Generating Tools to Investigate Infant Leukemia Disease Biology and Identify New Therapies Through Rapid Drug Screening.

by

Keon Collett

Submitted in partial fulfilment of the requirements for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
May 2020

© Copyright by Keon Collett, 2020

DEDICATION

For my sister, Neekoo.

TABLE OF CONTENTS

List of Tables ……………………………………………………………………...… vii

List of Figures …………………………………………………...……………..... viii

Abstract ………………………………………………………………………….….. ix

List of Abbreviations and Symbols Used …………………………………….….. x

Acknowledgements ………………………………………………………...……. xii

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

1.1 Infant Leukemia (IL) …………………………………...……………… 1

1.1.1 <u>IL Classification, Diagnosis and Symptomatic Presentation</u>……...…... 1

1.1.2 <u>IL Treatment, Challenges, and Prognosis</u> ………….…………..…... 3

1.2 Mixed Lineage Leukemia (MLL) ……………………………...………… 5

1.2.1 <u>Epigenetics</u> ………………………………………………...……... 5

1.2.2 *<u>Mixed Lineage Leukemia (MLL)</u>* …….……………………...………… 9

1.2.3 <u>MLL, HOX and Hematopoiesis</u> …...…………….…………...…... 13

1.2.4 <u>HOX cofactors: MENIN, LEDGF & MEIS1</u> ……………..………… 16

1.3 MLL-Fusions & Infant Leukemia (IL) ...………………………………. 18

1.3.1 <u>MLL-Fusions</u> ………………………………...………………… 18

1.3.2 <u>MLL-AF9 & MLL-ENL</u> ……………………………………….……. 21

1.3.3 <u>MLL-Fusions Bind MENIN & MEIS1</u> …………………...………. 22

1.3.4 <u>MLL-Fusions Recruit DOT1L</u> …………………………...….……. 24

1.3.5 <u>MLL-Fusions Activate HOX expression</u> …………………....…… 27

1.4 Germline KMT2C Mutations in IL ….……………………...…….…. 27

1.5 Current Clinical Trials & Drugs in Development ……………..………. 29

1.6 Murine Models Highlight Knowledge Gap in *MLL-r* Disease Biology …. 32

1.7 Zebrafish as a Model for Leukemia and Drugs Screening………..………. 33

1.7.1 Zebrafish Hematopoiesis …………….……...…………………... 33

1.7.2 Zebrafish as Pre-Clinical Models of Leukemia …………....………35

1.7.3 Zebrafish *mll* and Established *MLL-r* Leukemia models in Zebrafish………………………………………………………… 36

1.7.4 Zebrafish as a Platform for Drug Screening and Discovery ………. 37

1.8 Hypothesis and Objectives ….…………………..…………..………… 38

Chapter 2: Materials and Methods ……….……………………..……...…… 40

2.1 Study Approval ……………………………………………...…...…… 40

2.2 Zebrafish Husbandry, Embryo Collection, and Embryo Staging ………. 40

2.3 Zebrafish Embryo Injection ……………………..……………………. 40

2.4 RNA Isolation and cDNA Synthesis……………………………...……… 41

2.5 Vectors and Cloning…………………..………………………..……… 41

2.6 Generation of Transgenic Zebrafish Lines ……………..………...…..… 44

2.7 4-OHT Treatment …………………………………………………..…. 44

2.8 Fluorescent Imaging ………………………………………...…… 45

2.9 Capped and Un-Capped *in vitro* Transcription …………………...…… 45

2.10 sgRNA and Cas9 mRNA Synthesis ………………………………...…… 46

2.11 Generation of kmt2ca and kmt2cb Mutant Zebrafish Lines …..……..… 47

2.12 Fin-Clipping and gDNA Extraction ……………………………...…… 47

2.13 Frameshift Reporter Assay ……………………………………..…… 48

2.14 RNA Probe Synthesis and WISH ………..……………………..……. 48

2.15 Data Visualization and Statistical Analysis…………………….…… 49

Chapter 3: Results I – First Steps Toward a Transgenic Zebrafish Model of MLL-r IL……………………………………………………………………………. 50

3.1 Introduction ………………………………………………………………. 50

3.2 Generating an Inducible Transgenic Zebrafish Model of MLL-r Leukemia…………………………………………………………………… 50

3.3 Transgenic Larvae express CreERT2 when driven by promoters *ubi*, *cd45*, and *pu.1* but not *runx1+23*………………………………………………… 54

3.4 Chimeric, but not stable transgenic larvae expressing *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* are able to undergo recombination in the presence of Cre recombinase………………………… 56

3.5 Summary of Findings ……………………………………………………… 62

Chapter 4: Results II – Generating kmt2ca and kmt2cb mutant zebrafish to assess the role of KMT2C in hematopoiesis…………………………………………… 63

4.1 Introduction ………………………………………………………………. 63

4.2 Zebrafish Homologs kmt2ca and kmt2cb are Significantly Conserved with Human KMT2C…………………………………………………………… 63

4.3 Expression of *kmt2ca* and *kmt2cb* Is Not Organ Specific and Peaks During the First 24 Hours of Development…………………………………… 66

4.4 Generating *kmt2ca* and *kmt2cb* Mutant Zebrafish………………………… 66

4.5 Reporter Assay demonstrates efficacy of *kmt2ca* and *kmt2cb* frameshift mutations……………………………………………………………………… 71

4.6 *Kmt2cb-/-* Embryo's have Decreased Numbers of Neutrophils, Mast Cells, and other Leukocytes……………………………………………………… 75

4.7 Summary of Findings……………………………………………………… 78

Chapter 5: Discussion…………………………………………………………… 79

5.1 Overview of Thesis………………………………………………………… 79

5.2 Rational For Tissue Specific CreERT2 Transgenic Fish………………… 80

5.3 Limitations of Cre/lox System in Zebrafish……………………………… 81

5.3.1 Existing Cre/Lox lines Are Needed to Efficiently Generate New Models…………………………………………………………………… 81

5.3.2 Recombination at Lox Sites Requires Single Insertion Transgenesis……………………………………..…………………… 82

5.4 Permanent Recombination of Cre/lox System Provides Advantage Over Other Inducible Systems…………………………….…………………..….. 85

5.5 Drug Screening Application of the Fluorescent IL Model……………… 87

5.6 Understanding *kmt2ca* and *kmt2cb* Expression Patterns……………..… 89

5.7 Rational and Validation of *kmt2ca* and *kmt2cb* Mutant Zebrafish Model…………………………………………………………………..…. 90

5.8 Assessing Blood Development in *kmt2ca-/-* and *kmt2ca-/-* Zebrafish……………………………………...…………………….…….. 92

5.9 Observational Findings in *kmt2ca-/-* and *kmtcb-/-* Zebrafish………..…. 93

5.10 Limitations of the Zebrafish *KMT2C* Model …………………………… 96

5.11 Future Dirrections………………………………………………………96

5.12 Impact and Conclusions ………………………………..……………. 97

References ……………………………………………………………...…… 97

LIST OF TABLES

Table 1.2.1: List of Histone Methylation and Effects on Gene Expression …….....…7

Table 2.10: List of single guide RNA sequences used to generate zebrafish *kmt2ca* and *kmt2cb* mutants ……………………...……………………….……….….. 46

Table: 2.14.1: Primers used to create WISH probes ……………..……………… 48

Table 2.14.2: Hematopoietic Lineage Specific Whole-Mount In-Situ Probes …..… 49

Table 4.4.1: Table of different *kmt2ca* mutations characterized at the genomic DNA (gDNA) and compliment DNA (cDNA levels), before establishing the *kmt2ca(-/-)* line that contains mutation #13……….……………………………………...……..… 69

Table 4.4.2: Table of different *kmt2cb* mutations characterized at the genomic DNA (gDNA) and complimentary DNA (cDNA) levels, before establishing the *kmt2cb(-/-)* line that contains mutation #15……………….…………………………..…….... 69

LIST OF FIGURES

Figure 1.2.1: An overview of the 3 types of epigenetic regulation…………………..... 8

Figure 1.2.2: MLL interacts with a complex of proteins …………………………… 11

Figure 1.2.3: The role of HOX proteins in control of hematopoiesis ………….…… 15

Figure 1.3.1: Only 6 different fusion partners make up the majority of MLL-fusion genes, despite the characterization of >90 different fusion partners .. 20

Figure 1.3.4: MLL-Fusions recruit DOT1L to activate gene expression through histone lysine methylation in the absence of a SET domain ……………………..…… 26

Figure 3.2: Schema for the application of cre/lox inducible technology to generate transgenic zebrafish that express either human *MLL-AF9* or *MLL-ENL* fusion genes in a variety of cell populations ...………………………….…. 53

Figure 3.3: Transgenic larvae express *CreERT2* when driven by *ubi*, *Cd45*, and *pu.1* promoters but not *runx1+23* ……………………………………………….…. 55

Figure 3.4.1: *CreERT2* inducible fluorescent reporter transgenes undergo recombination *in vivo* ……………………………………………………………..…. 57

Figure 3.4.2: Transgenic larvae expressing *ubi:LBL-P2A-mCherry*, but not *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* are able to undergo Cre-lox recombination …………………………………………...… 60

Figure 4.2: Zebrafish kmt2ca and kmt2cb proteins are highly conserved with human KMT2C ………………………………………………………………………..… 65

Figure 4.3: Analysis of physiological *kmt2ca* and *kmt2cb* in early zebrafish development ………………………………………………………………………… 67

Figure 4.4: Generation and Mutation Characterization of *kmt2ca* and *kmt2cb* mutant zebrafish …………………………………………………………………… 70

Figure 4.5: Expression of mutation reporters for *kmt2ca* and *kmt2cb* indicate efficient frameshift mutations …………………………………………………..…. 73

Figure 4.6: *kmt2cb(-/-)* zebrafish larvae demonstrate reduced myelopoiesis during primitive hematopoiesis and reduced numbers of mast cells, but not neutrophils during definitive hematopoiesis……………………………………………..….. 76

Figure 5.5: Drug screening in the zebrafish can be partly automated using a Biosorter to detect levels of fluorescence…………………………………………..…. 88

Figure 5.9: Adult *kmt2ca-/-* zebrafish succumb to disease starting as early as ~7months of age………………………………………………………………...... 95

Abstract

Infant Leukemia (IL) is a rare and aggressive congenital leukemia that is frequently caused by fusion genes involving lysine methyltransferase *MLL (MLL1; KMT2A)* and has a dismal prognosis. Further investigation into MLL-fusion disease biology and the role of contributing genetic factors, such as predisposing germline mutations to lysine methyltransferase *KMT2C (MLL3)*, is needed. I genetically engineered two zebrafish models: a Cre/lox inducible transgenic model to express *MLL-AF9* or *MLL-ENL* within specific blood cell lineages, and a knockout model of *KMT2C* homologs, *kmt2ca* and *kmt2cb*. I successfully established three blood cell specific CreERT2 expressing transgenic lines. However, *MLL-AF9* and *MLL-ENL* transgenic fish failed to expression of the *MLL*-fusion genes and this model requires further development. Preliminary results indicate *kmt2cb-/-* larvae may have decreased myelopoiesis during primitive hematopoiesis suggesting a block in differentiation, while *kmt2ca-/-* fish have yet to be assessed. These models represent the development of new tools for IL research.

List of Abbreviations Used

4-OHT	4-hydroxytamoxifen
AA	amino acid
ALL	acute lymphoid leukemia
AML	acute myeloid leukemia
ATP	adenosine triphosphate
bp	base pairs
Cas9	CRISPR associated protein 9
cDNA	compliment DNA
CDS	coding domain sequence
CHT	caudal hematopoietic tissue
CNS	central nervous system
COMPASS	complex of proteins associated with Set1
CRISPR	clustered regularly interspersed short palindromic repeats
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
dpf	days post fertilization
eCFP	enhanced cyan fluorescent protein
eGFP	enhanced green fluorescent protein
gDNA	genomic DNA
floxed	flanking Lox sites
hpf	hours post fertilization
H3K4	histone 3 lysine 4
HOX	homeobox genes
HSC	hematopoietic stem cell
HSCT	hematopoietic stem cell transplantation
HSPC	hematopoietic stem and progenitor cells
IL	infant leukemia

ISH	in-situ hybridization
LOF	loss of function
Lox	locus of recombination
MLL	Mixed Lineage Leukemia
MLL-r	MLL- rearrangements
MPAL	mixed phenotype acute leukemia
mRNA	messenger RNA
miRNA	micro RNA
mTagBFP2	monomeric Tag blue fluorescent protein 2
MUT	mutant
PcG	Polycomb group proteins
PCR	polymerase chain reaction
PHD	plant homology domain
RBC	red blood cell
RNA	ribonucleic acid
sgRNA	single guide RNA
TagRFP	Tag red fluorescent protein
TrxG	Trithorax group proteins
ubi	ubiquitin C
WBC	white blood cell
WISH	whole-mount in-situ hybridization
WT	wild type

Acknowledgments

First and foremost, I would like to thank my supervisor Dr. Jason Berman for his enthusiastic and endless support. I am grateful to him for giving me the creative freedom to develop an ambitious project and for the trust he placed in me. He has been a source of inspiration, and a true mentor. I would also like to thank you to my Scientific Advisory Committee (Graham Dellaire, Clinton Campbell, Paola Marcato, and Todd Druley), who have provided advice and input throughout this project.

Thank you to all the members of the Berman lab for their support and encouragement, and for always taking the time to answer questions and teach me new methods or concepts. Your commitment to Berman Lab culture made my time in the lab extremely enjoyable. Sergey Prykhozhij and Vinothkumar Rajan, thank you for guiding me throughout this project and for all of your mentorship. Benno Orr & Anna Cordeiro-Santanach directly contributed to the development of the transgenic and mutant zebrafish within this study. To the cove: Nicole, Jaime, Olivia, and to Kim, thank you for constantly answering my questions and helping me through the highs and lows of graduate school.

Lastly, I would especially like to thank my siblings, Neekoo and Emon. Thank you for always listening to me complain and picking up the phone when I call. Truly none of this would be possible without your encouragement and endless support.

Chapter 1: Introduction

1.1 Infant Leukemia (IL)

1.1.1 IL Classification, Diagnosis and Symptomatic Presentation

Infant leukemia (IL) is an aggressive type of leukemia that presents in children before the age of one year. In the United States, IL has an incidence rate of 41 cases per million births, or roughly 160 cases each year (Brown, Pieters, & Biondi, 2019), making it a rare disease. Despite this infrequency, IL is an aggressive disease that affects an extremely vulnerable population and the pediatric oncology community has become increasingly focused on tackling challenges in IL treatment and poor outcomes. The onset of IL is rapid, and symptoms include hyperleukocytosis (high WBC counts), hepatosplenomegaly (enlargement of the spleen and liver), CNS infiltration and leukemia cutis (skin infiltration). Furthermore, anemia (decreased numbers of red blood cells) and neutropenia (decreased numbers of neutrophils) occurs when leukemia blasts take over the bone marrow, displacing healthy tissue and disrupting the production of normal blood cells. (Rubnitz, Gibson, & Smith, 2008). Patients with IL typically present in the clinic with lethargy, fever, rash, and a lack of appetite (Brown, Pieters, & Biondi, 2019).

IL can be either an acute myeloid leukemia (AML) or an acute lymphoid leukemia (ALL). Myeloid leukemias are the result of the uncontrollable expansion of immature myeloid precursor or progenitor cell populations. The WHO (World Health Organization) classifies AML disease based on significant cytogenetic and molecular genetic abnormalities, such as chromosomal translocations or specific gene mutations, rather than the by the morphology of leukemic blasts like the older FAB (French-American-British) system (Arber et al., 2016). Where AML entities are not able to be classified by genetic signature, the WHO system relies cell lineage classifications and levels of cell differentiation. However, these classifications are not widely used in IL. IL is not currently a discrete entity as defined by the WHO despite the frequent presence of genetic abnormalities and displaying unique and aggressive disease progression (Arber et al., 2016). Additionally, the morphological phenotype of IL AML does not currently inform treatment protocols (Brown, Pieters & Biondi, 2019). Higher rates of monoblastic, myelomonoblastic and megakaryoblastic AML is seen in infants as compared to older

children, and these morphological phenotypes make up ~70% of IL AML cases (Creutzig et al., 2012). ALL in infants is categorized by immunophenotype into either B-ALL (B-cell leukemia) or T-ALL (T-cell leukemia). Importantly, B-ALLs are tested for the presence or absence of the cell surface marker CD10. CD10 negative B-ALL represents a uniquely immature population of lymphoid cells (pre-pre B-ALL or pro-B-ALL) that is not typically seen in older pediatric B-ALL patients and is a marker of poor prognosis in infants (Brown, 2013). Within infant ALL, nearly all cases are B-cell lineage with <5% being T-ALL (Brown, Pieters, & Biondi, 2019). Overall, there is a slight predominance of ALL cases (~56%) over AML cases (~44%), but interestingly AML in infants occurs at twice the rate compared to older pediatric leukemia patients (Brown, 2013).

Although IL is typically reported as either AML or ALL, largely in part because this classification informs the treatment regimen, a third phenotypic categorization exists — mixed phenotype acute leukemia (MPAL). Like the name suggests, MPAL is a leukemia that displays characteristics of more than one of the following hematopoietic cell lineages: B-lineage, T-lineage, myeloid lineage. MPAL is further classified into two categories: bi-phenotypic or bi-lineal leukemia. Bi-phenotypic leukemia exists when morphological and/or genetic characteristics of both the myeloid and lymphoid cell lineages are present within a single leukemic population. (Charles & Boyer, 2017). In contrast, bi-lineal MPAL is when two distinct leukemia populations of different lineages are present within the same acute disease (Charles & Boyer, 2017).

The phenomenon of MPAL and 'lineage infidelity' within leukemia, and specifically infant leukemia, is largely associated with the aberrant expression of hematopoietic cell lineage associated genes. Typically, this is due to the presence of *MLL (Mixed Lineage Leukemia; MLL1; KMT2A)* rearrangements (*MLL-r*), which are present in ~80% of all IL cases (Rossi et al., 2012). Although, MPAL largely exists in the presence *MLL-r, MLL-r* do not always result in an MPAL phenotype and an *MLL-r* IL patients do not always develop MPAL (Charles & Boyer, 2017). Children diagnosed with MPAL are typically defined as a high-risk group, as these leukemias express high levels of multidrug resistant proteins and can undergo cell lineage switching in response to treatment, that renders therapy ineffective (Charles & Boyer, 2017; Rayes, McMasters, & O'Brien, 2016).

However, treatment of pediatric MPAL leukemia with ALL based therapies have resulted in outcomes similar to non-MPAL pediatric leukemias (Hrusak et al., 2018; Hunger et al., 2012; Orgel et al., 2019; Pui et al., 2009).

1.1.2 Treatment, Challenges, and Prognosis

Infants were initially treated with general protocols used for the treatment of acute leukemia in older children. To improve outcomes, unique IL protocols have since been developed, but with minimal improvement. Despite efforts to improve chemotherapy regimens, prognosis for infant leukemia patients is dismal with less than 50% surviving 5 years (Linabery & Ross, 2008).

Induction therapy uses large doses of chemotherapy to kill fast replicating leukemia cells, after which it is hoped the patient will achieve remission. In addition to intravenous administration, patients with CNS involvement can be treated with intrathecal doses of chemotherapy. Cranial radiation is largely avoided due to the severe impact on CNS development it causes in young children. An overwhelming majority of patients (~80-90%) achieve remission, however over half of these patients relapse (Reaman et al., 1999). The bone marrow is the most common site of relapse, but intensive consolidation approaches like bone marrow transplantation remains a controversial treatment approach due to high levels of toxicity and increased mortality associated with the procedure (Dreyer et al., 2011; Pieters et al., 2007; Pui et al., 2002).

Toxicity is a major treatment barrier in IL. Rapid physiological and developmental advancements occur in the first year of life and therefore, patients suffer from treatment-related toxicities and high rates of infection. The Children's Oncology Group (COG) has recorded rates of death from primary infections as high as 25% for ALL trials, with similar issues reported for the treatment of IL AML (Salzer et al., 2012). Therefore, further intensifying induction therapy to reduce rates of relapse, represents a high cost and low reward scenario (Hilden et al., 2006) and may not be the ideal solution. The alternative, bone marrow transplantation (allogenic hematopoietic stem cell transplantation; HSCT), has been shown by several studies to generally not improve outcomes and confers

unnecessary risks to the patient (Dreyer et al., 2011; Pieters et al., 2007; Pui et al., 2002). Maintenance therapy during the first complete remission (CR1) remains the preferred treatment approach, except in a subset of high-risk patients that may benefit for HSCT. The Interfant-99 trial showed patients diagnosed <6 months of age, and with either high WBC counts or the persistence of a small number of leukemia cells post-induction therapy (minimal residual disease), may benefit from allogenic HSCT during CR1 (Mann et al., 2010). Another recent approach to improve outcomes, has been to optimize the combination of agents within chemotherapy regimens. The Interfant-06 protocol hypothesized that infant *MLL-r* ALL may respond better to combinations of drugs typically used for AML such as cytarabine, daunorubicin/mitoxantrone and etoposide. This is because *MLL-r* ALL is thought to arise from or possess the characteristics of early progenitor cell populations with increased plasticity that are capable of myeloid differentiation (Pieters et al., 2019)

Both phenotype and the presence of *MLL-r* are large determinants of IL prognosis. For ALL, the prognosis for infants is worse than that of older children, with a 4-year event-free survival (EFS) of 47% compared to 85%, respectively (Pieters et al., 2007). One contributing factor is likely the presence of *MLL-r* in 70-80% of IL ALL patients but is only found in ~5% of older pediatric ALL patients (Behm et al., 1996). In contrast, outcomes for infant AML patients are similar to that of their older counterparts with 5-year EFS rates between 40-60% (Creutzig et al., 2012). Mirroring the relationship between phenotype and prognosis, the presence of *MLL*-r has a greater impact on infant ALL than AML. *MLL*-r are undoubtedly associated with poorer outcomes in infant ALL, with a 4-year EFS of 34% compared to 74% for *MLL*-wild-type patients (Pieters et al., 2007). For infants with AML, *MLL*-r do not greatly impact prognosis and have shown a slightly worse 5-year EFS (43%) compared to that of *MLL*-wild-type infant AML (52%) (Creutzig et al., 2012).

The treatment of infant leukemia has several challenges that are centered around the extremely young age of the patients and the aggressive characteristics of the disease. As a result, prognosis for IL is dismal, especially for patients with MLL-r who have worse outcomes than their MLL wild-type counterparts.

1.2 Mixed Lineage Leukemia (MLL)

1.2.1 Epigenetics

MLL (MLL1; KMT2A) regulates the expression of target genes by invoking epigenetic changes. Epigenetics is the process of modulating gene expression without altering the sequence of DNA that encodes the genetic information. Gene expression is regulated in three major ways: DNA methylation, histone modification, and through non-coding RNAs. The addition of epigenetic marks, such as the methylation of cytosine at CpG sites, directly onto DNA strands can modulate the binding of transcription factors and other enzymes to enhance promoter regions. In cancer, the hypermethylation of 5' regions, such as promoters and the first exons of tumor suppressors, is a central tumorigenic process that silences cancer-fighting genes (Esteller 2005). This phenomenon, coupled with global hypomethylation, contributes to drastic changes in gene expression that promote tumorigenesis (Hanahan & Weinberg, 2000). Histones have emerged as much more than a simple scaffold on which to package DNA, but instead are dynamic regulators of gene expression. Histone modifications, such as methylation, acetylation and ubiquitination of residues, influence which regions of DNA are tightly or loosely packaged within the structure of chromatin. Dense chromatin is made up of tightly packed and coiled DNA and does not allow transcription machinery access to genes for expression. In contrast, the most highly transcribed genes are found in regions of 'open' or loosely packaged chromatin. Lastly, non-coding RNAs modulate gene expression through several mechanisms at both the transcription and post-transcriptional levels (Wei et al., 2017). For example, micro RNAs (miRNA) and small interfering RNAs (siRNA) can interact with target mRNA through complimentary sequences, and block translation by initiating mRNA cleavage (Yekta et al., 2004). In some cases, miRNAs are also able to activate gene expression by blocking repressive proteins from binding to their gene targets (Eiring et al., 2010) (**Figure 1.2.1**).

Histone marks have 2 main categories: activators and repressors — referring to their effect on gene transcription — and can be represented by two antagonizing protein groups: Trithorax and Polycomb. These proteins are methyltransferases that add methyl groups to the side chains of lysine or arginine residues found within histone tails. Lysine

and arginine can both undergo mono- and di- methylation, while lysine can also undergo trimethylation (Cheung & So 2011; Labbé, Holowatyj & Yang, 2014). The Trithorax protein group (TrxG) methylates histone 3 lysine 4 (H3K4) residues and frequently results in upregulated gene expression (**Table 1.2.1**). Active and potentially active genes are generally associated with H3K4 dimethylation, while H3K4 trimethylation is predominantly a mark at the start site of transcription (Ng et al., 2003; Krogan et al., 2003). H3K36, and H3K79 methylation as well as the asymmetric dimethylation of histone 3 arginine 4 (H3R4) are also generally considered activator modifications (Di Lorenzo & Bedford, 2011; Huang & Zhu, 2018) (**Table 1.2.1**). In contrast, Polycomb protein groups (PcG) are responsible for H3K27 methylation — a mark commonly associated with the repression of gene expression (**Table 1.2.1**).

While histone methylation can have either activating or repressing effects on gene expression, histone acetylation only activates gene expression (reviewed by Rice & Allis, 2001). Histone acetyltransferases (HATs) are responsible for the addition of acetyl groups to histone lysine residues such as H3K9 and H3K27, while histone deacetylases (HDACs) remove histone lysine acetylation. Interestingly histone acetylation frequently opposes the role of histone methylation, for example at H3K9 and H3K27 where acetylation activates gene expression and methylation represses gene expression.

In addition to modulated gene expression, PcG and TrxG enable cells to uphold developmental states and lineage identity through the inheritability of transcriptional patterns (Ringrose & Paro, 2004). These marks enable cells to maintain a programmed developmental pathway, determined during early embryogenesis or early hematopoiesis, through the many mitotic events needed to reach a mature state. TrxG and PcG proteins are important for the regulation and maintenance of *Homeobox* (*HOX)* gene expression (Ringrose & Paro, 2004). *HOX* genes are central to developmental processes like embryogenesis, organogenesis, body planning, segmentation, and hematopoiesis (Abramovich & Humphries, 2005). Humans have 6 Trx-related proteins involved in H3K4 methylation through the catalytic activity of their SET (Su(var)3-9, enhancer of zeste, trithorax) domains: *SET1A*, *SET1B*, *MLL1*, *KMT2B*, *KMT2C*, and *KMT2D* (Shilatifard, 2008).

Table 1.2.1: List of Histone Methylation and Effects on Gene Expression

Histone Mark	Effect on Gene Expression
H3K4 me2/3	Activation
H3K9 me2/3	Repression
H3K27 me2/3	Repression
H3K36 me2/3	Activation
H3K79me2a	Activation

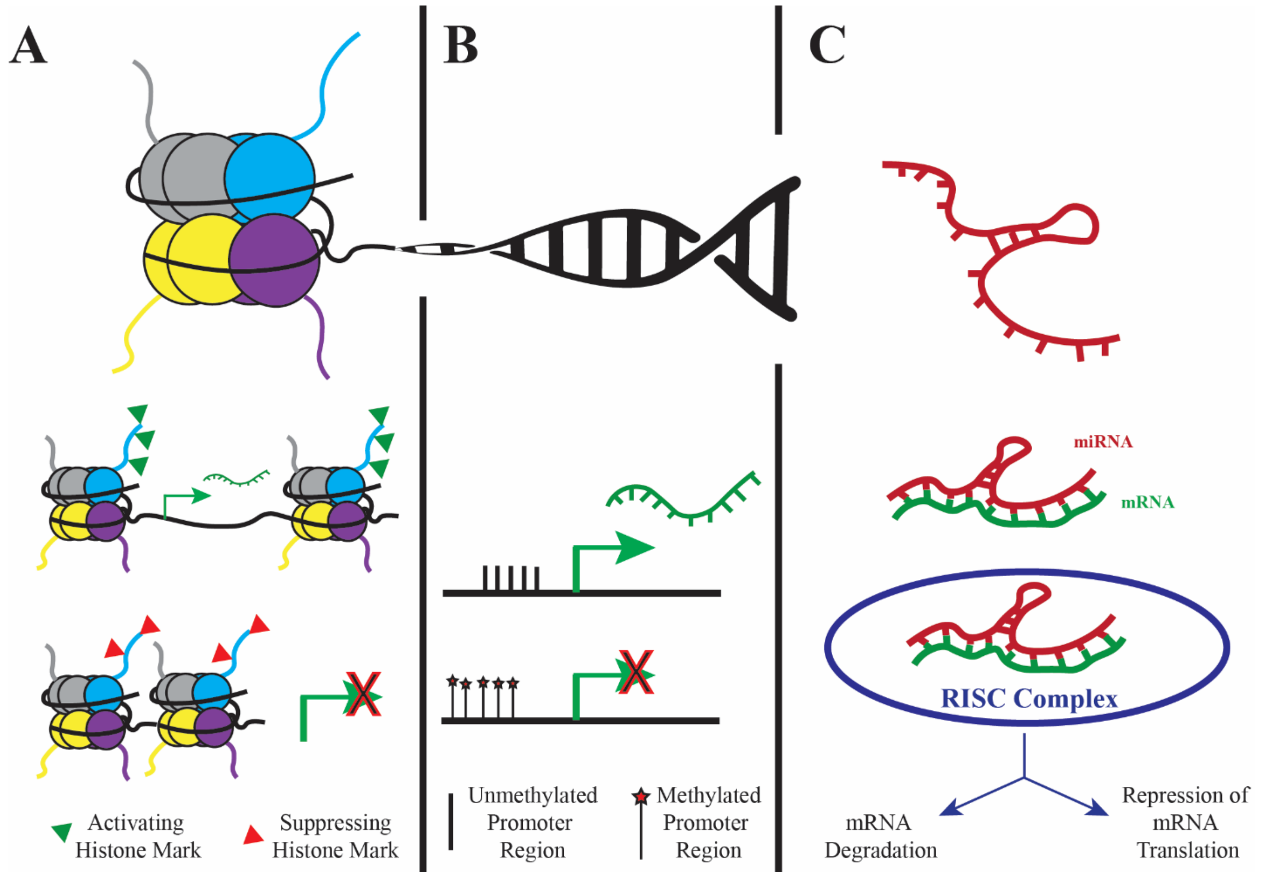


Figure 1.2.1: An overview of the 3 types of epigenetic regulation. A) Histone modification – histone residues can me modified by the addition or removal of methyl and acetyl groups, impacting the structure of chromatin. Activating marks, cause chromatin to become less tightly packed and allowing transcription factors to bind and transcribe genes. Suppressing histone marks cause histones to become tightly packed and repress gene expression. **B)** The methylation of DNA promoter regions can suppress gene expression. **C)** Micro RNAs (miRNA) can bind complementary mRNA, is processed through the RISC (RNA-Induced Silencing Complex) and will either result in mRNA degradation or the repression of mRNA translation.

1.2.2 *Mixed Lineage Leukemia (MLL)*

The Mixed Lineage Leukemia family (now known as — Lysine Methyltransferase 2; KMT2; *MLL*; *HRX*; *ALL-1*), is a family that contains 4 H3K4 histone methylases. MLL (MLL1), found at chromosome 11q23, is involved in three main aspects of histone biology: methylation, acetylation, and chromatin remodelling, but also impacts processes like DNA methylation and cell cycle progression through interactions with COMPASS (Complex Of Proteins ASsociated with Set1) proteins. Examining the many auxiliary ways, in addition to its primary role as an H3K4 methylator, in which *MLL* gene activity impacts essential cellular processes gives context to why MLL is referred to as a 'master epigenetic regulator'.

The SET domain of MLL is the major catalytic domain responsible for the mono-methylation of H3K4 and is located at the C-terminus of the gene (Milne et al., 2002). For example, WDR5 associates with MLL through its Win motif and recruits ASH2L/RBBP6 proteins to form a complex capable of catalyzing the di – and tri-methylation of H3K4 (Cao et al., 2010; Patel et al., 2009) (**Figure 1.2.2**). WDR5 plays a crucial role within the epigenetic regulation of MLL targeted *HOX* genes. *HOX* gene expression is decreased when WDR5 is knocked down *in vitro* and *in vivo*, which caused developmental abnormalities in chickens and tadpoles (Wysocka et al., 2005; Zhu et al., 2010). Interestingly, the knockdown of WDR5 decreased global levels of H3K4 tri-methylation (H3K4me3) but did not affect the binding of the MLL complex to the *HOX* promoters (Wysocka et al., 2005) and similar results have been shown for the knockdown of RBBP5 or ASH2, which also reduce *HOX* expression (Dou et al., 2005; Dou et al., 2006). These findings highlight the important role of WDR5, RBBP5 and ASH2 in the regulation of *HOX* expression through H3K4 di and tri-methylation (Steward et al., 2006).

Through interactions with COMPASS (COMplex of Proteins Associated with Set1) proteins, MLL evokes epigenetic changes beyond just H3K4 methylation, including the methylation of DNA. The protein has two DNA binding domains that help MLL find target genes. The first is a set of 3 AT hook motifs that bind AT rich DNA indiscriminately by interacting with the minor groove of the DNA double helix when orientated in certain geometries (Aravind & Landsman, 1998; Zeleznik-Le et al., 1994). The second is a CXXC

zinc finger domain homologous to cytosine DNA methyltransferase 1 (DNMT1) and is termed a transcriptional repression motif since it binds un-methylated CpG di-nucleotides preventing methylation (Birke, 2002). Through these domains MLL can target and bind DNA, blocking the methylation of and unhinging transcriptional regulation at important loci like *HOXA9* (Risner et al., 2013).

Plant homology domains (PHD) zinc fingers are important readers of H3K4 methylation. MLL has a cluster of PHD domains located downstream from its DNMT homology region, that are responsible for the reading of H3K4 methylation status and initiating the catalytic activity at the c-terminal SET domain (Baker et al., 2008). In addition to "reading" and facilitating the writing of H3K4 methylation, PHDs can recruit and bind effectors that activate gene expression, like CREB binding proteins (CPB, CREB = cAMP response element binding protein) or suppress gene expression, like prolyl-isomerase cyclophilin33 (CYP33), through histone acetylation (Ernst et al., 2001). CYP33 interacts with the zinc finger PHD-3 that is near the CXXC domain within MLL, termed the 'repression domain'. The MLL-CYP33 interaction induces a conformational change that recruits histone deacetylase 1 (HDAC1) as well as the Polycomb proteins (BMI1, CtBP and HPC2) to the RD1/2 domain of MLL and together function to repress the transcription of MLL target genes (Fair et al., 2001; Wang et al., 2010; Xia et al., 2003). To this end, the overexpression of CYP33 increases the association of MLL PHD-3, and its CXXC extension, with HDACs to repress the transcription of *HOX* genes, *in vitro* (Anderson et al., 2002; Fair et al., 2001).

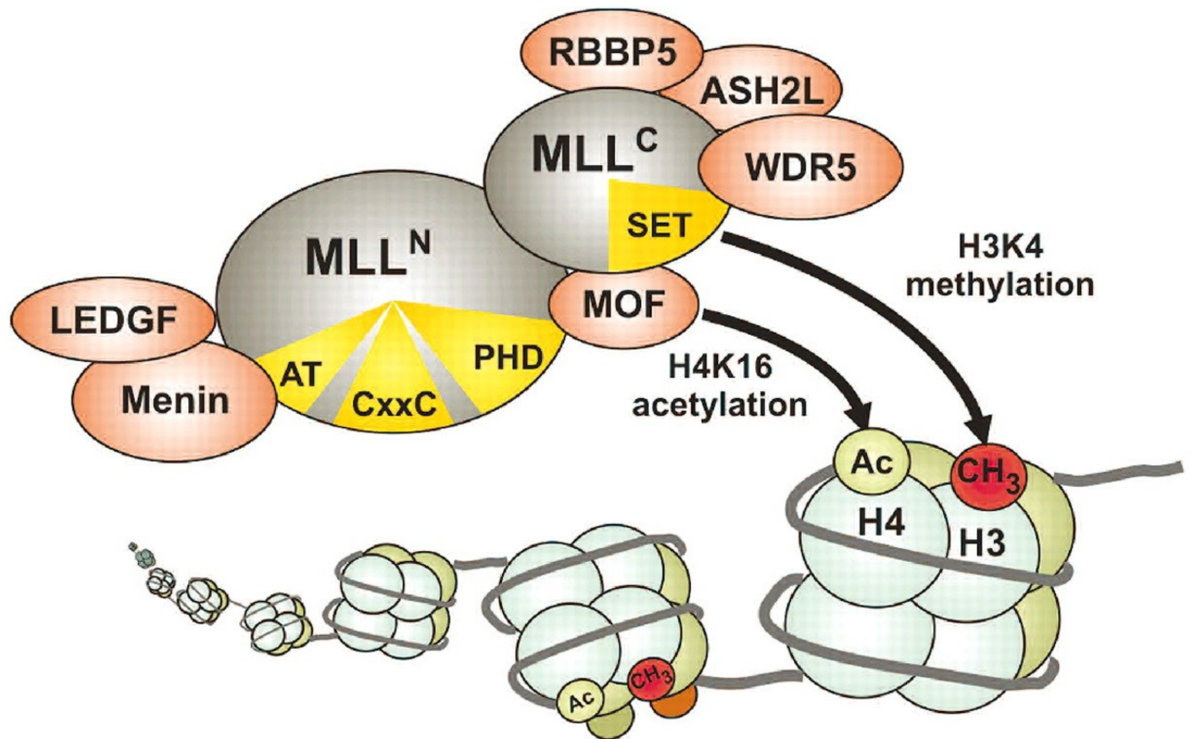


Figure 1.2.2: MLL interacts with a complex of proteins. After post-transcriptional proteolytic processing amino-terminal and carboxy-terminal portions of MLL are incorporated in a macromolecular complex with histone methyltransferase and histone acetyltransferase function. Functional domains in MLL are indicated in yellow. AT = AT-hooks, a DNA binding domain, CxxC = motif recognizing unmethylated CpG dinucleotides, PHD = plant homeodomain, SET = histone methyltransferase active site. Proteins associated with MLL are explained in the text. Reproduced with permission (Slany, 2009).

Bromodomains and extra terminal (BET) motifs within MLL recognize and bind acetylated lysine residues on histone tails through association with bromodomain (BRD) proteins. BRD proteins facilitate transcription through histone remodelling (Taverna et al., 2007), and been implicated in leukemia. Specifically, BRD4 has been shown to be an important player in the maintenance of AML (Dawson et al., 2011; Dawson et al., 2014; Delmore et al., 2011). Interestingly within AML, BRD4 is not typically mutated or overexpressed and the mechanism by which BDR4 targets promoter or enhancer regions is unknown (Mertz et al., 2011). However, BRD4 has been linked to the expression of several proto-oncoproteins, such as MYC, BCL2 and CDK6 (Delmore et al., 2011), making the inhibition of BET motif containing proteins a possible therapeutic option for AML.

MLL also plays a role in cell cycle regulation. MLL is cleaved by taspase-1 (threonine aspartase 1) at a single cleavage site to produce N- and C-terminal fragments that heterodimerize through interactions between PHD domains 1 and 4 as well as the FYRC and FYRN (Phenylalanine- and Tyrosine-rich N- and C-terminal) domains. This is the first step in MLL maturation, and the resulting heterodimer constitutes the active form of MLL (Yokoyama et al., 2013). Mutations in taspase-1 that prevent the cleavage and heterodimerization of MLL, result in the down regulation of Cyclin E, A and B and concurrently upregulate the activity of S-phase inhibitor p16 which decreases cell cycling and proliferation (Takeda et al., 2006). MLL normally binds the promoters of cyclins E1/E2, but within taspase-1 negative cells there is a reduction of MLL association with these promoters and decreased H3K4 trimethylation. This can account for the decreased expression of cyclins E1/E2. Additionally, MLL interacts with E2F transcription factors, to help regulate the cell cycle (Takeda et al., 2006).

