GLIS2 positive non-DS-AMKL may be immunogenic through identification of candidate, targetable antigens and a novel murine model to aid in identification of responsive CD8+ T cells that have the potential to be developed into a TCR-T immunotherapy. This research sheds light on the immunogenicity of CBFA2T3-GLIS2 mediated AMKL and the potential for TCR-T immunotherapy in the setting of minimal residual disease for this chemotherapy resistant malignancy.

Exploration of the surfaceome of a patient sample revealed two leukemia-associated antigens of particular interest because of their biological relevance, DDX3X and BMP2, among 141 candidate peptides not reported in the human ligand atlas. Although peptides from the CBFA2T3-GLIS2 fusion junction or its full-length partners were not found by immunopeptidomics, algorithmic predictions identified two peptides spanning the fusion junction with significant binding affinity. A novel, exact HLA class I matched PBMC-humanized CBFA2T3-GLIS2 positive PDX murine model showed a variable reduction in leukemia burden. The responders had a reduction in CD8+ T cell diversity in conjunction with an increase in granzyme A production compared to a PBMC only control. This is consistent with a clonal expansion of effector CD8+ T cells that may be directed against leukemia-specific and/or -associated antigens. Studies are ongoing to identify the TCR clone(s) responsible for the anti-leukemic response seen in the model and the leukemia antigens that are mediating this response.

Discussion

Although small in scale and still preliminary, the data presented here set the precedent for subsequent immunotherapy studies for pediatric non-DS-AMKL. It was

Table 3-4. Top Clonally Expanded CD8+ T Cells from the Responders Group.

Clone Size	CDR3β Amino Acid Seq	TCRdatabase
52	CASSLRPDSYNEQFF	Not present
41	CASSPRSGAWEQYF	Not present
23	CASTDFRTGELFF	Not present
18	CASTSDRANYEQYF	Not present
16	CASSLDRAQETQYF	Not present
14	CASSMLDEQFF	Not present
13	CASSFFPLGEQFF	Not present
10	CASSLPMGFQPQHF	Not present
10	CASSPGPNNEQFF	Not present

CDR3β, TCRβ chain complementarity-determining region 3
TCRdatabase, NCBI

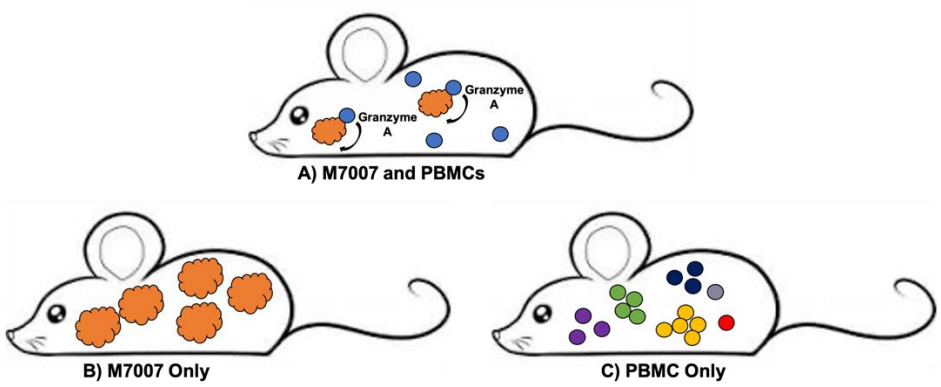


Figure 3-12. Summary of Response Seen in Responder Group.
A) At day 40, mice in the responder group showed a reduced leukemia burden (orange symbol) when exposed to PBMCs compared to **B)** M7007 only control group. This response is mediated by **A)** the release of granzyme A from a subset of clonally expanded, potentially leukemia-directed CD8+ T cells (blue circles) that are not present in the highly diverse TCR repertoire of the **C)** PBMC only control group (colored circles represent CD8+ T cells with different TCRs).

encouraging to find potential leukemia-associated antigens presented by surface HLA class I on the M7007 CBFA2T3-GLIS2 positive PDX cells by mass spectrometry; however, it was discouraging that we did not find peptides corresponding to any portion of the fusion junction as the fusion junction is the most attractive target due to its restriction to the AMKL cells. Although the fusion is the driver of the leukemia and is detectable by RNA-seq, the cell surface level expression of the junctional peptides may be below the threshold of sensitivity of detection. It is possible that targeted mass spectrometry such as multiple reaction monitoring (MRM-MS), which has a limit of detection down to one copy of peptide per cell will detect fusion junction peptides (Tan et al., 2011). The sensitivity of the pan analysis is estimated to be an order of magnitude lower than the targeted analysis. Although not yet widely used to probe the surface of patient leukemia cells, MRM-MS has been used to detect low abundance of biomarkers in kidney disease, which has assisted in disease staging criteria (MacMullan et al., 2019). Similarly, MRM-MS has been used to detect varying levels of HER2 protein in breast cancer samples (Zhang et al., 2019). We did not perform this analysis because we initially wanted to probe the entire surface peptidome of the M7007 AMKL cells without incorporating prior knowledge of any expected peptides. Now that we have performed pan immunopeptidomics, we plan to perform targeted mass spectrometry in more expansive studies.

The low surface HLA class I expression on the M7007 PDX cells in the absence of IFNγ (**Figure 3-4**) is a major challenge that will need to be addressed in pre-clinical studies. Unlike CAR-T cells that do not require HLA class I presentation of antigen, TCR-T immunotherapy is HLA-dependent. Therefore, high and persistent HLA class I presentation is essential for successful treatment. We have shown that rhIFNγ greatly increases expression *in vivo* with no observable side effects in the immunodeficient mice (**Figures 3-4, 3-9**); IFNγ administration has been given in the setting of pediatric chronic myeloid leukemia (CML) (Kurzrock et al., 1987) and high-grade glioma (Wolff et al., 2006), demonstrating feasibility. An ongoing early phase 1 clinical trial (NCT04628338) is assessing safe dose levels of IFNγ in patients with relapsed AML after allogeneic stem cell transplant to promote graft versus leukemia in patients mediated by alloreactive T cells from the donor graft which will provide further data on tolerability.

Ideally, I would have used PBMCs from patient M7007 in the PBMC-humanized PDX model, but these were unavailable. Instead, commercially available 6/6 HLA class I matched PBMCs from a healthy, unrelated donor were used. NSG mice can support human T cell development, in contrast to other immunodeficient strains (Najima et al., 2016). Najima et al. transplanted NSG mice with human cord blood containing antigen-presenting cells (APCs) and T cells. The mice were then immunized with Wilms Tumor 1 (WT1) peptides which resulted in high frequencies of WT1-specific CD8+ T cells within 7 days. This supports the model in which PDX cells can serve as APCs or, alternatively, mononuclear APCs from the HLA matched PBMCs can take up antigen from PDX cells undergoing apoptosis for presentation to T cells. Therefore, even though the PBMCs are from an unrelated donor, the model provides the proper microenvironment and time to stimulate an expansion of potential PDX-specific CD8+ T cells. Natural killer (NK) cells compose about 5-20% of PBMCs. Although the model is supportive of a CD8+ T cell

response due to the decrease in diversity in the responders group, NK cells may also be playing a role in antigen independent mediated killing. Depletion of CD8+ T cells from the PBMCs prior to humanization of the M7007 PDX model would allow assessment of the requirement for CD8+ T cells in leukemia reduction.

