USING THE ZEBRAFISH TO INVESTIGATE THE ROLE OF THE *NUP98-NSD1* ONCOGENE AND LOSS OF *NUP98* IN HIGH-RISK PEDIATRIC ACUTE MYELOID LEUKEMIA

by

Corey Lynn Aston Filiaggi

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

at

Dalhousie University
Halifax, Nova Scotia
July 2017

DEDICATION

I would like to dedicate this thesis to the people who helped me get through grad school: my parents, sister, John, Patricia, Ellie, and Emma. You made this process endurable and helped get me to the finish line.

I would also like to dedicate this thesis to Julia MacLeod: a cancer survivor, chemo nurse, best friend, and an inspiration to my research.

TABLE OF CONTENTS

List of Tables ix

List of Figures x

Abstract xii

List of Abbreviations and Symbols Used xiii

Acknowledgements xvii

Chapter 1: Introduction 1

1.1 Acute Myeloid Leukemia (AML) 1

1.1.1 AML Diagnosis and Epidemiology 1

1.1.2 AML Etiology 2

1.1.3 AML Classification and Treatment 3

1.1.4 Pediatric versus Adult AML 4

1.1.5 NUP98 Fusions in AML and Other Hematologic Malignancies 5

1.2 *NUP98*-NSD1 in AML 8

1.2.1 Discovery and Characterization of the Patient Group Defined by the NUP98-NSD1 Translocation 8

1.2.2 NSD1 in Health and Disease 9

1.2.3 Aberrant Expression of NUP98-NSD1 and the Effect on HOX Gene Expression 10

1.2.4 Previous Studies Involving Models of NUP98-NSD1-Induced Leukemia 13

1.3 NUP98 and Nucleoporins in Health and Disease 13

1.3.1 Nucleoporins and the Nuclear Pore Complex (NPC) 13

1.3.2 De-regulation of NUP98 and Disease Consequences in AML 14

1.4 Zebrafish Blood Development and the Use of Zebrafish as a Model Organism 16

1.4.1 Overview of the Zebrafish as a Model Organism **16**

1.4.2 Overview of Hematopoiesis in the Zebrafish **17**

1.5 Transgenic and CRISPR Technology in the Context of the Zebrafish 20

1.5.1 Forward and Reverse Genetics in the Zebrafish **20**

1.5.2 Transgenic Zebrafish and Their Use as Models **22**

1.5.3 The Use of CRISPR for Disease Modelling in Zebrafish **25**

1.6 Rationale 28

1.6.1 Previous Work Modelling the NUP98-HOXA9 (NHA9) Oncogene in AML Using the Zebrafish **28**

1.6.2 Modelling NUP98-NSD1-induced AML and decreased nup98 in the zebrafish **28**

Chapter 2: Materials and Methods 30

2.1 Zebrafish Husbandry, Embryo Collection, and Embryo Staging 30

2.2 Zebrafish Embryo Injections at the One-cell Stage 36

2.3 Creation of a Transgenic Zebrafish Line 36

2.3.1 Bacterial Cloning **36**

2.3.2 NUP98-NSD1 Plasmid **37**

2.3.3 NUP98-NSD1 Zebrafish Line Creation **39**

2.4 Creation of a Mutant Zebrafish Line 39

2.4.1 nup98 Knockout CRISPR Design **39**

2.4.2 Taq LongAmp® Polymerase Chain Reaction (PCR) **40**

2.4.3 Heteroduplex Mobility Assay (HMA) **42**

2.5 Fin Clipping and Genomic DNA (gDNA) Extraction 45

2.6 Whole-mount *in* situ Hybridization (WISH) 45

2.6.1 RNA Probe Synthesis **45**

2.6.2 Embryo Preparation **46**

2.6.3 WISH Protocol **48**

2.6.4 Analysis of in situs **49**

2.7 Western Blotting 51

2.7.1 Protein Extraction **51**

2.7.2 Western Blot Protocol **51**

2.8 Zebrafish Peripheral Blood Extraction and Whole Kidney Marrow Dissection 52

2.8.1 Peripheral Blood (PB) Extraction **52**

2.8.2 Whole Kidney Marrow (WKM) Dissection **54**

2.9 Fluorescently Activated Cell Sorting (FACS) and Flow Cytometry 54

2.9.1. FACS for Western Blotting **54**

2.9.2 Flow Cytometry and FACS for Peripheral Blood (PB) and Whole Kidney Marrow (WKM) **55**

2.10 RNA-sequencing (RNA-seq) 55

2.10.1 Preparing Embryos for RNA-seq **55**

2.10.2 RNA Extraction Using TRIzol™ Reagent **56**

2.10.3 RNA-seq Alignment and Analysis **56**

Chapter 3: Results 58

3.1: Transgenic *cd45:NND1:P2A:sGFP* Embryos and Adults Display Green Fluorescence in Hematopoietic Sites and Tissues 58