Furthermore, MLL associates with transcriptionally active chromatin in G1 but dissociates from condensed chromatin during mitosis and migrates to the cytoplasm before it returns to the nucleus at the end of telophase. MLL stays bound to chromatin during mitosis to regulate cell-cycle related genes and the absence of MLL results in cell cycle arrest at G2/M which ultimately inhibits cell growth (Mishra et al., 2009). Chromatin-bound MLL during cell division contributes to the ability of MLL to produce inheritable epigenetic modification and cellular memory. To this end, global levels of H3K4

methylation are not affected during cell cycle progression (Mishra et al., 2009), and the bi-phasic expression of MLL is thought to work as a universal cell-cycle clock that is ultimately disrupted by *MLL-r*.

1.2.3 *MLL, HOX* & Hematopoiesis

Blood is made up of many distinct functional cells that are being continuously produced throughout life. HSCs give rise to progenitor cells that are responsible for the production and maintenance of hematopoietic cells (Ogawa, 1993). This results in a hierarchical system of proliferation, differentiation, maturation and self-renewal that needs constant regulation. *HOX* genes are central to the proper development of human tissues, organs and other structures that comprise their bodies — including blood. Humans have 39 different *HOX* genes clustered into 4 groups A-D, which have a high degree of conserved identity, however only groups A-C are readily expressed in hematopoietic lineages (Argiropoulos & Humphries, 2007). *HOX* genes play a crucial role in establishing cellular identity and when deregulated, they play a critical role in oncogenesis, especially within hematopoietic stem and progenitor cells (Grier et al., 2005). *HOX* expression confers stem-like characteristics, such as self-renewal and proliferation and therefore *HOX* expression is highest in HSCs and multipotent progenitor cells (Forsberg et al., 2005; Phillips, et al., 2000). *HOX* expression within HSCs and progenitor populations is essential to the expansion and reconstitution of the hematopoietic system throughout the life of the organism (Kroon et al., 1998). However, for these precursor cells to differentiate into mature effector blood cell types, *HOX* expression needs to be down regulated and eventually terminated (Somervaille & Cleary, 2006). These principles have been demonstrated by several groups in murine models. For example, *Hoxa9* -/- mice had a 30-40% reduction in total leukocyte and lymphocyte counts – highlighting the role *Hox* genes play in maintaining progenitor populations (Lawrence et al., 1997). Similarly, the overexpression of *Hoxa9* in murine bone marrow led to enhanced HSC regeneration *in vivo*, increased numbers of granulocytes, but partially blocked lymphocyte differentiation (Thorsteinsdottir et al., 2002). The critical role of *HOX* regulators is evident and expression lapses can lead to abnormal blood production and disease. Deregulation that leads to continuous *HOX* expression, can block hematopoietic differentiation, leading to an ever-

expanding population of immature blood progenitors and a pre-leukemic proliferative state that can readily transform into acute leukemia (**Figure 1.2.3**) (Slany, 2009)

MLL is a master regulator of *HOX* genes and is an essential gene for embryonic development and hematopoiesis. *Mll* knockout studies in mice show that the initiation of stage specific *Hox* expression occurs but is not maintained during embryo development and ultimately results in embryonic lethality (Yu et. al, 1995; Yu et. al, 1998). Interestingly, when only the SET domain of *Mll* is knocked out, there is no embryonic lethality and mice are fertile but endure skeletal abnormalities exemplifying dysregulation of *Hox* expression (Terranova et al., 2006). The knockdown of Taspase-1 mRNA resulting in the loss of MLL cleavage and subsequent loss of function, also causes the loss of physiological *HOX* gene expression (Hsieh et al., 2003). Concerning blood development, MLL activates the expression of posterior *HOXA* genes within the hematopoietic cell lineages, specifically *HOXA9* (Jude et al., 2007). *Mll* -/- mice have decreased *Hoxa7*, *Hoxa9*, *Hoxa10* and *Hoxa4* expression and ultimately succumb to the fatal underdevelopment of yolk sac hematopoiesis and that in the fetal liver, establishing the necessity of MLL in the construction of hematopoietic tissues (Hess et al., 1997; Yagi et al., 1998). Furthermore, the overexpression of either *Hoxa9*, *Hoxa10*, *Hoxb3* or *Hoxb4* was adequate to rescue the hematopoietic-colony forming ability of *Mll* -/- embryonic bodies (Ernst et al., 2004). Changes in *Hox* gene expression have been linked to reduced H3K4 methylation and changes in DNA methylation at *Hox* loci, in *Mll* mutant mice (Terranova et al., 2006). This highlights the epigenetic regulation of *HOX* genes by MLL.

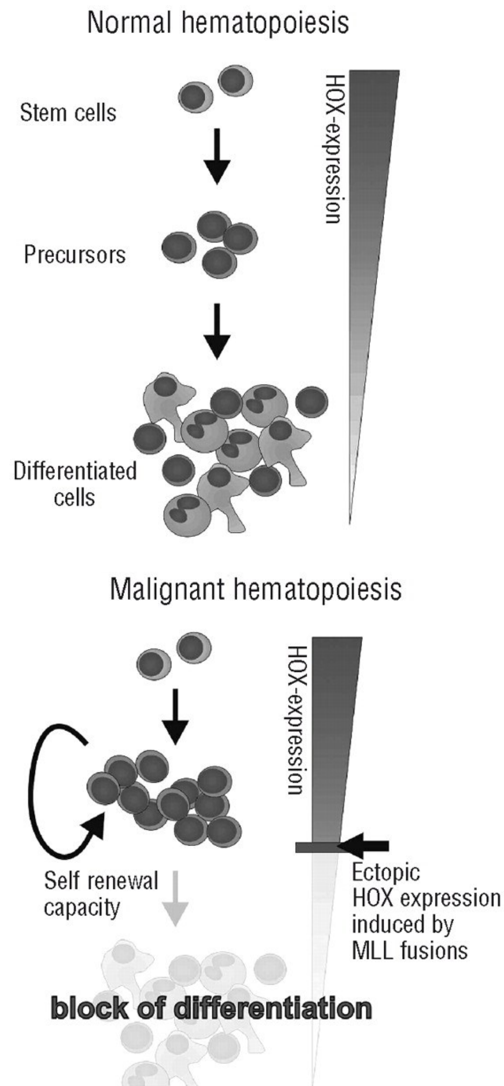


Figure 1.2.3: The role of HOX proteins in control of hematopoiesis. HOX transcription factors control hematopoietic differentiation. HOX expression must be terminated for maturation to occur, and therefore ectopic presence will block maturation and cause a population of self-renewing precursor cells to expand. Reproduced with permission. (Slany, 2009).

1.2.4 HOX cofactors: MENIN, LEDGF & MEIS1

Important cofactors for the activation of *HOX* genes by MLL are MENIN (MEN1) and MEIS1. *MEN1* is essential to embryonic development and has been implicated as a tumour suppressor gene, as the loss of *MEN1* in mice results in embryonic lethality whereas heterozygous *MEN1* mice develop a variety of endocrine tumours (Bertolino, Radovanovic, et al., 2003; Bertolino, Tong, et al., 2003). MENIN binds MLL through the RXRFP motif within the first 10 amino acids of MLL, and both MENIN and MLL associate with the *HOXA9* promoter to activate expression. Further investigations into the role of MENIN in *HOXA9* expression show the knockdown of *MEN1* in HeLa cells results in reduced *HOXA9* expression (Yokoyama et al., 2004), and that MENIN is a requirement for the majority of H3K4 methylation by endogenous MLL at *HOX* loci (Wang et al., 2009). Furthermore, the loss of *MENIN* phenocopies the loss of *MLL* in terms of *HOX* expression and cell cycle deregulation, confirming the essential role of the MENIN-MLL interaction (Milne et al., 2005; A Yokoyama et al., 2004)

MENIN also functions as a physical link, tethering MLL with lens epithelium derived growth factor (LEDGF) (Yokoyama & Cleary, 2008). LEDGF binds nucleosomes through the specific recognition of di- and trimethylated H3K36 marks that are relatively abundant modifications at promoter proximal regions (Ng et al., 2003). Furthermore, LEDGF can also bind DNA non-specifically (Eidahl et al., 2013; Botbol et al., 2007). H3K36me2 marks and un-methylated CpGs are generally enriched within the proximal promoter regions that are preferentially targeted by MLL and the Menin/MLL-LEDGF interaction is important for the association of MLL with target genes. (Kuo et al., 2011; Okuda et al., 2014). The majority of mice that lack LEDGF die perinatally and those that survive have craniofacial skeletal and developmental defects suggesting LEDGF has a contributing role in *HOX* expression (Sutherland et al., 2006). Conditional knockout mice that lack *Ledgf* within HSCs have WBC counts that are ~65% lower than their wild-type counterparts and show significant reductions in neutrophils and lymphocytes (El Ashkar et al., 2017). LEDGF has also been shown to colocalize at *HOXA* and *MEIS1* genes with MLL/MEN1 (Yokoyama & Cleary, 2008).

MEIS1 is a homeobox co-factor and dimerization partner of HOXA9. Along with *HOX* genes, *MEIS1* is expressed in early hematopoietic stem and progenitor cells and expression within these cells is decreased following differentiation (Imamura et al., 2002; Pineault et al., 2002). *MEIS1* is upregulated alongside *HOXA9* in leukemias with MLL-r, irrespective of fusion partner (Argiropoulos et decreased al., 2007). Knocking out *Meis1* in mice, results in extensive hemorrhaging and a failure to produce megakaryocytes that causes death soon after birth (Hisa et al., 2004). To further examine the role of *Meis1* in hematopoiesis, conditional knockout mice were generated that only lack *Meis1* within blood cells of adult mice (Ariki et al., 2014). These mice had supressed hematopoiesis due to a loss of HSCs and progenitor populations. HSPC populations continued to enter the cell cycle but were unable to undergo complete differentiation or maturation which lead to HSC exhaustion. Several MLL target *HOX* genes (HOXA10, HOXB7) are expressed differentially during cell cycle progression. For example, HOXA10 expression is significantly higher during S phase before decreasing throughout G2/Mitosis until it is absent in G1 (Mishra et al., 2009). When *Mll* deletions are introduced into myelo-erythroid progenitors *in vivo*, there is a reduction in proliferation and cell-cycle entry in response to cytokines and was linked to a decrease in expression of *Hox* genes (Jude et al., 2007). These findings highlight the role for *MEIS1* in the maintenance and self-renewal abilities of HSPCs, as and important cofactor for MLL mediated *HOX* expression.

These findings present compelling evidence that MLL drives the proliferation of hematopoietic stem and progenitor cells through the upregulation of *HOX* expression and interactions with MENIN, LEDGF, and MEIS1. Failure to inhibit *HOX* gene expression at the appropriate time point, can lead to the failure of differentiation of hematopoietic progenitor cells and lead to a pre-leukemic hyper-proliferative state. Considering the high expression of *MLL* in blood progenitor cells and its ability to expand these populations, it is evident that MLL plays a critical role in the architecture of the hematopoietic system (McMahon et al., 2007). This provides insight into the essential role of *MLL* and how aberrations to MLL cause leukemia in infants.

1.3 MLL-Fusions & Infant Leukemia (IL)

1.3.1 MLL-Fusions

One of the defining features of IL is the presence of *MLL-r* that result in the expression of *MLL*-fusion genes. Roughly 65-70% of IL patients have *MLL-r*, although *MLL-r* are more common in IL ALL (~70%) than IL AML (~50%) (Harrison et al., 2010; Hilden et al., 2006; Pieters et al., 2007). *MLL-r* are reciprocal translocations at the 11q23 locus that result in the fusion of *MLL* with a partner gene creating a fusion gene. *MLL*-fusions involve the first 8-11 exons of *MLL* being fused in frame with one of more than 90 identified fusion partners (Meyer et al., 2018). Interestingly, despite the promiscuous nature of *MLL*, only a select few partner genes make up the majority of IL *MLL-r* cases. In IL ALL, 93% of cases are made up by just 4 partner genes: *AF4* (49%), *ENL* (22%), *AF9* (17%), and *AF10* (5%) (Meyer et al., 2018). Similarly, *AF9* (22%), *AF10* (27%) and *ELL* (17%) make up 66% of IL AML cases (**Figure 1.3.1**) (Meyer et al., 2018). While most adult leukemias and other cancers are the result of many mutations that have been accumulated over time, the investigation of infant leukemia in twins and the retrospective analysis of neonatal blood spots (Guthrie cards) has revealed *MLL-r* occur *in utero*, likely contributing to the rapid onset of IL (Greaves, 2005). Identical clonotypic breakpoints in MLL have been found in twins with MLL-r ALL indicating a single clone of origin that must have spread through intraplacental vessels (Ford et al., 1993; Greaves, Wiemels & Ford, 2003).

MLL-r result in the loss of several important MLL domains, such as PHD fingers and the C-terminal SET domain, due to the location of its breakpoint cluster region — exons 9-11 (Gu et al., 1992; Meyer et al., 2018). Loss of the MLL-N/MLL-C interactions motifs signify the unlikely association of *MLL*-fusions with C-terminal MLL fragments. Evidence suggests fusion genes involving the C-terminal portion of MLL (Fusion-MLL) are able to express a 5' truncated MLL protein due to a gene internal promoter upstream of *MLL* exon 12, however, these are likely unimportant for leukemogenesis (Scharf et al., 2007). Expression of such 5' truncated MLL have been identified in *Mll(-/-)* mice and 5' truncated MLL proteins are able to associate with C-terminal MLL fragments, although the mutation is still embryonically lethal (Yu et al., 1995). This suggests that 5' MLL domains such as the AT Hook and PHD domains 1-3 are essential to the function of MLL (Yu et al., 1995).

N- and C-terminal fragments of MLL are individually unstable and thought to degrade if not stabilized through interactions with each other (Hsieh, Ernst et al., 2003). This has led to the hypothesis that fusion partners stabilize N-terminal MLL and may contribute to the large number of partner genes accepted by MLL. Importantly, *Mll(+/-)* mice do not develop leukemia, while knock-in models develop leukemia if *Mll* is fused to a partner gene but not if *Mll* is truncated (Corral et al., 1996; Dobson et al., 1999; Forster et al., 2003). Similar retroviral transduction experiments show that only *MLL*-fusions and not individual *MLL* N-terminal or C-terminal fragments can impact hematopoietic progenitor cells (DiMartino et al., 2000; Lavau et al., 1997, 2000; Slany et al., 1998). From these experiments, haploinsufficiency can be excluded from the mechanisms by which MLL aberrations cause leukemia.

MLL-fusions convert MLL into a potent epigenetic mediated transcriptional activator through (i) the fusion with partner genes that contain activation domains such as AF9, ENL, and AF4 (Zeisig et al., 2003), (ii) fusion directly to transcriptional co-activators (as is the case with MLL-CBP and MLL-p300) (Ida et al., 1997; Taki et al., 2018), and (iii) fusion to partners that recruit activators like ELL or ABI1 (García-Cuéllar et al., 2000; Simone et al., 2001). Like wild-type *MLL*, *MLL*-fusions affect gene expression through several different mechanisms and pathways beyond H3K4 methylation. To this end, Erfurth and colleagues have proposed the “MLL Web Hypothesis” to reconcile the fact that although some partner genes are structurally similar and can be grouped together (like *AF9* and *ENL*), there is a lack of common functional domains among all fusion partner genes (Erfurth et al., 2004). Instead, the authors suggest that the structure of the fusion partner is less important than its contribution to the protein super-complex with which MLL-fusions operate to perform transcriptional activation and chromatin modification (Erfurth et al., 2004). The following section will focus on the major mechanisms by which fusion partners *AF9* and *ENL* convert *MLL* into a potent oncogene.

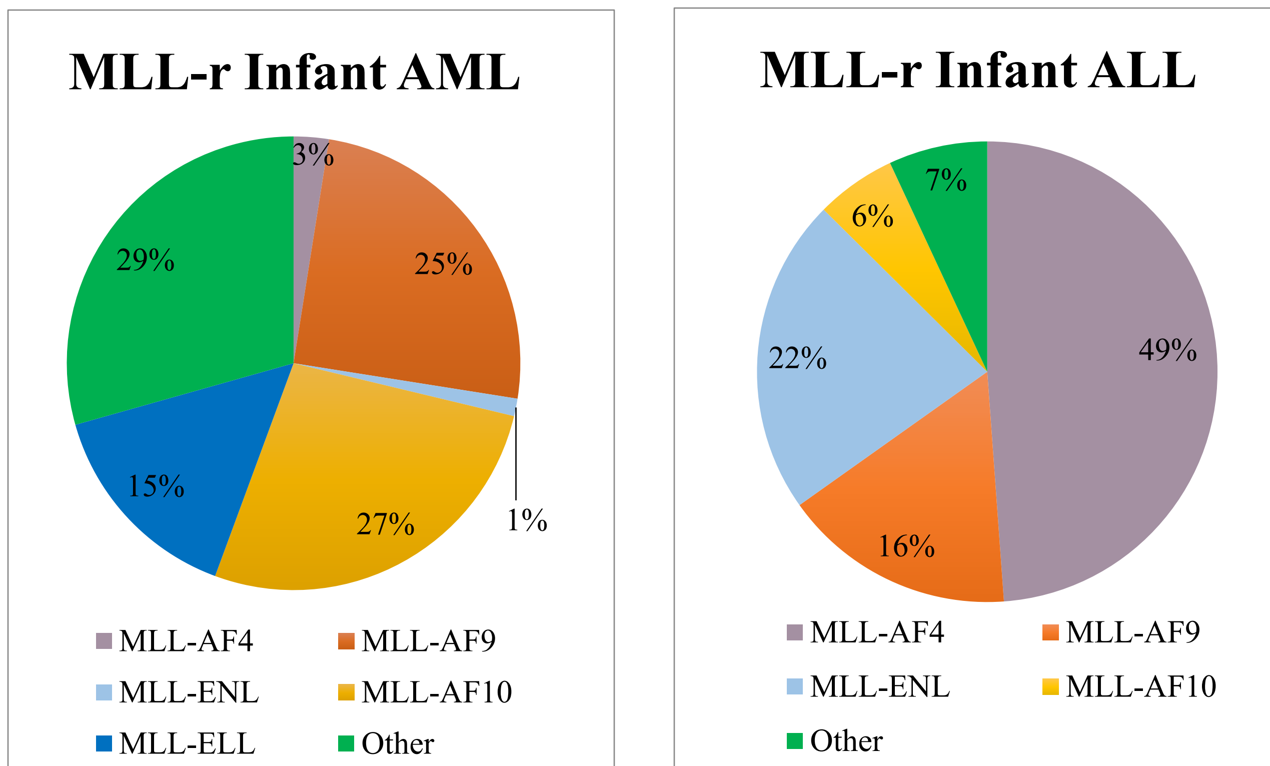


Figure 1.3.1: Only 6 different fusion partners make up the majority of MLL-fusion genes, despite the characterization of >90 different fusion partners. The pie charts represent the relative frequencies of different partner genes in infant ALL and infant AML, respectively. Figure produced using data from Meyer et al., 2018.

1.3.2 MLL-AF9 & MLL-ENL

ALL1-fused gene from chromosome 9 (*AF9*) and *eleven-nineteen leukemia* (*ENL*) are the second and third most common MLL fusion partners, respectively, with *AF4* being the most common (Brown, 2013). Here we will focus on the on the mechanisms of action of MLL-AF9 and MLL-ENL, as they are most relevant to this study.

AF9 and ENL are share significant homology and both contain YEAT domains important for histone reading and transcriptional regulation (Muntean & Hess, 2012). In the context of normal development, Af9 is important for proper skeletal development and body planning (Collins, et al., 2002). AF9 is highly expressed in the blood, where it is a regulator of erythroid and megakaryocyte lineages and for the maintenance of HSCs (Prina et al., 2008; Calvanese et al., 2019), as well as in the brain, where it promotes neural differentiation and development of the cerebral cortex in embryonic mice (Collins et al., 2002; Qiao et al., 2015; Vogel et al., 2009). Mutations to *AF9* have been implicated in neurological diseases, such as epilepsy and ataxia (Pramparo et al., 2005; Striano et al., 2005). ENL is a reader of histone acetylation and an important RNA POL II elongation factor (Muntean & Hess, 2012). ENL directly binds DOT1L and promotes gene expression (Bitoun et al., 2007), and both ENL and AF9 are key factors in the SEC (Super Elongation Complex) which is responsible for the elongation phase of gene transcription by RNA POL II (Mueller et al., 2007). Changes in SEC activity can result in aberrant gene expression, and deregulation of the SEC by MLL-fusions leads to abnormal *HOX* gene expression (Lin et al., 2010; Yokoyama et al., 2010). There has been limited investigation into the physiological role of ENL in development, but homozygous *Enl* mutations have proven embryonically lethal, implicating the gene as essential (Doty et al., 2002). Mutations to the YEATS domain of *ENL* have been implicated in Wilms' tumours, an embryonal neoplasm of the kidney, and render ENL unable to bind acetylated H3K9 residues and results in deregulated *HOX* gene expression and the upregulation of *MYC* (Perlman et al., 2015)

Despite AF9 and ENL having largely similar structures and protein interactions, MLL-AF9 *t(9,11)(p22,q23)* is much more prominent in infant AML where as MLL-ENL *t(11,19)(q23p13.3)* is common in both infant myeloid and lymphoid leukemia (Brown, 2013). AF9 and ENL contain important domains for gene transactivation and AF4 binding

— both of which are important for leukemic transformation by these fusions (Bitoun et al., 2007; Slany et al., 1998). These fusions bind AF4 and/or recruit DOT1L (for H3K79 methylation —discussed below) through the ANC1 homology domain (AHD) as well as interact with transcription factor, pTEFb (Mueller et al., 2007). pTEFb is a cyclin dependent kinase (CDK) that promotes efficient transcriptional elongation through the phosphorylation of RNA POL II within the SEC (Peterlin & Price, 2006). Additionally, MLL-AF9 and MLL-ENL recruit histone acetyltransferase Tip60 through interactions with CBX8, a component of Polycomb repressor complex 1 (PRC1), to promote activation of MLL-fusion targets (Mueller et al., 2007; Tan et al., 2011).

H3K16 histone acetyltransferase *Mof* (*Kat8*) is critical for MLL-AF9 leukemia growth *in vivo* and homozygous *Mof* loss resulted in reduced colony forming ability, reduced tumour burden and prolonged survival in mice (Valerio et al., 2017). Interestingly the loss of *Mof* did not result in the downregulation of *Meis1* or *Hoxa* cluster genes but rather downregulated DNA damage repair pathway genes, and increased γ-H2AX staining in *Mof -/- Mll-Af9* cells, indicating an increased sensitivity to DNA damage likely causing impaired growth (Valerio et al., 2017). The homozygous loss of *Mof*, resulted in the dramatic decrease of global H3K16 acetylation levels implicating HAT inhibitors and HDACs is potential therapies (Valerio et al., 2017).

MLL-ENL is involved in the direct recruitment of the SWI/SNF nucleosome remodelling complex, which results in the aberrant activation and transcription of HOXA7 (Nie et al., 2003). The SWI/SNF complex disrupts nucleosome structure by using ATP to slide nucleosomes along the DNA helix exposing targeted DNA to transcriptions factors that can ultimately increase gene expression.

1.3.3 MLL-Fusions bind MENIN & MEIS1

The MLL fusion complex acquires many 'gain of function' mechanisms through its vast catalogue of fusion partner genes. However, MLL fusions still activate the same genes and transcriptional programs regulated by wild-type MLL, such as *HOXA9* (Wang et al., 2011). Several studies have also shown that MLL-fusion proteins and WT MLL regulate a

common set of HSC program genes *in vivo* (Artinger & Ernst, 2013; Jude et al., 2007; Yagi et al., 1998). Although fusion partners provide several novel oncogenic mechanisms, the oncogenic potential of MLL-r function through established pathways important to wild-type MLL. The MLL-HOX regulatory pathway and important players Meis1 and Menin, remain ever important as their presence is consistent in a wide scope of fusion partners.

MLL-fusions retain important domains such as AT hooks, zinc finger CXXC motifs and MEIS1- and MENIN-interacting domains that are essential to the oncogenic success of *MLL*- fusions (Slany, Lavau & Clearly, 1998). The inhibition of the CXXC domain severely reduces the ability of MLL-ENL retrovirally transduced cells to grow in colony forming assays (Slany, Lavau & Clearly, 1998). Similarly, *Meis1*-/- murine fetal liver cells are incapable of transformation by the retroviral transduction of *MLL*-fusion genes (Wong et al., 2007), and in the absence of Meis1, MLL-fusions fail to up regulate expression of *Hoxa9* and therefore limit the oncogenic potential of MLL-fusions (Kumar et al., 2009; Wong et al., 2007). Mll-AF9 knock-in mice develop AML that closely resembles the human disease (Kumar et al., 2009). Cell lines derived from these knock-in mice show high levels of *Meis1* and are sensitive to *Meis1* knockdown, which resulted in cell cycle arrest and cell death (Kumar et al., 2009). These findings highlight not only the central role of MEIS1 in MLL-r leukemogenesis, but also the ability of MEIS1 to act as a rate limiting step. For instance, the constitutive expression of *Hoxa9* and *Meis1* in hematopoietic progenitors of primary bone marrow cells induces leukemia in mice within three months (Kroon et al., 1998). Conversely, the overexpression of either *Hoxa9* or *Meis1* alone, failed to transform the cells within six months (Kroon et al., 1998). When *MEIS1, HOXA7, HOXA9* and *HOXA10* genes, representing four of the most commonly upregulated genes in *MLL-r* leukemia, are knocked down separately within human B-cell leukemia cell line, RS4;11 (that is MLL-AF4 expressing), these mutant cell lines were unable to be engrafted into the bone marrow of NOD/SCID mice (Orlovsky et al., 2011)

Like MEIS1, interactions with MENIN are conserved in MLL-fusions and remain important to the sustained misexpression of *Hox* genes, which is central to MLL-r leukemogenesis (Caslini et al., 2007; Y. Chen et al., 2006; Yokoyama et al., 2005). Consistent with its normal physiological functions, MENIN binds MLL-fusions through a

conserved high-affinity menin binding motif (hMBM) and associates with *Hox* loci (Yokoyama et al., 2005). However, unlike MEIS1, whether MENIN is essential to the oncogenic abilities of MLL-fusions is still up for debate and has led to postulation of MENIN-dependent and MENIN-independent pathways. Investigation into the conserved interactions of MENIN and MLL-r has suggested that MENIN is in fact an essential cofactor. Such evidence includes the limited oncogenic ability of MLL-ENL to transform myeloid progenitor cells, and the reduced expression of *Hoxa* genes when the hMBM is mutated (Yokoyama et al., 2005). When Menin is removed from MLL-r transformed leukemia blasts, these cells show signs of maturation and decreased cell cycling, indicating that Menin is critical for the maintenance of stem-like characteristics and in its absence, the block in hematopoietic differentiation is lifted. The same group looked at the conserved role of LEDGF in MLL-r oncogenesis and demonstrated that MLL-ENL is also reliant on LEDGF (Yokoyama & Cleary, 2008). LEDGF binds MLL-fusions through conserved residues that make up the LEDGF binding domain (LBD), which is distinct from the hMBM. Mutations to the LBD rendered MLL-ENL incapable of transforming myeloid progenitor cells and the knockdown of *Ledgf* in MLL-ENL transformed cells, impaired *Hoxa9* expression, proving LEDGF to be essential to the transforming ability of MLL-fusions and the misregulation of *Hox* genes. These findings emphasize the conserved MLL-fusion-MENIN-LEDGF interaction. However, one study has shown that N-terminal MLL fragments that lack MENIN and LEDGF interaction remain able to target and associate with *Hoxa9 in vivo* (Milne et al., 2010). Not much is known mechanistically regarding this MENIN-independent pathway, but it is thought that *Hox* targeting likely relies on the H3K4me2/3 binding of PHD finger 3 (retained in MLL-r) (Milne et al., 2010).

1.3.4 MLL-Fusions Recruit DOT1L

Since the SET domain, responsible for the H3K4 methylation potential of MLL, is lost during the formation of MLL-r, it is unclear how the oncogenic activity of MLL-fusions relates to H3K4 methylation. Rather, the involvement of MLL-fusions with H3K36 methylation are facilitated through interactions with MEN1 and LEDGF. Similarly, H3K79 methylation occurs through the recruitment and binding of H3K79 methyltransferase

DOT1L. H3K79 methylation has been linked to the positive regulation of transcription through genome wide analysis studies (**Table 1.2.1**) (Guenther et al., 2007; Lachner et al., 2003), suggesting that the loss of H3K4 methylation capability is likely compensated for through the recruitment of DOT1L in MLL-fusions.

Prominent fusion proteins MLL-AF9 and MLL-AF4 bind DOT1L directly (Bitoun et al., 2007; Zhang et al., 2006). In cases where the MLL-fusion does not bind DOT1L, such as MLL-ENL, it has been found that MLL-ENL bind AF4 and AF10, both of which subsequently bind DOT1L (Zeisig et al., 2005). These findings indicate where direct binding through the MLL-fusions is unavailable, MLL-fusions are still able to exploit H3K79 methylation (an activating epigenetic mark) by binding DOT1L through other common binding proteins. MLL-fusion interactions with DOT1L prompted the investigation into the H3K79 methylation of known MLL target genes. Several groups have shown that over-expression of MLL fusions in leukemia cells enhances H3K79 methylation at genes typically upregulated in MLL-r leukemia, specifically at Hoxa and Meis1 loci (Guenther et al., 2008; Lachner et al., 2003). Similarly, suppression of Dot1L decreases the expression of *Hox* genes and Meis1 and murine bone marrow cells fail to be transformed by MLL-AF9 (Nguyen et al., 2011). These findings highlight the indispensable nature of DOT1L in MLL leukemogenesis. A correlation between H3K79me2 and MLL target genes has been shown and identified as a special epigenetic lesion in MLL leukemia (Guenther et al., 2008). Genome-wide analysis of MLL-r leukemia samples further identified unique transcriptional and H3K79 methylation profiles that can discriminate between MLL+ and MLL-germline leukemia (Krivtsov et al., 2008). The direct fusion of MLL to DOT1L is sufficient to activate and increase expression of *Hoxa* cluster genes (Okada et al., 2005). These findings implicate DOT1L as an essential player in MLL-r leukemia. (**Figure 1.3.4**)

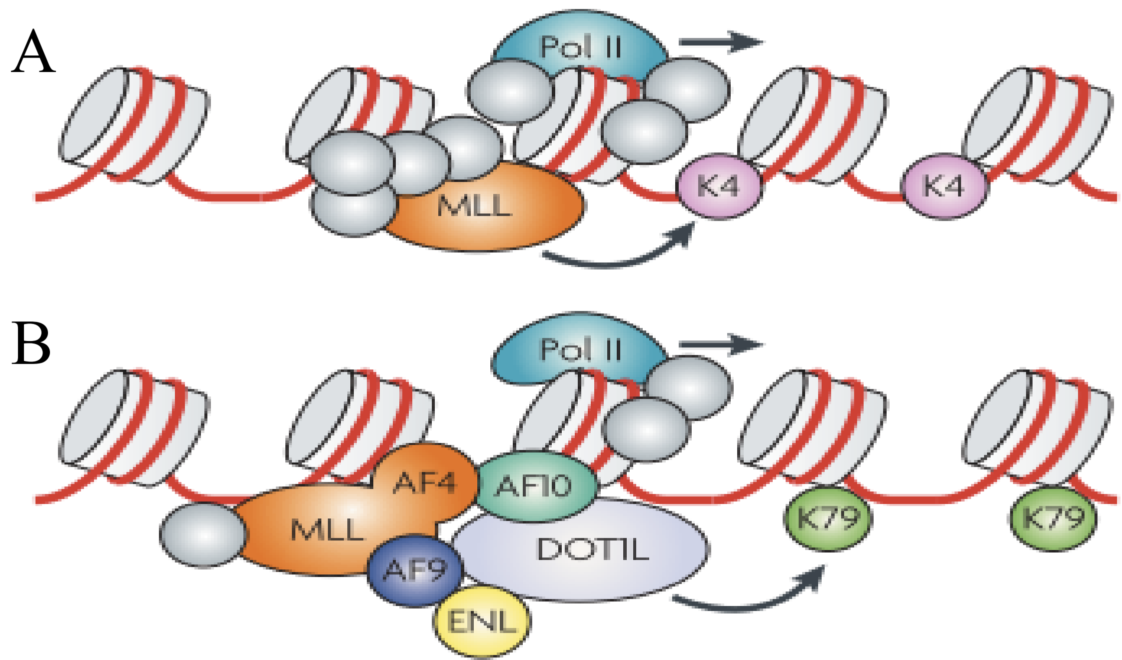


Figure 1.3.4: MLL-Fusions recruit DOT1L to activate gene expression through histone lysine methylation in the absence of a SET domain. A) Mixed lineage leukaemia (MLL) is a member of a multiprotein complex that mediates methylation of H3K4 within the promoter region of genes occupied by RNA polymerase II. **B)** MLL fusion(s) that lack the SET (Su(var)3-9, enchancer-of-zeste, trithorax) domain H3K4 methyltransferase activity may recruit the H3K79 methyltransferase DOT1L. MLL-fusion-mediated recruitment of DOT1L to promoters normally occupied by MLL, (such as the HoxA cluster) allows H3K79 methylation of the HoxA cluster, which may lead to aberrant expression of HoxA cluster genes. Reproduced with permission. (Krivtsov & Armstrong, 2007)

1.3.5 MLL-Fusions Activate HOX Expression

Ultimately, *MLL*-fusion gene expression upregulates *HOX* genes that prevent hematopoietic progenitor cells from differentiating. This leads to an ever-expanding population of immature blood progenitor cells and a pre-leukemic proliferative state that can readily transform into acute leukemia (Drynan et al., 2005). Numerous studies have outlined the key roles MEN1, MEIS1 and DOT1L play in the up regulation of *HOX* genes by MLL-fusions. Therefore, it is no surprise that of the small molecule inhibitors currently in development, MENIN and DOT1L are the most promising therapeutic targets. Leukemia, like other cancers, usually takes a long time develop and is the result of several genetic mutations that accumulate over time. In contrast, evidence shows that MLL-fusions originate *in utero* and are potent oncogenes that work through a variety of mechanisms and induce a complex disease with rapid onset. Whether or not MLL-fusions are sufficient alone to induce such an aggressive and rapidly progressing disease, is still up for debate and has surfaced as a central issue in *MLL-r* animal modelling.

1.4 Germline KMT2C Mutations in IL

Mouse studies of IL have indicated that the expression of MLL-fusions alone are insufficient to recapitulate the rapid onset of the human disease — indicating that additional environmental and genetic factors contribute to disease progression. The need for additional genetic factors is highlighted by an MLL-ENL chicken model that only produced a leukemia phenotype when coupled with an activating kinase mutation (Schulte et al., 2002). However, both genome and exome sequencing of *MLL-r* IL patients found few somatic mutations (Andersson et al., 2015; Chang et al., 2013). By the same token, there have not been any strong links between exposure to environmental agents and IL — except topoisomerase II inhibitors (Aplan et al., 1996; Ross et al., 1994). Topoisomerase II inhibitors are known to be a cause of *MLL-r* in treatment related AML (Super et al., 1993), and it has been suggested that maternal exposure during pregnancy contributes to development of infant AML (Ross, 1998; Ross, 2008). Further investigation revealed a statistically significant association between maternal exposure to topoisomerase II inhibitors and *MLL-r* infant AML but not *MLL-wt* infant AML or infant ALL (Spector et al., 2005). Given this history, Druley and colleagues set out to test whether IL occurred in

genetic backgrounds containing functional mutations to known leukemia-associated genes. The germline exome sequencing of non-cancerous samples from IL patients and their mothers found IL patients harbor significant enrichment of rare, non-synonymous germline variations (Valentine et al., 2014). Missense mutations in *KMT2C* (*MLL3*) were identified as a top candidate gene due to known associations of *KMT2C* mutations in myelodysplastic syndrome (MDS) and AML (Meeks & Shilatifard, 2017). Building on this initial finding, the Druley lab compared 115 germline IL patient and control exome sequences to publicly available exome data and found every IL patient possessed non-synonymous germline variation in *MLL/KMT2* genes (such as KMT2C) and their associated COMPASS proteins, regardless of whether they possessed an *MLL-r* or not (unpublished data).

KMT2C (MLL3) is a H3K4 methyltransferase and is part of the KMT2/MLL family but is not to be confused with MLL1/KMT2A which readily forms *MLL*-fusions in IL. KMT2C confers a positive epigenetic mark to regulate gene expression and interacts with COMPASS complex proteins (Hu et al., 2013). Mice with homozygous in frame deletions in the SET domain of *KMT2C*, show partial embryonic lethality, stunted growth, decreased cellular doubling rate and low fertility, suggesting that KMT2C function is essential to proper embryogenesis and development (Arcipowski et al., 2016; Lee et al., 2008). Beyond H3K4 methylation, KMT2C is involved in nuclear receptor signalling and transcriptional activation through the activating signal cointegrator-2 (ASC-2) complex, and is a co-activator of p53 target genes such as p21 (Lee et al., 2008; Lee et al., 2009). KMT2C also acts in the cellular response to DNA damage and DNA repair through mechanisms like homology directed repair (HDR) (Rampias et al., 2019). Decreased expression of genes *ATM* (*ataxia-telangiectasia mutated*), *ATR* (*ATM- and Rad3-Related*), *CHEK2* (*checkpoint kinase 1*), *BRCA1* (*breast cancer 1*) and *RAD51* was observed when *KMT2C* was knocked down in a variety of epithelial cancer cell lines and in primary tumour samples that contained mutant *KMT2C*. *KMT2C* knockdown epithelial cancer cell lines also showed increased DNA damage, chromosomal instability and a notable dependence on alternative end-joining (Alt-EJ) pathways for the repair of double-stranded DNA breaks (Rampias et al., 2019). PARP1/2 inhibition of *KMT2C* knockdown cells resulted in decreased cell viability *in vitro*, and decreased proliferation and tumour size *in vivo* (Rampias et al., 2019).

Mutations to *KMT2C* have been implicated in blood disorders such as MDS and AML, through the loss of chromosomal arm 7q, as well as several non-hematological malignancies (Chen et al., 2014). Missense and frameshift mutations to KMT2C that likely cause loss of function, are observed (14-15%) in non-invasive bladder and colorectal cancers (Hurst et al., 2017; Watanabe et al., 2011). Mutations have also been implicated in liver, pancreatic, breast and ovarian cancers, as well as osteosarcomas (Biankin et al., 2012; Fujimoto et al., 2012; Gala et al., 2018; Hurst et al., 2017; Kanchi et al., 2014). Therefore, KMT2C is considered a pan-cancer tumour suppressor, however the prognostic implications are not yet understood. For example, reduced *KMT2C* expression in pancreatic ductal carcinoma correlates to improved outcomes, whereas ER+ breast cancer patients tend to have shorter progression-free survival if they have a deleted or truncated Kmt2c (Dawkins et al., 2016; Gala et al., 2018).

Limited investigation has been conducted into the roles of KMT2C in blood development and leukemia. Germline mutations to the SET domain of *Kmt2c* in mice have been shown to promote a shift in hematopoiesis to favour granulocyte/myeloid progenitor cells (Arcipowski et al., 2016). A shift which mirrors the human transition from MDS to AML (Arcipowski et al., 2016). Additionally, the knockdown of *Kmt2c* and *Nf1* in p53-deficient mouse HSPCs led to the development of myeloid leukemia, and the single knockdown of *Kmt2c* in these cells led to increased proliferation and impaired differentiation (Chen et al., 2014). It is clear KMT2C plays a role in blood development and differentiation, however further investigation is needed. Elucidating the role of KMT2C in normal hematopoiesis may provide a better understanding of how *KMT2C* mutations influence the development of IL.

1.5 Current Clinical Trials & Drugs in Development

Overall, survival for IL has seen no major improvements within the last decade. This highlights an immense need for improved treatment options for IL with a focus on the inhibition of the major driving factor– MLL rearrangements – using targeted and epigenetic therapies.

Targeting the product of fusion genes in leukemia has shown great success in the past, as seen in the treatment of acute promyelocytic leukemia (APL) by targeting the driving PML-RARA fusion protein with all trans retinoic acid (ATRA), which boosted complete remission rates to over 90% (Jing, 2004). Similarly, the targeting of the BCR-ABL fusion in chronic myeloid leukemia (CML), by the small-molecule ABL kinase inhibitor, imatinib mesylate, has created a new treatment paradigm in this disease (Kalmanti et al., 2015). Although MLL-r have been well characterized, development of new therapies is an expensive process that can take decades to complete, and no drugs directly targeting MLL-fusions have been developed. There are however, several small molecule inhibitor therapies in development (Sun et al., 2018; Tsai & So, 2017) that target downstream effectors that are deregulated by MLL-fusion proteins, namely MENIN and DOT1L inhibitors, which have shown varying degrees of efficacy.