Within the scope of these studies, I did not attempt to evaluate the effectiveness of chemotherapy and/or check-point inhibitors in combination with the PBMCs, although the PBMC-humanized PDX murine model has the potential to be used for such trials. The variability in response seen in the initial cohort described here (**Figure 3-9**) suggests a potential immune escape by the non-responders. PDL1 expression on the PDX cells was assessed by flow cytometry, but there was no difference between the non-responders and responders (average CD45lo CD34+ PDL1+ for both groups = 1.13%). Further investigation of mechanisms of immune escape in this model are needed to guide the appropriate use of check-point inhibitor therapy.

CHAPTER 4. SUMMARY, LIMITATIONS, AND FUTURE DIRECTIONS

Summary

In summary, I have developed two novel murine models to aid in the understanding of the CBFA2T3-GLIS2 transcriptional complex and identification of new therapeutic targets. I have shown the fusion oncogene can act as a driver of AMKL by binding sites throughout the genome at GLIS consensus sequences and altering gene expression to block differentiation and promote proliferation. I found that NHR2 is the most critical component of the fusion as deletion of this region abrogates leukemogenesis *in vivo* and impairs ETO binding to the fusion that contributes to downregulation of oncogenic pathways. My preliminary data suggests that CBFA2T3-GLIS2 AMKL cells are immunogenic and can be targeted by leukemia-associated antigen specific CD8+ T cells. Ongoing studies to further elucidate the role of ETO in CBFA2T3-GLIS2 driven leukemogenesis and screen TCR:antigen pairs to identify leukemia-specific TCRs are in progress. The data generated from both novel murine models have, and will continue to, contribute toward the development of targeted therapies against the *CBFA2T3-GLIS2* driver mutation in AMKL.

Limitations and Future Directions

The Role of CBFA2T3-GLIS2

Due to the stability of disease onset in serial transplantations of the novel humanized mouse model (**Figure 2-3**), I infer there are no *de novo* acquired cooperating mutations. To confirm this, genomic stability and clonal diversity will be tracked by performing whole genome sequencing on primary megakaryoblasts derived from an individual patient's cord blood, to control for SNPs, before and after transduction with the CBFA2T3-GLIS2 fusion and after serial transplantations (**Figure 4-1**).

Ideally, I would knockdown *ETO* with an shRNA construct to confirm the data that suggests loss of ETO due to deletion of NHR2 is responsible for the abrogated leukemogenesis observed *in vivo*. However, due to the homology between *ETO* and *CBFA2T3*, shRNA knockdown of *ETO* also reduces expression of the fusion. These studies cannot be performed due to this off-target effect because it would not be possible to discern if the resultant effects are due to knock-down of *ETO* or reduction of fusion expression. To overcome this, CRISPR-Cas9 editing technology will be used to specifically knockout *ETO* without effecting the fusion in the primary human megakaryoblasts which will then be transplanted into immunodeficient mice to assess their leukemic potential. *ETO* will also be knocked-out in fusion negative megakaryoblasts to control for potential effects on other pathways unrelated to those regulated by the fusion.

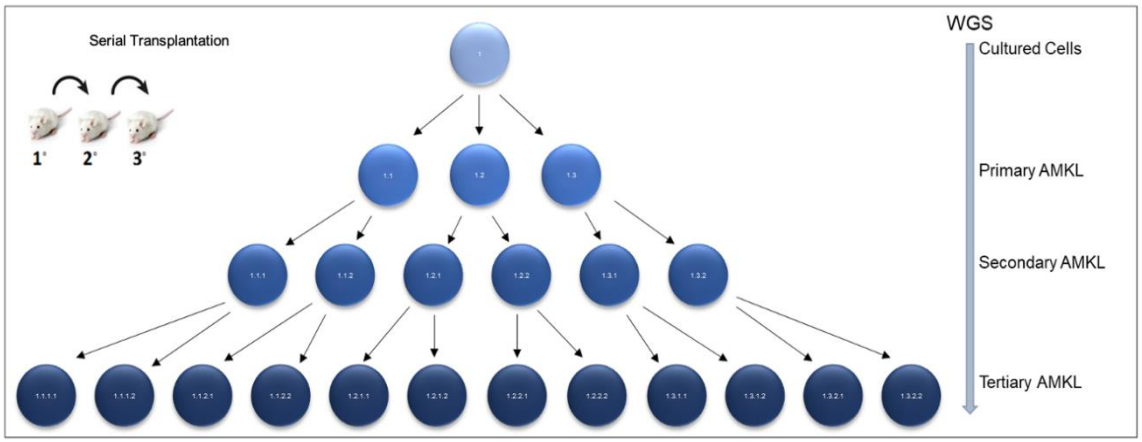


Figure 4-1. Genomic Stability Assessment.
Schematic of genomic stability and clonal diversity experimental design. CBFA2T3-GLIS2 transduced primary human megakaryoblasts derived from one individual human cord blood will be cultured and transplanted into three NSG-SGM3 immunodeficient mice (primary AMKL). Cells will be recovered from the moribund mouse and transplanted into six mice (secondary AMKL). Cells will be recovered from the secondary transplants and transplanted into twelve mice (tertiary AMKL). DNA will be extracted from the cultured cells and primary, secondary, and tertiary AMKL cells and whole genome sequencing (WGC) will be performed to track genomic stability and clonal diversity.

To better define the signaling pathways driving leukemogenesis, a CRISPR dropout screen will be employed. A lentivirus CRISPR-Cas library of siRNA guides (gRNA) targeting genes in the Notch, WNT, and Hedgehog pathways will be transduced into bulk C-G+ pMks at a low efficiency rate so that each cell will receive a single, unique gRNA. The cells will then be cultured *in vitro* in human methylcellulose media over the course of four weeks. If a targeted gene is essential for self-renewal, the cells containing that specific gRNA will not serially re-plate. Pooled DNA extraction and high-throughput sequencing will elucidate which genes are required and which are dispensable or advantageous by comparing representation of the gRNA library that initially was used for the transduction to the sequenced gRNAs after serial replating. gRNAs that are depleted indicate that gene is essential for self-renewal. gRNAs that are enriched suggest that gene is advantageous for self-renewal (Yau et al., 2018). C-G- pMks will also be transduced with the CRISPR library to control for loss-of-function effects unrelated to the expression of the *CBFA2T3-GLIS2* fusion.

CUT&RUN-seq will be done for p300 and H3K27ac to explore the role of p300 in the fusion transcriptional complex and the role of epigenetic regulation, respectively. Specifically, the sequencing data will be used to interrogate shifts in super enhancers between fusion negative and fusion positive primary megakaryoblasts. Additionally, acetylation patterns at promoters and co-occupation of target genes with p300 will inform the contribution to transcriptional dysregulation beyond the impact on enhancers.