3.1.1 Transgenic cd45:NND1:P2A:sGFP Zebrafish Embryos show Green Fluorescence in Sites of Hematopoiesis, and Express the NUP98-NSD1 Protein **58**

3.1.2 Adult cd45:NND1:P2A:sGFP Fish have GFP+ Cells Present in the Peripheral Blood and Whole Kidney Marrow **61**

3.2: Successful Knockout of the *nup98* Gene using CRISPR/Cas9 ..... **63**

3.2.1 The nup98 Gene was Efficiently Knocked out in Casper Embryos using CRISPR/Cas9 Gene Editing **63**

3.3: Transgenic Embryos with and without *nup98* Knockout, and Adult Transgenic Fish Display Disrupted Blood Development Compared to *Casper* Controls **65**

3.3.1 The NUP98-NSD1 Fusion and the nup98 Knockout (KO) Contribute to Disrupted Blood Development in Zebrafish Embryos **65**

3.3.2 Adult cd45:NND1:P2A:sGFP Transgenic Fish show Changes to the Proportions of Blood Cell Populations Present in the Peripheral Blood (PB) and Whole Kidney Marrow (WKM) **70**

3.4: RNA-sequencing shows Changes in Gene Expression of Transgenic Embryos Compared to *Casper* Embryos **73**

3.4.1 RNA Sequencing Results show a Developmental Delay in Transgenic Zebrafish Embryos **73**

Chapter 4: Discussion **75**

4.1 *cd45:NND1:P2A:sGFP* embryos display GFP in sites of hematopoiesis and have off-target expression **75**

4.1.1 Choice of Promoter Matters for Expression of NUP98-NSD1 in a Transgenic Line **75**

4.1.2 NUP98-NSD1 was Detected in the GFP+ Cells of Transgenic Embryos **77**

4.2 CRISPR/Cas9 was used to knockout *nup98* in zebrafish embryos **77**

4.2.1 Design of the nup98 Knockout and Taq LongAmp® PCR Assay **77**

4.2.2 The Use and Limitations of the HMA Assay to Identify Founder Fish with Mutated nup98 ..... **78**

4.3 Hematopoiesis is disrupted in embryos that express *NUP98-NSD1* and/or have a *nup98* knockout 79

4.3.1 Both NUP98-NSD1 Expression and nup98 Knockout Contribute to the Disruption of Blood Development in Zebrafish Embryos **79**

4.3.2 The Embryonic Phenotype of NUP98-NSD1 Transgenic Embryos Differs from NUP98-HOXA9 Transgenic Embryos **81**

4.3.3 The Embryonic Phenotype of nup98 KO Induced by CRISPR/Cas9 Differs from nup98 Knocked Down with Morpholino **82**

4.3.4 cd45:NND1:P2A:sGFP Embryos with nup98 KO may be Displaying an MDS Phenotype **83**

4.3.5 Ectopic Expression of pu.1 Seen in Transgenic Embryos **84**

4.4 Hematopoiesis is Disrupted in Adult Transgenic Fish 86

4.4.1 cd45:NND1:P2A:sGFP Adult Fish have Different Peripheral Blood and Kidney Marrow Blood Cell Populations Compared to Caspers **86**

4.5 RNA-sequencing Found Many Genes Downregulated in Transgenic Zebrafish Embryos Compared to *Casper* Embryos 87

4.5.1 RNA-seq Data Shows a Developmental Delay in cd45:NND1:P2A:sGFP Transgenic Embryos **87**

4.5.2 Limitations of the RNA-seq and Suggestions for Future Experiments **89**

4.6 Limitations 88

4.6.1 General Limitations of the Zebrafish Model **90**

4.6.2 Many Transgenic Approaches Can Be Taken to Model Cancer **90**

4.6.3 Off-target Expression May Confound Disease Phenotype **91**

4.7 Future Directions 92

4.7.1 Patients with a NUP98-NSD1 Translocation Often Harbour a FLT3-ITD Mutation **92**

4.7.2 Downstream Use of the Transgenic cd45:NND1:P2A:sGFP as a Platform for Drug Discovery **93**

4.8 Conclusions 93

References **95**

Appendix A: Study Approval for *NUP98-NSD1* Project **114**

LIST OF TABLES

Table 1.1.5.1: NUP98 fusion partner genes..... 7

Table 2.4.1.1: the six small guide RNAs (sgRNAs) used to knockout the *nup98* gene in zebrafish..... 41

Table 2.4.3.1: Gel mixes for 6, 8% polyacrylamide gels (15mL each)..... 43

Table 2.6.2.1: ProK permeabilization times for zebrafish embryos at various stages of development..... 47

Table 3.3.1.1: Summary of whole-mount *in situ* hybridization results analyzing changes in blood development of *cd45:NND1:P2A:sGFP* transgenic embryos and *casper* embryos with *nup98* knockout, alone and together, compared to the *casper* control..... 69

LIST OF FIGURES

Figure 1.2.3.1: Both parts of the NUP98-NSD1 fusion protein work together to activate transcription and over-express the *HOXA* genes 12