Two small molecule inhibitors of the MLL-MENIN interaction have been developed, MI-463 and MI-503 (Borkin et al., 2015; Daigle et al., 2013). Both these molecules have been shown to block the progression of MLL-r leukemia without impairing normal hematopoiesis in mice, representing a very attractive potential therapeutic for MLL-r leukemia patients. Additionally, DC_YM21 was developed through the modification of antidiarrheal loperamide (Imodium) to enhance the inhibition of MLL-MENIN interactions (Yue et al., 2016). Pre-clinical studies show DC_YM21 inhibits proliferation, induces cell cycle arrest and promotes differentiation when treating *MLL-r* leukemia cells.

EPZ4777 (Daigle et al., 2011) and EPZ5676 (pinometostat) (Stein et al., 2018) are HMT inhibitors that target DOT1L and have been shown to selectively inhibit H3K79 methylation within cells containing MLL-fusions. In murine models, the continuous infusion of DOT1L inhibitor EPZ4777 increased the overall survival (OS) of mice injected with the MV4-11 cell line expressing the MLL-AF4 fusion (Daigle et al., 2011). These findings led to the Phase I clinical trial of EPZ5676 (pinometostat), however only 3 of 51 patients finished the trial, with 41 patients discontinuing the study due to disease progression, lack of efficacy or adverse effects (Stein et al., 2018). The pharmokinetics of EPZ5676 (pinometostat) may limit its clinical development as demonstrated by its low oral bioavailability and relatively high clearance in mice and rats (Basavapathruni et al., 2014;

Daigle et al., 2011). Although challenging, continuous administration of drugs is feasible as evident by the successful use of blinatumomab to treat B-ALL (Topp et al., 2015). The lack of efficacy of EPZ5676, may suggest such drugs are better used in combination rather than as single agents (Dafflon et al., 2017).

The global gene-expression profile of MLL-r infant ALL is characterized by the overexpression of FLT3 (fms-like trosine kinase 3), by either activating mutations or autocrine signalling (Armstrong et al., 2002; Armstrong et al., 2003; Brown et al., 2005; Kang et al., 2012; Stam et al., 2010; Takentani et al., 2004). Infants with MLL-r ALL and high expression of FLT3 do especially poor (Chillón et al., 2012; Stam et al., 2007). A tyrosine kinase inhibitor, lestaurtinib, was incorporated into COG trial AALL0631 alongside intensive chemotherapy but failed to demonstrate any benefit (Brown et al., 2016).

Demethylating agents are also under investigation for use in MLL-r IL since the silencing of tumour suppressors by promoter region hypermethylation is a characteristic of the disease and poor survival correlates with greater extents of promoter hypermethylation (Schafer et al., 2010; Stumpel et al., 2009). The ongoing COG trial NCT02828358 has incorporated use of demethylating agent, azacytidine, in addition to chemotherapy, due to the effectiveness of azacytidine to treat MLL-r ALL cells in pre-clinical studies (Schafer et al., 2010; Stumpel et al., 2009; Stumpel et al., 2013).

The development of these novel epigenetic targeted compounds represents the progress that has been made in understanding MLL-r IL disease processes and gives hope that new treatment options and better outcomes for IL patients are near. However, to date, no targeted or epigenetic therapies have been incorporated as standard of care for IL. One challenge of epigenetic therapies is that although they target specific proteins involved in oncogenesis, they are not specific to cancer cells — meaning their effects will also impact healthy tissues. This phenomenon is not strictly unique to epigenetic agents, as many chemotherapies that target rapidly proliferating cells also affect tissues that frequently repopulate, such as tissues within the GI tract. However, unlike many chemotherapies, the global effects of epigenetic therapies are largely unknown and require significant investigation into both short- and long-term toxicities. Additionally, the development of

successful new drugs requires decades of research and data collection and multibillion-dollar budgets. The development of DC_YM21 demonstrates the utility of modifying or repurposing existing FDA approved drugs, which can dramatically shorten the time needed to get drugs to market and in a more cost-effective manner. Much of what we know about MLL-r leukemia to date has been informed by animal models, which also serve as essential platforms for the pre-clinical testing of new therapies.

1.6 Murine Models Highlight Knowledge Gap in MLL-r Disease Biology

The most common MLL-r modelled in mice are MLL-AF9 and MLL-ENL, which include both *in vivo* and *ex vivo* platforms. These models demonstrate the ability of MLL-fusions to transform mouse blood cells (DiMartino et al., 2000; DiMartino et al., 2002; Lavau et al., 1997; Lavau et al., 2000). Furthermore, the models have been used to study pathways in which MLL-fusions initiate leukemogenesis, such as those mentioned in the previous sections. Mouse translocator models of Mll-Af9 and Mll-Enl, which use Cre-Lox technology to recreate these reciprocal translocations *in vivo,* demonstrated two important characteristics of MLL-r leukemia (Drynan et al., 2005). Firstly, they revealed that the phenotype of leukemia is dependent on the fusion protein. This is evidenced by the development of strictly AML in mice expressing *MLL-AF9* in *Lmo2* expressing primitive blood progenitor cells. While mice expressing *MLL-ENL* in *Lck* expressing T cells could develop either a myeloid or lymphoid disease. These results are consistent with the phenotypes seen in humans. Secondly, these models showed that the phenotype of leukemia is also dependent on the cell type in which the MLL-fusion is being expressed. The expression of MLL-AF9 in *Lck* expressing T-cells had no ability to initiate leukemia, while MLL-ENL expression within primitive progenitor cells produced strictly AML.

A third factor, the hematopoietic niche, has also been found to influence the disease phenotype of MLL-r leukemia (Rowe et al., 2019). Within leukemia patient populations, ALL is more common in children, while adult patients overwhelmingly present with AML. It was noted that this shift from ALL to AML parallels a shift in the blood development of mice and humans from a lymphoid to myeloid lineage bias (Pang et al., 2011; Rebel et al., 1996). This hypothesis is supported by the fact that *MLL-AF9* transformed HPSCs

transplanted into neonatal mice resulted in a leukemia containing cells of both myeloid and lymphoid lineages, and is congruent with the ability of MLL-AF9 to produce either AML or ALL in humans (Meyer et al., 2018; Rowe et al., 2019). However, transplanting the same cells into adult mice produced exclusively a myeloid phenotype. While mouse models have identified important factors influencing the development *MLL-r* leukemia, these models have failed to consistently reproduce leukemias that recapitulate the rapid onset and extent of lineage phenotypes seen in the human disease, such as pro-B cell leukemia (CD10 negative) that carry the worst prognosis out of all *MLL-r* IL (Bueno et al., 2011). Evidence from *MLL-AF9* transformed HSPCs in the neonatal mice, indicate that the neonatal hematopoietic environment can support early pre/pro-B-cell lymphoid progenitor cells, and such shifting MLL-r modelling into 'infant' mice is an important step in generating models that better represent infant leukemia (Rowe et al., 2019).

Lastly, mouse models have failed to reproduce the short latency disease seen in IL patients and suggest that the expression of MLL-fusions alone are insufficient in producing a rapid onset disease that phenocopies clinical IL (Rowe et al., 2019). In MLL-AF9 and MLL-ENL murine translocation models, mice took upwards of 200 days (Valentine et al., 2014), and MLL-AF9 knock-in mice (Dobson et al., 1999) took on average 6 months before they began developing a leukemic disease. Similarly, the retroviral expression of MLL-AF9 in neonatal mice developed disease between 76 and 101 days, compared to adult mice who developed disease within an average 79 days (Rowe et al., 2019). These findings suggest that concurrent mutations may be needed to induce the rapid disease onset seen in IL patients and that the complexity and heterogeneity of *MLL-r* IL poses several challenges to animal modelling. There may be advantages to generating additional non-murine based models that can feasibly address these issues.

1.7 Zebrafish as a Model for Leukemia and Drug Screening

1.7.1 Zebrafish Hematopoiesis

Zebrafish have proven to be attractive models for studying blood development. Hematopoiesis in zebrafish is highly conserved with blood development in humans and recapitulates all the major blood cell types and lineages (Traver et al., 2003). Like humans

and other mammals, constitution of the hematopoietic tissues in zebrafish occurs in two waves: primitive and definitive (Galloway & Zon, 2003). The primitive wave of blood cell production occurs first and is responsible for the initial generation of myeloid and erythroid cells during the first 24 hours of development. Macrophage production occurs at the anterior lateral mesoderm (ALM), while erythroid cells are produced from the intermediate cell mass (ICM) (Davidson et al., 2003; Detrich et al., 1995). Proerythroblasts and endothelial cells that will form vasculature are produced from the ICM, which is formed at roughly 18 hpf, and circulation begins in the zebrafish embryo between 25-26 hpf (Long et al., 1997; Vogeli et al., 2006). Once circulation begins, the function of the ALM and ICM is transient and blood production to the posterior blood island (PBI) marks the switch from primitive to definitive hematopoiesis (Bertrand et al., 2007). The main goal of primitive hematopoiesis is to support rapid embryonic development and the transition to definitive blood development brings about hematopoietic stem cells (HSCs) that will maintain the hematopoietic tissues for the rest of the animal's life. The PBI produces erythromyeloid progenitor cells that can produce both myeloid and erythroid lineages but lack the ability to self-renew (Bertrand et al., 2007). HSCs first arise from ventral wall of the dorsal aorta a structure analogous to the mammalian aorta-gonad-mesonephros (AGM) around 36 hpf and migrate to two different sites (Murayama et al., 2006). Facilitated by macrophages, HSCs migrate to the caudal hematopoietic tissue (previously the PBI) a structure analogous to the mammalian fetal liver, and to the developing thymus to produce lymphoid lineages. Definitive erythroid and myeloid cells begin replacing their primitive counterparts around 3 dpf, and at 4 dpf HSCs migrate from the CHT to the kidney marrow (Bertrand et al., 2008; Jin et al., 2009; Murayama et al., 2006; Zhang & Rodaway, 2007). Analogous to the bone marrow, the zebrafish kidney marrow supports both larval and adult blood development. In addition to conserved cell types, analogous hematopoietic tissues, and similar waves of blood development, signalling pathways and gene regulation are also conserved between humans and zebrafish (Amatruda & Zon, 1999; Chen & Zon, 2009). This signifies that findings from zebrafish blood development studies provide valuable translational information have the potential to inform clinical decisions.

Several technologies have been developed to allow the characterization and monitoring of hematopoiesis in zebrafish. Investigation into gene expression in a variety

of blood cell lineages has led to characterization of specific gene markers for many important hematopoietic populations, and labelling these populations with RNA probes has become common place (Chitramuthu & Bennett, 2013; Shen et al., 2013; Song et al., 2004). This is partly due to the lack of zebrafish specific antibodies, which in many cases, prevents the use of immunohistochemistry and flow cytometry for lineage identification. Additionally, several blood and vasculature specific transgenic reporter lines have been developed that have proven useful in tracking blood cell migration and characterizing interactions between blood cells and their microenvironment (Stachura & Traver, 2016; Tamplin et al., 2015). Lastly, the development of flow cytometry techniques that use forward, side scatter, and fluorescence to quantify the abundance of the major blood cell lineages: HSCs, erythrocytes, lymphocytes and myeloid cells from the adult zebrafish kidney marrow, have proven invaluable since there are no specific antibodies for these cell populations (Traver et al., 2003).

1.7.2 Zebrafish as pre-clinical models for leukemia

The zebrafish has emerged as a highly efficient model system for studying human cancer. Many of the oncogenes and tumor suppressor genes identified as critical players in human malignancies have zebrafish homologs, and the key pathways regulating cell growth, proliferation, apoptosis and differentiation are well conserved. Specific to the modeling of leukemia, signaling pathways involved in zebrafish hematopoiesis and blood development are highly conserved with humans. To this end, several zebrafish models of leukemia have been generated, including models of T-ALL, B-ALL and AML that have been used to investigate disease biology and drug screening (Forrester et al., 2011; Langenau et al., 2003; Langenau et al., 2005; Sabaawy et al., 2006; Yeh et al., 2008). One such AML model generated by the Berman lab, expresses the human *NUP98-HOXA9* fusion gene within myeloid progenitor cells of the zebrafish using an inducible Cre-lox system (Forrester et al., 2011). Embryos expressing the *NUP98-HOXA9* transgene showed increased numbers of HSCs, myeloid expansion and anemia, indicative of a hox mediated block in hematopoietic differentiation (Deveau et al., 2015). It was further shown that NUP98-HOXA9 induced leukemogenesis was dependent on the downstream activation of

meis1, and that the knockdown of *meis1* or treatment with DNMT inhibitors reversed the leukemic phenotype (Deveau et al., 2015). Treatment with DNMT inhibitors restored global methylation levels from a hypermethylated state back to near normal levels. This highlights the way in which zebrafish can be used to model human leukemia and identify potential new treatments, including epigenetic therapies.

1.7.3 Zebrafish *mll* and established models of MLL-r Leukemia in Zebrafish

Zebrafish *mll* shows significant nucleotide identity (54.5%) and highly conserved functional domains (up to 88.5% nucleotide similarity) with its human orthologue, giving reason to believe its functionality is similarly conserved (Robinson et al., 2011). Zebrafish that lack *mll*, develop severe defects in hematopoietic development, including a lack of blood flow and a block in differentiation as evidenced by a lack of mature blood cells and the increased expression of progenitor cell markers (Wan et al., 2011). Gene expression changes mimic those seen in humans, as *mll* morphant embryos showed reduced expression of several *hox* genes, including *hoxa9* (Wan et al., 2011). Additionally, effectors central to *MLL-r* leukemogenesis like MENIN and DOT1L*,* are also highly conserved in zebrafish (Khodaei et al., 1999; Morello et al., 2012). Zebrafish *dot1l* morphants displayed significant increases in *pu.1* expression (a marker of erythromyeloid progenitor cells) and a downregulation of *hoxa9* and *meis1*, confirming the conservation of epigenetic hematopoietic regulators between zebrafish and humans (Morello et al., 2012).

A recent study showed the ability of human *MLL-AF9* mRNA to form a functional fusion protein within zebrafish, giving credibility to the use of zebrafish to model MLL-r leukemia (Tan et al., 2018). *Hox* cluster and *meis1* gene expression was significantly upregulated in hematopoietic tissues within 48 hours of injection, suggesting transcriptional programs essential to MLL-fusion leukemogenesis are highly conserved in zebrafish. Additionally, zebrafish injected with *MLL-AF9* mRNA developed a myeloproliferative disorder that could be alleviated through the treatment of MENIN inhibitor MI-2, further demonstrating the feasibility of using a zebrafish *MLL-r* model to screen for appropriate targeted therapies (Tan et al., 2018). However, the current MLL-AF9 mRNA injection model lacks the ability to restrict MLL-r expression to specific

hematopoietic cell types, or the ability to assess animals past 4-5 days of life. The ubiquitous expression and rapid degradation of injected mRNA, warrants the development of a more robust transgenic model to study MLL-r.

1.7.4 Zebrafish as a Platform for Drug Screening and Discovery

One advantage of modelling leukemia or other diseases in zebrafish, is the ability to leverage the model for medium to high throughput chemical drug screens. Zebrafish produce large numbers of offspring through external fertilization and are more cost-effective to house than mice, especially when large numbers of animals are needed for the screening of drug libraries. Zebrafish embryos develop rapidly and are optically clear, lending well to the use of phenotype-based screening using fluorescent labelling. Tissue structure and specific proteins can be labelled genetically using fluorescent protein tags or highly specific dyes. This allows for the examination and monitoring of *in vivo* cellular and proteomic processes without having to sacrifice the animal. These techniques allow longitudinal studies to be conducted by monitoring the same animal throughout development. It also provides an efficient way to monitor disease progression by assessing the production of specific proteins or the proliferation, migration or metastatic potential of cancer cells. Due to their small size, larvae can be arrayed in a variety of isolated well plates and in many cases, simply bathed in water containing the compounds of interest. To date, numerous chemical screens have been performed using zebrafish and have identified several potential novel anti-cancer agents (Murphey et al., 2006; Ridges et al., 2012; Yeh et al., 2009). Yeh and colleges were able to screen 2,000 bioactive compounds for potential inhibitors of the AML1-ETO fusion gene product that causes AML. AML1-ETO expression impedes hematopoietic differentiation and therefore fish were stained using in-situ hybridization post-treatment to identify compounds that increased levels of *gata1 (*a marker of erythrocytes). The screen identified COX-2 inhibitor nimesulide, and further showed that treatment with nimesulide was also able to reduce the number of myeloid progenitor cells and counteract the effects of AML1-ETO expression (Yeh et al., 2009). Similarly, the transgenic zebrafish reporter line (*lck:eGFP*) that has GFP expressing T-cells was used to screen the ChemBridge DIVERSet Library, for chemicals that could be

effective in treating T-ALL. Lenaldekar (LDK) was identified for its ability to eliminate immature T-cells without harming other cell-types or tissues and treatment with LDK induced long term remission in adult zebrafish with cMYC-induced T-ALL (Ridges et al., 2012).

It is likely that as the number and variety of different zebrafish cancer models increase, drug screens will become more commonplace and the identification of new potential therapies will continue to increase. Furthermore, combining the zebrafish disease model with new technology like the Union Biometrica Biosorter and LP sampler, can partly automate the screening process and make the rapid screening of large drug libraries more achievable. It does this by leveraging the transparency of zebrafish larvae and the use of fluorescent protein tags or dyes. The Biosorter with an LP sampler is able to remove fish from wells within a 96-well plate, measure levels of fluorescence within each fish and deposit back into the well of origin, in an automated fashion. This process removes the need for time intensive imaging and image processing. These advantages make zebrafish a unique, cost-effective platform for medium to high-throughput phenotype-based *in vivo* testing of small molecule inhibitors and targeted therapies, something that is not feasible in the mouse model. Ultimately, zebrafish provide an effective way to reduce large screens to select few “hit” compounds that could then be further studied in mice — streamlining pre-clinical drug discovery.

1.8 Hypothesis and Objectives

Infant leukemia research has focused heavily on two main issues: first, clinical trials to improve outcomes by optimizing treatment regimens using drugs already established for the treatment of leukemia. And secondly, on murine models that seek to understand the biological mechanisms by which *MLL-r* and their resulting fusion genes initiate and maintain leukemogenesis. As a result, the many epigenetic functions of MLL have become well understood, and it has become clear that MLL-r IL is a heterogenous and complex disease. However, there have not been any significant improvements to IL treatment and outcomes for these patients remain dismal — new treatment options are a necessity.

I have identified two major challenges to the current framework of MLL-r leukemia research. Firstly, murine models that have proven the oncogenic abilities of MLL-fusions and have highlighted key factors in IL oncogenesis, such as 1) the identity of the MLL-fusion, 2) the blood cell population harbouring the MLL-r and 3) the impact of the neonatal hematopoietic environment, still lack the ability to recapitulate a true early onset IL phenotype. Adding to the complexity of the disease, additional concurrent mutations and contribution of environmental factors have been shown to associate with IL development but their specific roles remain largely unclear. Secondly, despite the abundant need for new treatment options, current *in vivo* models are not designed to be easily applied to drug screening. Mice either take many months to develop a leukemic phenotype or require invasive transplantation procedures — both of which make drug testing in mice models expensive and slow. Zebrafish offer a unique solution to the aforementioned challenges due to the ease in which they can be genetically manipulated and applied to high-throughput phenotype-based drug screening. This leads me to two overarching hypotheses:

Firstly, I **hypothesize** that the zebrafish would provide a capable and versatile model of MLL-r leukemia and be a major asset for infant leukemia drug development. Therefore, my **objective** in Chapter 3 is to develop a transgenic zebrafish model with the capability to express the human *MLL-AF9* and *MLL-ENL* fusion genes within targeted blood cell populations. Fusion gene expression will start during larval development and continue throughout life and will incorporate fluorescent markers amendable for rapid drug screening.

Secondly, I **hypothesize** that germline mutations of the epigenetic regulator KMT2C predispose infants to developing leukemia and play a role in accelerating disease progression. Therefore, my **objective** in Chapter 4 is to generate knockout zebrafish of *KMT2C* homologs, *kmt2ca* and *kmt2cb*, and investigate the role of these genes in normal hematopoiesis and how they may affect IL development.

Chapter 2: Materials and Methods

2.1: Study Approval

Approval for all zebrafish studies were granted by the Dalhousie University Committee on Laboratory Animals (UCLA), under protocol #17-129, and by the University of Ottawa's Animal Care Committee (ACC), under protocol #3173, both in accordance with the Canadian Council for Animal Care (CCAC) guidelines.

2.2 Zebrafish Husbandry, Breeding, and Embryo Staging

Danio rerio (zebrafish) lines used in this study were maintained at Dalhousie University's Zebrafish CORE Facility and then subsequently at the University of Ottawa's Aquatics Facility. Zebrafish were housed on a recirculating system and water composition, light/dark cycle, temperature, feed, and feeding schedule remained at the discretion of facility and veterinary staff. Embryos were produced using both single pair and group mating practices, and maintained at 28°C (Westerfield, 2000). The double pigment mutant, *casper* (mitfa w2/w2; mpv17 a9/a9) line (White et al., 2008) was generously provided by the Zon Laboratory, Children's Hospital, Boston, MA under a material transfer agreement (MTA). Embryos were staged in accordance with a standard protocol (Kimmel et al., 1995).

2.3 Zebrafish Embryo Injection

DNA plasmids, Tol2 transposase, sgRNAs, and Cas9 mRNA were injected at the one-cell embryo stage into the cell. Synthetic mRNA expression studies were performed by injecting mRNA at the one-cell embryo stage into the yolk sac within proximity of the cell. Injections were performed in accordance with standard protocols (Rosen et al., 2009; Xu, 1999).

2.4 RNA Isolation and cDNA Synthesis

RNA Isolation

Embryos (50-100) were combined with 500uL of TRIzol Reagent (Thermo Fisher Scientific) and homogenized by drawing the embryos in and out of a 22G needle attached to a 1mL syringe 10-15 times. RNA extraction was performed using Phrasemaker™ Tubes (Thermo Fisher Scientific) in accordance with manufacturer's instructions.

cDNA Synthesis

cDNA was made by adding 10uL of total RNA to 2 µl of 100µM oligo-dT(18) (Integrated DNA Technologies), heating at 70°C for 10 min and cooling on ice. Next 2µl of M-MuLV buffer, 0.25µl of RNAse Inhibitor RNAsin (NEB), 0.25µL of M-MuLV reverse transcriptase (NEB) and 1.6µL of water were added to the reaction and incubated at 42°C for 1 h. Directly following cDNA synthesis, reverse transcriptase was inactivated by heating at 90°C for 10 min before cooling.

2.5 Vectors and Cloning

p3E-PacI-AscI-P2A-mCherry

In order generate p3E-PacI-AscI-P2A-*mCherry*, the PacI-AscI-P2A-*mCherry* amplicon was generated using PCR and deposited into a p3E plasmid backbone using TA-cloning techniques. P2A-mCherry was amplified using Q5 DNA polymerase (NEB) from the 'Dual Fucci' transgene (Bouldin & Kimelman, 2014), using a forward primer with flanking PacI and AscI restriction site sequences. The 8bp sequence "gcgcatgc" was included in the primer design between the PacI and AscI site to allow effective cutting at both sites during double digestions. The 802bp PacI-AscI-P2A-*mCherry* amplicon was purified using the QIAquick PCR Purification Kit (Qiagen).

A-tailing was performed by combining 8uL purified PCR product with 1µL of 10X PCR buffer with Mg^{2+} (ABM), 1µL of 10mM dATP, and 1µL of Taq polymerase (ABM), and incubating at 72°C for 15min followed by column purification. The A-tailed PCR

product was then ligated into XcmI digested p3E-TA (Miles & Verkade, 2014) using T4 DNA ligase (NEB) resulting in the plasmid p3E-PacI-AsI-P2A-*mCherry*.

<u>*MLL-AF9* and *MLL-ENL* Plasmids</u>

MLL-AF9 was amplified from pMIG-*MLL-AF9* (Addgene, #71443) using Q5 high fidelity polymerase (NEB, M0492S) and ligated into PacI and AscI digested p3E-PacI-AscI-P2A-*mCherry* using T4 DNA ligase (NEB). Initial difficulties amplifying *MLL-AF9* led to the division of the gene into two segments, the first (segment 1) spanned from position 0-2298bp and the second spanned from position 2225-4536bp.

Segment 1 was amplified using primers with flanking PacI and AscI sequences on the forward and revers primer respectively. Segment 1 was ligated into the digested p3E-PacI-AscI-P2A-*mCherry* resulting in the intermediate p3E-PacI-Segment1-AscI-P2A-*mCherry* plasmid. Next, segment 2 was amplified with the AscI sequence flanking the reverse primer. Segments 1 and 2 share a 73bp region that encompassed a BglII site. The segment 2 PCR product was ligated into the BglII and AscI digested intermediate p3E-PacI-Segment1-AscI-P2A-*mCherry* to create the complete p3E-*MLL-AF9-P2A-mCherry*.

The resulting p3E-*MLL-AF9-P2A-mCherry* was sequenced, revealing a 128bp deletion within the 5' region of MLL. This sequence error was traced back the PCR amplification process and could not be rectified through PCR troubleshooting or with different polymerase enzymes. To bypass any amplification steps, an EcoRI site was inserted between the PacI site and the kozak sequence of MLL within p3E-*MLL-AF9-P2A-mCherry*. To do this, p3E-*MLL-AF9-P2A-mCherry* was digested with EcoRI-HF and bunted with Quick Blunting™ Kit (NEB), to remove an existing EcoRI site. Next segment 1 was amplified using a forward primer with PacI and EcoRI flanking sequences. This product was ligated into the PacI and BglII digested p3E-PacI-*MLL-AF9-P2A-mCherry* to create p3E-PacI-EcoRI-*MLL-AF9-P2A-mCherry*. pMIG-*MLL-AF9* was digested using EcoRI-HF and BglII, and the EcoRI-MLL-AF9-BglII segment (corresponding to segment 1) inserted into p3E-PacI-EcoRI-*MLL-AF9-P2A-mCherry*, replacing the incorrectly amplified sequence.

p3E-*MLL-ENL-P2A-mCherry* was generated by digesting pMSCV-*FlagMLL-pl-ENL* (Addgene, #20873) with BglII and AscI and inserting this segment into the BglII and AscI digested p3E-PacI-EcoRI-*MLL-AF9-P2A-mCherry*, replacing *AF9* with *ENL* to create p3E-*MLL-ENL-P2A-mCherry*.

Expression vectors Ubi:Lox-BFP-STOP-Lox-MLL-AF9-P2A-mCherry (Ubi:LBL-MA9-P2A-mCh) and Ubi:Lox-*BFP*-STOP-Lox-*MLL-ENL-P2A-mCherry* (*Ubi:LBL-MENL-P2A-mCh*) were then assembled using Gateway™ Cloning (Thermo Fisher Scientific) techniques. Gateway compatible plasmids pENTR5'-Ubi (Addgene, #27320), pME-LoxP-mTagBFP2-LoxP (Addgene, #75158), p3E-PacI-EcoRI-*MLL-AF9-P2A-mCherry* or p3E-*MLL-ENL-P2A-mCherry*, and pDestTOl2PA2 (Tol2kit, #394) were assembled using an 'LR' reaction. 10fmol of each entry vector and 20fmol of the destination vector were combined with LR Clonase® II (Thermo Fisher Scientific), incubated over-night, and transformed according to manufacturer's instructions.

CreERT2 Plasmids

CreERT2 plasmids were generated using Gateway™ Cloning (Thermo Fisher Scientific) techniques. Plasmids pENTR5'-*ubi* (Addgene, #27320), p5E-*runx1+23* (Addgene, #69602), p5E-*Cd45* (generously provided by Dr. David Traver, University of California San Diego, San Diego, CA), and p5E-*pu.1*(*Spi1*) were used to drive expression of CreERT2. Each p5E promoter was combined with pENTR/D_*creERt2* (Addgene, #27321), p3E-polyA (Tol2kit, #302), and either pDestTol2cry-eCFP (Lopez et al., 2017) or pDestTol2CG2 (Tol2kit, #395) and assembled using 'LR' reactions and LR Clonase® II Plus (Thermo Fisher Scientific) in accordance with manufacturer's instructions. The *cry-eCFP* marker was used with *ubi:CreERT2*, *runx1+23:CreERT2*, and *pu.1:CreERT2*, while the *Cd45:CreERT2* transgene was generated with *cmlc2:eGFP* marker after previous experimentation showed the inappropriate activation of a *Cd45* driven transgene that was suspected to be caused by the *cry-eCFP* cassette.

Frameshift Reporter Plasmids

The coding sequence for the first 426 amino acids of *kmt2ca* were amplified from *kmt2ca* wild-type and mutant cDNA and ligated into digested pCS2+MCS-P2A-sfGFP

(Addgene, #74668) using restriction enzymes PacI and AscI. Similarly, the first 419 codons of *kmt2cb* were amplified from *kmt2cb* wild-type and mutant cDNA and cloned into pCS2+MCS-P2A*sfGFP* using identical processes.

2.6 Generation of Transgenic Zebrafish Lines

Transgenic zebrafish lines were generated using proven Tol2 transgenesis techniques (Suster et al., 2009). Tol2 transposase was synthesized from NotI-HF linearized pCS-*zT2TP* (Kawakami et al., 2004) using SP6 capped *in vitro* transcription. Embryos were injected with a solution of 75ng/µL plasmid, 25 ng/µL Tol2 transposase mRNA and 0.025% phenol red. Injected larvae were fluorescently screened at 3 dpf for the relevant markers: *cry-eCFP* for *runx1+23:CreERT2*, *pu.1:CreERT2*, and *ubi:CreERT2*, and *cmlc2:eGFP* for *Cd45:CreERT2*. For the *ubi:LBL-MA9-P2A-mCherry* and *ubi:LBL-MENL-P2A-mCherry* transgenes, larvae were screened for ubiquitous TagBFP2 expression. Positive chimeric larvae were then grown to adulthood and outcrossed. F1 larvae were fluorescently screened and positive larvae were used to establish stable transgenic lines. F1 larvae expressing mTagBFP2 were screened using the BioSorter (Union Biometrica Inc.), a large particle flow sorter, rather than fluorescent microscopy due to weak signals and significant autofluorescence. Larvae at 3 dpf were bulk sorted for enhanced TagBFP2 expression in the caudal 20% of the fish. This strategy was used to exclude any auto-fluorescent signal from the larval yolk sac, and two-tailed profiling was unavailable. Positive F1 larvae were grown to adulthood, fin-clipped, and genotyped to confirm presence of the transgene before breeding to produce F2 embryos for experimentation. For each transgene, two to three different founder lines were established. Expression of the transgene was assessed in each founding line and in some cases, lines of the same transgene were crossed together to increase copy number.

2.7 z-4-Hydroxytamoxifen Treatment

Zebrafish larvae were treated with 25µM z-4-Hydroxytamoxifen (4-OHT) in either a 10cm Petri dish, or a 6- or 12-well plate depending on the number of fish in each treatment

group. A 25mM 4-OHT 1000x stock solution was created by dissolving 4-OHT (Millimore Sigma, H7904) in 100% ethanol. To prevent degradation, 30µL aliquots were stored in an opaque container at -20°C. 25mM 4-OHT was diluted using E3 embryo media to the desired 25µM in a conical tube and mixed by vortex. Zebrafish larvae were treated overnight (~16hrs), exchanging their water for the 25µM 4-OHT solution. Following treatment, larvae were rinsed three times with E3 embryo media to efficiently remove any remaining 4-OHT.

2.8 Fluorescent Imaging

Fluorescent transgenic zebrafish larvae were imaged by immobilizing the larvae in a 0.2% tricaine solution in a 35mm glass bottom dish (VWR). Fish were imaged using the Zeiss Axio Observer inverted microscope or in a plastic bottom Petri dish using the Zeiss SteREO Discovery V16 epi-fluorescent microscope.

2. 9 Capped and Un-Capped *in vitro* Transcription

Capped RNA for embryo injection was synthesized using one of two kits: mMessage mMachine™ T3 kit or mMessage mMachine™ SP6 Transcription Kit (Thermo Fisher Scientific) and precipitated using LiCl according to manufacturer's instructions.

Un-capped sgRNA was synthesized using MEGAshortscript™ T7 kit (Thermo Fisher Scientific) and precipitated using ammonium acetate procedures outlined by the manufacturer.

DIG-labelled anti-sense RNA probes for WISH were synthesized using 2µL of either T3, T7, or Sp6 enzymes (Thermo Fisher Scientific), 2µL 10x DIG labelling solution (Roche), 1µL Protector RNase Inhibitor (Millipore Sigma), 2µL10X transcription buffer (Thermo Fisher Scientific) and up to 2µg template DNA. Reactions were incubated for 2-4 hours at 37°C. A TURBO™ DNase treatment was performed to digest template DNA, and probes were purified using NucAway™ Spin Columns (Thermo Fisher Scientific).

2.10 sgRNA and Cas9 Synthesis

The sgRNA spacer sequences targeting exon 5 of zebrafish *kmt2ca* and *kmt2cb* were generated using SSC (http://cistrome.org/SSC/)(Xu et al., 2015) and were ranked based on predicted efficiency (Table 2.10). Oligonucleotides were ordered (IDT) with a T7 promoter sequence flanking the 5' and an overlap region flanking the 3' end. Overlap extension PCR was used to generate the sgRNA templates as previously defined (Gagnon et al., 2014). sgRNAs were synthesized using T7 RNA polymerase. Zebrafish codon-optimized capped Cas9 mRNA was synthesized from an XbaI linearized pT3TS-*nCas9n* plasmid (Addgene, 46757) using T3 RNA polymerase.

Table 2.10: List of single guide RNA sequences used to generate zebrafish *kmt2ca* and *kmt2cb* mutants.

sgRNA	Spacer	Position within Exon 5	Strand	Theoretical Score
kmt2ca 5-1	GGGGGACGTAACCAGGCGT	65	-	0.94
kmt2ca 5-2	CTGCACATCCGGAACCGCCG	94	+	0.90
kmt2ca 5-3	TTTTCTCCACAGTAGCAGA	6	-	0.85
kmt2ca 5-4	TGACAACAACCGCAGCCCA	170	+	0.77
kmt2cb 5-1	TCGGCCACCACAGTAACAGA	6	-	1.1269
kmt2cb 5-2	GACCATAGCGACAAGACCG	126	+	0.9932
kmt2cb 5-3	GGCTGCCAGCCAATCAGGG	146	+	0.8934

2.11 Generation of kmt2ca and kmt2cb Mutant Zebrafish Lines

The zebrafish *kmt2ca* and *kmt2cb* genes were mutated using CRISPR/Cas9 technology. Embryos were injected with a cocktail of 4 sgRNA against *kmt2ca* or 3 sgRNA against *kmt2cb* (Table 1) at a concentration of 100ng/µL of each sgRNA and 200ng/µL *Cas9* mRNA. Mutants in *kmt2ca* and *kmt2cb* were generated as separate lines. Injected embryos were genotyped by amplifying 400-700bps of gDNA that encompassed the intended cut sites. This assay was used to assess cutting efficiency and later genotyping of heterozygous and homozygous mutants as wild type and mutant bands containing deletions were discernable on a 1.5-2% agarose gel. Injected embryos were then grown to adulthood and outcrossed. Clutches containing mutant F1 larvae were grown to adulthood, then genotyped and sequenced to identify male/female pairs with identical frameshift mutations that were then in-crossed to homozygosity. Three different mutations for each gene were assessed at both the gDNA and cDNA levels before establishing the mutant line for experimentation.

2.12 Fin-clipping and gDNA Extraction

When removing tissue from adult zebrafish for the purposes of genotyping, adult zebrafish were anesthetized in a 0.2% tricaine (MS-222) solution and placed onto a parafilm sheet inside a 10cm Petri dish. A small piece of tissue from the tail fin, 1-2mm in length, was then removed using forceps and a surgical blade.

Genomic DNA (gDNA) was extracted from whole embryos, larvae or tissue samples from fin-clipping in either one of two ways. The sample was placed in 50mM NaOH and heated at 95°C for 10 minutes before neutralizing the solution with a buffered solution of 1M Tris-HCl pH 8.0. When PCR failed to amplify intended targets using the gDNA extracted using NaOH, such was the case when amplifying *MLL*, Proteinase K digestion was used to isolate gDNA instead. Samples were placed in a lysis buffer of 10mM Tris-Hcl pH 8.0, 50mM KCL, 0.3% Tween-20, 0.3% NP40 and 1mM EDTA. 0.2µL of 20mg/µL Proteinase K was added to each sample and digested for at least 4 hours, then heat inactivated and diluted with water (Westerfield, 2000).

2.13 Frameshift Reporter Assay

The frameshift reporter vectors and pCS2+*TagRFP* were linearized using *NotI* and purified using QIAquick PCR Purification Kit (Qiagen). Capped RNA was synthesized using SP6 RNA Polymerase. Frameshift reporter RNA and TagRFP mRNA was injected into single cell embryos, imaged and analyzed as previously described (Prykhozhij et al., 2017).

2.14 RNA Probe Synthesis and WISH

Anti-sense digoxigenin-labeled riboprobes were created by performing *in vitro* transcription from linearized plasmids or PCR products containing a T7 promoter (**Table 2.14.1**). WISH was performed on stage-matched larvae fixed with 4% paraformaldehyde using a battery of hematopoietic cell specific probes (**Table 2.14.2**) and probes targeting *kmt2ca* and *kmt2cb* mRNAs using previously described protocols (Bennett et al., 2001; Dobson et al., 2008). Brightfield images of individual larvae were taken using a Zeiss SteREO Discovery V8 or V20 microscope (Carl Zeiss, Germany). Z-stack images were imported into the FIJI software (ImageJ version 1.52p) (Schindelin et al., 2012) and the number of positively stained cells was counted using the multi-point tool.

Table: 2.14.1: Primers used to create WISH probes.

Gene	Primer Sequence	Direction	Product Length (bp)
kmt2ca	ATGTAGAGCTGGCCCTGAGA	F	930
kmt2ca	TAATACGACTCACTATAGGG ATGGAGGGACACCGATCTCA	R	
kmt2cb	CGTCGGTCTACCAGATGACG	F	816
kmt2cb	TAATACGACTCACTATAGGG CCTTCACTCGGTCGTTCACA	R	

Table 2.14.2: Hematopoietic Lineage Specific Whole-Mount In-Situ Probes.

Probe	Lineage	Time point Analyzed (hpf)	
lcp1	Pan-leukocyte	24	48
mpx	Neutrophils	24	48
cpa5	Mast Cells	24	48

2.15 Data Visualization and Statistical Analysis

Data was collected within Microsoft Excel (Office 365, Microsoft) and imported into R Studio (Version 3.5.1) for visualization and statistical analysis. The 'ggplot2', 'dplyr', 'tidyr', 'ggsci', and 'readxl' packages were used within R software to process data and generate plots. For WISH experiments, each probe and timepoint were analyzed separately and when quantitative statistics were involved a student t-test was performed.

Chapter 3: Results I – First Steps Toward a Transgenic Zebrafish Model of MLL-r IL

3.1 Introduction

MLL-r IL is a largely heterogeneous disease and despite the success of murine models to characterize important oncogenic mechanisms of MLL-fusion proteins, new models are needed to address (1) how the cell of leukemic origin affects leukemic phenotype (AML vs ALL), (2) how the identity of the fusion partner gene relates to the ability of the MLL-fusion to transform certain blood cell lineages, and (3) the need for a platform that feasibly allows large-scale testing of drug libraries and novel therapies for IL treatment. Using the zebrafish model to addresses these questions requires an inducible genetic system that allows both temporal and spatial control of MLL-fusion expression, such that *MLL*-fusion genes can be expressed within cell lineages of interest in experimental fish while still maintain a healthy non-induced fish for line maintenance and breeding. Additionally, the use fluorescent protein tags within the transgenic system are needed for efficient phenotype-based drug screening.

3.2: Generating an Inducible Transgenic Zebrafish model of MLL-r Leukemia

In order to achieve both temporal and spatial control over *MLL*-fusion expression, I opted to use a tamoxifen inducible Cre recombinase (*CreERT2*) system. Cre is a site-specific recombinase, belonging to the integrase family and named so for is ability to Cause Recombination. Cre facilitates the recombination between two recognizable loxP (*locus of recombination)* sites (Hamilton & Abremski, 1984; Sternberg & Hamilton, 1981), and this technology has been well-established within zebrafish models, including those for ALL and myeloid diseases (Forrester et al., 2011; Langenau et al., 2005; Le et al., 2007; Yeh et al., 2008).