To further elucidate the effect of NHR2 deletion, CUT&RUN-seq for ETO and the fusion will be done using *in vitro* cultured primary megakaryoblasts since the deletion prevents expansion *in vivo*. I expect the fusion will still bind DNA since the DNA binding zinc fingers remain intact, but I hypothesize there will be a shift in bound sites due to the lack of ETO binding and the reduction of fusion dimerization. Identification of the altered DNA binding profile will further refine our understanding of the role of ETO in transcriptional regulation and the contribution of the fusion dimerization in DNA binding genome wide.

These studies will contribute to the characterization of the CBFA2T3-GLIS2 transcriptional complex and the identification of critical co-factors essential for leukemic transformation. A similar approach has been taken in the context of mixed lineage leukemia (MLL) translocations in acute leukemia. MLL rearrangements associate with protein complexes involved in transcriptional elongation, including DOT1L (disrupter or telomeric silencing 1-like), a histone methyltransferase (Bernt et al., 2011; Stein et al., 2018). Bernt et al. used a conditional knock-out mouse model for *Dot1l* to show that disruption of *Dot1l* with the MLL-AF9 transcriptional complex reduced leukemic blast colony formation in methylcellulose and decreased expression of MLL-AF9 target genes in *Dot1l* floxed mouse bone marrow co-transduced with MLL-AF9 and Cre, suggesting DOT1L is critical for MLL-AF9 leukemic transformation and may be a candidate therapeutic target (Bernt et al., 2011). Although Stein et al. reported modest efficacy of Pinometostat, a small molecule inhibitor of DOT1L, in a phase 1 study of adults with MLL gene rearrangement acute leukemias (Stein et al., 2018), there was evidence of response in a subset of patients. The predominance of non-responders has been attributed

to the poor pharmacokinetic properties of the inhibitor as opposed to the quality of the target and its requirement in leukemia initiation and maintenance.

I expect the knowledge gained through the studies presented and proposed in this dissertation will follow a similar progression to identify a candidate co-factor, such as ETO, that can be targeted by a small molecular inhibitor and provide efficacy in a clinical setting to improve overall survival for patients with CBFA2T3-GLIS2 driven AMKL. This framework will also set the foundation for similar studies on other AMKL sub-types.

The Immunogenicity of CBFA2T3-GLIS2

As mentioned in Chapter 3, the lack of detection of peptides derived from the CBFA2T3-GLIS2 fusion junction may be due to an absence of surface presentation or the level of presentation is below the range of detection by the pan immunopeptidomics approach. Targeted mass spectrometry analysis will be performed on all four CBFA2T3-GLIS2 PDX samples (**Figure 3-1**) and four NUP98 rearranged pediatric AML specimens, another poor prognosis subtype. This will expand the surface peptidome studies to include additional HLA alleles, increased antigen sensitivity, and include a second high-risk pediatric AML subtype.

The supply of M7007 PDX cells and matched PBMCs is highly limited and, therefore, I am unable to expand the PBMC-humanized murine model. To identify the TCR clone(s) responsible for the anti-leukemic response seen in the initial cohort and to ascertain the antigens that are mediating the response, I have generated stable T cell lines that express the top expanded TCRs from the responders (**Table 3-4**). TCR-null Jurkat 76 cells were transduced with a lentiviral vector encoding the TCR transgenes. These cell lines will be used for *in vitro* antigen screening including tandem minigene assay, peptide stimulation assay, and tetramer staining.

To screen the antigens predicted by NetMHCcons and identified by mass spectrometry from the M7007 PDX cells, all 143 candidate antigens (2 fusion junction antigens predicted by NetMHCcons and 141 mass spectrometry identified antigens) will be evaluated by tandem minigene constructs (TMG). The use of TMGs will allow us to screen many putative neoantigens at once, similar to a peptide pool, to narrow down potential immunogenic antigens that can then be individually validated. TMGs will be synthesized that contain up to seven minigenes encoding seven peptides that are restricted to the same HLA. Each TMG will be electroporated into artificial APC (aAPC) cell lines (K562 HLA-null cells, ATCC) expressing a single TMG corresponding HLA molecule. If a peptide is predicted to be presented by multiple HLA molecules, those peptides will be screened using both aAPC lines. TMG encoded HIV peptides will serve as a negative control. This system will allow presentation of seven HLA-restricted peptides at one time in aAPC cells which can then be used to screen the TCR cell lines with peptide stimulation assays. TMGs that stimulate a functional T cell response will be further

studies through peptide stimulation assays and tetramer staining of each individual peptide encoded in the TMG (Zamora et al., 2019).

To assess the ability of the candidate TCRs to target the leukemia cells *in vivo*, the TCR constructs will be overexpressed in primary human CD8+ T cells from a healthy, unrelated donor. An exact HLA match will be attempted, but, due to limited availability, this may not be achievable. I have used the mouse alpha and beta constant regions in place of human sequences in the TCR transgene sequences to reduce mispairing with endogenous TCRs. The engineered CD8+ T cells will be transplanted individually into M7007 engrafted NSG-SGM3 mice to assess the effectiveness of each clone in lowering the leukemia burden.

Lastly, diagnostic bone marrow from patient M7007 will be screened to see if it also caries clonally expanded T cells, and whether these overlap with those identified in the PBMC-humanized PDX model. The diagnostic marrow was not completely replaced with leukemia (70% blasts and 30% other mononuclear cells); allowing screening of the sample for potential exhausted, leukemia-directed T cell clones that can be compared to the clonally expanded T cells from the model. CD8+ T cells from the sample will be flow sorted and then 5'V(D)J single cell TCR sequencing with paired gene expression analysis (10X Technologies) will be performed to characterize the T cell clonotypes and compare these to the clonotypes identified in the model (**Table 3-4**). If overlapping TCRs from clonally expanded T cells in the patient bone marrow and model are identified, this would strongly suggest these T cells are reactive to the leukemia. An absence of shared clones, however, does not preclude that one or more clone(s) are leukemia specific; therefore, any additional clonally expanded T cells will be included in the *in vitro* and *in vivo* screening experiments outlined above.

These studies will further evaluate the immunogenicity of low mutational burden, fusion-gene driven AMKL. I recognize that finding an antigen:TCR pair is extremely challenging, but data from the model described here provides support for the presence of a leukemia-directed response. If successful, these studies will contribute to the long-term goal of developing TCR-T immunotherapy for AMKL patients to improve outcomes and will support the notion that even low mutational burden leukemias can be effectively treated with immunotherapy. Even if these studies are unable to identify an immunogenic antigen:TCR pair, this research is still of value to the immunotherapy research community as it will provide novel surfaceome and TCR repertoire analyses, in addition to a novel PBMC-humanized PDX model, in the context of fusion-gene driven pediatric leukemia, a broad field of study. Specifically, the framework set by these studies will be applied to the additional three established CBFA2T3-GLIS2 PDX models and will be expanded to other high risk, low mutational burden leukemias such as those driven by NUP98 rearrangements.

LIST OF REFERENCES

Allard, B., Aspeslagh, S., Garaud, S., Dupont, F. A., Solinas, C., Kok, M., Routy, B., Sotiriou, C., Stagg, J., & Buisseret, L. (2018). Immuno-oncology-101: overview of major concepts and translational perspectives. *Seminars in cancer biology*, *52*(Pt 2), 1–11. https://doi.org/10.1016/j.semcancer.2018.02.005

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., Bloomfield, C. D., Cazzola, M., & Vardiman, J. W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, *127*(20), 2391–2405. https://doi.org/10.1182/blood-2016-03-643544

Attanasio, M., Uhlenhaut, N. H., Sousa, V. H., O'Toole, J. F., Otto, E., Anlag, K., Klugmann, C., Treier, A. C., Helou, J., Sayer, J. A., Seelow, D., Nürnberg, G., Becker, C., Chudley, A. E., Nürnberg, P., Hildebrandt, F., & Treier, M. (2007). Loss of GLIS2 causes nephronophthisis in humans and mice by increased apoptosis and fibrosis. *Nature genetics*, *39*(8), 1018–1024. https://doi.org/10.1038/ng2072

Benbarche, S., Lopez, C. K., Salataj, E., Aid, Z., Thirant, C., Laiguillon, M. C., Lecourt, S., Belloucif, Y., Vaganay, C., Antonini, M., Hu, J., da Silva Babinet, A., Ndiaye-Lobry, D., Pardieu, B., Petit, A., Puissant, A., Chaumeil, J., Mercher, T., & Lobry, C. (2022). Screening of ETO2-GLIS2-induced Super Enhancers identifies targetable cooperative dependencies in acute megakaryoblastic leukemia. *Science advances*, *8*(6), eabg9455. https://doi.org/10.1126/sciadv.abg9455

Bernt, K. M., Zhu, N., Sinha, A. U., Vempati, S., Faber, J., Krivtsov, A. V., Feng, Z., Punt, N., Daigle, A., Bullinger, L., Pollock, R. M., Richon, V. M., Kung, A. L., & Armstrong, S. A. (2011). MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer cell*, *20*(1), 66–78. https://doi.org/10.1016/j.ccr.2011.06.010

Bertuccio, S. N., Boudia, F., Cambot, M., Lopez, C. K., Lordier, L., Donada, A., Robert, E., Thirant, C., Aid, Z., Serravalle, S., Astolfi, A., Indio, V., Locatelli, F., Pession, A., Vainchenker, W., Masetti, R., Raslova, H., & Mercher, T. (2020). The Pediatric Acute Leukemia Fusion Oncogene ETO2-GLIS2 Increases Self-Renewal and Alters Differentiation in a Human Induced Pluripotent Stem Cells-Derived Model. *HemaSphere*, *4*(1), e319. https://doi.org/10.1097/HS9.0000000000000319

Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M., & Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated

with a list of genes. *Bioinformatics (Oxford, England)*, *20*(18), 3710–3715. https://doi.org/10.1093/bioinformatics/bth456

Chamberlain, M., Hanamsagar, R., Nestle, F.O., de Rinaldis, E., & Savova, V. (2021). Cell type classification and discovery across diseases, technologies and tissues reveals conserved gene signatures and enables standardized single-cell readouts. *Cold Spring Habor Laboratory*, *10*(1101), e429207. https://doi.org/10.1101/2021.02.01.429207

Chapuis, A. G., Egan, D. N., Bar, M., Schmitt, T. M., McAfee, M. S., Paulson, K. G., Voillet, V., Gottardo, R., Ragnarsson, G. B., Bleakley, M., Yeung, C. C., Muhlhauser, P., Nguyen, H. N., Kropp, L. A., Castelli, L., Wagener, F., Hunter, D., Lindberg, M., Cohen, K., Seese, A., … Greenberg, P. D. (2019). T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nature medicine*, *25*(7), 1064–1072. https://doi.org/10.1038/s41591-019-0472-9

Chapuis, A. G., Ragnarsson, G. B., Nguyen, H. N., Chaney, C. N., Pufnock, J. S., Schmitt, T. M., Duerkopp, N., Roberts, I. M., Pogosov, G. L., Ho, W. Y., Ochsenreither, S., Wölfl, M., Bar, M., Radich, J. P., Yee, C., & Greenberg, P. D. (2013). Transferred WT1-reactive CD8+ T cells can mediate antileukemic activity and persist in post-transplant patients. *Science translational medicine*, *5*(174), 174ra27. https://doi.org/10.1126/scitranslmed.3004916

Chastagner, P., Rubinstein, E., & Brou, C. (2017). Ligand-activated Notch undergoes DTX4-mediated ubiquitylation and bilateral endocytosis before ADAM10 processing. *Science signaling*, *10*(483), eaag2989. https://doi.org/10.1126/scisignal.aag2989

Chen, G., Liu, A., Xu, Y., Gao, L., Jiang, M., Li, Y., Lv, N., Zhou, L., Wang, L., Yu, L., & Li, Y. (2019). The RUNX1-ETO fusion protein trans-activates c-KIT expression by recruiting histone acetyltransferase P300 on its promoter. *The FEBS journal*, *286*(5), 901–912. https://doi.org/10.1111/febs.14751

Dang, J., Nance, S., Ma, J., Cheng, J., Walsh, M. P., Vogel, P., Easton, J., Song, G., Rusch, M., Gedman, A. L., Koss, C., Downing, J. R., & Gruber, T. A. (2017). AMKL chimeric transcription factors are potent inducers of leukemia. *Leukemia*, *31*(10), 2228–2234. https://doi.org/10.1038/leu.2017.51

Dash, P., Fiore-Gartland, A. J., Hertz, T., Wang, G. C., Sharma, S., Souquette, A., Crawford, J. C., Clemens, E. B., Nguyen, T., Kedzierska, K., La Gruta, N. L., Bradley, P., & Thomas, P. G. (2017). Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature*, *547*(7661), 89–93. https://doi.org/10.1038/nature22383

Daver, N., Schlenk, R. F., Russell, N. H., & Levis, M. J. (2019). Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia*, *33*(2), 299–312. https://doi.org/10.1038/s41375-018-0357-9

de Rooij, J. D., Branstetter, C., Ma, J., Li, Y., Walsh, M. P., Cheng, J., Obulkasim, A., Dang, J., Easton, J., Verboon, L. J., Mulder, H. L., Zimmermann, M., Koss, C., Gupta, P., Edmonson, M., Rusch, M., Lim, J. Y., Reinhardt, K., Pigazzi, M., Song, G., … Gruber, T. A. (2017). Pediatric non-Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. *Nature genetics*, *49*(3), 451–456. https://doi.org/10.1038/ng.3772

de Rooij, J. D., Zwaan, C. M., & van den Heuvel-Eibrink, M. (2015). Pediatric AML: From Biology to Clinical Management. *Journal of clinical medicine*, *4*(1), 127–149. https://doi.org/10.3390/jcm4010127

Debela, D. T., Muzazu, S. G., Heraro, K. D., Ndalama, M. T., Mesele, B. W., Haile, D. C., Kitui, S. K., & Manyazewal, T. (2021). New approaches and procedures for cancer treatment: Current perspectives. *SAGE open medicine*, *9*, 20503121211034366. https://doi.org/10.1177/20503121211034366