Two types of transgenic zebrafish lines were generated for use in this system: (1) fish that express *CreERT2* within blood lineages and (2) fish that express the human *MLL*-fusion gene with flanking Lox sites ('floxed'). *CreERT2* expressing transgenic zebrafish are crossed to ones that express the inducible *MLL*-fusion transgene to produces double-transgenic larvae. Double-transgenic larvae are then treated with tamoxifen metabolite 4-hydroxytamoxifen (4-OHT) to induce cre/lox recombination and express the MLL-fusion

protein within the population of *CreERT2* expressing cells. In this manner the Cre/Lox system provides temporal and spatial control as recombination only occurs once tamoxifen is administered and only within *CreERT2* expressing cells.

The human *MLL*-fusion genes were used to better allow the zebrafish model to recapitulate what happens in human patients and provides the practical advantage of using anti-human *MLL* antibodies. When investigating disease biology and oncogenic processes within the cell, anti-MLL antibodies would be very useful for a variety of experimental methodologies such as imaging, western blotting, protein pull-down, and flow cytometry. Additionally, when considering using this model for the screening of novel therapies expressing the human MLL-fusion protein may be more representative of human drug-protein interactions as compared to zebrafish protein-drug interactions.

In order to investigate how the cell of leukemic origin impacts disease phenotype (AML vs. ALL), I needed to express CreERT2 within different populations of blood cells and lineages and generated 4 different *CreERT2* expressing transgenic lines. I used the *ubi, runx1+23*, *Cd45*, and *pu.1,* gene promoters to drive CreERT2 transgene expression ubiquitously, in HSCs, myeloid and lymphoid progenitors, or exclusively myeloid progenitors, respectively (Bertrand et al., 2008; Hsu et al., 2004; Mosimann et al., 2011; Tamplin et al., 2015)

In order to assess how the identity of the *MLL*-fusion gene impacts phenotype (AML vs. ALL) and establish a model that would allow new therapies to be tested on different *MLL*-fusion genes, I generated three inducible reporter lines. These transgenic lines ubiquitously express a ‘floxed’ *mTagBFP2* cassette (*Lox-mTagBFP2-stop-lox*; denoted as *LBL*) followed by either *MLL-AF9*, *MLL-ENL*, or no *MLL*-fusion gene (as a control) tagged with *mCherry*. These transgenes are referred to as *ubi:LBL-MLL-AF9-P2A-mCherry*, *ubi:LBL-MLL-ENL-P2A-mCherry* and *ubi:LBL-P2A-mCherry*, respectively. These *MLL-AF9* and *MLL-ENL* fusion genes were selected because they were readily available for cloning and allowed findings to be easily compared to existing mouse models. I employed fluorescence as a readout. Cells that undergo recombination and begin expressing either *MLL-AF9*, *MLL-ENL*, or no *MLL*-fusion should switch from expressing *BFP* (blue) to *mCherry* (red). Cre/Lox recombination is a permanent DNA excision event,

and therefore once *MLL*-fusion expression is induced within a population of cells, any daughter cell that arises from that population will also express the *MLL*-fusion (**Figure 3.2**). The temporal control added by using tamoxifen inducible CreERT2 is important because it allows for double-transgenic zebrafish lines to be maintained in a non-induced state without the toxic effects of oncogene expression.

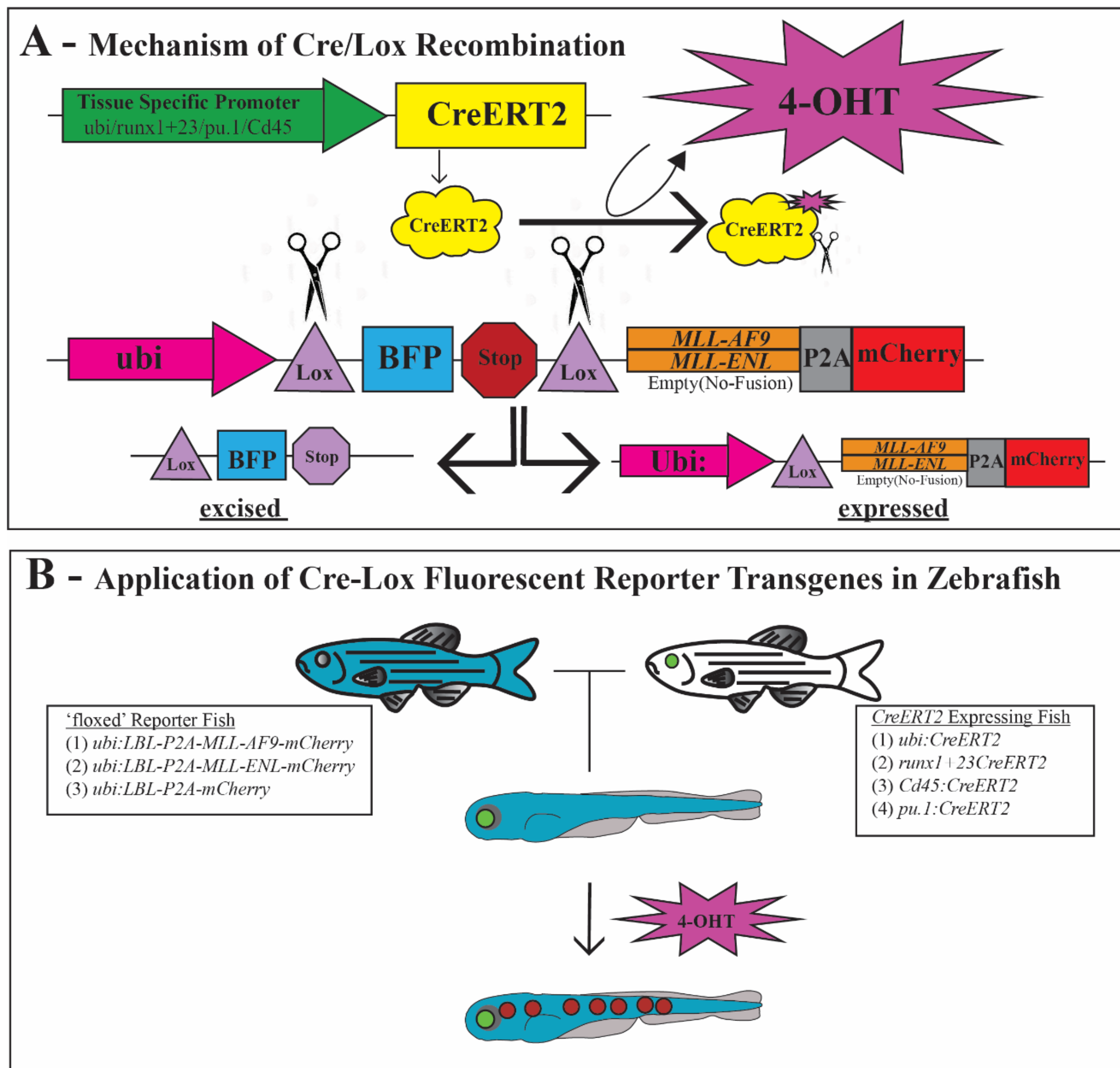


Figure 3.2: Schema for the application of cre/lox inducible technology to generate transgenic zebrafish that express either human *MLL-AF9* or *MLL-ENL* fusion genes in a variety of cell populations. A) *CreERT2* transgene expression is driven by either the *ubi, runx1+23, cd45,* or *pu.1* gene promoters that allow for expression either ubiquitously, or within hematopoietic stem cells, myeloid and lymphoid progenitors populations or exclusively myeloid progenitors. CreERT2 binding of 4-hydroxytamoxifen (4-OHT) induces a conformational change that activates the CreERT2 protein. Activated CreERT2 can recognize and bind locus of recombination (lox) sites and cause homologous recombination at these regions. Ultimately this excises *mTagBFP2* (*BFP*) and a stop codon (stop) and allows the expression of the *MLL*-fusion gene (*MLL-AF9* or *MLL-ENL*, or none) and *mCherry*. The *mCherry* tag is linked to the MLL-fusion gene using a self-cleavage P2A sequence. **B)** Transgenic fish expressing the 'floxed' reporter transgene can be crossed to *CreERT2* expressing lines to produce double-transgenic larvae, capable of undergoing Cre/lox recombination when induced with 4-OHT.

3.3: Transgenic Larvae express *CreERT2* when driven by promoters, *ubi*, *cd45*, and *pu.1*, but not *runx1+23*

To assess the specificity of *CreERT2* expression in each of the *ubi:CreERT2*, *cd45:CreERT2*, *pu.1:CreERT2*, and *runx1+23:CreERT2* transgenic lines, I used whole-mount in-situ hybridization (WISH) to label *CreERT2* expressing cells using a *Cre* specific riboprobe. As expected, *ubi:CreERT2* larvae showed ubiquitous *CreERT2* expression (**Figure 3.3A**), while both *Cd45:CreERT2* and *pu.1:CreERT2* larvae showed *CreERT2* expression within the intermediate cell mass (ICM) and posterior blood island (PBI) at 24 hours post-fertilization (hpf) in association with the production of primitive myeloid and erythroid cells (**Figure 3.3B & 3.3C**). *Runx1+23:CreERT2* fish were assessed at 80 hpf in accordance with the developmental timing of migration of definitive HSCs to the caudal hematopoietic tissue (CHT) and established expression of this promoter in zebrafish (Tamplin et al., 2015), but no *CreERT2* positive cells in the CHT were identified (**Figure 3.3D**). All 4 *CreERT2* lines contained diffuse and non-specific staining within the head, which likely represents off target expression or non-specific staining of the CNS (Figures 2A-D).

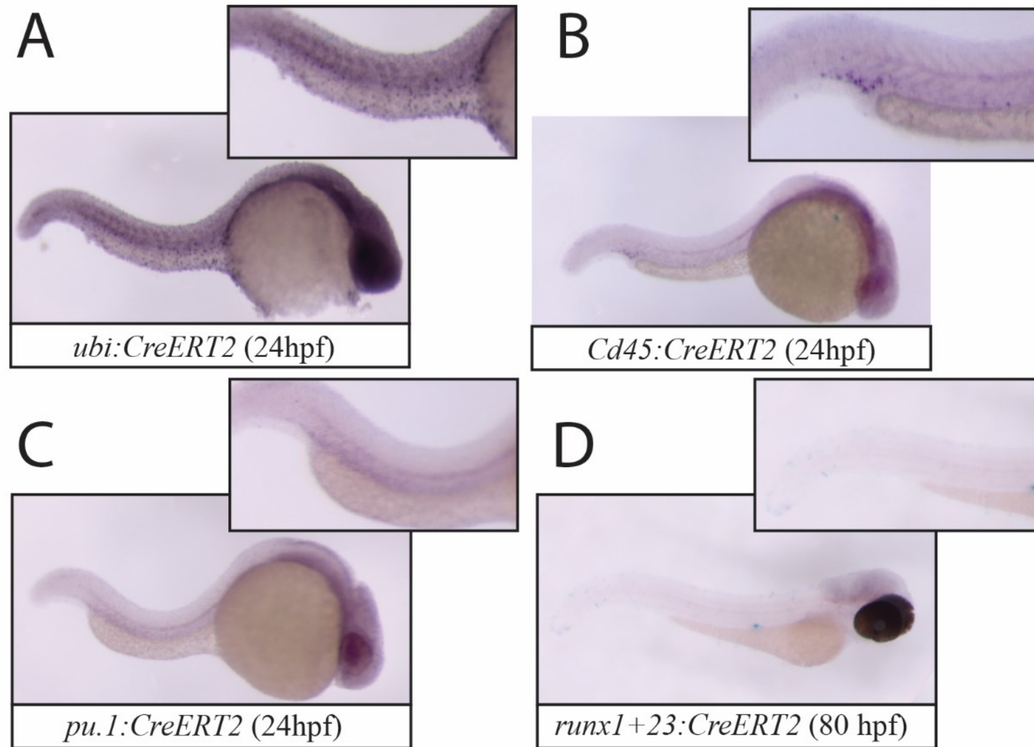
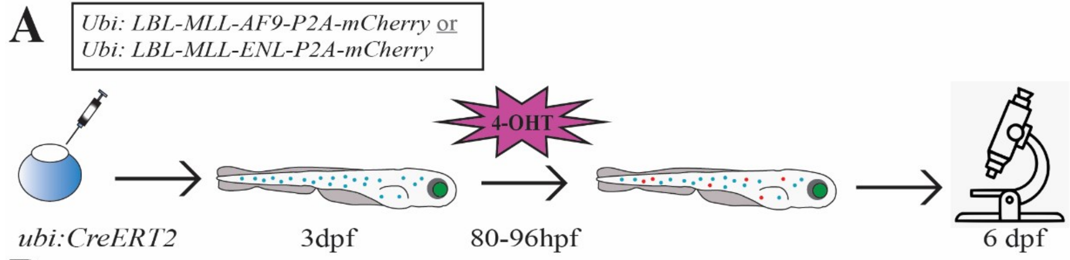


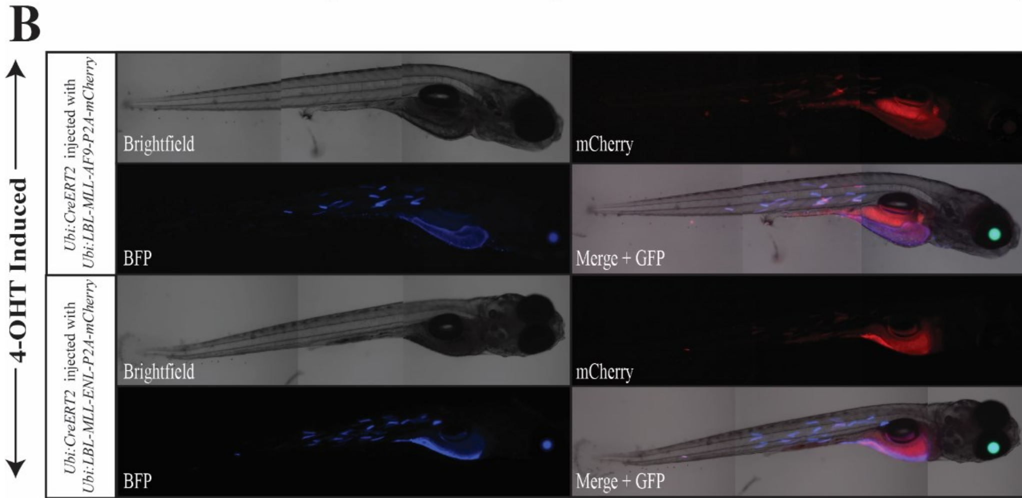
Figure 3.3: Transgenic larvae express *CreERT2* when driven by *ubi*, *Cd45*, and *pu.1* promoters but not *runx1+23.* Larvae were stained using a digoxigenin labeled *Cre* riboprobe using whole-mount in-situ hybridization. Stained larvae were immersed in 70% glycerol and imaged using brightfield microscopy. **A)** *ubi:CreERT2* transgenic larvae show ubiquitous *CreERT2* expression at 24 hours post fertilization (hpf) **B)** *Cd45:CreERT2* and **C)** *pu.1:CreERT2* transgenic larvae show *CreERT2* expression within the intermediate cell mass (ICM) and posterior blood island (PBI) regions at 24hpf **D)** *runx1+23:CreERT2* transgenic larvae do not show any *CreERT2* expression. All larvae show a degree of non-specific and diffuse staining within the head region. WISH was performed on both founder lines of *each Cd45:CreERT2*, *pu.1:CreERT2*, and *runx1+23:CreERT2* with similar results. Experiment was performed once with n=20 fish per group. These lines have since been in-crossed to increase transgene copy number and *CreERT2* expression levels.

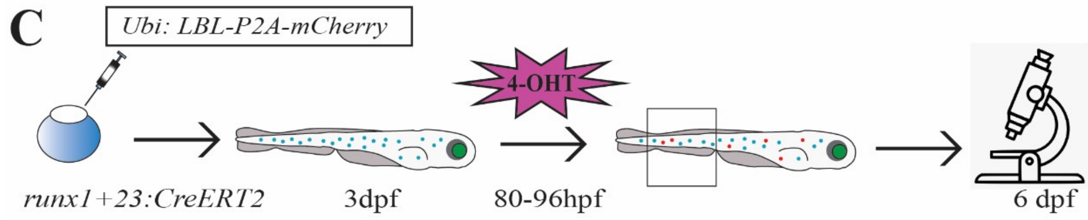
3.4: Chimeric, but not stable transgenic larvae expressing *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* are able to undergo recombination in the presence of Cre recombinase

To test the ability of the *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* transgenes to undergo cre/lox recombination, I injected plasmids containing either transgene into *ubi:CreERT2* transgenic larvae (**Figure 3.4.1A**). Injected larvae were sorted for BFP expression, indicating chimeric expression of the MLL-fusion transgene, at 3 days post-fertilization (dpf). These larvae were treated with 25uM 4-OHT for 16 hours between 80-96 hpf and at 6 dpf larvae were immobilized using tricaine and imaged. Evidence of Cre-lox recombination was present in both *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* injected *ubi:CreERT2* embryos which had low levels of mCherry fluorescence post-treatment (**Figure 3.4.1B**).

Additionally, I tested the functionality of the 'empty' reporter transgene, *ubi:LBL-P2A-mCherry*, by injecting it into *runx1+23:CreERT2* transgenic embryos and administered the same 4-OHT treatment, timeline and imaging procedures as above (**Figure 3.4.1C**). I chose to inject into *runx1+23:CreERT2* embryos to demonstrate how the spatial control of the Cre/lox system allows for expression of the 'floxed' reporter transgenes within blood cells only. This experiment was actually performed before WISH analysis of *runx1+23:CreERT2* fish failed to identify CreERT2 expression. Successful Cre-Lox recombination was evident within the *ubi:LBL-P2A-mCherry* injected *runx1+23CreERT2* larvae as blood cells in circulation and peripheral tissues as well as muscle fibers expressed mCherry (**Figure 3.4.1D**).


A
Ubi: LBL-MLL-AF9-P2A-mCherry or
Ubi: LBL-MLL-ENL-P2A-mCherry
4-OHT
ubi:CreERT2
3dpf
80-96hpf
6 dpf


B
4-OHT Induced
Ubi:CreERT2 injected with Ubi:LBL-MLL-AF9-P2A-mCherry
Ubi:CreERT2 injected with Ubi:LBL-MLL-ENL-P2A-mCherry
Brightfield
mCherry
BFP
Merge + GFP


C
Ubi: LBL-P2A-mCherry
4-OHT
runx1+23:CreERT2
3dpf
80-96hpf
6 dpf

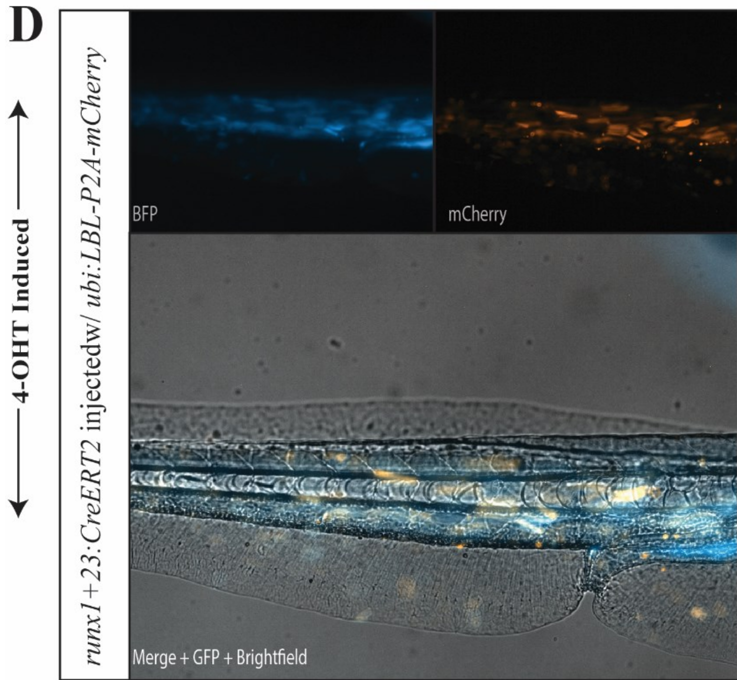
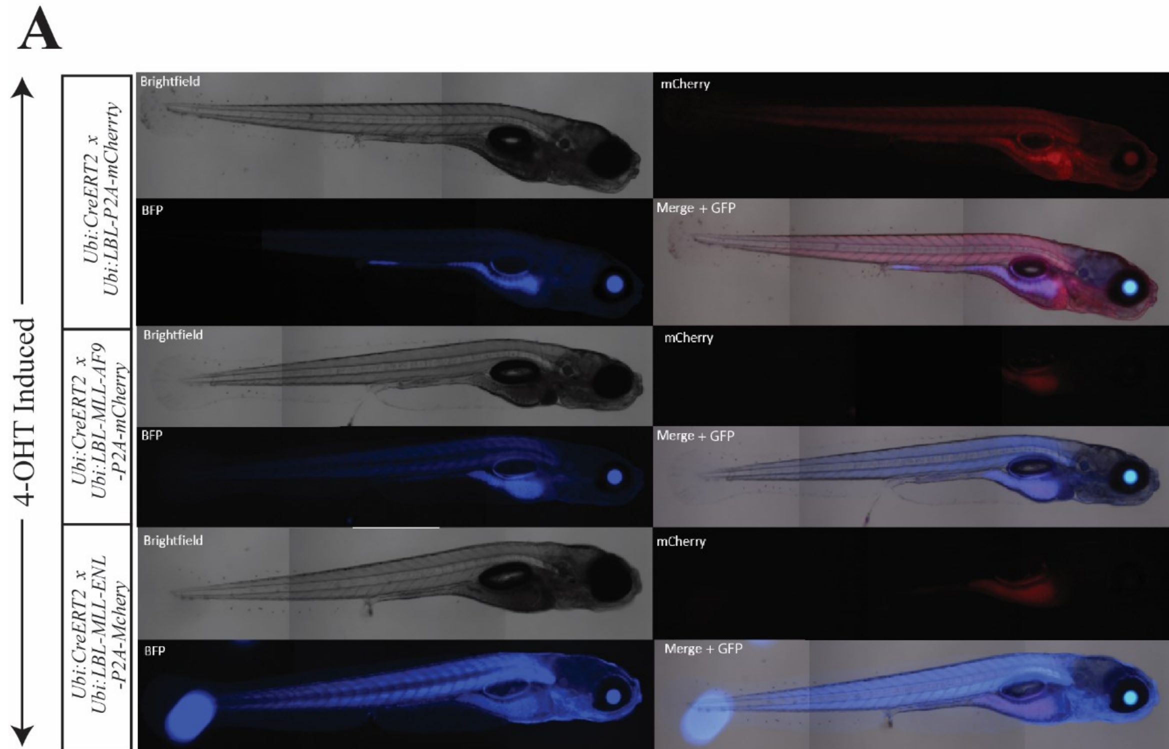

D
4-OHT Induced
runx1+23:CreERT2 injectedw/ ubi:LBL-P2A-mCherry
BFP
mCherry
Merge + GFP + Brightfield

Figure 3.4.1: *CreERT2* inducible fluorescent reporter transgenes undergo recombination *in vivo.* A) Experimental timeline for the injection of plasmids containing the *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* transgenes into *ubi:CreERT2* transgenic embryos. *BFP* expressing larvae are sorted at 3 days post-fertilization (dpf), treated with 25uM 4-OHT for 16 hours, and then imaged at 6dpf. **B)** Representative images of induced larvae with low levels of *mCherry* fluorescence that overlaps *BFP* expression indicating Cre/lox recombination occurred. Two replicates preformed each with n=20 fish per group **C)** Experimental timeline for the injection of plasmid containing the *ubi:LBL-P2A-mcherry* transgene — sorting, treatment and imaging conditions are consistent with panel A. **D)** Image of an induced larvae with muscle fibers and blood cells that express *mCherry* indicating Cre/lox recombination. One replicate performed with n=20 fish.

Once F1 transgenic *ubi:LBL-MLL-AF9-P2A-mCherry*, *ubi:LBL-MLL-ENL-P2A-mCherry* and *ubi:LBL-P2A-mCherry* lines were established, I again sought to validate the Cre/lox system. Transgenic *ubi:CreERT2* fish were crossed to either transgenic *ubi:LBL-MLL-AF9-P2A-mCherry*, *ubi:LBL-MLL-ENL-P2A-mCherry* or *ubi:LBL-P2A-mCherry* lines. Larvae were sorted at 3dpf for the expression of both transgenes using *cry-eCFP* expression as a marker for the *ubi:CreERT2* and *BFP* expression for the *MLL*-fusion reporter transgenes. Double-transgenic larvae were treated with 25uM 4-OHT for 16 hours from 80-96 hpf, then imaged at 6dpf — the same conditions used for chimeric animals in **Figure 3.4.1**. Only the *ubi:LBL-P2A-mCherry* control line showed evidence of recombination, while neither the *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* had any *mCherry* expression (**Figure 3.4.2A**). This finding was surprising, given both *MLL-AF9* and *MLL-ENL* transgenes underwent recombination in chimeric animals using the same 4-OHT induction protocol.

I decided to check the ability of each floxed report line using an additional method, to determine if the lack of recombination was due an insufficient amount of CreERT2 expression or an issue with the *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* transgenes. I injected zebrafish optimized *Cre* mRNA into either *ubi:LBL-MLL-AF9-P2A-mCherry*, *ubi:LBL-MLL-ENL-P2A-mCherry* or *ubi:LBL-P2A-mCherry* transgenic embryos. This *Cre* mRNA did not need to be induced by tamoxifen to initiate recombination. At 28 hpf, larvae were immobilized with tricaine and imaged. Again only the control transgenic fish, *ubi:LBL-P2A-mCherry*, had *mCherry* expression indicating successful recombination (**Figure 3.4.2B**).


A
4-OHT Induced
Ubi:CreERT2 x Ubi:LBL-P2A-mCherrty
Ubi:CreERT2 x Ubi:LBL-MLL-AF9 -P2A-mCherry
Ubi:CreERT2 x Ubi:LBL-MLL-ENL -P2A-Mchery
Brightfield
mCherry
BFP
Merge + GFP

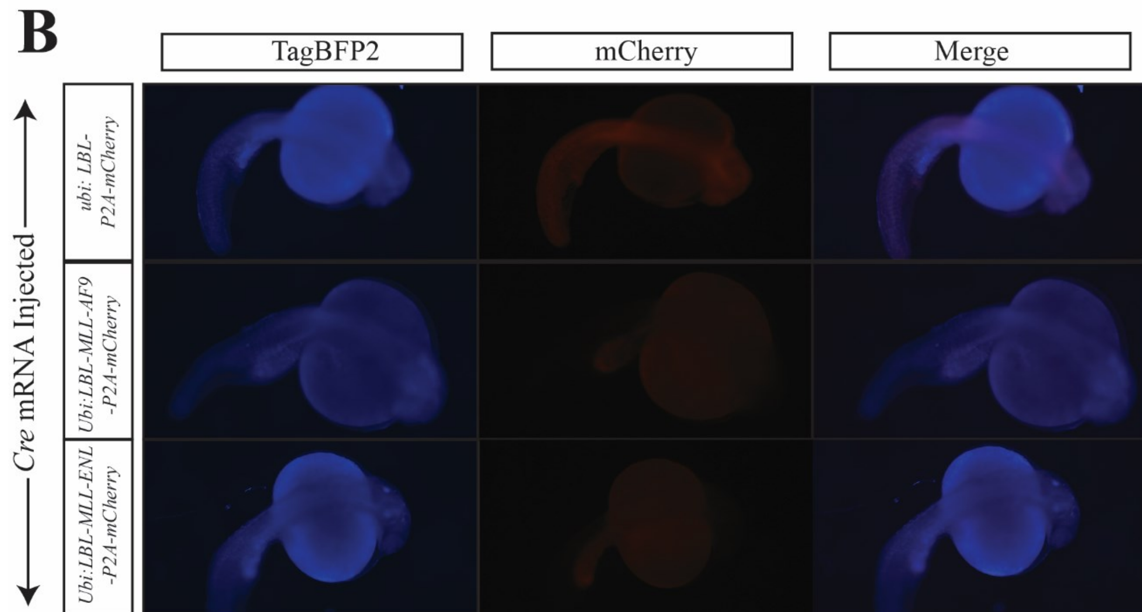

B
TagBFP2
mCherry
Merge
Cre mRNA Injected
ubi: LBL-P2A-mCherry
Ubi:LBL-MLL-AF9 -P2A-mCherry
Ubi:LBL-MLL-ENL -P2A-mCherry

Figure 3.4.2: Transgenic larvae expressing *ubi:LBL-P2A-mCherry*, but not *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* are able to undergo Cre-lox recombination. A) Double transgenic larvae expressing *ubi:CreERT2* and either *ubi:LBL-P2A-mCherry*, *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* were induced with 25uM 4-OHT for 16 hours from 80-96 hpf, then imaged at 6dpf. Only *ubi:CreERT2*;*ubi:LBL-P2A-mCherry* larvae show *mCherry* expression and Cre/lox recombination post-induction. Two replicates performed with n=20 fish per group. **B)** Transgenic embryos expressing either *ubi:LBL-P2A-mCherry, ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* were injected with roughly 500ng of zebrafish codon optimized *Cre* mRNA at the one-cell stage, and imaged at approximately 28 hpf. Only *ubi:LBL-P2A-mCherry* transgenic animals had *mCherry* expression post-induction. Experiment performed once with N=20 fish per group.

3.5: Summary of Findings

In summary, I generated a number of transgenic zebrafish lines as part of a modular, Cre/lox inducible model of MLL-r IL. I demonstrated through WISH analysis that *ubi:CreERT2*, *Cd45:CreERT2*, and *pu.1:CreERT2* transgenic larvae express *CreERT2* within appropriate tissues, although diffuse staining in the head region of these fish indicate there may also be some off-target expression (**Figure 3.3A-C**). *Runx1+23:CreERT2* transgenic larvae did not show any *CreERT2* expression through WISH analysis (**Figure 3.3D**), however, when these embryos were injected with a 'floxed' reporter transgene 4-OHT induction surprisingly produced *mCherry* fluorescence (**Figure 3.4.1D**). This finding indicates that perhaps the WISH staining protocol used was inadequate for the detection of such a small population of HSCs, and more work is needed to assess *CreERT2* expression in *runx1+23:CreERT2* transgenic fish. I also generated 3 'floxed' fluorescent reporter transgenic zebrafish lines, two of which contain either *MLL-AF9* or *MLL-ENL,* and one 'empty' reporter as a control. The chimeric expression of these transgenes was enough to enable Cre/Lox recombination *in vivo* and the expression of the *mCherry* protein tag (**Figure 3.4.1B**). However, F2 transgenic animals expressing either the *ubi:LBL-MLL-AF9-P2A-mCherry* or *ubi:LBL-MLL-ENL-P2A-mCherry* transgenes were unable to undergo such recombination when crossed to Ubi:CreERT2 trangenics (**Figure 3.4.2A**) and furthermore *mCherry* expression could not be induced by injection of synthetic *Cre* mRNA (**Figure 3.4.2B**). The functionality of this transgenic model remains uncertain.

Chapter 4: Results II – Generating *kmt2ca* and *kmt2cb* mutant zebrafish to assess the role of *KMT2C* in hematopoiesis

4.1 Introduction

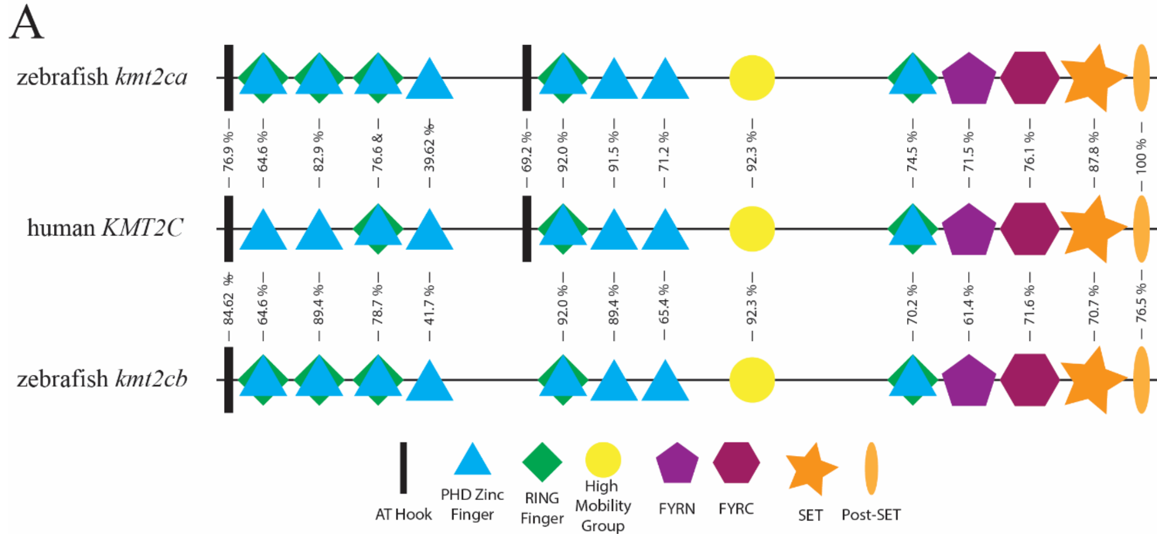
Mouse models of MLL-r leukemia have shown that the expression of *MLL*-fusion genes alone is insufficient to produce rapid onset disease that models IL. Inherited germline mutations, rather than cooperating somatic mutations or environmental factors, are thought to have a significant impact on the development of the disease and neonatal onset (Andersson et al., 2015; Chang et al., 2013; Spector et al., 2005; Valentine et al., 2014). With this in mind, I developed a zebrafish model of loss-of-function (LOF) mutations of the epigenetic regulator, KMT2C, to evaluate the impact on blood development and how that might contribute to leukemogenesis.

4.2: Zebrafish Homologs *kmt2ca* and *kmt2cb* are Significantly Conserved with Human *KMT2C*

Zebrafish underwent an additional whole-genome duplication event during evolution compared to humans and as a result, two zebrafish *KMT2C* homologs exist, *kmt2ca* and *kmt2cb*. I first wanted to assess the level of conservation between *KMT2C* and zebrafish homologs *kmt2ca* and *kmt2cb*, to ensure a zebrafish model could accurately inform the human condition. For this I used alignment software Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) to align the amino acid sequences of each gene to human *KMT2C*. The percent identify of kmt2ca and kmt2cb to the human protein was 48% and 44%, respectively, and indicates a significant level of conservation. I used domain annotations from SMART (Simple Modular Architecture Research Tool; http://smart.embl-heidelberg.de/help/smart_about.shtml) and found that zebrafish kmt2ca and kmt2cb have very similar protein primary structure with human KMT2C, including the types and numbers of domains and their locations within the overall protein structure. Comparing amino acid sequences that encoding major functional domains such as PHD fingers, that are important for the reading of histone methylation and recruitment of COMPASS proteins (Baker et al., 2008; Ernst et al., 2001), and the SET domain (methyltransferase domain) (Milne et al., 2002) reached levels of conservation between

70-100% identity (**Figure 4.2**). Domain sequences were highly conserved between the zebrafish and human homologs, especially within the SET domain.

A



B

SET DOMAIN

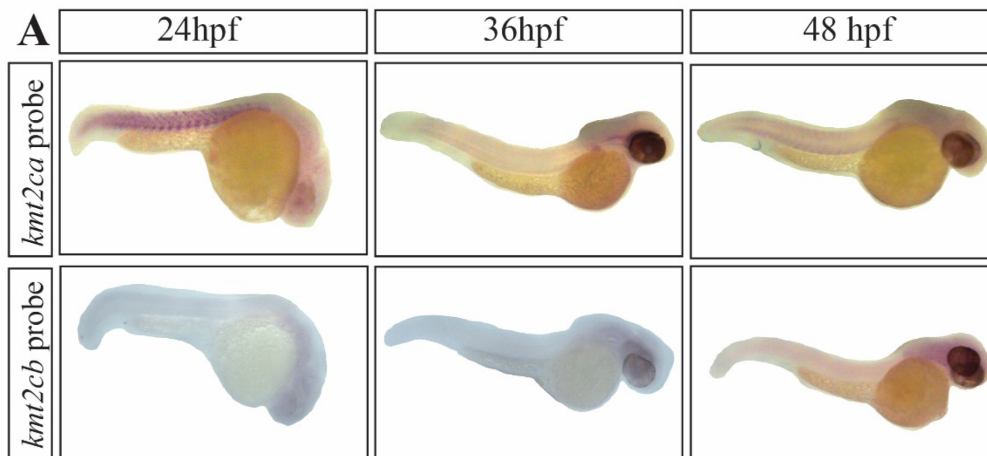
```
kmt2ca    4806    SNVYLARSRIQGLGLYAARDIEKYTMVIEYIGTIIRSEVANRKEKMYEAQNRGVYMFRID    4866
KMT2C     4771    SNVYLARSRIQGLGLYAARDIEKHTMVIEYIGTIIRNEVANRKEKLYESQNRGVYMFRMD    4831
kmt2cb    4739    SNVYLAHSRIQGLGLFAARAIEKQTMVIEYMGDILRTEVAMRRELLYKAKNRPAYMFCID    4799
                  ******:********:*** *** ******:* *:*.*** *:* :*:::** .*** :*

kmt2ca    4866    SEHVIDATITGGPARYINHSCAPNCITEVVALERGHKIIISSNRRIQRGEELCYDYKFDLEDD    4929
KMT2C     4831    NDHVIDATLTGGPARYINHSCAPNCVAEVVTFERGHKIIISSSRRIQKGEELCYDYKFDFEDD    4894
kmt2cb    4799    SERVIDATNSGSPARYINHSCSPNCVAEVVTFERGYKIIISAACRIERGEELCYDYKLTPVND    4862
                  .::***** :*.*********:***::***::***:*****:  **::*********:   :*
```

Figure 4.2: Zebrafish kmt2ca and kmt2cb proteins are highly conserved with human KMT2C. A) Graphical representation of the general location of major functional domains within the primary structure of each protein. Domain annotations were taken from SMART (http://smart.embl-heidelberg.de/). Amino acid (AA) identity of major domains between kmt2ca and KMT2C and between kmt2cb and KMT2C is indicated as a percentage. Only the order and general location of domains is accurately depicted, the size of domain symbol does not equate to the actual size of the domain nor does the space between domain symbols accurately represent of the distance between domains in the primary protein structure due to the large size of the proteins and limited size of the figure. **B)** Alignment of the catalytic SET domain. '*' indicates a single, fully conserved residue. ':' indicated conservation between residues with strongly similar properties (roughly equivalent to a score of >0.5 in the Gonnet PAM 250 matrix). '.' Indicates conservation of residues with weakly similar properties (roughly equivalent to a score of <0.5 in the Gonnet PAM 250 Matrix).

4.3: Expression of *kmt2ca* and *kmt2cb* Is Not Tissue Specific and Peaks During the First 24 Hours of Development.

Before generating mutant zebrafish lines of either *kmt2ca* or *kmt2cb*, I wanted to compare and contrast their relative expression patterns. More specifically, I wanted to know if one homolog was more highly expressed than the other, and if there was a high level of expression of either gene within hematopoietic tissues. To understand the spatial pattern of *kmt2ca* and *kmt2cb* expression during early zebrafish development, I subjected wild-type fish to WISH using either *kmt2ca* of *kmt2cb* specific riboprobes. I found that expression of both genes was diffuse and widespread throughout the developing zebrafish. Expression of both genes was most intense within the head region, likely corresponding to the development of the CNS, with weaker expression throughout the rest of the body and tail, with the exception of *kmt2ca* staining at 24 hpf which had more intense staining in the tail region than the head (**Figure 4.3A**). Expression of neither gene was specific to hematopoietic tissues. Larvae stained with the *kmt2cb* probe at 24 & 48hpf appear a different color because they were taken on a different microscope and white-balance settings were unable to be replicated. To more accurately understand how the relative expression levels of *kmt2ca* and *kmt2cb* relate to each other, I analyzed publicly available RNA-sequencing data from the first 5 days of zebrafish development (White et al., 2017). I imported sequencing data for *kmt2ca* and *kmt2cb* into R software and normalized the data by dividing the RNA counts for each gene by the total number of sequencing reads to account for variation. I found that *kmt2ca* expression spikes within the first ~3 hours of development indicating that it could be maternally deposited and important for early embryogenesis. Starting at ~11 hpf *kmt2cb* expression predominates over that of *kmt2ca*, but by 5 dpf there expression reaches more similar levels (**Figure 4.3B**). Based on this expression analysis, it is not clear if one homolog likely plays a greater role in zebrafish development over the other. Given these findings, in order to make a *KMT2C* knockout model, I needed to make mutations to both zebrafish homologs.