Drenberg, C. D., Shelat, A., Dang, J., Cotton, A., Orwick, S. J., Li, M., Jeon, J. Y., Fu, Q., Buelow, D. R., Pioso, M., Hu, S., Inaba, H., Ribeiro, R. C., Rubnitz, J. E., Gruber, T. A., Guy, R. K., & Baker, S. D. (2019). A high-throughput screen indicates gemcitabine and JAK inhibitors may be useful for treating pediatric AML. *Nature communications*, *10*(1), 2189. https://doi.org/10.1038/s41467-019-09917-0

Farber, M., Chen, Y., Arnold, L., Möllmann, M., Boog-Whiteside, E., Lin, Y. A., Reinhardt, H. C., Dührsen, U., & Hanoun, M. (2021). Targeting CD38 in acute myeloid leukemia interferes with leukemia trafficking and induces phagocytosis. *Scientific reports*, *11*(1), 22062. https://doi.org/10.1038/s41598-021-01300-8

Gamis, A. S., Alonzo, T. A., Meshinchi, S., Sung, L., Gerbing, R. B., Raimondi, S. C., Hirsch, B. A., Kahwash, S. B., Heerema-McKenney, A., Winter, L., Glick, K., Davies, S. M., Byron, P., Smith, F. O., & Aplenc, R. (2014). Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves event-free survival by reducing relapse risk: results from the randomized phase III Children's Oncology Group trial AAML0531. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, *32*(27), 3021–3032. https://doi.org/10.1200/JCO.2014.55.3628

Gelmetti, V., Zhang, J., Fanelli, M., Minucci, S., Pelicci, P. G., & Lazar, M. A. (1998). Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Molecular and cellular biology*, *18*(12), 7185–7191. https://doi.org/10.1128/MCB.18.12.7185

Gill, P. S., & Rosenblum, N. D. (2006). Control of murine kidney development by sonic hedgehog and its GLI effectors. *Cell cycle (Georgetown, Tex.)*, *5*(13), 1426–1430. https://doi.org/10.4161/cc.5.13.2928

Gruber, T. A., & Downing, J. R. (2015). The biology of pediatric acute megakaryoblastic leukemia. *Blood*, *126*(8), 943–949. https://doi.org/10.1182/blood-2015-05-567859

Gruber, T. A., Larson Gedman, A., Zhang, J., Koss, C. S., Marada, S., Ta, H. Q., Chen, S. C., Su, X., Ogden, S. K., Dang, J., Wu, G., Gupta, V., Andersson, A. K., Pounds, S., Shi, L., Easton, J., Barbato, M. I., Mulder, H. L., Manne, J., Wang, J., … Downing, J. R. (2012). An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer cell*, *22*(5), 683–697. https://doi.org/10.1016/j.ccr.2012.10.007

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., 3rd, Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., Hoffman, P., Stoeckius, M., Papalexi, E., Mimitou, E. P., Jain, J., Srivastava, A., Stuart, T., Fleming, L. M., Yeung, B., Rogers, A. J., … Satija, R. (2021). Integrated analysis of multimodal single-cell data. *Cell*, *184*(13), 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048

Higuchi, M., O'Brien, D., Kumaravelu, P., Lenny, N., Yeoh, E. J., & Downing, J. R. (2002). Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer cell*, *1*(1), 63–74. https://doi.org/10.1016/s1535-6108(02)00016-8

Hildebrand, D., Tiefenbach, J., Heinzel, T., Grez, M., & Maurer, A. B. (2001). Multiple\ regions of ETO cooperate in transcriptional repression. *The Journal of biological chemistry*, *276*(13), 9889–9895. https://doi.org/10.1074/jbc.M010582200

Huang, S., Zhang, B., Fan, W., Zhao, Q., Yang, L., Xin, W., & Fu, D. (2019). Identification of prognostic genes in the acute myeloid leukemia microenvironment. *Aging*, *11*(22), 10557–10580. https://doi.org/10.18632/aging.102477

Hug, B. A., & Lazar, M. A. (2004). ETO interacting proteins. *Oncogene*, *23*(24), 4270 4274. https://doi.org/10.1038/sj.onc.1207674

Jappe, E. C., Garde, C., Ramarathinam, S. H., Passantino, E., Illing, P. T., Mifsud, N. A., Trolle, T., Kringelum, J. V., Croft, N. P., & Purcell, A. W. (2020). Thermostability profiling of MHC-bound peptides: a new dimension in

immunopeptidomics and aid for immunotherapy design. *Nature communications*, *11*(1), 6305. https://doi.org/10.1038/s41467-020-20166-4

Jin, W., Scotto, K. W., Hait, W. N., & Yang, J. M. (2007). Involvement of CtBP1 in the transcriptional activation of the MDR1 gene in human multidrug resistant cancer cells. *Biochemical pharmacology*, *74*(6), 851–859. https://doi.org/10.1016/j.bcp.2007.06.017

Karosiene, E., Lundegaard, C., Lund, O., & Nielsen, M. (2012). NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics*, *64*(3), 177–186. https://doi.org/10.1007/s00251-011-0579-8

Khan, I., Malinge, S., & Crispino, J. (2011). Myeloid leukemia in Down syndrome. *Critical reviews in oncogenesis*, *16*(1-2), 25–36. https://doi.org/10.1615/critrevoncog.v16.i1-2.40

Kim, J. H., Cho, E. J., Kim, S. T., & Youn, H. D. (2005). CtBP represses p300-mediated transcriptional activation by direct association with its bromodomain. *Nature structural & molecular biology*, *12*(5), 423–428. https://doi.org/10.1038/nsmb924

Kim, S. C., Kim, Y. S., & Jetten, A. M. (2005). Krüppel-like zinc finger protein Gli similar 2 (Glis2) represses transcription through interaction with C-terminal binding protein 1 (CtBP1). *Nucleic acids research*, *33*(21), 6805–6815. https://doi.org/10.1093/nar/gki985

Kim, Y. S., Kang, H. S., & Jetten, A. M. (2007). The Krüppel-like zinc finger protein Glis2 functions as a negative modulator of the Wnt/beta-catenin signaling pathway. *FEBS letters*, *581*(5), 858–864. https://doi.org/10.1016/j.febslet.2007.01.058

Kurzrock, R., Talpaz, M., Kantarjian, H., Walters, R., Saks, S., Trujillo, J. M., & Gutterman, J. U. (1987). Therapy of chronic myelogenous leukemia with recombinant interferon-gamma. *Blood*, *70*(4), 943–947.

Li, L., & Wang, Y. (2020). Recent updates for antibody therapy for acute lymphoblastic leukemia. *Experimental hematology & oncology*, *9*(1), 33. https://doi.org/10.1186/s40164-020-00189-9

Lieberman, J. (2010). Granzyme A activates another way to die. *Immunological reviews*, *235*(1), 93–104. https://doi.org/10.1111/j.0105-2896.2010.00902.x

Lin, S., Luo, R. T., Ptasinska, A., Kerry, J., Assi, S. A., Wunderlich, M., Imamura, T., Kaberlein, J. J., Rayes, A., Althoff, M. J., Anastasi, J., O'Brien, M. M., Meetei, A. R., Milne, T. A., Bonifer, C., Mulloy, J. C., & Thirman, M. J. (2016). Instructive Role of MLL-Fusion Proteins Revealed by a Model of t(4;11) Pro-B Acute

Lymphoblastic Leukemia. *Cancer cell*, *30*(5), 737–749. https://doi.org/10.1016/j.ccell.2016.10.008

Liu, Y., Cheney, M. D., Gaudet, J. J., Chruszcz, M., Lukasik, S. M., Sugiyama, D., Lary, J., Cole, J., Dauter, Z., Minor, W., Speck, N. A., & Bushweller, J. H. (2006). The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. *Cancer cell*, *9*(4), 249–260. https://doi.org/10.1016/j.ccr.2006.03.012

Liu, Y., Yan, X., Zhang, F., Zhang, X., Tang, F., Han, Z., & Li, Y. (2022). TCR-T Immunotherapy: The Challenges and Solutions. *Frontiers in oncology*, *11*, 794183. https://doi.org/10.3389/fonc.2021.794183

Lopez, C. K., Noguera, E., Stavropoulou, V., Robert, E., Aid, Z., Ballerini, P., Bilhou Nabera, C., Lapillonne, H., Boudia, F., Thirant, C., Fagnan, A., Arcangeli, M. L., Kinston, S. J., Diop, M., Job, B., Lecluse, Y., Brunet, E., Babin, L., Villeval, J. L., Delabesse, E., … Mercher, T. (2019). Ontogenic Changes in Hematopoietic Hierarchy Determine Pediatric Specificity and Disease Phenotype in Fusion Oncogene-Driven Myeloid Leukemia. *Cancer discovery*, *9*(12), 1736–1753. https://doi.org/10.1158/2159-8290.CD-18-1463

Lutterbach, B., Westendorf, J. J., Linggi, B., Patten, A., Moniwa, M., Davie, J. R., Huynh, K. D., Bardwell, V. J., Lavinsky, R. M., Rosenfeld, M. G., Glass, C., Seto, E., & Hiebert, S. W. (1998). ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Molecular and cellular biology*, *18*(12), 7176–7184. https://doi.org/10.1128/MCB.18.12.7176

Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., & Lajoie, G. (2003). PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM*, *17*(20), 2337–2342. https://doi.org/10.1002/rcm.1196

MacMullan, M. A., Dunn, Z. S., Graham, N., Yang, L., & Wang, P. (2019). Quantitative Proteomics and Metabolomics Reveal Biomarkers of Disease as Potential Immunotherapy Targets and Indicators of Therapeutic Efficacy. *Theranostics*, *9*(25), 7872–7888. https://doi.org/10.7150/thno.37373

Marcu, A., Bichmann, L., Kuchenbecker, L., Kowalewski, D. J., Freudenmann, L. K., Backert, L., Mühlenbruch, L., Szolek, A., Lübke, M., Wagner, P., Engler, T., Matovina, S., Wang, J., Hauri-Hohl, M., Martin, R., Kapolou, K., Walz, J. S., Velz, J., Moch, H., Regli, L., … Neidert, M. C. (2021). HLA Ligand Atlas: a benign reference of HLA-presented peptides to improve T-cell-based cancer immunotherapy. *Journal for immunotherapy of cancer*, *9*(4), e002071. https://doi.org/10.1136/jitc-2020-002071

Marshall, J. S., Warrington, R., Watson, W., & Kim, H. L. (2018). An introduction to immunology and immunopathology. *Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology*, *14*(Suppl 2), 49. https://doi.org/10.1186/s13223-018-0278-1

Marvin-Peek, J., Savani, B. N., Olalekan, O. O., & Dholaria, B. (2022). Challenges and Advances in Chimeric Antigen Receptor Therapy for Acute Myeloid Leukemia. *Cancers*, *14*(3), 497. https://doi.org/10.3390/cancers14030497

Matthijssens, F., Sharma, N. D., Nysus, M., Nickl, C. K., Kang, H., Perez, D. R., Lintermans, B., Van Loocke, W., Roels, J., Peirs, S., Demoen, L., Pieters, T., Reunes, L., Lammens, T., De Moerloose, B., Van Nieuwerburgh, F., Deforce, D. L., Cheung, L. C., Kotecha, R. S., Risseeuw, M. D., … Matlawska-Wasowska, K. (2021). RUNX2 regulates leukemic cell metabolism and chemotaxis in high-risk T cell acute lymphoblastic leukemia. *The Journal of clinical investigation*, *131*(6), e141566. https://doi.org/10.1172/JCI141566

Milne, T. A. (2017). Mouse models of MLL leukemia: recapitulating the human disease. *Blood*, *129*(16), 2217–2223. https://doi.org/10.1182/blood-2016-10-691428

Mo, J., Liang, H., Su, C., Li, P., Chen, J., & Zhang, B. (2021). DDX3X: structure, physiologic functions and cancer. *Molecular cancer*, *20*(1), 38. https://doi.org/10.1186/s12943-021-01325-7

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D., … Groop, L. C. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics*, *34*(3), 267–273. https://doi.org/10.1038/ng1180

Najima, Y., Tomizawa-Murasawa, M., Saito, Y., Watanabe, T., Ono, R., Ochi, T., Suzuki, N., Fujiwara, H., Ohara, O., Shultz, L. D., Yasukawa, M., & Ishikawa, F. (2016). Induction of WT1-specific human CD8+ T cells from human HSCs in HLA class I Tg NOD/SCID/IL2rgKO mice. *Blood*, *127*(6), 722–734. https://doi.org/10.1182/blood-2014-10-604777

Ortega, E., Rengachari, S., Ibrahim, Z., Hoghoughi, N., Gaucher, J., Holehouse, A. S., Khochbin, S., & Panne, D. (2018). Transcription factor dimerization activates the p300 acetyltransferase. *Nature*, *562*(7728), 538–544. https://doi.org/10.1038/s41586-018-0621-1

Pan, C., Liu, H., Robins, E., Song, W., Liu, D., Li, Z., & Zheng, L. (2020). Next generation immuno-oncology agents: current momentum shifts in cancer immunotherapy. *Journal of hematology & oncology*, *13*(1), 29. https://doi.org/10.1186/s13045-020-00862-w

Pott, S., & Lieb, J. D. (2015). What are super-enhancers?. *Nature genetics*, *47*(1), 8–12. https://doi.org/10.1038/ng.3167

Prakash, A., Mahoney, K. E., & Orsburn, B. C. (2021). Cloud Computing Based Immunopeptidomics Utilizing Community Curated Variant Libraries Simplifies and Improves Neo-Antigen Discovery in Metastatic Melanoma. *Cancers*, *13*(15), 3754. https://doi.org/10.3390/cancers13153754

Purcell, A. W., Ramarathinam, S. H., & Ternette, N. (2019). Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nature protocols*, *14*(6), 1687–1707. https://doi.org/10.1038/s41596-019-0133-y

Radtke, I., Mullighan, C. G., Ishii, M., Su, X., Cheng, J., Ma, J., Ganti, R., Cai, Z., Goorha, S., Pounds, S. B., Cao, X., Obert, C., Armstrong, J., Zhang, J., Song, G., Ribeiro, R. C., Rubnitz, J. E., Raimondi, S. C., Shurtleff, S. A., & Downing, J. R. (2009). Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 12944–12949. https://doi.org/10.1073/pnas.0903142106