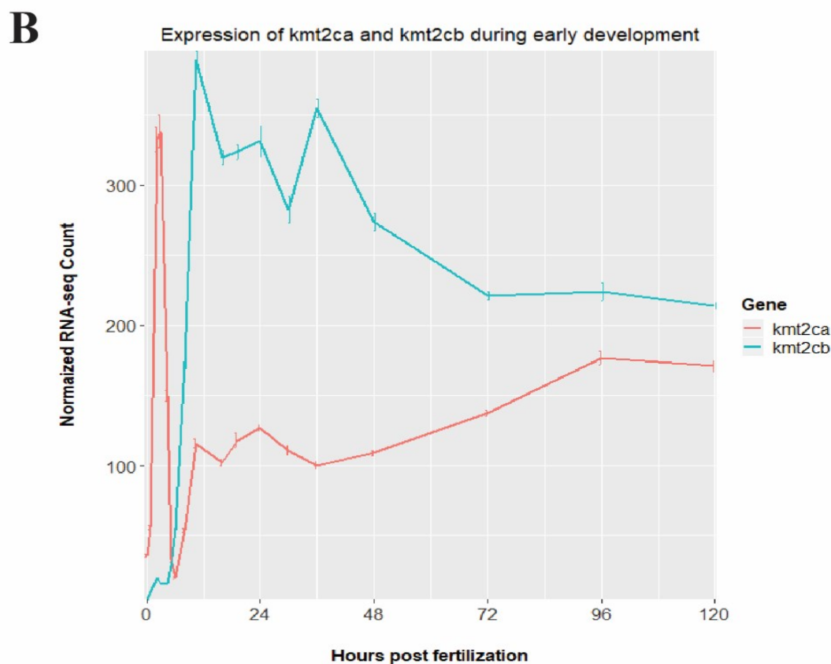


Figure 4.3: Analysis of physiological *kmt2ca* and *kmt2cb* in early zebrafish development. A) Whole-mount in-situ hybridization using *kmt2ca* and *kmt2cb* specific probes at 24, 36, and 48 hpf. Larvae show diffuse staining throughout the head region and staining of the somite region. Larvae stained with the *kmt2cb* probe at 24 & 48hpf appear a different color because they were taken on different microscopes and white-balance settings were unable to be replicated. **B)** RNA sequencing of zebrafish embryos and larvae between 0-5 days post-fertilization. Expression of *kmt2ca* is highest during the first few hours of embryo development but is largely reduced by 6hpf when *kmt2cb* expression is initiated. Expression of *kmt2cb* peaks at ~11hpf and remains expressed at higher levels than *kmt2ca* up to 5dpf. RNA counts were obtained from the publicly available database generated by White and colleagues and normalized to account for variance by dividing the count for either gene by total number of sequencing reads (White et al., 2017).

4.4: Generating *kmt2ca* and *kmt2cb* Mutant Zebrafish

To generate mutations in both *kmt2ca* and *kmt2cb*, I used CRISPR/Cas9 technology to target exon 5 of each gene, since exon 5 precedes the first predicted PHD domain in each gene and a frameshift at such location would likely produce null proteins. Zebrafish embryos were injected at the one-cell stage with either 3 sgRNAs against *kmt2ca* or 4 sgRNAs against *kmt2cb* and *Cas9* mRNA to generate deletions between 50-150bps that could be detected using standard PCR genotyping and agarose gel electrophoresis (**Figure 4.4A & 4.4B**). In order to identify frameshift mutations that produced pre-mature stop codons and were not removed by exon skipping, the genomic and compliment DNA (cDNA) of 3 mutations for each gene were sequenced (**Tables 4.4.1 & 4.4.2**). The 113bp deletion of *kmt2ca* mutation #13 satisfied such criteria and male/female pairs of zebrafish heterozygous for the mutation were in-crossed to generate the maternal zygotic homozygous *kmt2ca(-/-)* line (**Figure 4.4C**). Similarly, *kmt2cb* mutation #15 produced a frameshift mutation present within cDNA and fish heterozygous for this mutation were in-crossed to produce the *kmt2cb(-/-)* line (**Figure 4.4D**). These *kmt2ca(-/-)* and *kmt2cb(-/-)* lines were used for experimentation going forward.

Table 4.4.1: Table of different *kmt2ca* mutations characterized at the genomic DNA (gDNA) and compliment DNA (cDNA levels), before establishing the *kmt2ca(-/-)* line that contains mutation #13.

***Kmt2ca* mutation**	**#9**	**#13**	**#21**
gDNA	111bp deletion within exon 5 Creates in- frame mutation	113bp deletion within exon 5 Creates a frameshift mutation that is predicted to produce premature stop codons	119 bp deletion and 2bp insertion within exon 5 Creates a frameshift mutation that is predicted to produce premature stop codons
cDNA	Not assessed	Frameshift mutation retained	Exon 5 skipped

Table 4.4.2: Table of different *kmt2cb* mutations characterized at the genomic DNA (gDNA) and complimentary DNA (cDNA) levels, before establishing the *kmt2cb(-/-)* line that contains mutation #15.

***Kmt2cb* mutation**	**#15**	**#17**	**#21**
Mutant gDNA	1bp replacement, 11bp insertion & 59bp deletion within exon 5 Creates frame shift that is predicted to produce premature stop codons	138bp deletion within exon 5 Creates in-frame mutation	174bp deletion spanning exon 5 intron/exon boundary Creates frame shift that is predicted to produce premature stop codons
Mutant cDNA	Frameshift mutation retained	Exon 5 skipped	Exon 5 skipped

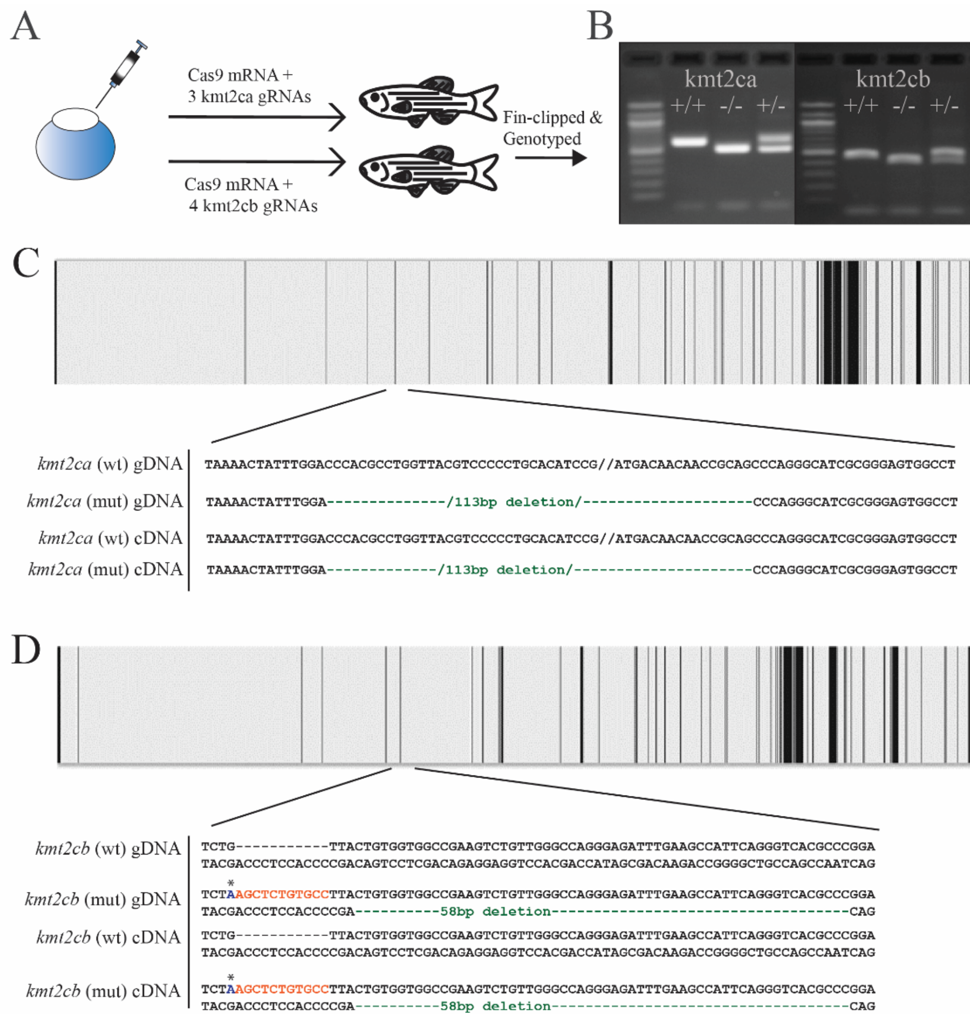


Figure 4.4: Generation and Mutation Characterization of *kmt2ca* and *kmt2cb* mutant zebrafish. A) Embryos were injected with either 3sgRNAs against exon 5 of *kmt2ca* or 4 sgRNAs against exon 5 of *kmt2cb* and Cas9 mRNA at the one-cell stage, then grown to adulthood and bred to homozygosity. To identify mutants, gDNA was extracted from tail-fin clippings and PCR was used to amplify exon 5 of either gene then ran on an agarose gel. **B)** Mutant fish containing deletions to either gene were identified using agarose gel electrophoresis. **C)** Graphical map of the *kmt2ca* gene, where exons are represented as black bars and introns are represented as grey bars. A 113bp deletion (green) was made within exon 5 of *kmt2ca* which resulted in the formation of pre-mature stop codons downstream within exon 6 (not shown). **D)** Graphical map of the *kmt2cb* gene, where exons are represented as black bars and introns are represented as grey bars. A 1bp replacement (blue), 11 bp insertion (orange) and a 58bp deletion (green) was made within exon 5 of *kmt2ca* which resulted in the formation of pre-mature stop indicated by the *, as well as additional pre-mature stop codons downstream (not shown).

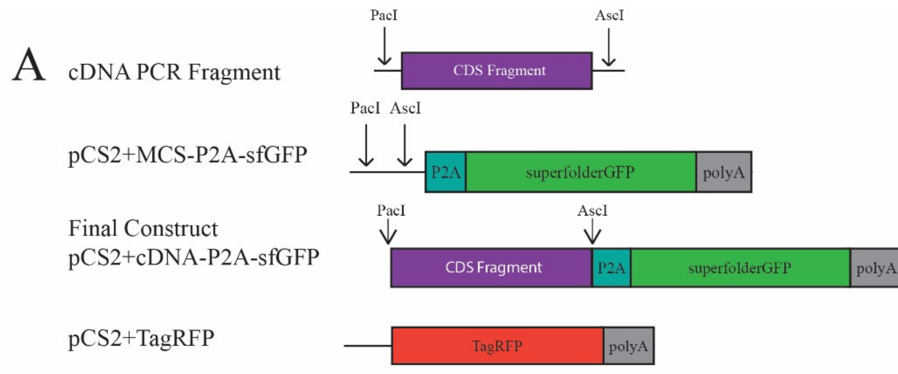
4.5: Reporter Assay demonstrates efficacy of *kmt2ca* and *kmt2cb* frameshift mutations

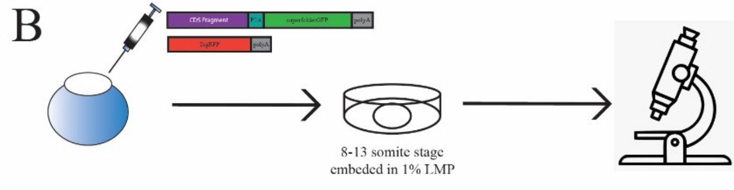
Sequencing studies revealed the generation of indels within *kmt2a* and *kmt2cb* that are predicted to produce frameshift mutations. However, the presence of alternative start sites down-stream of these mutations could ultimately result in an in-frame protein being produced. Due to the lack of zebrafish specific antibodies against *kmt2ca* or *kmt2cb* that could be used to confirm the translation of non-functional proteins, I used a fluorescent reporter assay previously developed in the Berman lab for this purpose (Prykhozhij et al., 2017) to test the efficiency of the *kmt2ca* and *kmt2cb* mutations to produce loss-of-function proteins.

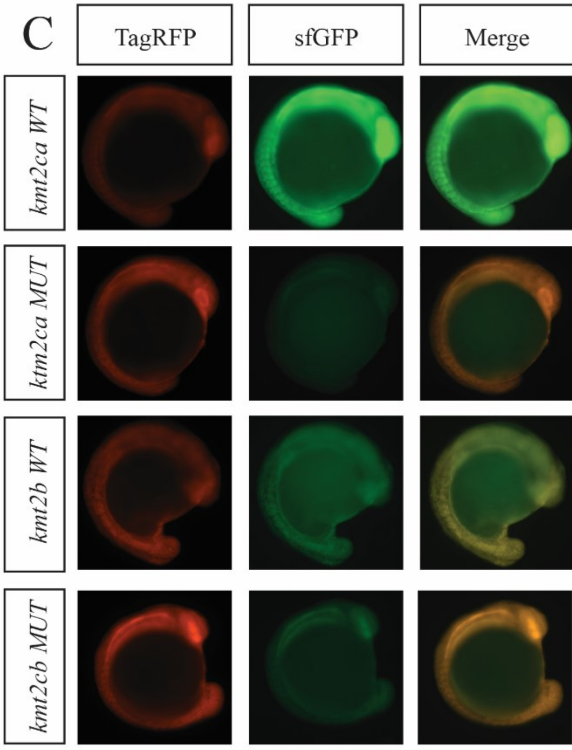
I cloned CDS fragments that excluded the 5'UTR and extended ~200 codons past the exon 5 mutation sites, from wild-type and mutant *kmt2ca* and *kmt2cb* cDNA. I then deposited these CDS fragments in-frame with a P2A-sfGFP fluorescent tag within a pCS2+ plasmid backbone (**Figure 4.5A**). Reporter mRNA synthesized from each plasmid was injected into zebrafish embryos at the one-cell stage, using *TagRFP* mRNA as an injection control, and embryos were imaged the following day (**Figure 4.5B & 4.5C**). To quantify changes in sfGFP expression, the average level of sfGFP and TagRFP expression of regions within the embryo somites was calculated using FIJI software (Schindelin et al., 2012). Values were normalized by subtracting average background fluorescence from the red and green channels used for each image and plotted as a ratio of sfGFP:TagRFP expression **(Figure 4.5D).** Embryos injected with mutant *kmt2ca* RNA had almost no sfGFP expression indicating the *kmt2ca* mutation produced a highly efficient frameshift mutation capable of producing mutant proteins that are out-of-frame. Similarly, embryos injected with mutant *kmt2cb* RNA had a significant decrease in sfGFP production as compared to wild-type, although the *kmt2cb* mutation impeded in-frame sfGFP translation to a lesser extent that the *kmt2ca* mutation.

To better understand why the *kmt2cb* frameshift mutation was less efficient at producing a null protein as compared to the *kmt2ca* frameshift mutation, I used a previously developed algorithm (Nishikawa et al., 2000) to assess the likelihood of alternative translational start sites downstream from the *kmt2ca* and *kmt2cb* mutations. Analysis of *kmt2ca* returned a 4528 AA truncated protein with a low reliability of 53%. If translated,

this protein would include the 4th PHD domain (encoded by exon 11), as well as the SET and post-SET domains. A similar 4275 AA truncated protein was predicted for *kmt2cb*, that would include domains spanning from PHD domain #5 (encoded by exon 18) through to the SET and post-SET domains, but with a lower reliability of 26%. Although it is plausible that alternative proteins may be produced from the *kmt2ca* and *kmt2cb* mutant alleles I generated, the predicted alternative start codons downstream from these mutations were absent from the CDS fragments used in the reporter assay and it is still unclear why the *kmt2cb* frameshift mutation seemingly performed worse than the *kmt2ca* frameshift.


A
cDNA PCR Fragment
PacI
AscI
CDS Fragment
pCS2+MCS-P2A-sfGFP
P2A
superfolderGFP
polyA
Final Construct
pCS2+cDNA-P2A-sfGFP
pCS2+TagRFP
TagRFP


B
8-13 somite stage
embeded in 1% LMP


C
TagRFP
sfGFP
Merge
kmt2ca WT
ktm2ca MUT
kmt2b WT
kmt2cb MUT

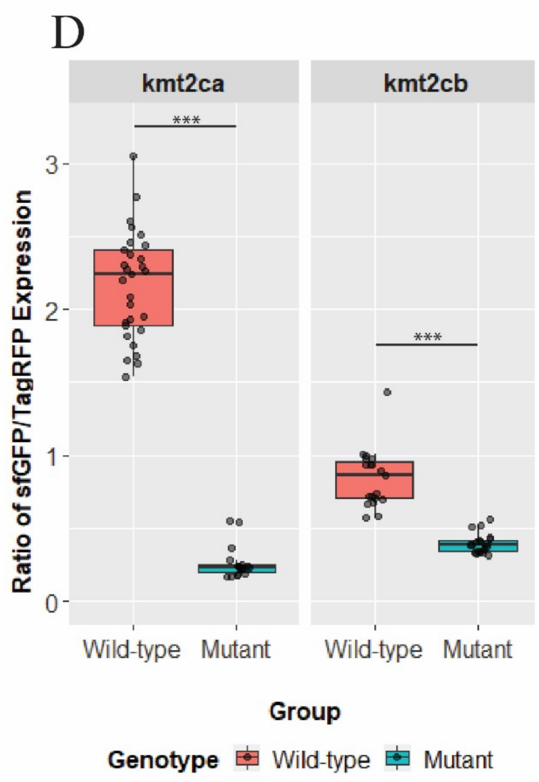
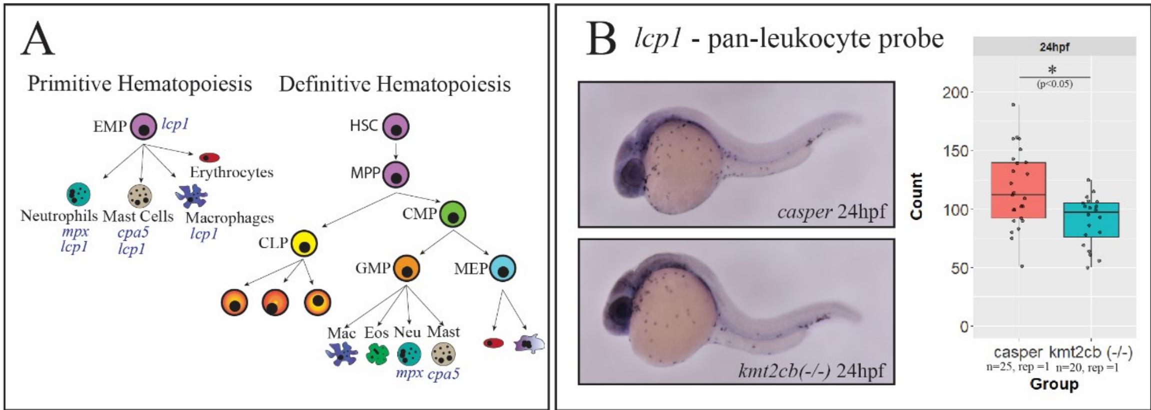
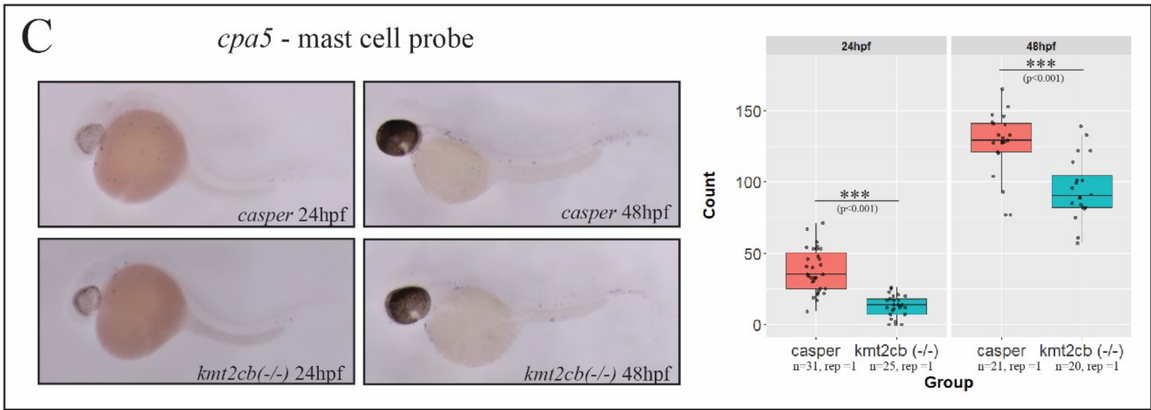

D
kmt2ca
kmt2cb
***
Ratio of sfGFP/TagRFP Expression
0
1
2
3
Wild-type
Mutant
Group
Genotype
Wild-type
Mutant

Figure 4.5: Expression of mutation reporters for *kmt2ca* and *kmt2cb* indicate efficient frameshift mutations. A) Schematic showing the cloning process of the mutation reporter vectors. **B)** Experimental timeline. **C)** Representative images of 10-13 somite stage embryos injected with the indicated mutation reporter RNA. Embryos injected with *kmt2ca-mut* and *kmt2cb-mut* reporter RNA have decreased expression of GFP as compared to *kmt2ca-wt* and *kmt2cb-wt*, respectively **D)** Fluorescent expression was quantified by measuring the average fluorescent intensities of both channels within the somite region and adjusting for background levels of fluorescence. Ratios of sfGFP intensity/TagRFP intensity are plotted and a t-test was performed between wild-type and mutant groups for each gene. Embryos injected with *kmt2ca* mutant (n=18) compared to wild-type (n=29) reporter RNA had a large reduction in sfGFP expression ($p<0.001$). Whereas embryos injected with *kmt2ca* mutant (n=18) reporter RNA had a more modest but still significant ($p<0.001$) reduction in sfGFP expression compared to wild-type injected embryos (n=29). Injections were performed once.

4.6: *Kmt2cb-/-* Embryos have Decreased Numbers of Neutrophils, Mast Cells, and other Leukocytes

To assess how the loss of *kmt2cb* impacts both primitive and definitive hematopoiesis I used WISH techniques that employ genetic markers to assess the abundance of myeloid cell lineages at 24hpf and 48hpf, respectively (**Figure 4.6A**). At the time these studies were performed, *kmt2ca(-/-)* mutants had not yet reached reproductive maturity and so only *kmt2cb(-/-)* embryos were assessed. I first used pan-leukocyte marker, *lcp1*, and found that *kmt2cb(-/-)* embryos had reduced numbers of myeloid cells as compared to wild-type at 24 hpf. (**Figure 4.6B**). At 48 hpf the number of *lcp1* positive cells in both *kmt2cb(-/-)* and wild-type larvae were so large that quantification was unreliable. Next, using marker *cpa5*, I found that *kmt2cb(-/-)* mutant larvae had decreased numbers of mast cell at both 24 and 48 hpf (**Figure 4.6C**). To assess the neutrophil population, I used the *mpx* probe and found reduced expression in *kmt2cb(-/-)* mutants at 24 hpf, but not 48 hpf when compared to wild-type larvae (**Figure 4.6D**). These preliminary results indicate that the loss of *kmt2cb* impedes myelopoiesis, but additional replicates and analysis of HSC and progenitor populations are needed. Assessing changes to lymphoid populations can be difficult in larval fish as T-cells and B-cell populations are not established until 7dpf and 21dpf, respectively, and would be better analyzed in adult fish (Page et al., 2013).


A
Primitive Hematopoiesis
Definitive Hematopoiesis
EMP
lcp1
Erythrocytes
Neutrophils
mpx
lcp1
Mast Cells
cpa5
lcp1
Macrophages
lcp1
HSC
MPP
CMP
CLP
GMP
MEP
Mac
Eos
Neu
Mast
mpx
cpa5
B lcp1 - pan-leukocyte probe
casper 24hpf
kmt2cb(-/-) 24hpf
24hpf
*
(p<0.05)
Count
casper
kmt2cb (-/-)
n=25, rep =1
n=20, rep =1
Group


C
cpa5 - mast cell probe
casper 24hpf
casper 48hpf
kmt2cb(-/-) 24hpf
kmt2cb(-/-) 48hpf
24hpf
48hpf
***
(p<0.001)
Count
casper
kmt2cb (-/-)
n=31, rep =1
n=25, rep =1
n=21, rep =1
n=20, rep =1
Group

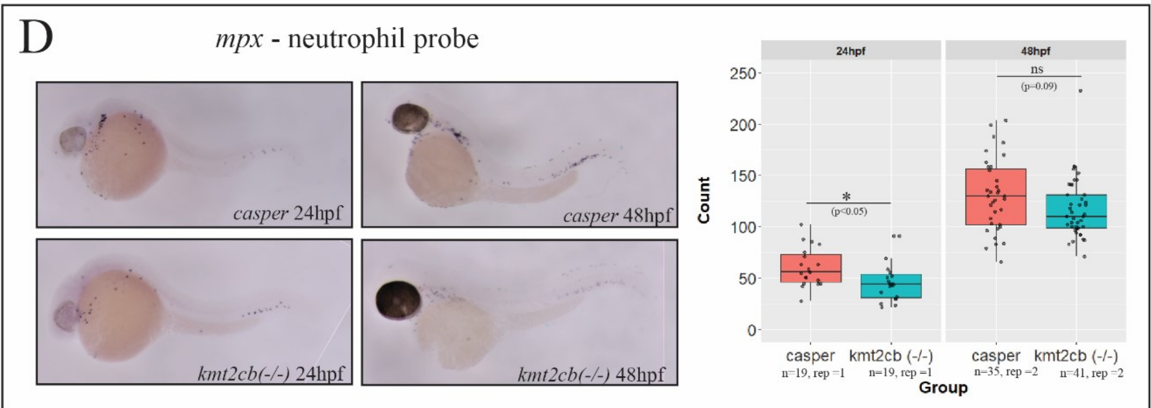

D
mpx - neutrophil probe
casper 24hpf
casper 48hpf
kmt2cb(-/-) 24hpf
kmt2cb(-/-) 48hpf
24hpf
48hpf
*
(p<0.05)
ns
(p=0.09)
Count
casper
kmt2cb (-/-)
n=19, rep =1
n=19, rep =1
n=35, rep =2
n=41, rep =2
Group

Figure 4.6: *kmt2cb(-/-)* zebrafish larvae demonstrate reduced myelopoiesis during primitive hematopoiesis and reduced numbers of mast cells, but not neutrophils during definitive hematopoiesis. A) Gene markers used for whole-mount in-situ hybridization mark various myeloid lineages during primitive hematopoiesis (24 hpf) and definitive hematopoiesis (48hpf). **B)** *kmt2cb(-/-)* larvae (n=20) have reduced numbers of myeloid cells compared to wild-type (n=25) during primitive hematopoiesis ($p<0.05$). **C)** *kmt2cb(-/-)* larvae (n=25) have reduced numbers of mast cells compared to wild-type (n=31) during primitive ($p<0.001$) and definitive (n=35 for wild-type and n=20 for *kmt2cb(-/-)*) ($p<0.001$) hematopoiesis. **D)** *kmt2cb(-/-)* larvae (n=19) have reduced numbers of neutrophils compared to wild-type (n=19) during primitive ($p<0.05$) but not definitive (n= 35 for wild-type and n=41 for *kmt2cb(-/-)*(p= 0.09) hematopoiesis. Two replicates were performed for *mpx* probe at 48hpf, one replicate was performed for all other groups.

4.7 Summary of Findings

In summary, I first assessed the level of conservation between human *KMT2C* and zebrafish homologs *kmt2ca* and *kmt2cb* and found these genes highly conserved at both DNA and protein levels, including domains important to the function of the genes (**Figure 4.2**). Next, I assembled a profile of *kmt2ca* and *kmt2cb* expression during early zebrafish development and found that expression of both genes was diffuse and not organ specific, and expression levels of either gene varied during the first 5 days of development (**Figure 4.3**). Given these findings, I opted to generate *kmt2ca* and *kmt2cb* mutant zebrafish lines and used CRISPR/Cas9 technology to generate frameshift mutations early in either gene (**Figure 4.4**). In the absence of a zebrafish specific kmt2ca or kmt2cb antibodies, I used a previously developed fluorescent reporter assay (Prykhozhij et al., 2017) to test the effectiveness of either mutation to produce null proteins (**Figure 4.5A & 4.5B**). I found that both *kmt2ca* and *kmt2cb* frameshift mutations were effective in reducing the level of sfGFP fluorescence (**Figure 4.5C & 4.5D**), indicating the mutations produced null proteins. Further, I used an existing algorithm (Nishikawa et al., 2000) to identify the likelihood of alternative translational start sites downstream from the mutations I generated. The algorithm identified translational start sites that would produce a truncated kmt2ca protein that contained important interacting and catalytic domains with a reliability of 53% and similar truncated kmt2cb protein with a low reliability of 26%. Lastly, I assessed the abundance of different myeloid blood cell populations in *kmt2cb(-/-)* mutants at 24 and 48 hpf and found reduced numbers of neutrophils and mast cells at 24 hpf compared to wild-type (**Figure 4.6C & 4.6D**). At 48 hpf only the number of mast cells and not neutrophils were reduced in *kmt2cb(-/-)* mutants, indicating that the loss of *kmt2cb* may have a greater impact on myelopoiesis during primitive than definitive blood development.

Chapter 5 — Discussion

5.1 Overview of Thesis

Infant leukemia (IL) is an aggressive disease that has a particularly dismal prognosis (Brown, Pieters, & Biondi, 2019). IL is generally characterized by the presence of more immature leukemic cells compared to older pediatric counterparts and a high prevalence of *MLL-r* (Brown, 2013). Together, these characteristics coupled with the unique physiology and critical developmental state of infants, makes the treatment of IL especially difficult (Dreyer et al., 2011; Pieters et al., 2007; Reaman et al., 1999; Salzer et al., 2012). *MLL-fusion* genes have largely been characterized as the driving mutations of IL, however it is not fully understood how the identity of the partner gene in *MLL-fusions* impacts the disease phenotype or the role of additional genetic factors in IL development (Drynan et al., 2005; Rowe et al., 2019; Valentine et al., 2014). To better understand the complex disease biology of IL, this thesis largely focused on generating new tools that address gaps in our understanding of *MLL-r* disease biology and provide an effective platform for the rapid screening of drugs to identify new potential therapies. Ease of genetic manipulation and the optical clarity of larvae makes the zebrafish and advantageous platform for the generation of complex genetic models and *in vivo* drug screening.

In chapter 3, I took the first steps in developing inducible Cre/lox zebrafish models that express the human *MLL-AF9* and *MLL-ENL* fusion genes. I generated three *CreERT2* transgenic zebrafish lines that express *CreERT2* either ubiquitously, within myeloid and lymphoid progenitor populations, or within myeloid progenitor populations only. A fourth line, expressing *CreERT2* within HSCs, requires further characterization. I also generated, three 'floxed' transgenic reporter zebrafish lines that when crossed with a *CreERT2* expressing line and when treated with 4-OHT, would switch from expressing BFP (blue) to expressing mCherry (red) and either *MLL-AF9*, *MLL-ENL*, or no *MLL-fusion* gene. Although the Cre-lox two-transgene system, was validated *in vivo* using chimeric animals, I was unable to generate stable transgenic 'floxed' reporter fish that could successfully undergo Cre/lox recombination and express MLL-AF9 or MLL-ENL. Rather, the only functional floxed line was *ubi:LBL-P2A-mCherry*, which was generated as a control and does not express an MLL-fusion. Generating these models posed a number of significant

challenges, and as such the discussion of this model will largely focus on tool creation methodology, limitations and alternative strategies.

Murine models have highlighted the inability of MLL-fusions alone to reproduce the aggressive disease found in patients. This indicates that additional genetic factors, such as recently identified germline variation in *KMT2C*, likely contribute to IL development. In chapter 4, I showed zebrafish *KMT2C* homologs: *kmt2ca* and *kmt2cb,* are significantly conserved with humans and generated *kmt2ca* and *kmt2cb* loss-of-function zebrafish models. I presented preliminary phenotypic data that suggested the loss of *kmt2cb* may result in decreased numbers of WBCs during primitive hematopoiesis. I will discuss these findings as well as provide additional observations, while acknowledging that subsequent replicates and further experimentation are needed to increase the confidence of the phenotypic data presented.

5.2 Rationale For Tissue Specific CreERT2 Transgenic Fish

Within this Cre/Lox system, the expression of the *CreERT2* transgene dictates the blood cell lineage for which the *MLL-fusion* gene will be expressed **(Figure 3.2).** I chose to generate transgenic lines expressing CreERT2 under the control of the following promoters, *Cd45* (myeloid & lymphoid progenitors), *pu.1* (myeloid progenitors), and *runx1+23* (HSCs), because these promoters were readily available within the lab. Notably, lymphoid specific expression of *MLL*-fusions would be most relevant for the downstream use of this model as a drug screening platform, given that infant MLL-r ALL has worse outcomes compared to infant MLL-r AML. I decided not to generate a *rag2* driven lymphoid specific CreERT2 line because while *rag2* is expressed in T- and B-cell progenitor populations, it is unclear if it is expressed at an early enough stage to represent the pro-B-cell leukemia seen in IL (Liu et al., 2017). Furthermore, existing zebrafish models using *rag2* to drive oncogenic fusion genes, develop solely T-cell acute leukemias (Chen et al., 2007; Feng et al., 2010; Gutierrez et al., 2011; Langenau et al., 2003). Expression of *rag2* begins at 72 hpf in the zebrafish, but the first mature B-cells do not appear until roughly 3 weeks post-fertilization (Page et al., 2013). In contrast, T-cells are evident at 9-10 dpf, and the earlier constitution of T-cell populations may promote the

development of T-cell disease over that of B-cell in transgenic models where expression of the oncogenic gene is induced during larval development (Page et al., 2013). Over 95% of IL ALL are B-ALLs (Brown, Pieters & Biondi, 2019). To date only one B-cell specific zebrafish promoter, cd79b promoter (Liu et al., 2017) has been developed for transgene expression, but was not readily available for use in this study. Development of additional blood cell specific CreERT2 transgenic lines should focus on B-cell expression using the cd79b promoter.

5.3 Limitations of Cre/Lox System in Zebrafish

The spatiotemporal control that the Cre/lox system provides is very advantageous for disease modelling and developmental studies that involve lineage tracing, making it popular within the zebrafish research community. However, the generation of new models using Cre/lox is laborious and suffers from a few key limitations.

5.3.1 Existing Cre/Lox lines Are Needed to Efficiently Generate New Models

Generating new Cre/lox zebrafish lines is a particularly labour-intensive process because multiple founders for each line need to be screened for the appropriate expression of each transgene. Once transgenic lines with appropriate transgene expression are established, each line must then be functionally tested for the ability to induce (for *CreERT2* line) or undergo recombination (for floxed lines).

Screening for appropriate expression was especially important for the *CreERT2* lines because these lines confer tissue specificity to the model. I needed to eliminate founder lines with either off-target expression or insufficient levels of expression. Using Tol2 techniques to generate transgenic lines, will often result in the inheritance of multiple insertions of the transgene by F1 progeny and therefore display varying levels and patterns of expression (Kawakami, 2007). The regulatory environment surrounding transgene insertion sites can lead to substantial variability in transgene expression, referred to as positional effects, and can lead to non-specific expression or silencing (Jaenisch et al., 1981; Wilson et al., 1990; Rossant et al., 2011). In order to effectively identify transgene positional effects that negatively impact expression, multiple founders for each line should

be outcrossed until they carry only a single transgene insertion, evidenced by a 50% transmission rate to offspring (Felker & Mosimann, 2016). This step can take multiple generations, spanning over a year's time, and can place extra demands on space within aquatics facilities. Although tagging *CreERT2* with a fluorescent protein would make it easy to screen new *CreERT2* drivers, it is generally avoided as it would limit the universal use of such driver to only 'floxed' transgenes that lack that fluorescent protein. Additionally, as the number of fluorescent proteins increases (floxed reporters typically contain two fluorescent proteins), so does the risk that the spectrum of emission for a fluorescent protein will bleed into the spectrum of another, which makes quantifying the fluorescence from individual proteins difficult.

I assessed CreERT2 expression within two founder lines for each of the *cd45:CreERT2*, *pu.1:CreET2*, and *runx1+23:CreERT2* transgenes and only one founder for the *ubi:CreERT2* transgene. This was largely due constraints on time and space, and it would be advisable to generate new CreERT2 lines in sequence rather than in parallel in the future. All the transgenic *CreERT2* larvae I assessed, using in-situ hybridization (**Figure 3.3**), had non-specific staining within the head region. This likely indicates some degree of CreERT2 expression in the CNS and could be due to positional effects at the transgene loci or presence of multiple transgene insertions and the need for further out-crossing. Additionally, I failed to identity any CreERT2 expression within hematopoietic tissues of *runx1+23:CreERT2* larvae (**Figure 3.3D**), despite evidence of recombination during functional studies (**Figure 3.4.1D**). I suspect at 80 hpf, the larvae were not sufficiently permeabilized, for the identification of such a small cell population (8-10 cells) and further optimization of my ISH procedure is needed. However, even with optimization, assessment by ISH will only help rule out unsuitable transgenic lines, and functional studies are still needed to validate the efficacy of new CreERT2 lines to induce recombination. The floxed reporter lines I generated were screened adequate levels of ubiquitous BFP expression using fluorescent microscopy or the BioSorter, therefore eliminating the need for ISH analysis.

The gold standard for characterizing new Cre/lox lines is to cross them to established lines and assess recombination (Felker & Mosimann, 2016). This method is far more

efficient than the laborious process of validation with in-situ hybridization and assesses the true function of newly generated lines. To this end, I generated the *ubi:LBL-P2A-mCherry* transgenic line and *ubi:CreERT2* lines. Crossing the *ubi:CreERT2* and *ubi:LBL-P2A-mCherry* lines and treating the fish with 4-OHT induced efficient recombination and ubiquitous expression of mCherry. New *CreERT2* transgenic lines being developed in the Berman Lab can now be tested against *ubi:LBL-P2A-mCherry* and similarly new floxed lines (such as the inducible *MLL-AF9* and *MLL-ENL* lines) can be tested against *ubi:CreERT2* expressing fish.

A *ubi:CreERT2* transgenic line was previously established in the Zon Lab (Mosimann et al., 2011), however, due to the challenges with importing fish and the necessary quarantine of imported animals, I decided to generate a similar line for this project.

5.3.2 Recombination at Lox Sites Requires Single Insertion Transgenesis

I demonstrated the successful cloning and assembly of each of the *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* transgenes by injecting them into *ubi:CreERT2* fish and induced recombination with 4-OHT, which produced a mosaic pattern of mCherry expression (**Figure 3.4.1B**). However, despite establishing stable transgenic lines that expressed either transgene, I was unable to induce recombination in F2 larvae (**Figure 3.4.2A**). There are two possible explanations for this finding. Firstly, it has been proposed that unfavorable histone modification and chromatin effects can limit CreERT2 access to LoxP sites, reducing the efficiency of recombination (Mosimann & Zon, 2011). Gene silencing due to chromatin effects and methylation have long been established (Mutskov & Felsenfeld, 2003), although there has not yet been experimental evidence to show how these changes affect the access of CreERT2 to Lox sites. It is unclear how this relates to the development of *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* transgenic zebrafish, given they readily express the transgenes as evidenced by BFP expression. Interestingly, it is thought that CpG-enriched regions in transgenes may be targeted by CpG binding proteins and recruit repressor complexes that contain histone deacetylases and histone methyltransferases, causing changes to chromatin structure (Bird & Wolffe, 1999). MLL contains a 128bp CG rich region only 58 nucleotides

downstream from its ATG start site, that caused several challenges during the cloning process (see methods section), and ironically may also promote histone modification that could inhibit recombination of the transgene.

Secondly, the integration of linear DNA concatemers during transgenesis that result in tandem arrays of multiple floxed transgenes can cause unproductive recombination to occur (Mosimann & Zon, 2011). This is because recombination can occur at any two lox sites and could occur between lox sites of two different cassettes that are in tandem. One example of this would be the removal of *MLL-AF9-P2A-mCherry* and the *ubi* promoter of the adjacent cassette. This type of recombination would not remove the *BFP-STOP* sequence needed for the expression of the *MLL-fusion* and *mCherry*. It has been reported that more than 10 different founders for a given floxed reporter line need to be screened before a line with efficient recombination is established (Felker & Mosimann, 2016). Due to time restraints, and the Berman Lab moving institutions, I was only able to screen two different founders for each of the floxed reporter lines.

Efforts to improve the efficiency of recovering functional floxed lines focus on single insertion transgenesis (Felker & Mosimann, 2016). Using Tol2 transgenesis techniques, this is achieved by performing several outcrosses of each founder line (as outlined above). However, this process is very labour, space, and time intensive and alternative transgenesis strategies that promote single insertions, such as the use of meganucleases would likely streamline this process.

The use of meganuclease I-SceI to generate transgenic zebrafish is a proven and efficient method (Thermes et al., 2002). I-SceI is an endonuclease that recognizes an 18-bp sequence with high specificity and induces a double strand break with a 4bp overhand (sticky ends) on either side (Monteilhet et al., 1990). Transgenes are designed with flanking I-SceI sequences that allow the integration of the transgene into the host genome when injected alongside *I-SceI* mRNA or protein, using universal microinjection techniques (Babaryka, Kuhn & Koster, 2009). I-SceI promotes single transgene insertions in two ways. Firstly, the 18bp target sequence is expected to occur only once in $7x10^{10}$ bp of random sequence, limiting the number of sites for where transgenesis can occur in the host genome (Grabher, Joly, & Wittbrodt, 2004). This is in direct contrast to use of transposase

(Tol2) where integration is random (Kawakami, 2005). Secondly, it is believed that I-SceI remains linked to the cleaved recognition sites and protects linear monomers from degradation while also counteracting concatamer formation (Colleaux et al., 1988; Thermes et al., 2002). One study identified seven F1 founders, out of 60 embryos injected with the *her3:GalTA* and *I-SceI* mRAN. Of these seven founders, six carried single insertions as evidenced by transgene transmission rates to F2 larvae (Babaryka, Kuhn & Koster, 2009). This demonstrates the efficiency of the I-SceI technology and its particular use for single insertion transgenesis of Cre/Lox transgenes. Interestingly, the cre/lox inducible *NU98-HOXA9* zebrafish model of AML was generated in the Berman lab using I-SceI technology (Forrester et al., 2011). One drawback of the I-SceI system is the low germline transmission rates achieved, typically only reaching transmission rates of 10% (Collas and Alestrom, 1998, Culp et al., 1991, Lin et al., 1994, Stuart et al., 1988, Stuart et al., 1990, Tanaka and Kinoshita, 2001). This may also inadvertently increase the number of animals needed to be screened and is a major reason why Tol2 techniques is more commonly used for zebrafish transgenesis today.

5.4 Permanent Recombination of Cre/lox System Provides Advantages Over Other Inducible Systems

Alternative inducible systems to Cre/lox that allow spatiotemporal control of transgene expression, such as *Gal4-UAS* (Scheer et al., 2001) and *LexPR* (Emelyanov & Parinov, 2008), have been developed. Both *Gal4-UAS* and *LexPR* employ a two-transgene system that uses the production of a trans activator by one transgene to stimulate the expression of cargo genes (genes of interest) in the other. However, neither of these systems offer a suitable alternative to Cre/lox for the development of an IL model.

In the *Gal4-UAS* system, transgenic zebrafish that express the *Gal4* trans activator in a tissue specific manner are crossed to transgenic fish that contain the cargo genes under control of the Upstream Activating Sequence (UAS). In double transgenic fish that carry both transgenes, the expression of cargo genes will only be activated in tissues or cells that express *Gal4* (Davison et al., 2007, Scott et al., 2007). Although this system offers the same control as Cre/lox and has been widely utilized amongst zebrafish models, reports suggest Gal4 expression can be toxic at high levels and causes developmental defects (Habets et

al., 2003, Kramer and Staveley, 2003, Scott et al., 2007). Given this, the *Gal4-UAS* system can cause additionally challenges for zebrafish disease modelling, and toxicity related to this genetic system could confound phenotypes seen in MLL-fusion expressing zebrafish.

The *LexPR* system is similar to that of *Gal4-UAS* but operates in a chemically induced manner. The *LexPR* trans activator is expressed under the control of a tissue specific promoter, while the cargo genes are controlled by the *LexA* operator in either a downstream cassette or separate transgene (Emelyanov & Parinov, 2008). Only in the presence of mifepristone will the LexPR activator bind the *LexA* operator and induce expression of the cargo genes (Emelyanov & Parinov, 2008). However, the activation of the *LexA* operator is not a permanent event, and once mifepristone is removed, expression of the cargo genes will be turned off (Emelyanov & Parinov, 2008). Such a system is not suitable for the long-term expression of fusion oncogenes, like MLL-fusions, and the use of the *LexPR* system to develop an *MLL-r* zebrafish model would limit experimentation to the larval phase. Additionally, the constant presence of mifepristone within the fishes water may limit the utility of the model as a drug screening platform due to possible interactions between mifepristone and the drugs being tested. Using the Cre/lox system, *MLL-fusions* are expressed continuously after a single short treatment with 4-OHT, and the progression of a leukemia disease can be assessed in adult zebrafish.

One of the major advantages of developing a model that expresses *MLL-AF9* or *MLL-ENL* under a Cre/lox system, is the ability to maintain healthy populations for breeding and only induce expression of the potent oncogenes in experimental fish. To date, one non-inducible MLL-r transgenic zebrafish model has been reported to be in development. Saberi and colleagues established chimeric zebrafish that expressed MLL-AF9 after embryos were injected with a *runx1+23:MLL-AF9-IRES-eGFP* or *runx1+23:MLL-AF9-IRES-mCherry* transgene (Saberi et al., 2019). They found that 30% of adult chimeric fish showed signs of a leukemia-like disease as evidenced by expansion of hematopoietic progenitor cell populations and enlarged kidneys and splenomegaly. While this line, could be applied to phenotype-based drug screening, the ability to express *MLL-AF9* in only one blood cell lineage would not be useful in investigating the effects of *MLL-fusion* identity

on IL phenotype. It is not clear if this group has successfully established a stable transgenic line that expresses MLL-AF9.

5.5 Drug Screening Application of the Fluorescent IL Model

From the beginning, one major goal of this project was to develop a zebrafish model of IL that would provide an efficient platform for phenotype-based *in vivo* drug screening to identify new potential therapies for IL. It was for this reason that the transgene system developed would signify MLL-AF9 or MLL-ENL expression by switching from BFP to mCherry expression (**Figure 3.2**).

A phenotype of increased mCherry fluorescent white blood cells in the MLL-AF9 or MLL-ENL expressing larvae would provide a robust readout that can be exploited by fluorescent microscopy and robotic automation to identify compounds that restore normal blood cell ratios. Medium to high-throughput chemical screens could be carried out by arraying induced larvae that express either MLL-AF9 or MLL-ENL into well of a 96-well plate. Candidate compounds or chemicals from a drug library can be added to each well and administered through submersion therapy. After being treated with the candidate compounds or drug library, mCherry fluorescence within larvae from each well can be measured using the Union Biometrica Biosorter and LP sampler. The LP sampler removes fish from each well and passes them individually through a flow cell in the Biosorter, which records the fluorescence levels of each fish. The larvae can then be deposited back into their original well and possibly screened again at an additional timepoint. Using this system, MLL-AF9 or MLL-ENL expressing larvae that show decreased levels of mCherry fluorescence after the drug treatment, as compared to an untreated control, would indicate a 'hit' (**Figure 5.5**). Such a compound could then be followed up with additional experimentation to validate its therapeutic properties. Using this model to screen libraries that contain FDA and Health Canada approved drugs, like the Sigma LOPAC 1280 library, would facilitate translation of promising "hits" to Phase I clinical trials as a re-purposed therapeutic compound. Zebrafish screens considers drug toxicity and bioavailability as well as effects due to the tissue microenvironment and therefore it is an ideal platform for the consolidating a large drug library into a few promising agents (Carradice & Lieschke

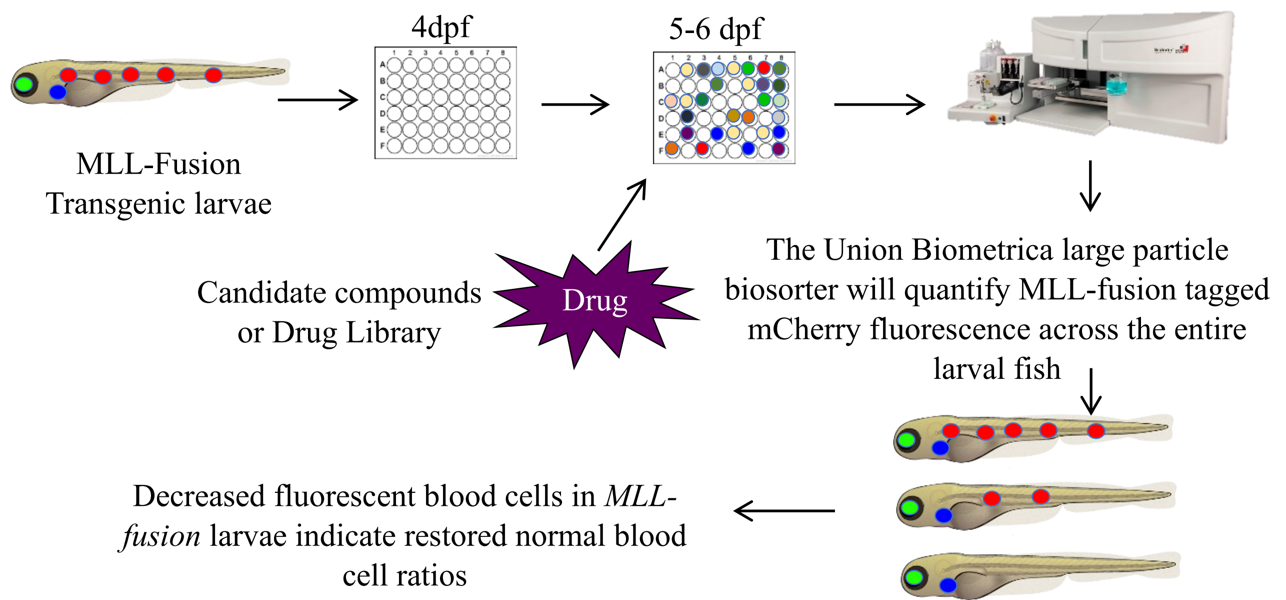


Figure 5.5: Drug screening in the zebrafish can be partly automated using a Biosorter to detect levels of fluorescence. Induced larvae that express an MLL-fusion (either MLL-AF9 or MLL-ENL) can be arrayed in 96-well plates and treated with either candidate compounds or a drug library. After drug treatment, larvae are passed through the Biosorter, which measures the level of MLL-fusion tagged mCherry. Drugs that result in a decrease of mCherry fluorescence as compared to controls, would indicate a 'hit' compound that could then undergo further testing.

2008; Zon & Peterson 2005). However, it is not a replacement for mammalian-based testing were the efficacy and safety of a drug can be better evaluated for human use.

5.6 Understanding *kmt2ca* and *kmt2cb* Expression Patterns

As a result of an extra whole genome duplication event, as compared to humans, zebrafish have two homologs of *KMT2C*: *kmt2ca* and *kmt2cb.* Before establishing a knockout model of either gene, I wanted to first assess the expression of both genes to determine that they were significantly expressed in hematopoietic tissues. This was important as I would be investigating the roles of these genes in hematopoietic development. I also wanted to identify if *kmt2ca* and *kmt2cb* had different expression patterns, which might indicate that they had unique functions. The first step was to perform in-situ hybridization on wild-type larvae using probes specific to *kmt2ca* and *kmt2cb*. Staining for either gene was diffuse and rather non-specific with generally more staining in the head/CNS region and less staining within the tail of the fish (**Figure 4.2A**). Expression of both genes did not appear organ or tissue specific. This pattern of expression is congruent with the expression *Kmt2c* in mice, and *KMT2C* in humans. Whole-mount in-situ hybridization (WISH) studies in the developing mouse embryo showed high levels of expression within the CNS, specifically within the telencephalon, ventral neural tube, and somites (Brun et al., 2006). PCR amplification of cDNAs from different human tissues revealed *KMT2C* expression within the heart, brain, lung, liver, pancreas, peripheral blood, testes and thymus (Ruault et al., 2002). Furthermore, gene expression databases that compile gene expression data from many human studies, such as the EMBL Expression Atlas (https://www.ebi.ac.uk/gxa/home) and The Human Protein Atlas (https://www.proteinatlas.org/) indicate *KMT2C* is highly expressed within regions of the brain, bone marrow and heart.

I analyzed previously published RNA sequencing data from several timepoints within early zebrafish development (White et al., 2017), and found that the expression of *kmt2ca* peaks around 6 hpf then subsequently, expression of *kmt2cb* takes over and peaks at approximately 12 hpf (**Figure 4.2B**). This may suggest that *kmt2ca* is maternally deposited in the zebrafish embryo and may play a critical role during the first few hours of

development. The KMT2 protein family has been well established as regulators of the expression of *HOX* genes that are important for embryogenesis, body planning, and organ development. However, no obvious morphological defects were observed during this study in either *kmt2ca-/-* or *kmt2cb-/-* developing embryos, possibly due to compensation between homologs. Given large-scale transcriptome changes during the differentiation of embryonic stem cells is mediated by KMT2-family and COMPASS proteins (Ang et al., 2011), and the significant expression of *kmt2ca* and *kmt2cb* in the CNS, it is likely that the expression of these genes play critical roles in the development of these structures. Not much is known about the role of *KMT2C* in nervous system development, however, Mll-/- embryonic mice showed abnormal neural patterning and extensive fusion of spinal ganglia (Yu et al., 1998). It was observed that *kmt2ca-/-;kmt2cb-/-* double mutant zebrafish, generated during this study, displayed abnormal swimming patterns and constantly swam on a sideways tilt, which may suggest defects in the development of their sensory systems. Abnormal swimming patterns were not observed within either of the single mutants (*kmt2ca-/-* or *kmt2cb-/-)*.

5.7 Rationale and Validation of *kmt2ca* and *kmt2cb* Mutant Zebrafish Model

I decided to establish *kmt2ca* and *kmt2cb* mutant zebrafish as a tool to investigate the role of *KMT2C* in blood development and how mutations in *KMT2C* may impact IL development. To do this, I used CRISPR/Cas9 technology to generate frameshift mutations early within the coding sequence and before the first PHD domain of either *kmt2ca* or *kmt2cb* (**Figure 4.4**) Using a previously established assay, I showed that both frameshift mutations produce functionally null proteins (**Figure 4.5**). This knock-out strategy was chosen for its simplicity and because a null mutation would likely result in a more robust phenotype than a hypomorphic mutation generated within the SET domain. This is evidenced in mice as the ubiquitous knock-out of *Kmt2c* resulted in death shortly following birth (Lee et al., 2013), whereas mice that harboured deletions within the SET domain of *Kmt2c* demonstrated only partial embryonic lethality (Lee et al., 2008). These mice also displayed evidence of stunted growth and decreased white adipose tissue (discussed below). However, IL patients are not deficient in KMT2C, but rather frequently possess

germline missense or nonsense hypomorphic mutations, typically within C-terminal regions and include the FYRN, FYRC and SET domains (Druley, unpublished; Valentine et al., 2014). Given this, the *kmt2ca* and *kmt2cb* knockout model I created may be a better model of mutations found in adult AML, like the deletion of chromosome 7q (Dohner et al., 1998; Kotini et al., 2015; Kuhn et al., 2012; Ruault et al., 2002), than the hypomorphic mutations found in IL (Druley, unpublished; Valentine et al., 2014). In fact, 25% of somatic *KMT2C* mutations found in cancers have frameshift or non-sense mutations with PHD domains that likely lead to truncated proteins (Rao & Dou, 2015), similar to the zebrafish model I generated. Importantly, in **Figure 4.2**, I demonstrated that zebrafish homologs *kmt2ca* and *kmt2cb* show significant conservation with human *KMT2C*, especially within major functional domains. This provides the opportunity to generate patient-specific hypomorphic mutations in the future that would better represent mutations found in IL.

Due to the difference in efficiencies of *kmt2ca* and *kmt2cb* frameshift mutations to produce null proteins (**Figure 4.5**), I used a previously established algorithm (Nishikawa et al., 2000) to identify possible alternative translational start sites downstream of each mutation. Although I did not identify any downstream translational start sites that would account for the difference in sfGFP:TagRFP ratios between embryos injected with the *kmt2ca* and *kmt2cb* mutation reporter RNA, the program did identify possible translational start sites that existed outside of the CDS cloned for the reporter assay. Analysis of *kmt2ca* returned a 5' truncated protein that would begin in exon 11. A similar 5' truncated protein was predicted for *kmt2cb*, that would begin in exon 18. Interestingly, the location of these alternative translational starts resembles the location of an internal promoter identified in *Mll* (*Kmt2a*) that gives rise to a transcript that starts in exon 12, within a *Mll-/-* embryonic fibroblast cell line (Scharf et al., 2007). Translation of the 5' truncated transcript, termed *Mll**, likely starts in either exon, 17, 18, or 19 (Scharf et al., 2007). *Mll-/-* mice that expressed the *Mll** transcript, die in utero, suggesting that the N-terminal domain of MLL is critical to its function (Yu et al., 1995). Given these findings, it is possible that even if alternative in-frame transcripts are initiated downstream of the *kmt2ca* or *kmt2cb* frameshift mutations, the resulting proteins would not maintain normal physiology.

For the further validation of this mutant model, it is necessary to assess genetic compensations by related genes (transcriptional activation). A correlation between levels of mutant mRNA decay and the upregulation of related genes has been demonstrated in several models (El Brolosy et al., 2017; Rossi et al., 2015; Tondeleir et al., 2012) and highlights the need to assess mutant mRNA stability in newly generated models (El Brolosy et al., 2018). In addition, Kmt2c has been shown to compensate for the loss of Kmt2d in murine cell culture, and Kmt2c and Kmt2d have partially redundant functions (Kim et al., 2014; Lee et al., 2008). Therefore, it is likely that *kmt2d* may compensate for the loss of *kmt2ca* and *kmt2cb* in zebrafish and should be investigated alongside compensation between *kmt2ca* and *kmt2cb* within our model using qPCR analysis.

5.8 Assessing Blood Development in *kmt2ca -/-* and *kmt2cb-/-* Zebrafish

To examine the role of *kmt2ca* and *kmt2cb* in blood development, I decided to use established in-situ hybridization techniques and a battery of probes against different hematopoietic lineages to assess the abundance of different blood cell populations (**Figure 4.6A).** However, due to stunted growth and poor fertility (discussed below) of *kmt2ca-/-* and *kmt2cb-/-* zebrafish, I was only able to perform limited experimentation on *kmt2cb-/-* fish. Preliminary results indicated that *kmt2cb-/-* embryos had reduced numbers of white blood cells at 24 hpf (primitive hematopoiesis) **(Figure 4.6)**. Although reduced levels of mast cells (*cpa5)* were also seen at 48 hpf, in *kmt2cb-/-* embryos, the number of neutrophils (*mpx)* and leukocytes (labelled by pan-leukocyte marker (*lcp1*)) were comparable to wild-type at 48 hpf. It is not surprising *kmt2cb* may have a more important role in primitive hematopoiesis than definitive hematopoiesis given it is more highly expressed during this period **(Figure 4.2B)** and the importance of the *kmt2cb* in early embryonic development. Going forward, it will be important to assess populations of HSCs (using gene markers *c-myb* & *runx1*) as well as myeloid progenitor populations (using gene marker *pu.1/spi1*) in order to establish a more complete analysis of any defects in hematopoiesis caused by the loss of *kmt2cb*. Additional replicates are also needed to increase confidence in this data.

I cautiously suspect that the loss of *kmt2cb* induces a block in hematopoietic differentiation, however literature to date has only provided evidence for this hypothesis when examining definitive hematopoiesis. Whereas preliminary evidence from this study saw decreased myelopoiesis during primitive hematopoiesis but not definitive hematopoiesis, as expression of *mpx* (neutrophils) was comparable between *kmt2cb-/-* and wild-type at 48hpf **(Figure 4.6D)**. The somatic loss of *KMT2C* is frequently found in adult AML (Dohner et al., 1998; Kotini et al., 2015; Kuhn et al., 2012; Ruault et al., 2002), which is a disease characterized by the expansion of immature myeloid blasts. In mice, the inactivation of *Kmt2c* promotes a shift in hematopoiesis to favour granulocyte/myeloid progenitor cells, which is also typically seen in AML (Arcipowski et al., 2016). Additionally, mono- and bi-allelic deletions of *Kmt2c*, resulted in the modest increase of HSCs and a significant reduction in committed hematopoietic progenitor populations. Interestingly, it was shown that *Kmt2c* mutations sustain the self-renewal capacity of HSCs by dampening interleukin-1 (IL-1) driven myeloid commitment (Chen et al., 2019). Further experimentation is needed to better understand the role of *kmt2ca* and *kmt2cb* in zebrafish hematopoiesis.

5.9 Observational Findings in *kmt2ca-/-* and kmtcb-/- Zebrafish

One challenge that hindered my ability to assess the impacts of *kmt2ca* and *kmt2cb* mutations on blood development, was the poor fertility and developmental delay of the *kmt2ca-/-* and *kmt2cb-/-* fish. Although I was not able to conduct size or mass quantification for this thesis, observationally, both the *kmt2ca-/-* and *kmt2cb-/-* mutant lines took 1-2 months longer than wild-type lines to reach reproductive maturity and successfully produce offspring. Additionally, both mutant lines also took months longer than wild-type lines to reach their full size. These observational findings are consistent with *Kmt2c-/-* mice that show reduced fertility and retarded development (Lee et al., 2009). These mice had significant depletion of white fat and revealed vital roles for *Kmt2c* in adipogenesis (Lee et al., 2009). In-crossing *kmt2ca+/-;kmt2cb+/-* zebrafish and genotyping 5 day old larvae, revealed the presence of *kmt2ca-/-;kmt2cb-/-* double mutants in mendelian ratios. However, when 78 fish from the *kmt2ca+/-;kmt2cb+/-* in-cross were grown to 3 months

of age, no *kmt2ca-/-;kmt2cb-/-* double mutants were recovered. By performing larval fin-clipping to identify *kmt2ca-/-;kmt2cb-/-* double mutants, I was able to grow these fish in isolation and observed abnormal swimming behaviour (swimming on a tilt, and failure to maintain buoyancy) and growth defects more severe than either *kmt2ca-/-* or *kmt2cb-/-* mutants. Several *kmt2ca-/-;kmt2cb-/-* fish have reached adulthood (3 months of age) but it is unclear if these fish are fertile. A lack of genetic compensation between *kmt2ca* and *kmt2cb* likely contributes to this more severe phenotype.

In addition to growth and fertility defects, adult *kmt2ca-/-* show signs of disease. Out of 15 *kmt2ca-/-* fish, 8 died between 7-13 months of age. A few days before death, *kmt2ca-/-* fish showed signs of distention, sloughing of scales, protruding mouth, bulging eyes, as well as difficulty breathing and maintaining buoyancy **(Figure 5.9).** One fish showed obvious signs of hemorrhage, which was also observed in *Kmt2c-/-* mice (Lee et al., 2008). *Kmt2d* is largely responsible for H3K4me1 (activating mark) at myocyte specific enhancers (Lee et al., 2013) and regulates specific programs in heart development (Ang et al., 2016). Given the partial redundancy of *Kmt2c* and *Kmt2d*, I suspect *kmt2ca-/-* fish may have heart defects such as tetralogy of fallot (Borland et al., 2019), or decreased vascular integrity that results in hemorrhage.

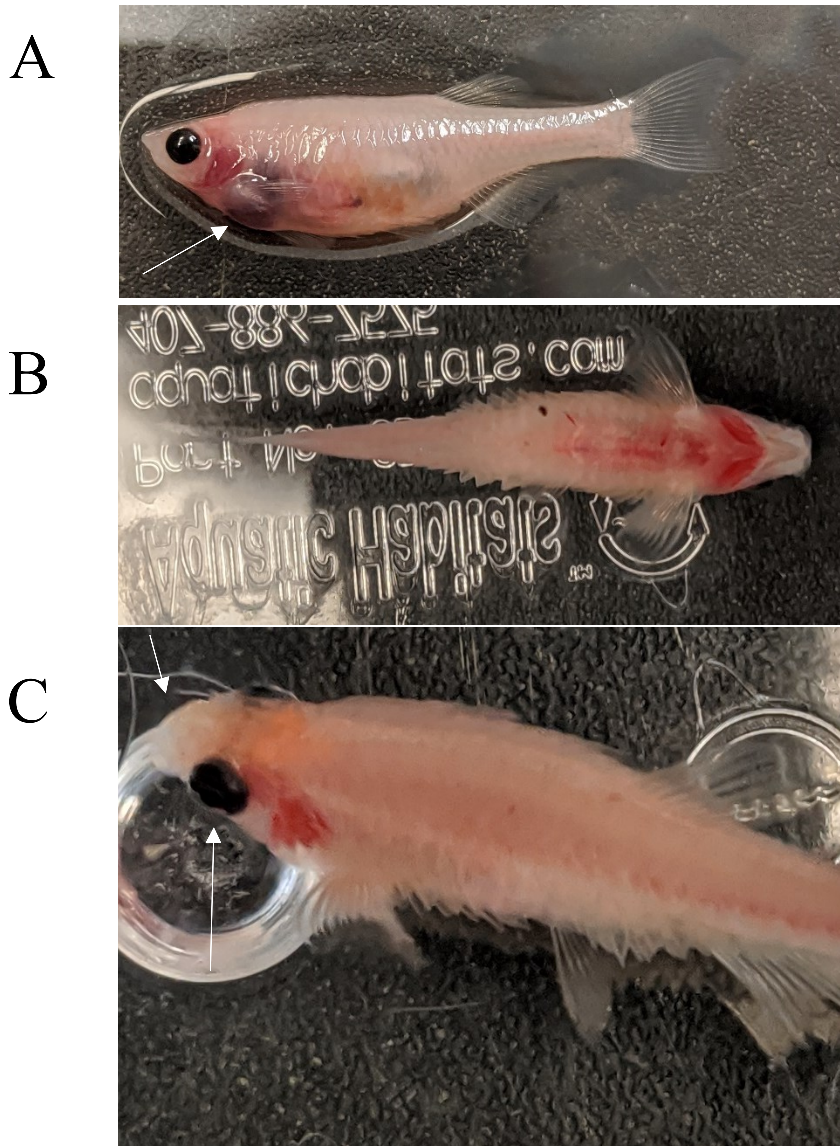


Figure 5.9: Adult *kmt2ca-/-* zebrafish succumb to disease starting as early as ~7months of age. A) Hemorrhage is evident (indicated by arrow) **B &C)** Sloughing of skin/scales is common among sick *kmt2ca-/-* fish. Protruding mouth and bulging eyes are indicated by arrows.

5.10 Limitations of the Zebrafish *KMT2C* Model

One major limitation of using zebrafish to model *KMT2C* mutations, is the presence of two zebrafish homologs. Generating any *KMT2C* mutant model in zebrafish requires the generation of two models – one for each homolog – and a series of breeding steps to generate double mutants that would model the human genetic condition. Additionally, there is no guarantee that *kmt2ca;kmt2cb* double mutants will be viable/fertile and in single mutation lines there is likely compensation by the other homolog. This undoubtably increases the time and validation needed to establish new *KMT2C* mutant zebrafish lines and may limit the utility of this model for disease modelling.

5.11 Future Directions

Despite challenges establishing implementing Cre/lox technology, it remains an advantageous system for developing a zebrafish model of *MLL-r* IL. It offers a modular system with the spatial and temporal control needed to investigate the impact of *MLL-fusion* identity on phenotype and patient outcomes and incorporates fluorescent protein tags that can be utilized for phenotype-based drug screening. Given these advantages and the fact only a limited number of founders for each *MLL-fusion* lines were screened during this study, I suggest future efforts in establishing this model are warranted.

In order to further validate the functionality of the *CreERT2* expressing transgenic lines I created, I would cross each line to *ubi:LBL-P2A-mCherry* transgenic fish, and assess for adequate mCherry expression within the intended tissues post-induction.

To determine if the lack of mCherry fluorescence in 4-OHT treated *ubi:LBL-MLL-AF9-P2A-mCherry* and *ubi:LBL-MLL-ENL-P2A-mCherry* larvae is due to the inability to undergo Cre/lox or due to transcription/translation issues post-recombination, I would cross each line to *ubi:CreERT2* transgenic fish and isolate gDNA and cDNA from 4-OHT treated double-transgenic larvae and use a PCR assay to detect the presence or absence of transgene recombination.

Additionally, the established *ubi:CreERT2* lines now available within the Berman Lab, will make screening additional floxed *MLL-fusion* founders for transgene insertions that allow functional recombination at the F1 stage more efficient. To generate new founder lines, I would consider injecting two cohorts of embryos, using Tol2 mediated transgenesis in one cohort and I-SceI with the other. This would allow comparison between technologies specifically for the insertion of Lox containing transgenes into the zebrafish genome.

I discussed the need to continue WISH assays in *kmt2ca-/-* and *kmt2cb-/-* mutant zebrafish to assess defects in hematopoiesis as well as the use of qPCR to identify any genetic compensation between *kmt2ca*, *kmt2cb,* and *kmt2d*. Due to the severe phenotype that has prohibited *kmt2ca-/-;kmt2cb-/-* double homozygous line from being established and the characterization of *KMT2C* as a haploinsufficient tumour suppressor in AML (Chen et al., 2014), it will be important to also assess blood development within *kmt2ca+/-kmt2cb+/-* double heterozygous mutants. Doing so will identify the mutant line that can best be used to investigate the impact of *KMT2C* mutations on leukemia development. To specifically assess the impact of *KMT2C* mutation in MLL-r IL, *MLL-AF9* and *MLL-ENL* RNA can be injected into *KMT2C* mutant zebrafish embryos and any leukemic phenotypes can be assessed using the same WISH assays and blood cell specific probes used in this study. Furthermore, RNA sequencing and methylation sequencing studies in *KMT2C* mutant zebrafish using whole larvae and adult kidney marrow would provide insight into the global impact due to the loss of zebrafish *KMT2C* homologs and may provide possible explanation for phenotypes observed during this study as well as identify deregulated genetic pathways for further experimentation.

5.12 Impact and Conclusions

This project focussed heavily on developing new tools that can be used to provide insights into IL disease biology and address the contribution of additional genetic mutations in IL development. I took the first steps in developing a Cre/lox inducible MLL-r zebrafish model that can be used to investigate the impact of MLL-fusion identity on IL phenotype (AML vs ALL). Several challenges prohibited the development of a functional MLL-r transgenic model, and additional MLL-fusion expressing transgenic zebrafish need

to be generated in order to identify a functional model. Given the successful generation of the MLL-fusion transgene constructs, I am hopeful that a functional Cre/lox MLL-fusion line will soon be established. With further validation, the *cd45:CreERT2*, *pu.1:CreERT2* and *runx1+23:CreERT2* transgenic lines I created represent exciting new tools for the zebrafish leukemia modeling community. The project sought to examine the impact of germline *KMT2C* mutation on hematopoiesis and IL development by generating loss-of-function models in zebrafish. The *kmt2ca-/-* and *kmt2cb-/-* models can be used in the future to examine the impact of MLL-fusions in the presence of *KMT2C* mutations. Together the development of these new zebrafish models represent exciting progress towards powerful tools in understanding IL disease latency and represents a potential platform for high throughput drug screening.

References

Abramovich, C., & Humphries, R. K. (2005). Hox regulation of normal and leukemic hematopoietic stem cells. *Current opinion in hematology*, *12*(3), 210-216.

Amatruda, J. F., & Zon, L. I. (1999). Dissecting Hematopoiesis and Disease Using the Zebrafish. *Developmental Biology*, *216*(1), 1–15. https://doi.org/10.1006/dbio.1999.9462

Andersson, A. K., Ma, J., Wang, J., Chen, X., Gedman, A. L., Dang, J., Nakitandwe, J., Holmfeldt, L., Parker, M., Easton, J., Huether, R., Kriwacki, R., Rusch, M., Wu, G., Li, Y., Mulder, H., Raimondi, S., Pounds, S., Kang, G., … Downing, J. R. (2015). The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature Genetics*, *47*(4), 330–337. https://doi.org/10.1038/ng.3230

Anderson, M., Fair, K., Amero, S., Nelson, S., Harte, P. J., & Diaz, M. O. (2002). A new family of cyclophilins with an RNA recognition motif that interact with members of the trx/MLL protein family in Drosophila and human cells. *Development genes and evolution*, *212*(3), 107-113.

Ang, Y. S., Tsai, S. Y., Lee, D. F., Monk, J., Su, J., Ratnakumar, K., ... & Wang, J. (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. Cell, 145(2), 183-197.

Aplan, P., Chervinsky, D., Stanulla, M., & Burhans, W. (1996). Site-specific DNA cleavage within the MLL breakpoint cluster region induced by topoisomerase II inhibitors. *Blood*, *87*(7), 2649–2658. https://doi.org/10.1182/blood.V87.7.2649.bloodjournal8772649

Aravind, L., & Landsman, D. (1998). AT-hook motifs identified in a wide variety of DNA-binding proteins. Nucleic acids research, 26(19), 4413-4421.

Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. 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.

Arcipowski, K. M., Bulic, M., Gurbuxani, S., & Licht, J. D. (2016). Loss of Mll3 catalytic function promotes aberrant myelopoiesis. *PLoS ONE*, *11*(9), 1–16. https://doi.org/10.1371/journal.pone.0162515

Ariki, R., Morikawa, S., Mabuchi, Y., Suzuki, S., Nakatake, M., Yoshioka, K., Hidano, S., Nakauchi, H., Matsuzaki, Y., Nakamura, T., & Goitsuka, R. (2014). Homeodomain Transcription Factor Meis1 Is a Critical Regulator of Adult Bone Marrow Hematopoiesis. PLoS ONE, 9(2), e87646. https://doi.org/10.1371/journal.pone.0087646

Argiropoulos, B., & Humphries, R. K. (2007). Hox genes in hematopoiesis and leukemogenesis. *Oncogene*, *26*(47), 6766–6776. https://doi.org/10.1038/sj.onc.1210760

Argiropoulos, B., Yung, E., & Humphries, R. K. (2007). Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis. Genes & Development, 21(22), 2845–2849. https://doi.org/10.1101/gad.1619407

Armstrong, S. A., Kung, A. L., Mabon, M. E., Silverman, L. B., Stam, R. W., Den Boer, M. L., Pieters, R., Kersey, J. H., Sallan, S. E., Fletcher, J. A., Golub, T. R., Griffin, J. D., & Korsmeyer, S. J. (2003). Inhibition of FLT3 in MLL. *Cancer Cell*, *3*(2), 173–183. https://doi.org/10.1016/S1535-6108(03)00003-5

Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., Minden, M. D., Sallan, S. E., Lander, E. S., Golub, T. R., & Korsmeyer, S. J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics*, *30*(1), 41–47. https://doi.org/10.1038/ng765

Artinger, E. L., & Ernst, P. (2013). Cell context in the control of self-renewal and proliferation regulated by MLL1. *Cell Cycle*, *12*(18), 2969–2972. https://doi.org/10.4161/cc.26032

Babaryka, A., Kühn, E., & Köster, R. W. (2009). In vivo synthesis of meganuclease for generating transgenic zebrafish Danio rerio. *Journal of Fish Biology*, *74*(2), 452–457. https://doi.org/10.1111/j.1095-8649.2008.02075.x

Baker, L. A., Allis, C. D., & Wang, G. G. (2008). PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *647*(1-2), 3-12.

Basavapathruni, A., Olhava, E. J., Daigle, S. R., Therkelsen, C. A., Jin, L., Boriack-Sjodin, P. A., Allain, C. J., Klaus, C. R., Raimondi, A., Scott, M. P., Dovletoglou, A., Richon, V. M., Pollock, R. M., Copeland, R. A., Moyer, M. P., Chesworth, R., Pearson, P. G., & Waters, N. J. (2014). Nonclinical pharmacokinetics and metabolism of EPZ-5676, a novel DOT1L histone methyltransferase inhibitor. *Biopharmaceutics & Drug Disposition*, *35*(4), 237–252. https://doi.org/10.1002/bdd.1889

Behm, F. G., Raimondi, S. C., Frestedt, J. L., Liu, Q., Crist, W. M., Downing, J. R., ... & Pui, C. H. (1996). Rearrangement of the MLL gene confers a poor prognosis in childhood acute lymphoblastic leukemia, regardless of presenting age.

Bennett, C. M., Kanki, J. P., Rhodes, J., Liu, T. X., Paw, B. H., Kieran, M. W., Langenau, D. M., Delahaye-Brown, a, Zon, L. I., Fleming, M. D., & Look, a T. (2001). Myelopoiesis in the zebrafish, Danio rerio. Blood, 98(3), 643–651. https://doi.org/10.1182/blood.V98.3.643

Bertolino, P., Radovanovic, I., Casse, H., Aguzzi, A., Wang, Z.-Q., & Zhang, C.-X. (2003). Genetic ablation of the tumor suppressor menin causes lethality at mid-gestation with defects in multiple organs. Mechanisms of Development, 120(5), 549–560. https://doi.org/10.1016/S0925-4773(03)00039-X

Bertolino, P., Tong, W.-M., Galendo, D., Wang, Z.-Q., & Zhang, C.-X. (2003). Heterozygous Men1 Mutant Mice Develop a Range of Endocrine Tumors Mimicking Multiple Endocrine Neoplasia Type 1. Molecular Endocrinology, 17(9), 1880–1892. https://doi.org/10.1210/me.2003-0154

Bertrand, J. Y., Kim, A. D., Teng, S., & Traver, D. (2008). CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development*, *135*(10), 1853–1862. https://doi.org/10.1242/dev.015297

Bertrand, J. Y., Kim, A. D., Violette, E. P., Stachura, D. L., Cisson, J. L., & Traver, D. (2007). Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. *Development*, *134*(23), 4147–4156. https://doi.org/10.1242/dev.012385

Biankin, A. V, Waddell, N., Kassahn, K. S., Gingras, M., Muthuswamy, L. B., Johns, A. L., Miller, D. K., Wilson, P. J., Patch, A.-M., Wu, J., Chang, D. K., Cowley, M. J., Gardiner, B. B., Song, S., Harliwong, I., Idrisoglu, S., Nourse, C., Nourbakhsh, E., Manning, S., … Grimmond, S. M. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature*, *491*(7424), 399–405. https://doi.org/10.1038/nature11547

Bird, A. P., & Wolffe, A. P. (1999). Methylation-Induced Repression— Belts, Braces, and Chromatin. *Cell*, *99*(5), 451–454. https://doi.org/10.1016/S0092-8674(00)81532-9

Birke, M., Schreiner, S., García-Cuéllar, M. P., Mahr, K., Titgemeyer, F., & Slany, R. K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic acids research*, *30*(4), 958-965.

Bitoun, E., Oliver, P. L., & Davies, K. E. (2007). The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Human Molecular Genetics, 16(1), 92–106. https://doi.org/10.1093/hmg/ddl444

Borkin, D., He, S., Miao, H., Kempinska, K., Pollock, J., Chase, J., Purohit, T., Malik, B., Zhao, T., Wang, J., Wen, B., Zong, H., Jones, M., Danet-Desnoyers, G., Guzman, M. L., Talpaz, M., Bixby, D. L., Sun, D., Hess, J. L., … Grembecka, J. (2015). Pharmacologic inhibition of the menin-MLL interaction blocks progression of MLL leukemia invivo. *Cancer Cell*, *27*(4), 589–602. https://doi.org/10.1016/j.ccell.2015.02.016

Borland, S., Tenin, G., Williams, S., Monaghan, R., Baxter, M., Ray, D., Abraham, S., & Keavney, B. (2019). BS9 KMT2C- a tetralogy of fallot candidate gene. *Heart*, *105*(Suppl 6), A145 LP-A146. https://doi.org/10.1136/heartjnl-2019-BCS.173

Botbol, Y., Raghavendra, N. K., Rahman, S., Engelman, A., & Lavigne, M. (2007). Chromatinized templates reveal the requirement for the LEDGF/p75 PWWP domain during HIV-1 integration in vitro. Nucleic Acids Research, 36(4), 1237–1246. https://doi.org/10.1093/nar/gkm1127

Bouldin, C. M., & Kimelman, D. (2014). Dual Fucci: A new transgenic line for studying the cell cycle from embryos to adults. Zebrafish, 11(2), 182–183. https://doi.org/10.1089/zeb.2014.0986

Brown, P. (2013). Treatment of infant leukemias: challenge and promise. Hematology 2013, the American Society of Hematology Education Program Book, 2013(1), 596-600.