Raskov, H., Orhan, A., Christensen, J. P., & Gögenur, I. (2021). Cytotoxic $CD8^+$ T cells in cancer and cancer immunotherapy. *British journal of cancer*, *124*(2), 359–367. https://doi.org/10.1038/s41416-020-01048-4

Reya, T., Morrison, S. J., Clarke, M. F., & Weissman, I. L. (2001). Stem cells, cancer, and cancer stem cells. *Nature*, *414*(6859), 105–111. https://doi.org/10.1038/35102167

Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic acids research*, *48*(W1), W449–W454. https://doi.org/10.1093/nar/gkaa379

Richards, R. M., Zhao, F., Freitas, K. A., Parker, K. R., Xu, P., Fan, A., Sotillo, E., Daugaard, M., Oo, H. Z., Liu, J., Hong, W. J., Sorensen, P. H., Chang, H. Y., Satpathy, A. T., Majzner, R. G., Majeti, R., & Mackall, C. L. (2021). NOT-Gated CD93 CAR T Cells Effectively Target AML with Minimized Endothelial Cross-Reactivity. *Blood cancer discovery*, *2*(6), 648–665. https://doi.org/10.1158/2643-3230.BCD-20-0208

Roth, M. D., & Harui, A. (2015). Human tumor infiltrating lymphocytes cooperatively regulate prostate tumor growth in a humanized mouse model. *Journal for immunotherapy of cancer*, *3*, 12. https://doi.org/10.1186/s40425-015-0056-2

Said, F., Shafik, R.E. & Hassan, N.M. *RUNX1* gene expression in Egyptian acute myeloid leukemia patients: may it have therapeutic implications?. *Egypt J Med Hum Genet* **22,** 58 (2021). https://doi.org/10.1186/s43042-021-00179-4

Salmikangas, P., Kinsella, N., & Chamberlain, P. (2018). Chimeric Antigen Receptor T Cells (CAR T-Cells) for Cancer Immunotherapy - Moving Target for Industry?. *Pharmaceutical research*, *35*(8), 152. https://doi.org/10.1007/s11095-018-2436-z

Schumacher, T. N., & Schreiber, R. D. (2015). Neoantigens in cancer immunotherapy. *Science (New York, N.Y.)*, *348*(6230), 69–74. https://doi.org/10.1126/science.aaa4971

Sheykhhasan, M., Manoochehri, H., & Dama, P. (2022). Use of CAR T-cell for acute lymphoblastic leukemia (ALL) treatment: a review study. *Cancer gene therapy*, 10.1038/s41417-021-00418-1. Advance online publication. https://doi.org/10.1038/s41417-021-00418-1

Shimasaki, N., Jain, A., & Campana, D. (2020). NK cells for cancer immunotherapy. *Nature reviews. Drug discovery*, *19*(3), 200–218. https://doi.org/10.1038/s41573-019-0052-1

Stein, E. M., Garcia-Manero, G., Rizzieri, D. A., Tibes, R., Berdeja, J. G., Savona, M. R., Jongen-Lavrenic, M., Altman, J. K., Thomson, B., Blakemore, S. J., Daigle, S. R., Waters, N. J., Suttle, A. B., Clawson, A., Pollock, R., Krivtsov, A., Armstrong, S. A., DiMartino, J., Hedrick, E., Löwenberg, B., … Tallman, M. S. (2018). The DOT1L inhibitor pinometostat reduces H3K79 methylation and has modest clinical activity in adult acute leukemia. *Blood*, *131*(24), 2661–2669. https://doi.org/10.1182/blood-2017-12-818948

Steinauer, N., Guo, C., Huang, C., Wong, M., Tu, Y., Freter, C. E., & Zhang, J. (2019). Myeloid translocation gene CBFA2T3 directs a relapse gene program and determines patient-specific outcomes in AML. *Blood advances*, *3*(9), 1379–1393. https://doi.org/10.1182/bloodadvances.2018028514

Steinauer, N., Guo, C., & Zhang, J. (2017). Emerging Roles of MTG16 in Cell-Fate Control of Hematopoietic Stem Cells and Cancer. *Stem cells international*, *2017*, 6301385. https://doi.org/10.1155/2017/6301385

Steinauer, N., Guo, C., & Zhang, J. (2020). The transcriptional corepressor CBFA2T3 inhibits all-*trans*-retinoic acid-induced myeloid gene expression and differentiation in acute myeloid leukemia. *The Journal of biological chemistry*, *295*(27), 8887–8900. https://doi.org/10.1074/jbc.RA120.013042

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Tan, C. T., Croft, N. P., Dudek, N. L., Williamson, N. A., & Purcell, A. W. (2011). Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. *Proteomics*, *11*(11), 2336–2340. https://doi.org/10.1002/pmic.201000531

Thirant, C., Ignacimouttou, C., Lopez, C. K., Diop, M., Le Mouël, L., Thiollier, C., Siret, A., Dessen, P., Aid, Z., Rivière, J., Rameau, P., Lefebvre, C., Khaled, M., Leverger, G., Ballerini, P., Petit, A., Raslova, H., Carmichael, C. L., Kile, B. T., Soler, E., … Mercher, T. (2017). ETO2-GLIS2 Hijacks Transcriptional Complexes to Drive Cellular Identity and Self-Renewal in Pediatric Acute Megakaryoblastic Leukemia. *Cancer cell*, *31*(3), 452–465. https://doi.org/10.1016/j.ccell.2017.02.006

Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., & Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry*, *75*(8), 1895–1904. https://doi.org/10.1021/ac0262560

Tijssen, M. R., Cvejic, A., Joshi, A., Hannah, R. L., Ferreira, R., Forrai, A., Bellissimo, D. C., Oram, S. H., Smethurst, P. A., Wilson, N. K., Wang, X., Ottersbach, K., Stemple, D. L., Green, A. R., Ouwehand, W. H., & Göttgens, B. (2011). Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Developmental cell*, *20*(5), 597–609. https://doi.org/10.1016/j.devcel.2011.04.008

Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., 2nd, Treacy, D., Trombetta, J. J., Rotem, A., Rodman, C., Lian, C., Murphy, G., Fallahi-Sichani, M., Dutton-Regester, K., Lin, J. R., Cohen, O., Shah, P., Lu, D., Genshaft, A. S., Hughes, T. K., Ziegler, C. G., Kazer, S. W., … Garraway, L. A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (New York, N.Y.)*, *352*(6282), 189–196. https://doi.org/10.1126/science.aad0501

Uy, G. L., Aldoss, I., Foster, M. C., Sayre, P. H., Wieduwilt, M. J., Advani, A. S., Godwin, J. E., Arellano, M. L., Sweet, K. L., Emadi, A., Ravandi, F., Erba, H. P., Byrne, M., Michaelis, L., Topp, M. S., Vey, N., Ciceri, F., Carrabba, M. G., Paolini, S., Huls, G. A., … DiPersio, J. F. (2021). Flotetuzumab as salvage immunotherapy for refractory acute myeloid leukemia. *Blood*, *137*(6), 751–762. https://doi.org/10.1182/blood.2020007732

van der Lee, D. I., Koutsoumpli, G., Reijmers, R. M., Honders, M. W., de Jong, R., Remst, D., Wachsmann, T., Hagedoorn, R. S., Franken, K., Kester, M., Harber, K. J., Roelofsen, L. M., Schouten, A. M., Mulder, A., Drijfhout, J. W., Veelken, H., van Veelen, P. A., Heemskerk, M., Falkenburg, J., & Griffioen, M. (2021). An HLA-A*11:01-Binding Neoantigen from Mutated NPM1 as Target for TCR Gene Therapy in AML. *Cancers*, *13*(21), 5390. https://doi.org/10.3390/cancers13215390