Brown, P., Kairalla, J., Wang, C., Dreyer, Z., Salzer, W., Sorenson, M., ... & Gore, L. (2016, November). Addition of FLT3 inhibitor lestaurtinib to post-induction chemotherapy does not improve outcomes in Mll-rearranged infant acute lymphoblastic leukemia (ALL): AALL0631, a Children's Oncology Group Study. In Pediatric Blood & Cancer (Vol. 63, pp. S7-S7).

Brown, P., Levis, M., Shurtleff, S., Campana, D., Downing, J., & Small, D. (2005). FLT3 inhibition selectively kills childhood acute lymphoblastic leukemia cells with high levels of FLT3 expression. *Blood*, *105*(2), 812–820. https://doi.org/10.1182/blood-2004-06-2498

Brown, P., Pieters, R., & Biondi, A. (2019). How I treat infant leukemia. Blood, 133(3), 205-214.

Brun, M. E., Gasca, S., Girard, C., Bouton, K., De Massy, B., & De Sario, A. (2006). Characterization and expression analysis during embryo development of the mouse ortholog of MLL3. Gene, 371(1), 25-33.

Bueno, C., Montes, R., Catalina, P., Rodríguez, R., & Menendez, P. (2011). Insights into the cellular origin and etiology of the infant pro-B acute lymphoblastic leukemia with MLL-AF4 rearrangement. *Leukemia*, *25*(3), 400–410. https://doi.org/10.1038/leu.2010.284

Calvanese, V., Nguyen, A. T., Bolan, T. J., Vavilina, A., Su, T., Lee, L. K., ... & Kurdistani, S. K. (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. Nature, 576(7786), 281-286.

Cao, F., Chen, Y., Cierpicki, T., Liu, Y., Basrur, V., Lei, M., & Dou, Y. (2010). An Ash2L/RbBP5 heterodimer stimulates the MLL1 methyltransferase activity through coordinated substrate interactions with the MLL1 SET domain. *PloS one*, *5*(11).

Carradice, D., & Lieschke, G. J. (2008). Zebrafish in hematology: sushi or science?. Blood, The Journal of the American Society of Hematology, 111(7), 3331-3342.

Caslini, C., Yang, Z., El-Osta, M., Milne, T. A., Slany, R. K., & Hess, J. L. (2007). Interaction of MLL Amino Terminal Sequences with Menin Is Required for Transformation. *Cancer Research*, *67*(15), 7275–7283. https://doi.org/10.1158/0008-5472.CAN-06-2369

Chang, V. Y., Basso, G., Sakamoto, K. M., & Nelson, S. F. (2013). Identification of somatic and germline mutations using whole exome sequencing of congenital acute lymphoblastic leukemia. *BMC Cancer*, *13*(1), 55. https://doi.org/10.1186/1471-2407-13-55

Charles, N. J., & Boyer, D. F. (2017). Mixed-phenotype acute leukemia: diagnostic criteria and pitfalls. Archives of pathology & laboratory medicine, 141(11), 1462-1468.

Chen, A. T., & Zon, L. I. (2009). Zebrafish blood stem cells. *Journal of Cellular Biochemistry*, *108*(1), 35–42. https://doi.org/10.1002/jcb.22251

Chen, C., Liu, Y., Rappaport, A. R., Kitzing, T., Schultz, N., Zhao, Z., Shroff, A. S., Dickins, R. A., Vakoc, C. R., Bradner, J. E., Stock, W., LeBeau, M. M., Shannon, K. M., Kogan, S., Zuber, J., & Lowe, S. W. (2014). MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. *Cancer Cell*, *25*(5), 652–665. https://doi.org/10.1016/j.ccr.2014.03.016

Chen, Y.-X., Yan, J., Keeshan, K., Tubbs, A. T., Wang, H., Silva, A., Brown, E. J., Hess, J. L., Pear, W. S., & Hua, X. (2006). The tumor suppressor menin regulates hematopoiesis and myeloid transformation by influencing Hox gene expression. *Proceedings of the National Academy of Sciences*, *103*(4), 1018–1023. https://doi.org/10.1073/pnas.0510347103

Chen, J., Jette, C., Kanki, J. P., Aster, J. C., Look, A. T., & Griffin, J. D. (2007). NOTCH1-induced T-cell leukemia in transgenic zebrafish. *Leukemia*, *21*(3), 462–471. https://doi.org/10.1038/sj.leu.2404546

Chen, R., Owuor, T. O., Patel, R. M., Casey, E., & Magee, J. A. (2019). Kmt2c Limits the Self-Renewal Capacity of Multiply Divided HSCs By Promoting Sensitivity to Interleukin-1. *Blood*, *134*(Supplement_1), 3711. https://doi.org/10.1182/blood-2019-126229

Cheung, N., & So, C. W. E. (2011). Transcriptional and epigenetic networks in haematological malignancy. *FEBS letters*, *585*(13), 2100-2111.

Chillón, M. C., Gómez-Casares, M. T., López-Jorge, C. E., Rodriguez-Medina, C., Molines, A., Sarasquete, M. E., Alcoceba, M., Miguel, J. D. G.-S., Bueno, C., Montes, R., Ramos, F., Rodríguez, J. N., Giraldo, P., Ramírez, M., García-Delgado, R., Fuster, J. L., González-Díaz, M., & Menendez, P. (2012). Prognostic significance of FLT3 mutational status and expression levels in MLL-AF4+ and MLL-germline acute lymphoblastic leukemia. *Leukemia*, *26*(11), 2360–2366. https://doi.org/10.1038/leu.2012.161

Collas, P., & Aleström, P. (1998). Nuclear localization signals enhance germline transmission of a transgene in zebrafish. Transgenic research, 7(4), 303-309.

Colleaux, L., d'Auriol, L., Galibert, F., & Dujon, B. (1988). Recognition and cleavage site of the intron-encoded omega transposase. Proceedings of the National Academy of Sciences, 85(16), 6022-6026.

Collins, E. C., Appert, A., Ariza-McNaughton, L., Pannell, R., Yamada, Y., & Rabbitts, T. H. (2002). Mouse Af9 is a controller of embryo patterning, like Mll, whose human homologue fuses with Af9 after chromosomal translocation in leukemia. Molecular and cellular biology, 22(20), 7313-7324.

Corral, J., Lavenir, I., Impey, H., Warren, A. J., Forster, A., Larson, T. A., Bell, S., McKenzie, A. N. ., King, G., & Rabbitts, T. H. (1996). An Mll–AF9 Fusion Gene Made by Homologous Recombination Causes Acute Leukemia in Chimeric Mice: A Method to Create Fusion Oncogenes. *Cell*, *85*(6), 853–861. https://doi.org/10.1016/S0092-8674(00)81269-6

Creutzig, U., Zimmermann, M., Bourquin, J. P., Dworzak, M. N., Kremens, B., Lehrnbecher, T., ... & Starý, J. (2012). Favorable outcome in infants with AML after intensive first-and second-line treatment: an AML-BFM study group report. *Leukemia*, *26*(4), 654-661.

Culp, P., Nüsslein-Volhard, C., & Hopkins, N. (1991). High-frequency germ-line transmission of plasmid DNA sequences injected into fertilized zebrafish eggs. Proceedings of the National Academy of Sciences, 88(18), 7953-7957.

Dafflon, C., Craig, V. J., Méreau, H., Gräsel, J., Schacher Engstler, B., Hoffman, G., Nigsch, F., Gaulis, S., Barys, L., Ito, M., Aguadé-Gorgorió, J., Bornhauser, B., Bourquin, J. P., Proske, A., Stork-Fux, C., Murakami, M., Sellers, W. R., Hofmann, F., Schwaller, J., & Tiedt, R. (2017). Complementary activities of DOT1L and Menin inhibitors in MLL-rearranged leukemia. *Leukemia*, *31*(6), 1269–1277. https://doi.org/10.1038/leu.2016.327

Daigle, S. R., Olhava, E. J., Therkelsen, C. A., Basavapathruni, A., Jin, L., Boriack-Sjodin, P. A., Allain, C. J., Klaus, C. R., Raimondi, A., Scott, M. P., Waters, N. J., Chesworth, R., Moyer, M. P., Copeland, R. A., Richon, V. M., & Pollock, R. M. (2013). Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood*, *122*(6), 1017–1025. https://doi.org/10.1182/blood-2013-04-497644

Daigle, S. R., Olhava, E. J., Therkelsen, C. A., Majer, C. R., Sneeringer, C. J., Song, J., Johnston, L. D., Scott, M. P., Smith, J. J., Xiao, Y., Jin, L., Kuntz, K. W., Chesworth, R., Moyer, M. P., Bernt, K. M., Tseng, J.-C., Kung, A. L., Armstrong, S. A., Copeland, R. A., … Pollock, R. M. (2011). Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. *Cancer Cell*, *20*(1), 53–65. https://doi.org/10.1016/j.ccr.2011.06.009

Davison, J. M., Akitake, C. M., Goll, M. G., Rhee, J. M., Gosse, N., Baier, H., ... & Parsons, M. J. (2007). Transactivation from Gal4-VP16 transgenic insertions for tissue-specific cell labeling and ablation in zebrafish. Developmental biology, 304(2), 811-824.

Dawkins, J. B. N., Wang, J., Maniati, E., Heward, J. A., Koniali, L., Kocher, H. M., Martin, S. A., Chelala, C., Balkwill, F. R., Fitzgibbon, J., & Grose, R. P. (2016). Reduced Expression of Histone Methyltransferases KMT2C and KMT2D Correlates with Improved Outcome in Pancreatic Ductal Adenocarcinoma. *Cancer Research*, *76*(16), 4861–4871. https://doi.org/10.1158/0008-5472.CAN-16-0481

Dawson, M. A., Prinjha, R. K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W. I., ... & Huthmacher, C. (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature*, *478*(7370), 529-533.

Dawson, M. A., Gudgin, E. J., Horton, S. J., Giotopoulos, G., Meduri, E., Robson, S., ... & Fong, C. Y. (2014). Recurrent mutations, including NPM1c, activate a BRD4-dependent core transcriptional program in acute myeloid leukemia. *Leukemia*, *28*(2), 311-320.

Delmore, J. E., Issa, G. C., Lemieux, M. E., Rahl, P. B., Shi, J., Jacobs, H. M., ... & Chesi, M. (2011). BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*, *146*(6), 904-917.

Detrich, H. W., Kieran, M. W., Chan, F. Y., Barone, L. M., Yee, K., Rundstadler, J. A., Pratt, S., Ransom, D., & Zon, L. I. (1995). Intraembryonic hematopoietic cell migration during vertebrate development. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(23), 10713–10717. https://doi.org/10.1073/pnas.92.23.10713

Deveau, A. P., Forrester, A. M., Coombs, A. J., Wagner, G. S., Grabher, C., Chute, I. C., Léger, D., Mingay, M., Alexe, G., Rajan, V., Liwski, R., Hirst, M., Stegmaier, K., Lewis, S. M., Look, a T., & Berman, J. N. (2015). Epigenetic therapy restores normal hematopoiesis in a zebrafish model of NUP98–HOXA9-induced myeloid disease. *Leukemia*, *29*(10), 2086–2097. https://doi.org/10.1038/leu.2015.126

Di Lorenzo, A., & Bedford, M. T. (2011). Histone arginine methylation. *FEBS Letters*, *585*(13), 2024–2031. https://doi.org/10.1016/j.febslet.2010.11.010

DiMartino, J. F., Ayton, P. M., Chen, E. H., Naftzger, C. C., Young, B. D., & Cleary, M. L. (2002). The AF10 leucine zipper is required for leukemic transformation of myeloid progenitors by MLL-AF10. *Blood*, *99*(10), 3780–3785. https://doi.org/10.1182/blood.V99.10.3780

DiMartino, J. F., Miller, T., Ayton, P. M., Landewe, T., Hess, J. L., Cleary, M. L., & Shilatifard, A. (2000). A carboxy-terminal domain of ELL is required and sufficient for immortalization of myeloid progenitors by MLL-ELL. Blood, 96(12), 3887–3893. https://doi.org/10.1182/blood.V96.12.3887.h8003887_3887_3893

Dobson, C. L., Warren, A. J., Pannell, R., Forster, A., Lavenir, I., Corral, J., Smith, A. J., & Rabbitts, T. H. (1999). The mll-AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis. *The EMBO Journal*, *18*(13), 3564–3574. https://doi.org/10.1093/emboj/18.13.3564

Dobson, J. T., Seibert, J., Teh, E. M., Da'as, S., Fraser, R. B., Paw, B. H., Lin, T. J., & Berman, J. N. (2008). Carboxypeptidase A5 identifies a novel mast cell lineage in the zebrafish providing new insight into mast cell fate determination. Blood, 112(7), 2969–2972. https://doi.org/10.1182/blood-2008-03-145011

Döhner, K., Brown, J., Hehmann, U., Hetzel, C., Stewart, J., Lowther, G., ... & Lichter, P. (1998). Molecular cytogenetic characterization of a critical region in bands 7q35-q36 commonly deleted in malignant myeloid disorders. Blood, The Journal of the American Society of Hematology, 92(11), 4031-4035.

Doty, R. T., Vanasse, G. J., Disteche, C. M., & Willerford, D. M. (2002). The leukemia-associated gene Mllt1/ENL: characterization of a murine homolog and demonstration of an essential role in embryonic development. Blood Cells, Molecules, and Diseases, 28(3), 407-417.

Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., ... & Roeder, R. G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nature structural & molecular biology, 13(8), 713-719.

Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., ... & Roeder, R. G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell*, *121*(6), 873-885.

Dreyer, Z. E., Dinndorf, P. A., Camitta, B., Sather, H., La, M. K., Devidas, M., ... & Willman, C. L. (2011). Analysis of the role of hematopoietic stem-cell transplantation in infants with acute lymphoblastic leukemia in first remission and MLL gene rearrangements: a report from the Children's Oncology Group. Journal of clinical oncology, 29(2), 214.

Drynan, L. F., Pannell, R., Forster, A., Chan, N. M. M., Cano, F., Daser, A., & Rabbitts, T. H. (2005). Mll fusions generated by Cre-loxP-mediated de novo translocations can induce lineage reassignment in tumorigenesis. *The EMBO Journal*, *24*(17), 3136–3146. https://doi.org/10.1038/sj.emboj.7600760

Eidahl, J. O., Crowe, B. L., North, J. A., McKee, C. J., Shkriabai, N., Feng, L., Plumb, M., Graham, R. L., Gorelick, R. J., Hess, S., Poirier, M. G., Foster, M. P., & Kvaratskhelia, M. (2013). Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. Nucleic Acids Research, 41(6), 3924–3936. https://doi.org/10.1093/nar/gkt074

Eiring, A. M., Harb, J. G., Neviani, P., Garton, C., Oaks, J. J., Spizzo, R., Liu, S., Schwind, S., Santhanam, R., Hickey, C. J., Becker, H., Chandler, J. C., Andino, R., Cortes, J., Hokland, P., Huettner, C. S., Bhatia, R., Roy, D. C., Liebhaber, S. A., … Perrotti, D. (2010). miR-328 Functions as an RNA Decoy to Modulate hnRNP E2 Regulation of mRNA Translation in Leukemic Blasts. *Cell*, *140*(5), 652–665. https://doi.org/10.1016/j.cell.2010.01.007

El Ashkar, S., Schwaller, J., Pieters, T., Goossens, S., Demeulemeester, J., Christ, F., Van Belle, S., Juge, S., Boeckx, N., Engelman, A., Van Vlierberghe, P., Debyser, Z., & De Rijck, J. (2017). LEDGF/p75 is dispensable for hematopoiesis but essential for MLL-rearranged leukemogenesis. Blood, 131(1), blood-2017-05-786962. https://doi.org/10.1182/blood-2017-05-786962

El-Brolosy, M., Rossi, A., Kontarakis, Z., Kuenne, C., Guenther, S., Fukuda, N., ... & Gerri, C. (2018). Genetic compensation is triggered by mutant mRNA degradation. bioRxiv 328153.

El-Brolosy, M. A., & Stainier, D. Y. (2017). Genetic compensation: A phenomenon in search of mechanisms. PLoS genetics, 13(7).

Emelyanov, A., & Parinov, S. (2008). Mifepristone-inducible LexPR system to drive and control gene expression in transgenic zebrafish. *Developmental Biology*, *320*(1), 113–121. https://doi.org/10.1016/j.ydbio.2008.04.042

Erfurth, F., Hemenway, C. S., de Erkenez, A. C., & Domer, P. H. (2004). MLL fusion partners AF4 and AF9 interact at subnuclear foci. Leukemia, 18(1), 92–102. https://doi.org/10.1038/sj.leu.2403200

Ernst, P., Mabon, M., Davidson, A. J., Zon, L. I., & Korsmeyer, S. J. (2004). An Mll-Dependent Hox Program Drives Hematopoietic Progenitor Expansion. Current Biology, 14(22), 2063–2069. https://doi.org/10.1016/j.cub.2004.11.012

Ernst, P., Wang, J., Huang, M., Goodman, R. H., & Korsmeyer, S. J. (2001). MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Molecular and cellular biology*, *21*(7), 2249-2258.

Esteller, M. (2005). Aberrant DNA methylation as a cancer-inducing mechanism. *Annu. Rev. Pharmacol. Toxicol.*, *45*, 629-656.

Fair, K., Anderson, M., Bulanova, E., Mi, H., Tropschug, M., & Diaz, M. O. (2001). Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Molecular and cellular biology*, *21*(10), 3589-3597.

Felix, C. A., & Lange, B. J. (1999). Leukemia in infants. *The Oncologist*, *4*(3), 225–240. https://doi.org/10.1056/NEJM189408091310603

Felker, A., & Mosimann, C. (2016). Contemporary zebrafish transgenesis with Tol2 and application for Cre/lox recombination experiments. In *Methods in Cell Biology* (Vol. 135, pp. 219–244). Elsevier Ltd. https://doi.org/10.1016/bs.mcb.2016.01.009

Feng, H., Langenau, D. M., Madge, J. A., Quinkertz, A., Gutierrez, A., Neuberg, D. S., Kanki, J. P., & Thomas Look, A. (2007). Heat-shock induction of T-cell lymphoma/leukaemia in conditional Cre/lox-regulated transgenic zebrafish. *British Journal of Haematology*, *138*(2), 169–175. https://doi.org/10.1111/j.1365-2141.2007.06625.x

Ford, A. M., Ridge, S. A., Cabrera, M. E., Mahmoud, H., Steel, C. M., Chan, L. C., & Greaves, M. (1993). In utero rearrangements in the trithorax-related oncogene in infant leukaemias. *Nature*, *363*(6427), 358–360. https://doi.org/10.1038/363358a0

Forsberg, E. C., Prohaska, S. S., Katzman, S., Heffner, G. C., Stuart, J. M., & Weissman, I. L. (2005). Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS genetics*, *1*(3).

Forster, A., Pannell, R., Drynan, L. F., McCormack, M., Collins, E. C., Daser, A., & Rabbitts, T. H. (2003). Engineering de novo reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. *Cancer Cell*, *3*(5), 449–458. https://doi.org/10.1016/S1535-6108(03)00106-5

Forrester, A. M., Grabher, C., McBride, E. R., Boyd, E. R., Vigerstad, M. H., Edgar, A., Kai, F.-B., Da'as, S. I., Payne, E., Look, A. T., & Berman, J. N. (2011). NUP98-HOXA9-transgenic zebrafish develop a myeloproliferative neoplasm and provide new insight into mechanisms of myeloid leukaemogenesis. *British Journal of Haematology*, *155*(2), 167–181. https://doi.org/10.1111/j.1365-2141.2011.08810.x

Fujimoto, A., Totoki, Y., Abe, T., Boroevich, K. A., Hosoda, F., Nguyen, H. H., Aoki, M., Hosono, N., Kubo, M., Miya, F., Arai, Y., Takahashi, H., Shirakihara, T., Nagasaki, M., Shibuya, T., Nakano, K., Watanabe-Makino, K., Tanaka, H., Nakamura, H., … Nakagawa, H. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nature Genetics*, *44*(7), 760–764. https://doi.org/10.1038/ng.2291

Gala, K., Li, Q., Sinha, A., Razavi, P., Dorso, M., Sanchez-Vega, F., Chung, Y. R., Hendrickson, R., Hsieh, J. J., Berger, M., Schultz, N., Pastore, A., Abdel-Wahab, O., & Chandarlapaty, S. (2018). KMT2C mediates the estrogen dependence of breast cancer through regulation of ERα enhancer function. *Oncogene*, *37*(34), 4692–4710. https://doi.org/10.1038/s41388-018-0273-5

Galloway, J. L., & Zon, L. I. (2003). *Ontogeny of hematopoiesis: Examining the emergence of hematopoietic cells in the vertebrate embryo*. *53*, 139–158. https://doi.org/10.1016/S0070-2153(03)53004-6

García-Cuéllar, M. P., Schreiner, S. A., Birke, M., Hamacher, M., Fey, G. H., & Slany, R. K. (2000). ENL, the MLL fusion partner in t(11;19), binds to the c-Abl interactor protein 1 (ABI1) that is fused to MLL in t(10;11)+. Oncogene, 19(14), 1744–1751. https://doi.org/10.1038/sj.onc.1203506

Grabher, C., Joly, J.-S., & Wittbrodt, J. (2004). Highly Efficient Zebrafish Transgenesis Mediated by the Meganuclease I-SceI. In *Methods in Cell Biology* (Vol. 2004, Issue 77, pp. 381–401). https://doi.org/10.1016/S0091-679X(04)77021-1

Greaves, M. (2005). In utero origins of childhood leukaemia. *Early Human Development*, *81*(1), 123–129. https://doi.org/10.1016/j.earlhumdev.2004.10.004

Greaves, M. F., Maia, A. T., Wiemels, J. L., & Ford, A. M. (2003). Leukemia in twins: lessons in natural history. *Blood*, *102*(7), 2321–2333. https://doi.org/10.1182/blood-2002-12-3817

Grier, D. G., Thompson, A., Kwasniewska, A., McGonigle, G. J., Halliday, H. L., & Lappin, T. R. (2005). The pathophysiology of HOX genes and their role in cancer. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, *205*(2), 154-171.

Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., ... & Canaani, E. (1992). The t (4; 11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell, 71(4), 701-708.

Guenther, M. G., Lawton, L. N., Rozovskaia, T., Frampton, G. M., Levine, S. S., Volkert, T. L., Croce, C. M., Nakamura, T., Canaani, E., & Young, R. A. (2008). Aberrant chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. *Genes & Development*, *22*(24), 3403–3408. https://doi.org/10.1101/gad.1741408

Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R., & Young, R. A. (2007). A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell*, *130*(1), 77–88. https://doi.org/10.1016/j.cell.2007.05.042

Gutierrez, A., Grebliunaite, R., Feng, H., Kozakewich, E., Zhu, S., Guo, F., Payne, E., Mansour, M., Dahlberg, S. E., Neuberg, D. S., Hertog, J. Den, Prochownik, E. V., Testa, J. R., Harris, M., Kanki, J. P., & Look, A. T. (2011). Pten mediates Myc oncogene dependence in a conditional zebrafish model of T cell acute lymphoblastic leukemia. *The Journal of Experimental Medicine*, *208*(8), 1595–1603. https://doi.org/10.1084/jem.20101691

Habets, P. E., Clout, D. E., Deprez, R. H. L., Van Roon, M. A., Moorman, A. F., & Christoffels, V. M. (2003). Cardiac expression of Gal4 causes cardiomyopathy in a dose-dependent manner. Journal of Muscle Research & Cell Motility, 24(2-3), 205-209.

Hamilton, D. L., & Abremski, K. (1984). Site-specific recombination by the bacteriophage P1 lox-Cre system. Journal of Molecular Biology, 178(2), 481–486. https://doi.org/10.1016/0022-2836(84)90154-2

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. cell, 100(1), 57-70.

Harrison, C. J., Hills, R. K., Moorman, A. V., Grimwade, D. J., Hann, I., Webb, D. K. H., Wheatley, K., de Graaf, S. S. N., van den Berg, E., Burnett, A. K., & Gibson, B. E. S. (2010). Cytogenetics of Childhood Acute Myeloid Leukemia: United Kingdom Medical Research Council Treatment Trials AML 10 and 12. Journal of Clinical Oncology, 28(16), 2674–2681. https://doi.org/10.1200/JCO.2009.24.8997

Hess, J. L., Yu, B. D., Li, B., Hanson, R., & Korsmeyer, S. J. (1997). Defects in Yolk Sac Hematopoiesis in Mll-Null Embryos. Blood, 90(5), 1799–1806. https://doi.org/10.1182/blood.V90.5.1799

Hilden, J. M., Dinndorf, P. A., Meerbaum, S. O., Sather, H., Villaluna, D., Heerema, N. A., ... & Johnstone, H. S. (2006). Analysis of prognostic factors of acute lymphoblastic leukemia in infants: report on CCG 1953 from the Children's Oncology Group. Blood, 108(2), 441-451.

Hisa, T., Spence, S. E., Rachel, R. A., Fujita, M., Nakamura, T., Ward, J. M., Devor-Henneman, D. E., Saiki, Y., Kutsuna, H., Tessarollo, L., Jenkins, N. A., & Copeland, N. G. (2004). Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. The EMBO Journal, 23(2), 450–459. https://doi.org/10.1038/sj.emboj.7600038

Hrusak, O., de Haas, V., Stancikova, J., Vakrmanova, B., Janotova, I., Mejstrikova, E., Capek, V., Trka, J., Zaliova, M., Luks, A., Bleckmann, K., Möricke, A., Irving, J., Konatkowska, B., Alexander, T. B., Inaba, H., Schmiegelow, K., Stokley, S., Zemanova, Z., … Stary, J. (2018). International cooperative study identifies treatment strategy in childhood ambiguous lineage leukemia. *Blood*, *132*(3), 264–276. https://doi.org/10.1182/blood-2017-12-821363

Hsieh, J. J. D., Cheng, E. H. Y., & Korsmeyer, S. J. (2003). Taspase1: A threonine aspartase required for cleavage of MLL and proper HOX gene expression. Cell, 115(3), 293–303. https://doi.org/10.1016/S0092-8674(03)00816-X

Hsieh, J. J.-D., Ernst, P., Erdjument-Bromage, H., Tempst, P., & Korsmeyer, S. J. (2003). Proteolytic Cleavage of MLL Generates a Complex of N- and C-Terminal Fragments That Confers Protein Stability and Subnuclear Localization. *Molecular and Cellular Biology*, *23*(1), 186–194. https://doi.org/10.1128/MCB.23.1.186-194.2003

Hsu, K., Traver, D., Kutok, J. L., Hagen, A., Liu, T.-X., Paw, B. H., Rhodes, J., Berman, J. N., Zon, L. I., Kanki, J. P., & Look, A. T. (2004). The pu.1 promoter drives myeloid gene expression in zebrafish. *Blood*, *104*(5), 1291–1297. https://doi.org/10.1182/blood-2003-09-3105

Hu, D., Gao, X., Morgan, M. A., Herz, H.-M., Smith, E. R., & Shilatifard, A. (2013). The MLL3/MLL4 Branches of the COMPASS Family Function as Major Histone H3K4 Monomethylases at Enhancers. *Molecular and Cellular Biology*, *33*(23), 4745–4754. https://doi.org/10.1128/MCB.01181-13

Huang, C., & Zhu, B. (2018). Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. *Biophysics Reports*, *4*(4), 170–177. https://doi.org/10.1007/s41048-018-0063-1

Hunger, S. P., Lu, X., Devidas, M., Camitta, B. M., Gaynon, P. S., Winick, N. J., Reaman, G. H., & Carroll, W. L. (2012). Improved Survival for Children and Adolescents With Acute Lymphoblastic Leukemia Between 1990 and 2005: A Report From the Children's Oncology Group. *Journal of Clinical Oncology*, *30*(14), 1663–1669. https://doi.org/10.1200/JCO.2011.37.8018

Hurst, C. D., Alder, O., Platt, F. M., Droop, A., Stead, L. F., Burns, J. E., Burghel, G. J., Jain, S., Klimczak, L. J., Lindsay, H., Roulson, J.-A., Taylor, C. F., Thygesen, H., Cameron, A. J., Ridley, A. J., Mott, H. R., Gordenin, D. A., & Knowles, M. A. (2017). Genomic Subtypes of Non-invasive Bladder Cancer with Distinct Metabolic Profile and Female Gender Bias in KDM6A Mutation Frequency. *Cancer Cell*, *32*(5), 701-715.e7. https://doi.org/10.1016/j.ccell.2017.08.005

Ida, K., Kitabayashi, I., Taki, T., Taniwaki, M., Noro, K., Yamamoto, M., Ohki, M., & Hayashi, Y. (1997). Adenoviral E1A-Associated Protein p300 Is Involved in Acute Myeloid Leukemia With t(11; 22)(q23; q13). *Blood*, *90*(12), 4699–4704. https://doi.org/10.1182/blood.V90.12.4699.4699_4699_4704

Imamura, T., Morimoto, A., Takanashi, M., Hibi, S., Sugimoto, T., Ishii, E., & Imashuku, S. (2002). Frequent co-expression of HoxA9 and Meis1 genes in infant acute lymphoblastic leukaemia with MLL rearrangement. British Journal of Haematology, 119(1), 119–121. https://doi.org/10.1046/j.1365-2141.2002.03803.x

Jaenisch, R., Jähner, D., Nobis, P., Simon, I., Löhler, J., Harbers, K., & Grotkopp, D. (1981). Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell, 24(2), 519-529.

Jin, H., Sood, R., Xu, J., Zhen, F., English, M. A., Liu, P. P., & Wen, Z. (2009). Definitive hematopoietic stem/progenitor cells manifest distinct differentiation output in the zebrafish VDA and PBI. *Development*, *136*(4), 647–654. https://doi.org/10.1242/dev.029637

Jing, Y. (2004). The PML-RARα fusion protein and targeted therapy for acute promyelocytic leukemia. *Leukemia and Lymphoma*, *45*(4), 639–648. https://doi.org/10.1080/10428190310001609933

Jude, C. D., Climer, L., Xu, D., Artinger, E., Fisher, J. K., & Ernst, P. (2007). Unique and Independent Roles for MLL in Adult Hematopoietic Stem Cells and Progenitors. Cell Stem Cell, 1(3), 324–337. https://doi.org/10.1016/j.stem.2007.05.019

Kalmanti, L., Saussele, S., Lauseker, M., Müller, M. C., Dietz, C. T., Heinrich, L., Hanfstein, B., Proetel, U., Fabarius, A., Krause, S. W., Rinaldetti, S., Dengler, J., Falge, C., Oppliger-Leibundgut, E., Burchert, A., Neubauer, A., Kanz, L., Stegelmann, F., Pfreundschuh, M., … Hehlmann, R. (2015). Safety and efficacy of imatinib in CML over a period of 10 years: data from the randomized CML-study IV. *Leukemia*, *29*(5), 1123–1132. https://doi.org/10.1038/leu.2015.36

Kanchi, K. L., Johnson, K. J., Lu, C., McLellan, M. D., Leiserson, M. D. M., Wendl, M. C., Zhang, Q., Koboldt, D. C., Xie, M., Kandoth, C., McMichael, J. F., Wyczalkowski, M. A., Larson, D. E., Schmidt, H. K., Miller, C. A., Fulton, R. S., Spellman, P. T., Mardis, E. R., Druley, T. E., … Ding, L. (2014). Integrated analysis of germline and somatic variants in ovarian cancer. *Nature Communications*, *5*(1), 3156. https://doi.org/10.1038/ncomms4156

Kang, H., Wilson, C. S., Harvey, R. C., Chen, I.-M., Murphy, M. H., Atlas, S. R., Bedrick, E. J., Devidas, M., Carroll, A. J., Robinson, B. W., Stam, R. W., Valsecchi, M. G., Pieters, R., Heerema, N. A., Hilden, J. M., Felix, C. A., Reaman, G. H., Camitta, B., Winick, N., … Willman, C. L. (2012). Gene expression profiles predictive of outcome and age in infant acute lymphoblastic leukemia: a Children's Oncology Group study. *Blood*, *119*(8), 1872–1881. https://doi.org/10.1182/blood-2011-10-382861

Kawakami, K. (2007). Tol2: a versatile gene transfer vector in vertebrates. *Genome Biology*, *8*(Suppl 1), S7. https://doi.org/10.1186/gb-2007-8-s1-s7

Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N., & Mishina, M. (2004). A Transposon-Mediated Gene Trap Approach Identifies Developmentally Regulated Genes in Zebrafish. 7, 133–144.

Kim, D. H., Rhee, J. C., Yeo, S., Shen, R., Lee, S. K., Lee, J. W., & Lee, S. (2015). Crucial roles of mixed-lineage leukemia 3 and 4 as epigenetic switches of the hepatic circadian clock controlling bile acid homeostasis in mice. Hepatology, 61(3), 1012-1023.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. Developmental Dynamics, 203(3), 253–310. https://doi.org/10.1002/aja.1002030302

Khodaei, S., O'Brien, K. P., Dumanski, J., Wong, F. K., & Weber, G. (1999). Characterization of the MEN1 ortholog in Zebrafish. *Biochemical and Biophysical Research Communications*, *264*(2), 404–408. https://doi.org/10.1006/bbrc.1999.1529

Kotini, A. G., Chang, C. J., Boussaad, I., Delrow, J. J., Dolezal, E. K., Nagulapally, A. B., ... & Huangfu, D. (2015). Functional analysis of a chromosomal deletion associated with myelodysplastic syndromes using isogenic human induced pluripotent stem cells. Nature biotechnology, 33(6), 646.

Kramer, J. M., & Staveley, B. E. (2003). GAL4 causes developmental defects and apoptosis when expressed in the developing eye of Drosophila melanogaster. Genet Mol Res, 2(1), 43-47.

Krivtsov, A. V., Feng, Z., Lemieux, M. E., Faber, J., Vempati, S., Sinha, A. U., Xia, X., Jesneck, J., Bracken, A. P., Silverman, L. B., Kutok, J. L., Kung, A. L., & Armstrong, S. A. (2008). H3K79 Methylation Profiles Define Murine and Human MLL-AF4 Leukemias. *Cancer Cell*, *14*(5), 355–368. https://doi.org/10.1016/j.ccr.2008.10.001

Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., ... & Greenblatt, J. F. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Molecular cell*, *11*(3), 721-729.

Kroon, E., Krosl, J., Thorsteinsdottir, U., Baban, S., Buchberg, A. M., & Sauvageau, G. (1998). Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *The EMBO journal*, *17*(13), 3714-3725.

Kühn, M. W., Radtke, I., Bullinger, L., Goorha, S., Cheng, J., Edelmann, J., ... & Krauter, J. (2012). High-resolution genomic profiling of adult and pediatric core-binding factor acute myeloid leukemia reveals new recurrent genomic alterations. Blood, The Journal of the American Society of Hematology, 119(10), e67-e75.

Kumar, A. R., Li, Q., Hudson, W. A., Chen, W., Sam, T., Yao, Q., Lund, E. A., Wu, B., Kowal, B. J., & Kersey, J. H. (2009). A role for MEIS1 in MLL-fusion gene leukemia. *Blood*, *113*(8), 1756–1758. https://doi.org/10.1182/blood-2008-06-163287

Kuo, A. J., Cheung, P., Chen, K., Zee, B. M., Kioi, M., Lauring, J., Xi, Y., Park, B. H., Shi, X., Garcia, B. A., Li, W., & Gozani, O. (2011). NSD2 Links Dimethylation of Histone H3 at Lysine 36 to Oncogenic Programming. Molecular Cell, 44(4), 609–620. https://doi.org/10.1016/j.molcel.2011.08.042

Labbé, R. M., Holowatyj, A., & Yang, Z. Q. (2014). Histone lysine demethylase (KDM) subfamily 4: structures, functions and therapeutic potential. *American journal of translational research*, *6*(1), 1.

Lachner, M., O'Sullivan, R. J., & Jenuwein, T. (2003). An epigenetic road map for histone lysine methylation. *Journal of Cell Science*, *116*(11), 2117–2124. https://doi.org/10.1242/jcs.00493

Langenau, D. M. (2003). Myc-Induced T Cell Leukemia in Transgenic Zebrafish. *Science*, *299*(5608), 887–890. https://doi.org/10.1126/science.1080280

Langenau, D. M., Feng, H., Berghmans, S., Kanki, J. P., Kutok, J. L., & Look, A. T. (2005). Cre/lox-regulated transgenic zebrafish model with conditional myc-induced T cell acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences*, *102*(17), 6068–6073. https://doi.org/10.1073/pnas.0408708102

Lavau, C., Luo, R. T., Du, C., & Thirman, M. J. (2000). Retrovirus-mediated gene transfer of MLL-ELL transforms primary myeloid progenitors and causes acute myeloid leukemias in mice. Proceedings of the National Academy of Sciences, 97(20), 10984–10989. https://doi.org/10.1073/pnas.190167297

Lavau, C., Szilvassy, S. J., Slany, R., & Cleary, M. L. (1997). Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX–ENL. The EMBO journal, 16(14), 4226-4237.

Lawrence, H. J., Helgason, C. D., Sauvageau, G., Fong, S., Izon, D. J., Humphries, R. K., & Largman, C. (1997). Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis. Blood, The Journal of the American Society of Hematology, 89(6), 1922-1930.

Le, X., Langenau, D. M., Keefe, M. D., Kutok, J. L., Neuberg, D. S., & Zon, L. I. (2007). Heat shock-inducible Cre/Lox approaches to induce diverse types of tumors and hyperplasia in transgenic zebrafish. *Proceedings of the National Academy of Sciences*, *104*(22), 9410–9415. https://doi.org/10.1073/pnas.0611302104

Lee, J., Kim, D.-H., Lee, S., Yang, Q.-H., Lee, D. K., Lee, S.-K., Roeder, R. G., & Lee, J. W. (2009). A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. *Proceedings of the National Academy of Sciences*, *106*(21), 8513–8518. https://doi.org/10.1073/pnas.0902873106

Lee, J., Saha, P. K., Yang, Q. H., Lee, S., Jung, Y. P., Suh, Y., Lee, S. K., Chan, L., Roeder, R. G., & Lee, J. W. (2008). Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19229–19234. https://doi.org/10.1073/pnas.0810100105

Lee, J. E., Wang, C., Xu, S., Cho, Y. W., Wang, L., Feng, X., ... & Ge, K. (2013). H3K4 mono-and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Elife, 2, e01503.

Lin, C., Smith, E. R., Takahashi, H., Lai, K. C., Martin-Brown, S., Florens, L., ... & Shilatifard, A. (2010). AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Molecular cell, 37(3), 429-437.

Lin, S., Yang, S., & Hopkins, N. (1994). lacZ expression in germline transgenic zebrafish can be detected in living embryos. Developmental biology, 161(1), 77-83.

Linabery, A. M., & Ross, J. A. (2008). Trends in childhood cancer incidence in the US (1992–2004). Cancer: Interdisciplinary International Journal of the American Cancer Society, 112(2), 416-432.

Liu, X., Li, Y.-S., Shinton, S. A., Rhodes, J., Tang, L., Feng, H., Jette, C. A., Look, A. T., Hayakawa, K., & Hardy, R. R. (2017). Zebrafish B Cell Development without a Pre–B Cell Stage, Revealed by CD79 Fluorescence Reporter Transgenes. *The Journal of Immunology*, *199*(5), 1706–1715. https://doi.org/10.4049/jimmunol.1700552

Long, Q., Meng, A., Wang, M., Jessen, J. R., Farrell, M. J., & Lin, S. (1997). GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene. *Development*, *124*(20), 4105–4111.

Mann, G., Attarbaschi, A., Schrappe, M., De Lorenzo, P., Peters, C., Hann, I., ... & Szczepanski, T. (2010). Improved outcome with hematopoietic stem cell transplantation in a poor prognostic subgroup of infants with mixed-lineage-leukemia (MLL)–rearranged acute lymphoblastic leukemia: results from the Interfant-99 Study. Blood, The Journal of the American Society of Hematology, 116(15), 2644-2650.

McMahon, K. A., Hiew, S. Y. L., Hadjur, S., Veiga-Fernandes, H., Menzel, U., Price, A. J., Kioussis, D., Williams, O., & Brady, H. J. M. (2007). Mll Has a Critical Role in Fetal and Adult Hematopoietic Stem Cell Self-Renewal. Cell Stem Cell, 1(3), 338–345. https://doi.org/10.1016/j.stem.2007.07.002

Meeks, J. J., & Shilatifard, A. (2017). Multiple Roles for the MLL/COMPASS Family in the Epigenetic Regulation of Gene Expression and in Cancer. *Annual Review of Cancer Biology*, *1*(1), 425–446. https://doi.org/10.1146/annurev-cancerbio-050216-034333

Mertz, J. A., Conery, A. R., Bryant, B. M., Sandy, P., Balasubramanian, S., Mele, D. A., ... & Sims, R. J. (2011). Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proceedings of the National Academy of Sciences*, *108*(40), 16669-16674.

Meyer, C., Burmeister, T., Gröger, D., Tsaur, G., Fechina, L., Renneville, A., Sutton, R., Venn, N. C., Emerenciano, M., Pombo-de-Oliveira, M. S., Barbieri Blunck, C., Almeida Lopes, B., Zuna, J., Trka, J., Ballerini, P., Lapillonne, H., De Braekeleer, M., Cazzaniga, G., Corral Abascal, L., … Marschalek, R. (2018). The MLL recombinome of acute leukemias in 2017. Leukemia, 32(2), 273–284. https://doi.org/10.1038/leu.2017.213

Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., Allis, C. D., & Hess, J. L. (2002). MLL Targets SET Domain Methyltransferase Activity to Hox Gene Promoters. *Molecular Cell*, *10*(5), 1107–1117. https://doi.org/10.1016/S1097-2765(02)00741-4

Milne, T. A., Hughes, C. M., Lloyd, R., Yang, Z., Rozenblatt-Rosen, O., Dou, Y., Schnepp, R. W., Krankel, C., LiVolsi, V. A., Gibbs, D., Hua, X., Roeder, R. G., Meyerson, M., & Hess, J. L. (2005). Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. Proceedings of the National Academy of Sciences, 102(3), 749–754. https://doi.org/10.1073/pnas.0408836102

Milne, T. A., Kim, J., Wang, G. G., Stadler, S. C., Basrur, V., Whitcomb, S. J., Wang, Z., Ruthenburg, A. J., Elenitoba-Johnson, K. S. J., Roeder, R. G., & Allis, C. D. (2010). Multiple Interactions Recruit MLL1 and MLL1 Fusion Proteins to the HOXA9 Locus in Leukemogenesis. *Molecular Cell*, *38*(6), 853–863. https://doi.org/10.1016/j.molcel.2010.05.011

Mishra, B. P., Ansari, K. I., & Mandal, S. S. (2009). Dynamic association of MLL1, H3K4 trimethylation with chromatin and Hox gene expression during the cell cycle. *The FEBS journal*, *276*(6), 1629-1640.

Morello, G., Porazzi, P., Moro, E., Argenton, F., Basso, G., Felix, C. A., & Germano, G. (2012). Zebrafish Ortholog of Human DOT1L Regulates Primitive and Transient Definitive Hematopoiesis and Controls hoxa9 and meis1 Expression. *Blood*, *120*(21), 849–849. https://doi.org/10.1182/blood.V120.21.849.849

Mosimann, C., Kaufman, C. K., Li, P., Pugach, E. K., Tamplin, O. J., & Zon, L. I. (2011). Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish. *Development*, *138*(1), 169–177. https://doi.org/10.1242/dev.059345

Mosimann, C., & Zon, L. I. (2011). Advanced zebrafish transgenesis with Tol2 and application for Cre/lox recombination experiments. *Methods in Cell Biology*, *104*, 173–194. https://doi.org/10.1016/B978-0-12-374814-0.00010-0

Mueller, D., Bach, C., Zeisig, D., Garcia-Cuellar, M.-P., Monroe, S., Sreekumar, A., Zhou, R., Nesvizhskii, A., Chinnaiyan, A., Hess, J. L., & Slany, R. K. (2007). A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood, 110(13), 4445–4454. https://doi.org/10.1182/blood-2007-05-090514

Muntean, A. G., & Hess, J. L. (2012). The pathogenesis of mixed-lineage leukemia. Annual Review of Pathology: Mechanisms of Disease, 7, 283-301.

Murayama, E., Kissa, K., Zapata, A., Mordelet, E., Briolat, V., Lin, H., Handin, R. I., & Herbomel, P. (2006). Tracing Hematopoietic Precursor Migration to Successive Hematopoietic Organs during Zebrafish Development. *Immunity*, *25*(6), 963–975. https://doi.org/10.1016/j.immuni.2006.10.015

Murphey, R. D., Stern, H. M., Straub, C. T., & Zon, L. I. (2006). A Chemical Genetic Screen for Cell Cycle Inhibitors in Zebrafish Embryos. *Chemical Biology & Drug Design*, *68*(4), 213–219. https://doi.org/10.1111/j.1747-0285.2006.00439.x

Mutskov, V., & Felsenfeld, G. (2004). Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *The EMBO Journal*, *23*(1), 138–149. https://doi.org/10.1038/sj.emboj.7600013

Nie, Z., Yan, Z., Chen, E. H., Sechi, S., Ling, C., Zhou, S., Xue, Y., Yang, D., Murray, D., Kanakubo, E., Cleary, M. L., & Wang, W. (2003). Novel SWI/SNF Chromatin-Remodeling Complexes Contain a Mixed-Lineage Leukemia Chromosomal Translocation Partner. *Molecular and Cellular Biology*, *23*(8), 2942–2952. https://doi.org/10.1128/MCB.23.8.2942-2952.2003

Nishikawa, T., Ota, T., & Isogai, T. (2000). Prediction whether a human cDNA sequence contains initiation codon by combining statistical information and similarity with protein sequences. *Bioinformatics*, *16*(11), 960–967. https://doi.org/10.1093/bioinformatics/16.11.960

Ng, H. H., Robert, F., Young, R. A., & Struhl, K. (2003). Targeted Recruitment of Set1 Histone Methylase by Elongating Pol II Provides a Localized Mark and Memory of Recent Transcriptional Activity. Molecular Cell, 11(3), 709–719. https://doi.org/10.1016/S1097-2765(03)00092-3

Nguyen, A. T., Taranova, O., He, J., & Zhang, Y. (2011). DOT1L, the H3K79 methyltransferase, is required for MLL-AF9–mediated leukemogenesis. *Blood*, *117*(25), 6912–6922. https://doi.org/10.1182/blood-2011-02-334359

Ogawa, M. (1993). Differentiation and proliferation of hematopoietic stem cells.

Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V. M., Su, L., Xu, G., & Zhang, Y. (2005). hDOT1L Links Histone Methylation to Leukemogenesis. *Cell*, *121*(2), 167–178. https://doi.org/10.1016/j.cell.2005.02.020

Okuda, H., Kawaguchi, M., Kanai, A., Matsui, H., Kawamura, T., Inaba, T., Kitabayashi, I., & Yokoyama, A. (2014). MLL fusion proteins link transcriptional coactivators to previously active CpG-rich promoters. Nucleic Acids Research, 42(7), 4241–4256. https://doi.org/10.1093/nar/gkt1394

Orgel, E., Alexander, T. B., Wood, B. L., Kahwash, S. B., Devidas, M., Dai, Y., ... & Raetz, E. A. (2020). Mixed‐phenotype acute leukemia: A cohort and consensus research strategy from the Children's Oncology Group Acute Leukemia of Ambiguous Lineage Task Force. Cancer, 126(3), 593-601.

Orlovsky, K., Kalinkovich, A., Rozovskaia, T., Shezen, E., Itkin, T., Alder, H., Ozer, H. G., Carramusa, L., Avigdor, A., Volinia, S., Buchberg, A., Mazo, A., Kollet, O., Largman, C., Croce, C. M., Nakamura, T., Lapidot, T., & Canaani, E. (2011). Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation. *Proceedings of the National Academy of Sciences*, *108*(19), 7956–7961. https://doi.org/10.1073/pnas.1103154108

Page, D. M., Wittamer, V., Bertrand, J. Y., Lewis, K. L., Pratt, D. N., Delgado, N., Schale, S. E., McGue, C., Jacobsen, B. H., Doty, A., Pao, Y., Yang, H., Chi, N. C., Magor, B. G., & Traver, D. (2013). An evolutionarily conserved program of B-cell development and activation in zebrafish. *Blood*, *122*(8), e1–e11. https://doi.org/10.1182/blood-2012-12-471029

Pang, W. W., Price, E. A., Sahoo, D., Beerman, I., Maloney, W. J., Rossi, D. J., Schrier, S. L., & Weissman, I. L. (2011). Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20012–20017. https://doi.org/10.1073/pnas.1116110108

Patel, A., Dharmarajan, V., Vought, V. E., & Cosgrove, M. S. (2009). On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *Journal of Biological Chemistry*, *284*(36), 24242-24256.

Perlman, E. J., Gadd, S., Arold, S. T., Radhakrishnan, A., Gerhard, D. S., Jennings, L., ... & Meerzaman, D. (2015). MLLT1 YEATS domain mutations in clinically distinctive Favourable Histology Wilms tumours. Nature communications, 6(1), 1-10.

Peterlin, B. M., & Price, D. H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. *Molecular Cell*, *23*(3), 297–305. https://doi.org/10.1016/j.molcel.2006.06.014

Phillips, R. L., Ernst, R. E., Brunk, B., Ivanova, N., Mahan, M. A., Deanehan, J. K., ... & Lemischka, I. R. (2000). The genetic program of hematopoietic stem cells. *Science*, *288*(5471), 1635-1640.

Pieters, R., De Lorenzo, P., Ancliffe, P., Aversa, L. A., Brethon, B., Biondi, A., ... & Kotecha, R. S. (2019). Outcome of infants younger than 1 year with acute lymphoblastic leukemia treated with the Interfant-06 Protocol: results from an international phase III randomized study. Journal of Clinical Oncology, 37(25), 2246-2256.

Pieters, R., Schrappe, M., De Lorenzo, P., Hann, I., De Rossi, G., Felice, M., ... & Janka, G. (2007). A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. The Lancet, 370(9583), 240-250.

Pina, C., May, G., Soneji, S., Hong, D., & Enver, T. (2008). MLLT3 regulates early human erythroid and megakaryocytic cell fate. Cell stem cell, 2(3), 264-273.

Pineault, N., Helgason, C. D., Lawrence, H. J., & Humphries, R. K. (2002). Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. Experimental Hematology, 30(1), 49–57. https://doi.org/10.1016/S0301-472X(01)00757-3

Pramparo, T., Grosso, S., Messa, J., Zatterale, A., Bonaglia, M. C., Chessa, L., ... & Giorda, R. (2005). Loss-of-function mutation of the AF9/MLLT3 gene in a girl with neuromotor development delay, cerebellar ataxia, and epilepsy. Human genetics, 118(1), 76-81.

Prykhozhij, S. V., Steele, S. L., Razaghi, B., & Berman, J. N. (2017). A rapid and effective method for screening, sequencing and reporter verification of engineered frameshift mutations in zebrafish. Disease Models & Mechanisms, 10(6), 811–822. https://doi.org/10.1242/dmm.026765

Pui, C.-H., Campana, D., Pei, D., Bowman, W. P., Sandlund, J. T., Kaste, S. C., Ribeiro, R. C., Rubnitz, J. E., Raimondi, S. C., Onciu, M., Coustan-Smith, E., Kun, L. E., Jeha, S., Cheng, C., Howard, S. C., Simmons, V., Bayles, A., Metzger, M. L., Boyett, J. M., … Relling, M. V. (2009). Treating Childhood Acute Lymphoblastic Leukemia without Cranial Irradiation. *New England Journal of Medicine*, *360*(26), 2730–2741. https://doi.org/10.1056/NEJMoa0900386

Pui, C. H., Gaynon, P. S., Boyett, J. M., Chessells, J. M., Baruchel, A., Kamps, W., ... & Schrappe, M. (2002). Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. The Lancet, 359(9321), 1909-1915.

Qiao, Y., Wang, X., Wang, R., Li, Y., Yu, F., Yang, X., ... & Jing, N. (2015). AF9 promotes hESC neural differentiation through recruiting TET2 to neurodevelopmental gene loci for methylcytosine hydroxylation. Cell discovery, 1(1), 1-19.

Rampias, T., Karagiannis, D., Avgeris, M., Polyzos, A., Kokkalis, A., Kanaki, Z., Kousidou, E., Tzetis, M., Kanavakis, E., Stravodimos, K., Manola, K. N., Pantelias, G. E., Scorilas, A., & Klinakis, A. (2019). The lysine-specific methyltransferase KMT 2C/ MLL 3 regulates DNA repair components in cancer . *EMBO Reports*, *20*(3), 1–20. https://doi.org/10.15252/embr.201846821

Rao, R. C., & Dou, Y. (2015). Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. Nature Reviews Cancer, 15(6), 334-346.

Rayes, A., McMasters, R. L., & O'Brien, M. M. (2016). Lineage switch in MLL-rearranged infant leukemia following CD19-directed therapy. Pediatric blood & cancer, 63(6), 1113-1115.

Reaman, G. H., Sposto, R., Sensel, M. G., Lange, B. J., Feusner, J. H., Heerema, N. A., ... & Johnstone, H. S. (1999). Treatment outcome and prognostic factors for infants with acute lymphoblastic leukemia treated on two consecutive trials of the Children's Cancer Group. Journal of Clinical Oncology, 17(2), 445-445.

Rebel, V. I., Miller, C. L., Eaves, C. J., & Lansdorp, P. M. (1996). The repopulation potential of fetal liver hematopoietic stem cells in mice exceeds that of their adult bone marrow counterparts. *Blood*, *87*(8), 3500–3507. https://doi.org/10.1182/blood.v87.8.3500.bloodjournal8783500

Rice, J. C., & Allis, C. D. (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. Current opinion in cell biology, 13(3), 263-273.

Ridges, S., Heaton, W. L., Joshi, D., Choi, H., Eiring, A., Batchelor, L., Choudhry, P., Manos, E. J., Sofla, H., Sanati, A., Welborn, S., Agarwal, A., Spangrude, G. J., Miles, R. R., Cox, J. E., Frazer, J. K., Deininger, M., Balan, K., Sigman, M., … Trede, N. S. (2012). Zebrafish screen identifies novel compound with selective toxicity against leukemia. *Blood*, *119*(24), 5621–5631. https://doi.org/10.1182/blood-2011-12-398818

Ringrose, L., & Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.*, *38*, 413-443.

Risner, L. E., Kuntimaddi, A., Lokken, A. A., Achille, N. J., Birch, N. W., Schoenfelt, K., ... & Zeleznik-Le, N. J. (2013). Functional specificity of CpG DNA-binding CXXC domains in mixed lineage leukemia. *Journal of Biological Chemistry*, *288*(41), 29901-29910.

Robinson, B. W., Germano, G., Song, Y., Abrams, J., Scott, M., Guariento, I., Tiso, N., Argenton, F., Basso, G., Rhodes, J., Kanki, J. P., Look, A. T., Balice-Gordon, R. J., & Felix, C. A. (2011). mll ortholog containing functional domains of human MLL is expressed throughout the zebrafish lifespan and in haematopoietic tissues. *British Journal of Haematology*, *152*(3), 307–321. https://doi.org/10.1111/j.1365-2141.2010.08398.x

Rosen, J. N., Sweeney, M. F., & Mably, J. D. (2009). Microinjection of zebrafish embryos to analyze gene function. In Journal of Visualized Experiments (Issue 25). https://doi.org/10.3791/1115

Ross, J. A., Potter, J. D., & Robison, L. L. (1994). Infant Leukemia, Topoisomerase II Inhibitors, and the MLL Gene. *JNCI Journal of the National Cancer Institute*, *86*(22), 1678–1680. https://doi.org/10.1093/jnci/86.22.1678

Ross, J. A. (2008). Environmental and Genetic Susceptibility to MLL-Defined Infant Leukemia. *JNCI Monographs*, *2008*(39), 83–86. https://doi.org/10.1093/jncimonographs/lgn007

Ross, J. A. (1998). Maternal diet and infant leukemia: a role for DNA topoisomerase II inhibitors? *Int J Cancer Suppl*, *11*, 26–28.

Ross, J. A., Linabery, A. M., Blommer, C. N., Langer, E. K., Spector, L. G., Hilden, J. M., Heerema, N. A., Gretchen, A., Tower, R. L., & Davies, S. M. (2013). Genetic Variants Modify Susceptibility to Leukemia in Infants: A Children' s Oncology Group Report. *Pediatr Blood Cancer*, *60*(1), 31–34. https://doi.org/10.1002/pbc.24131.Genetic

Rossant, J., Nutter, L. M. J., & Gertsenstein, M. (2011). Engineering the embryo. *Proceedings of the National Academy of Sciences*, *108*(19), 7659–7660. https://doi.org/10.1073/pnas.1104844108

Rossi, J. G., Bernasconi, A. R., Alonso, C. N., Rubio, P. L., Gallego, M. S., Carrara, C. A., Guitter, M. R., Eberle, S. E., Cocce, M., Zubizarreta, P. A., & Felice, M. S. (2012). Lineage switch in childhood acute leukemia: An unusual event with poor outcome. *American Journal of Hematology*, *87*(9), 890–897. https://doi.org/10.1002/ajh.23266

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., & Stainier, D. Y. (2015). Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature, 524(7564), 230-233.

Rowe, R. G., Lummertz da Rocha, E., Sousa, P., Missios, P., Morse, M., Marion, W., Yermalovich, A., Barragan, J., Mathieu, R., Jha, D. K., Fleming, M. D., North, T. E., & Daley, G. Q. (2019). The developmental stage of the hematopoietic niche regulates lineage in MLL-rearranged leukemia. *Journal of Experimental Medicine*, *216*(3), 527–538. https://doi.org/10.1084/jem.20181765

Ruault, M., Brun, M. E., Ventura, M., Roizès, G., & De Sario, A. (2002). MLL3, a new human member of the TRX/MLL gene family, maps to 7q36, a chromosome region frequently deleted in myeloid leukaemia. Gene, 284(1-2), 73-81.

Rubnitz, J. E., Gibson, B., & Smith, F. O. (2008). Acute myeloid leukemia. Pediatric clinics of North America, 55(1), 21-51.

Sabaawy, H. E., Azuma, M., Embree, L. J., Tsai, H.-J., Starost, M. F., & Hickstein, D. D. (2006). TEL-AML1 transgenic zebrafish model of precursor B cell acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences*, *103*(41), 15166–15171. https://doi.org/10.1073/pnas.0603349103

Saberi, M.1; Delfi, O.1; Browett, P.J.1; Kakadiya, P.M.1; Bohlander, S.1 PF243 AN MLL-AF9 ZEBRAFISH LEUKEMIA MODEL, HemaSphere: June 2019 - Volume 3 - Issue S1 - p 73 doi: 10.1097/01.HS9.0000559188.32534.e0

Salzer, W. L., Jones, T. L., Devidas, M., Hilden, J. M., Winick, N., Hunger, S., ... & Dreyer, Z. E. (2012). Modifications to induction therapy decrease risk of early death in infants with acute lymphoblastic leukemia treated on Children's Oncology Group P9407. Pediatric blood & cancer, 59(5), 834-839.

Schafer, E. S., Irizarry, R., Negi, S., McIntyre, E., Small, D., Figueroa, M. E., Melnick, A., & Brown, P. (2009). Promoter Hypermethylation in MLL-r Leukemia: Biology and Therapeutic Targeting. *Blood*, *114*(22), 3472. https://doi.org/10.1182/blood.V114.22.3472.3472

Scharf, S., Zech, J., Bursen, A., Schraets, D., Oliver, P. L., Kliem, S., ... & Marschalek, R. (2007). Transcription linked to recombination: a gene-internal promoter coincides with the recombination hot spot II of the human MLL gene. Oncogene, 26(10), 1361-1371.

Scheer, N., Groth, A., Hans, S., & Campos-Ortega, J. A. (2001). An instructive function for Notch in promoting gliogenesis in the zebrafish retina. Development, 128(7), 1099-1107.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. Nature Methods, 9(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schulte, C. E. (2002). MLL-ENL cooperates with SCF to transform primary avian multipotent cells. *The EMBO Journal*, *21*(16), 4297–4306. https://doi.org/10.1093/emboj/cdf429

Scott, E. K., Mason, L., Arrenberg, A. B., Ziv, L., Gosse, N. J., Xiao, T., ... & Baier, H. (2007). Targeting neural circuitry in zebrafish using GAL4 enhancer trapping. Nature methods, 4(4), 323-326.

Shilatifard, A. (2008). Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. Current opinion in cell biology, 20(3), 341-348.

Simone, F., Polak, P. E., Kaberlein, J. J., Luo, R. T., Levitan, D. A., & Thirman, M. J. (2001). EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein. Blood, 98(1), 201–209. https://doi.org/10.1182/blood.V98.1.201

Slany, R. K., Lavau, C., & Cleary, M. L. (1998). The Oncogenic Capacity of HRX-ENL Requires the Transcriptional Transactivation Activity of ENL and the DNA Binding Motifs of HRX. Molecular and Cellular Biology, 18(1), 122–129. https://doi.org/10.1128/MCB.18.1.122

Slany, R. K. (2009). The molecular biology of mixed lineage leukemia. *Haematologica*, *94*(7), 984–993. https://doi.org/10.3324/haematol.2008.002436

Somervaille, T. C., & Cleary, M. L. (2006). Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer cell*, *10*(4), 257-268.

Spector, L. G. (2005). Maternal Diet and Infant Leukemia: The DNA Topoisomerase II Inhibitor Hypothesis: A Report from the Children's Oncology Group. *Cancer Epidemiology Biomarkers & Prevention*, *14*(3), 651–655. https://doi.org/10.1158/1055-9965.EPI-04-0602

Stam, R. W., Schneider, P., de Lorenzo, P., Valsecchi, M. G., den Boer, M. L., & Pieters, R. (2007). Prognostic significance of high-level FLT3 expression in MLL-rearranged infant acute lymphoblastic leukemia. *Blood*, *110*(7), 2774–2775. https://doi.org/10.1182/blood-2007-05-091934

Stam, R. W., Schneider, P., Hagelstein, J. A. P., van der Linden, M. H., Stumpel, D. J. P. M., de Menezes, R. X., de Lorenzo, P., Valsecchi, M. G., & Pieters, R. (2010). Gene expression profiling–based dissection of MLL translocated and MLL germline acute lymphoblastic leukemia in infants. *Blood*, *115*(14), 2835–2844. https://doi.org/10.1182/blood-2009-07-233049

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., … 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

Sternberg, N., & Hamilton, D. (1981). Bacteriophage P1 site-specific recombination. Journal of Molecular Biology, 150(4), 467–486. https://doi.org/10.1016/0022-2836(81)90375-2

Steward, M. M., Lee, J. S., O'Donovan, A., Wyatt, M., Bernstein, B. E., & Shilatifard, A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature structural & molecular biology*, *13*(9), 852-854.

Striano, P., Elia, M., Castiglia, L., Galesi, O., Pelligra, S., & Striano, S. (2005). A t (4; 9)(q34; p22) translocation associated with partial epilepsy, mental retardation, and dysmorphism. Epilepsia, 46(8), 1322-1324.

Stuart, G. W., McMurray, J. V., & Westerfield, M. O. N. T. E. (1988). Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. Development, 103(2), 403-412.

Stuart, G. W., Vielkind, J. R., McMurray, J. V., & Westerfield, M. (1990). Stable lines of transgenic zebrafish exhibit reproducible patterns of transgene expression. Development, 109(3), 577-584.

Stumpel, D. J. P. M., Schneider, P., van Roon, E. H. J., Boer, J. M., de Lorenzo, P., Valsecchi, M. G., de Menezes, R. X., Pieters, R., & Stam, R. W. (2009). Specific promoter methylation identifies different subgroups of MLL-rearranged infant acute lymphoblastic leukemia, influences clinical outcome, and provides therapeutic options. *Blood*, *114*(27), 5490–5498. https://doi.org/10.1182/blood-2009-06-227660

Stumpel, D. J. P. M., Schneider, P., van Roon, E. H. J., Pieters, R., & Stam, R. W. (2013). Absence of global hypomethylation in promoter hypermethylated Mixed Lineage Leukaemia-rearranged infant acute lymphoblastic leukaemia. *European Journal of Cancer*, *49*(1), 175–184. https://doi.org/10.1016/j.ejca.2012.07.013

Sun, Y., Chen, B.-R., & Deshpande, A. (2018). Epigenetic Regulators in the Development, Maintenance, and Therapeutic Targeting of Acute Myeloid Leukemia. *Frontiers in Oncology*, *8*(FEB), 1–16. https://doi.org/10.3389/fonc.2018.00041

Super, H., McCabe, N., Thirman, M., Larson, R., Le Beau, M., Pedersen-Bjergaard, J., Philip, P., Diaz, M., & Rowley, J. (1993). Rearrangements of the MLL gene in therapy-related acute myeloid leukemia in patients previously treated with agents targeting DNA- topoisomerase II. *Blood*, *82*(12), 3705–3711. https://doi.org/10.1182/blood.V82.12.3705.3705

Suster, M. L., Kikuta, H., Urasaki, A., Asakawa, K., & Kawakami, K. (2009). Transgenesis in Zebrafish with the Tol2 Transposon System (Issue September 2018). https://doi.org/10.1007/978-1-60327-019-9

Sutherland, H. G., Newton, K., Brownstein, D. G., Holmes, M. C., Kress, C., Semple, C. A., & Bickmore, W. A. (2006). Disruption of Ledgf/Psip1 Results in Perinatal Mortality and Homeotic Skeletal Transformations. Molecular and Cellular Biology, 26(19), 7201–7210. https://doi.org/10.1128/MCB.00459-06

Takeda, S., Chen, D. Y., Westergard, T. D., Fisher, J. K., Rubens, J. A., Sasagawa, S., ... & Hsieh, J. J. D. (2006). Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes & development*, *20*(17), 2397-2409.

Taketani, T. (2003). FLT3 mutations in the activation loop of tyrosine kinase domain are frequently found in infant ALL with MLL rearrangements and pediatric ALL with hyperdiploidy. *Blood*, *103*(3), 1085–1088. https://doi.org/10.1182/blood-2003-02-0418

Taki, T., Sako, M., Tsuchida, M., & Hayashi, Y. (1997). The t(11;16)(q23;p13) Translocation in Myelodysplastic Syndrome Fuses the MLL Gene to the CBP Gene. *Blood*, *89*(11), 3945–3950. https://doi.org/10.1182/blood.V89.11.3945

Tamplin, O. J., Durand, E. M., Carr, L. A., Childs, S. J., Hagedorn, E. J., Li, P., Yzaguirre, A. D., Speck, N. A., & Zon, L. I. (2015). Hematopoietic Stem Cell Arrival Triggers Dynamic Remodeling of the Perivascular Niche. *Cell*, *160*(1–2), 241–252. https://doi.org/10.1016/j.cell.2014.12.032

Tan, J., Jones, M., Koseki, H., Nakayama, M., Muntean, A. G., Maillard, I., & Hess, J. L. (2011). CBX8, a Polycomb Group Protein, Is Essential for MLL-AF9-Induced Leukemogenesis. *Cancer Cell*, *20*(5), 563–575. https://doi.org/10.1016/j.ccr.2011.09.008

Tan, J., Zhao, L., Wang, G., Li, T., Li, D., Xu, Q., Chen, X., Shang, Z., Wang, J., & Zhou, J. (2018). Human MLL-AF9 Overexpression Induces Aberrant Hematopoietic Expansion in Zebrafish. *BioMed Research International*, *2018*, 1–9. https://doi.org/10.1155/2018/6705842

Tanaka, M., & Kinoshita, M. (2001). Recent progress in the generation of transgenic medaka (Oryzias latipes). Zoological Science, 18(5), 615-622.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., & Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology*, *14*(11), 1025-1040.

Terranova, R., Agherbi, H., Boned, A., Meresse, S., & Djabali, M. (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. Proceedings of the National Academy of Sciences, 103(17), 6629-6634.

Thermes, V., Grabher, C., Ristoratore, F., Bourrat, F., Choulika, A., Wittbrodt, J., & Joly, J. S. (2002). I-SceI meganuclease mediates highly efficient transgenesis in fish. Mechanisms of development, 118(1-2), 91-98.

Thorsteinsdottir, U., Mamo, A., Kroon, E., Jerome, L., Bijl, J., Lawrence, H. J., ... & Sauvageau, G. (2002). Overexpression of the myeloid leukemia–associated Hoxa9 gene in bone marrow cells induces stem cell expansion. Blood, The Journal of the American Society of Hematology, 99(1), 121-129.

Tondeleir, D., Lambrechts, A., Müller, M., Jonckheere, V., Doll, T., Vandamme, D., ... & Decaestecker, C. (2012). Cells lacking β-actin are genetically reprogrammed and maintain conditional migratory capacity. Molecular & cellular proteomics, 11(8), 255-271.

Topp, M. S., Gökbuget, N., Stein, A. S., Zugmaier, G., O'Brien, S., Bargou, R. C., Dombret, H., Fielding, A. K., Heffner, L., Larson, R. A., Neumann, S., Foà, R., Litzow, M., Ribera, J.-M., Rambaldi, A., Schiller, G., Brüggemann, M., Horst, H. A., Holland, C., … Kantarjian, H. M. (2015). Safety and activity of blinatumomab for adult patients with relapsed or refractory B-precursor acute lymphoblastic leukaemia: a multicentre, single-arm, phase 2 study. *The Lancet Oncology*, *16*(1), 57–66. https://doi.org/10.1016/S1470-2045(14)71170-2

Traver, D., Paw, B. H., Poss, K. D., Penberthy, W. T., Lin, S., & Zon, L. I. (2003). Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nature Immunology*, *4*(12), 1238–1246. https://doi.org/10.1038/ni1007

Tsai, C. T., & So, C. W. E. (2017). Epigenetic therapies by targeting aberrant histone methylome in AML: molecular mechanisms, current preclinical and clinical development. *Oncogene*, *36*(13), 1753–1759. https://doi.org/10.1038/onc.2016.315

Valentine, M. C., Linabery, A. M., Chasnoff, S., Hughes, A. E. O., Mallaney, C., Sanchez, N., Giacalone, J., Heerema, N. A., Hilden, J. M., Spector, L. G., Ross, J. A., & Druley, T. E. (2014). Excess congenital non-synonymous variation in leukemia-associated genes in MLL− infant leukemia: a Children's Oncology Group report. *Leukemia*, *28*(6), 1235–1241. https://doi.org/10.1038/leu.2013.367

Valerio, D. G., Xu, H., Chen, C. W., Hoshii, T., Eisold, M. E., Delaney, C., ... & Zheng, Y. G. (2017). Histone acetyltransferase activity of MOF is required for MLL-AF9 leukemogenesis. Cancer research, 77(7), 1753-1762.

Vogel, T., & Gruss, P. (2009). Expression of Leukaemia associated transcription factor Af9/Mllt3 in the cerebral cortex of the mouse. Gene Expression Patterns, 9(2), 83-93.

Vogeli, K. M., Jin, S.-W., Martin, G. R., & Stainier, D. Y. R. (2006). A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. *Nature*, *443*(7109), 337–339. https://doi.org/10.1038/nature05045

Wan, X., Hu, B., Liu, J., Feng, X., & Xiao, W. (2011). Zebrafish mll Gene Is Essential for Hematopoiesis. *Journal of Biological Chemistry*, *286*(38), 33345–33357. https://doi.org/10.1074/jbc.M111.253252

Wang, P., Lin, C., Smith, E. R., Guo, H., Sanderson, B. W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L. M., Ge, K., Krumlauf, R., & Shilatifard, A. (2009). Global Analysis of H3K4 Methylation Defines MLL Family Member Targets and Points to a Role for MLL1-Mediated H3K4 Methylation in the Regulation of Transcriptional Initiation by RNA Polymerase II. Molecular and Cellular Biology, 29(22), 6074–6085. https://doi.org/10.1128/MCB.00924-09

Wang, Q., Wu, G., Mi, S., He, F., Wu, J., Dong, J., Luo, R. T., Mattison, R., Kaberlein, J. J., Prabhakar, S., Ji, H., & Thirman, M. J. (2011). MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood*, *117*(25), 6895–6905. https://doi.org/10.1182/blood-2010-12-324699

Wang, Z., Song, J., Milne, T. A., Wang, G. G., Li, H., Allis, C. D., & Patel, D. J. (2010). Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell*, *141*(7), 1183-1194.

Watanabe, Y., Castoro, R. J., Kim, H. S., North, B., Oikawa, R., Hiraishi, T., Ahmed, S. S., Chung, W., Cho, M.-Y., Toyota, M., Itoh, F., Estecio, M. R. H., Shen, L., Jelinek, J., & Issa, J.-P. J. (2011). Frequent Alteration of MLL3 Frameshift Mutations in Microsatellite Deficient Colorectal Cancer. *PLoS ONE*, *6*(8), e23320. https://doi.org/10.1371/journal.pone.0023320

Wei, J. W., Huang, K., Yang, C., & Kang, C. S. (2017). Non-coding RNAs as regulators in epigenetics. *Oncology reports*, *37*(1), 3-9.

Westerfield, M. (2000). The Zebrafish Book : A Guide for the Laboratory Use of Zebrafish. Http://Zfin.Org/Zf_info/Zfbook/Zfbk.Html. http://ci.nii.ac.jp/naid/10029409142/en/

White, R. J., Collins, J. E., Sealy, I. M., Wali, N., Dooley, C. M., Digby, Z., Stemple, D. L., Murphy, D. N., Billis, K., Hourlier, T., Füllgrabe, A., Davis, M. P., Enright, A. J., & Busch-Nentwich, E. M. (2017). A high-resolution mRNA expression time course of embryonic development in zebrafish. *ELife*, *6*, 1–32. https://doi.org/10.7554/eLife.30860

White, R. M., Sessa, A., Burke, C., Bowman, T., LeBlanc, J., Ceol, C., Bourque, C., Dovey, M., Goessling, W., Burns, C. E., & Zon, L. I. (2008). Transparent Adult Zebrafish as a Tool for In Vivo Transplantation Analysis. Cell Stem Cell, 2(2), 183–189. https://doi.org/10.1016/j.stem.2007.11.002

Wilson, C., Bellen, H. J., & Gehring, W. J. (1990). Position effects on eukaryotic gene expression. Annual review of cell biology, 6(1), 679-714.

Wong, P., Iwasaki, M., Somervaille, T. C. P., So, C. W. E., & Cleary, M. L. (2007). Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes & Development*, *21*(21), 2762–2774. https://doi.org/10.1101/gad.1602107

Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., ... & Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell, 121(6), 859-872.

Xia, Z. B., Anderson, M., Diaz, M. O., & Zeleznik-Le, N. J. (2003). MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proceedings of the National Academy of Sciences*, *100*(14), 8342-8347.

Xu, H., Xiao, T., Chen, C. H., Li, W., Meyer, C. A., Wu, Q., Wu, D., Cong, L., Zhang, F., Liu, J. S., Brown, M., & Liu, X. S. (2015). Sequence determinants of improved CRISPR sgRNA design. Genome Research, 25(8), 1147–1157. https://doi.org/10.1101/gr.191452.115

Xu, Q. (1999). Microinjection into zebrafish embryos. In Methods in molecular biology (Clifton, N.J.) (Vol. 127, pp. 125–132). https://doi.org/10.1385/1-59259-678-9:125

Yagi, H., Deguchi, K., Aono, A., Tani, Y., Kishimoto, T., & Komori, T. (1998). Growth Disturbance in Fetal Liver Hematopoiesis of Mll-Mutant Mice. Blood, 92(1), 108–117. https://doi.org/10.1182/blood.V92.1.108.413k11_108_117

Yeh, J.-R. J., Munson, K. M., Chao, Y. L., Peterson, Q. P., MacRae, C. A., & Peterson, R. T. (2008). AML1-ETO reprograms hematopoietic cell fate by downregulating scl expression. *Development*, *135*(2), 401–410. https://doi.org/10.1242/dev.008904

Yeh, J.-R. J., Munson, K. M., Elagib, K. E., Goldfarb, A. N., Sweetser, D. A., & Peterson, R. T. (2009). Discovering chemical modifiers of oncogene-regulated hematopoietic differentiation. *Nature Chemical Biology*, *5*(4), 236–243. https://doi.org/10.1038/nchembio.147

Yekta, S. (2004). MicroRNA-Directed Cleavage of HOXB8 mRNA. *Science*, *304*(5670), 594–596. https://doi.org/10.1126/science.1097434

Yokoyama, A., & Cleary, M. L. (2008). Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. Cancer Cell, 14(1), 36–46. https://doi.org/10.1016/j.ccr.2008.05.003

Yokoyama, A., Ficara, F., Murphy, M. J., Meisel, C., Hatanaka, C., Kitabayashi, I., & Cleary, M. L. (2013). MLL becomes functional through intra-molecular interaction not by proteolytic processing. *PloS one*, *8*(9).

Yokoyama, A., Lin, M., Naresh, A., Kitabayashi, I., & Cleary, M. L. (2010). A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. Cancer cell, 17(2), 198-212.

Yokoyama, A., Somervaille, T. C. P., Smith, K. S., Rozenblatt-Rosen, O., Meyerson, M., & Cleary, M. L. (2005). The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell*, *123*(2), 207–218. https://doi.org/10.1016/j.cell.2005.09.025

Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D. J., Kitabayashi, I., Herr, W., & Cleary, M. L. (2004). Leukemia Proto-Oncoprotein MLL Forms a SET1-Like Histone Methyltransferase Complex with Menin To Regulate Hox Gene Expression. Molecular and Cellular Biology, 24(13), 5639–5649. https://doi.org/10.1128/MCB.24.13.5639-5649.2004

Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E., & Korsmeyer, S. J. (1998). MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(18), 10632–10636. https://doi.org/10.1073/pnas.95.18.10632

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. J., & Korsmeyer, S. J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature*, *378*(6556), 505–508. https://doi.org/10.1038/378505a0

Yue, L., Du, J., Ye, F., Chen, Z., Li, L., Lian, F., Zhang, B., Zhang, Y., Jiang, H., Chen, K., Li, Y., Zhou, B., Zhang, N., Yang, Y., & Luo, C. (2016). Identification of novel small-molecule inhibitors targeting menin–MLL interaction, repurposing the antidiarrheal loperamide. *Organic & Biomolecular Chemistry*, *14*(36), 8503–8519. https://doi.org/10.1039/C6OB01248E

Zeisig, B. B., Schreiner, S., García-Cuéllar, M.-P., & Slany, R. K. (2003). Transcriptional activation is a key function encoded by MLL fusion partners. *Leukemia*, *17*(2), 359–365. https://doi.org/10.1038/sj.leu.2402804

Zeisig, D. T., Bittner, C. B., Zeisig, B. B., García-Cuéllar, M.-P., Hess, J. L., & Slany, R. K. (2005). The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. *Oncogene*, *24*(35), 5525–5532. https://doi.org/10.1038/sj.onc.1208699

Zeleznik-Le, N. J., Harden, A. M., & Rowley, J. D. (1994). 11q23 translocations split the" AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proceedings of the National Academy of Sciences*, *91*(22), 10610-10614.

Zhang, W., Xia, X., Reisenauer, M. R., Hemenway, C. S., & Kone, B. C. (2006). Dot1a-AF9 Complex Mediates Histone H3 Lys-79 Hypermethylation and Repression of ENaC α in an Aldosterone-sensitive Manner. *Journal of Biological Chemistry*, *281*(26), 18059–18068. https://doi.org/10.1074/jbc.M601903200

Zhang, X. Y., & Rodaway, A. R. F. (2007). SCL-GFP transgenic zebrafish: In vivo imaging of blood and endothelial development and identification of the initial site of definitive hematopoiesis. *Developmental Biology*, *307*(2), 179–194. https://doi.org/10.1016/j.ydbio.2007.04.002

Zhu, S., Zhu, E. D., Provot, S., & Gori, F. (2010). Wdr5 is required for chick skeletal development. Journal of Bone and Mineral Research, 25(11), 2504-2514.

Zon, L. I., & Peterson, R. T. (2005). In vivo drug discovery in the zebrafish. Nature reviews Drug discovery, 4(1), 35-44.