Vasanth, S., ZeRuth, G., Kang, H. S., & Jetten, A. M. (2011). Identification of nuclear localization, DNA binding, and transactivating mechanisms of Kruppel-like zinc finger protein Gli-similar 2 (Glis2). *The Journal of biological chemistry*, *286*(6), 4749–4759. https://doi.org/10.1074/jbc.M110.165951

Wang, L., Gural, A., Sun, X. J., Zhao, X., Perna, F., Huang, G., Hatlen, M. A., Vu, L., Liu, F., Xu, H., Asai, T., Xu, H., Deblasio, T., Menendez, S., Voza, F., Jiang, Y., Cole, P. A., Zhang, J., Melnick, A., Roeder, R. G., … Nimer, S. D. (2011). The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science (New York, N.Y.)*, *333*(6043), 765–769. https://doi.org/10.1126/science.1201662

Wang, L., Sun, X., He, J., & Liu, Z. (2021). Functions and Molecular Mechanisms of Deltex Family Ubiquitin E3 Ligases in Development and Disease. *Frontiers in cell and developmental biology*, *9*, 706997. https://doi.org/10.3389/fcell.2021.706997

Wang, P., Chen, Y., & Wang, C. (2021). Beyond Tumor Mutation Burden: Tumor Neoantigen Burden as a Biomarker for Immunotherapy and Other Types of Therapy. *Frontiers in oncology*, *11*, 672677. https://doi.org/10.3389/fonc.2021.672677

Wichmann, C., Chen, L., Heinrich, M., Baus, D., Pfitzner, E., Zörnig, M., Ottmann, O. G., & Grez, M. (2007). Targeting the oligomerization domain of ETO interferes with RUNX1/ETO oncogenic activity in t(8;21)-positive leukemic cells. *Cancer research*, *67*(5), 2280–2289. https://doi.org/10.1158/0008-5472.CAN-06-3360

Wieczorek, M., Abualrous, E. T., Sticht, J., Álvaro-Benito, M., Stolzenberg, S., Noé, F., & Freund, C. (2017). Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen

Presentation. *Frontiers in immunology*, *8*, 292. https://doi.org/10.3389/fimmu.2017.00292

Wieder, C., Frainay, C., Poupin, N., Rodríguez-Mier, P., Vinson, F., Cooke, J., Lai, R. P., Bundy, J. G., Jourdan, F., & Ebbels, T. (2021). Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis. *PLoS computational biology*, *17*(9), e1009105. https://doi.org/10.1371/journal.pcbi.1009105

Wolff, J. E., Wagner, S., Reinert, C., Gnekow, A., Kortmann, R. D., Kühl, J., & Van Gool, S. W. (2006). Maintenance treatment with interferon-gamma and low-dose cyclophosphamide for pediatric high-grade glioma. *Journal of neuro-oncology*, *79*(3), 315–321. https://doi.org/10.1007/s11060-006-9147-8

Wu, D., Yang H., Xue, X., Liang, W., Miao, X., Chen S., & Pang, H. (2005). Oligomerization study of NHR3 and NHR4 domains from ETO protein involved in t(8;12)-assoicated acute myeloid leukemia. *Chinese Science Bulletin*, *50*(9), 875-879. https://doi.org/10.1360/982004-820

Yan, M., Kanbe, E., Peterson, L. F., Boyapati, A., Miao, Y., Wang, Y., Chen, I. M., Chen, Z., Rowley, J. D., Willman, C. L., & Zhang, D. E. (2006). A previously unidentified alternatively spliced isoform of t(8;21) transcript promotes leukemogenesis. *Nature medicine*, *12*(8), 945–949. https://doi.org/10.1038/nm1443

Yau, E. H., & Rana, T. M. (2018). Next-Generation Sequencing of Genome-Wide CRISPR Screens. *Methods in molecular biology (Clifton, N.J.)*, *1712*, 203–216. https://doi.org/10.1007/978-1-4939-7514-3_13

Zamora, A. E., Crawford, J. C., Allen, E. K., Guo, X. J., Bakke, J., Carter, R. A., Abdelsamed, H. A., Moustaki, A., Li, Y., Chang, T. C., Awad, W., Dallas, M. H., Mullighan, C. G., Downing, J. R., Geiger, T. L., Chen, T., Green, D. R., Youngblood, B. A., Zhang, J., & Thomas, P. G. (2019). Pediatric patients with acute lymphoblastic leukemia generate abundant and functional neoantigen-specific $CD8^+$ T cell responses. *Science translational medicine*, *11*(498), eaat8549. https://doi.org/10.1126/scitranslmed.aat8549

Zhang, B., Whiteaker, J. R., Hoofnagle, A. N., Baird, G. S., Rodland, K. D., & Paulovich, A. G. (2019). Clinical potential of mass spectrometry-based proteogenomics. *Nature reviews. Clinical oncology*, *16*(4), 256–268. https://doi.org/10.1038/s41571-018-0135-7

Zhang, W., Ma, Q., Long, B., Sun, Z., Liu, L., Lin, D., & Zhao, M. (2021). Runt-Related Transcription Factor 3 Promotes Acute Myeloid Leukemia Progression. *Frontiers in oncology*, *11*, 725336. https://doi.org/10.3389/fonc.2021.725336

Zhou, S., Liu, M., Ren, F., Meng, X., & Yu, J. (2021). The landscape of bispecific T cell engager in cancer treatment. *Biomarker research*, *9*(1), 38. https://doi.org/10.1186/s40364-021-00294-9

VITA

Elizabeth A. Garfinkle was born in 1994 and grew up in Santa Barbara, California. In 2016, she earned a Bachelor of Science in Microbiology with a minor in Chemistry *cum laude* with Departmental Honors from California State University, Long Beach (CSULB). As a junior at CSULB, she received the Glenn M. Nagel Undergraduate Research Fellowship. In 2016, she was accepted into the Cancer and Developmental Biology Track of the Biomedical Sciences PhD Program at the University of Tennessee Health Science Center (UTHSC). She joined the lab of Dr. Tanja Gruber, formally of St. Jude Children's Research Hospital and currently at Stanford University, in 2017. During her graduate career, Elizabeth presented at multiple local, national, and international scientific conferences and earned several awards including the UTHSC College of Graduate Health Sciences Outstanding Student Award, a Young Investigators Abstract Travel Award from the Society for Immunotherapy of Cancer, and an Abstract Achievement Award from the American Society of Hematology. Elizabeth served as the Vice President of the Graduate Student Executive Council for the 2019-2020 academic year and served over 190 hours as a volunteer at St. Jude Children's Research Hospital. Elizabeth expects to complete her Doctor of Philosophy in Biomedical Sciences with a concentration in Cancer and Developmental Biology in June of 2022.