Figure 1.4.2.1: An overview of zebrafish hematopoiesis and similarities to other vertebrates 19

Figure 1.5.2.1: Creation and establishment of a stable transgenic zebrafish line 24

Figure 1.5.3.1: Creating a stable mutant zebrafish line using CRISPR/Cas9 gene knockout 27

Figure 2.1.1: Stages of early development of the zebrafish 31

Figure 2.3.2.1: Assembly of Gateway® (Invitrogen) cloning vectors to obtain an expression clone 38

Figure 2.4.3.1: Overview of the process and analysis of a heteroduplex mobility assay (HMA) 44

Figure 2.6.4.1: Qualitative and quantitative analysis methods for *in situs* 50

Figure 2.8.1.1: Schematic representation of a zebrafish peripheral blood extraction 53

Figure 3.1.1.1 (*next page*): Transgenic *cd45:NUP98-NSD1:P2A:sGFP* embryos display green fluorescence at sites of hematopoiesis and express the NUP98-NSD1 protein 59

Figure 3.1.2.1: Adult *cd45:NND1:P2A:sGFP* fish have cells positive for green fluorescence in the peripheral blood (PB) (a) and whole kidney marrow (WKM) 62

Figure 3.2.1.1: Assays demonstrating the efficiency of *nup98* knockout (KO) 64

Figure 3.3.1.1 (*continued on next 2 pages*): whole-mount *in situ* hybridization (WISH) assays show that blood development is disrupted in the transgenic *cd45:NND1:P2A:sGFP* and in fish injected with *nup98* knockout (KO) 66

Figure 3.3.2.1: Peripheral blood from adult *casper* and *cd45:NND1:P2A:sGFP* zebrafish show different amounts of blood cell populations 71

Figure 3.3.2.2: Whole kidney marrow from adult *casper* and *cd45:NND1:P2A:sGFP* zebrafish show different amounts of blood cell populations 72

Figure 3.4.1.1: Heat map showing the top genes with increased (red) and decreased (green) expression from RNA sequencing analysis 74

Figure 4.3.4.1: Hematopoietic stem and progenitor cell populations in a healthy individual (a) versus an individual with a low (b) or high (c) risk myelodysplastic syndrome (MDS) 85

ABSTRACT

Many genetic mutations lead to acute myeloid leukemia (AML), including the translocation *NUP98-NSD1* (*NND1*), which is found primarily in pediatric AML and causes high-risk disease. There is currently no animal model of the *NND1* translocation and consequential loss of endogenous NUP98, even though the loss of NUP98 may also affect leukemogenesis. I genetically engineered two zebrafish: a transgenic expressing human *NND1*, and another with decreased *nup98*. Zebrafish embryos with both genetic aberrations displayed disrupted blood development akin to myelodysplastic syndrome (MDS): decreased erythrocyes, decreased early and differentiated myeloid cells, and increased hemtoapoietic stem cells (HSCs). Adult transgenic fish also showed disrupted blood development, similar to MDS and AML, with decreased erythrocytes and lymphocytes, and increased myeloid cells and precursor cells. These results suggest that *NUP98-NSD1* is causing impaired cellular differentiation, which may be manifesting as a myelodysplastic syndrome (MDS). These zebrafish models provide new preclinical platforms to test targeted therapies.

LIST OF ABBREVIATIONS AND SYMBOLS USED

Gene and protein symbol conventions, as represented by the "*Nucleoporin 98"* gene

Species	Gene Symbol	Protein Symbol
Homo sapiens (human)	*NUP98*	NUP98
Mus musculus (mouse), *Xenopus spp.* (frog)	*Nup98*	Nup98
Danio rerio (zebrafish)	*nup98*	nup98

Adapted from Wikipedia, https://en.wikipedia.org/wiki/Gene_nomenclature

Units and Symbols

bp	base pairs
dpf	days post-fertilization
g	gram
x *g*	gravitational acceleration (centrifugation unit)
hpf	hours post-fertilization
k_	kilo_ (10^3)
kb	kilo base pairs
kDa	kilo Dalton (atomic mass unit; 1 kDa = 1.660 538 782(83)$\times 10^{-24}$g)
L	litre
M_	milli_ (10^{-3})
µ_	micro_ (10^{-6})
n_	nano_ (10^{-9})
nt	nucleotide
%	percentage concentration of solution; solid in solvent (1% = 1g per 100mL [w/v], or liquid in solvent (1mL per 100mL [v/v])
rpm	Revolutions per minute
v/v	volume by volume percentage solution
w/v	mass by volume percentage solution

Abbreviations

1°	primary
2°	secondary
AGM	aorta-gonad-mesonephros
ALL	acute lymphoblastic leukemia
ALPM	anterior lateral plate mesoderm
AML	acute myeloid leukemia
ANOVA	analysis of variance

AP	animal polar
APL	acute promyelocytic leukemia
araC	cytarabine
AYA	adolescent and young adult
BSA	bovine serum albumin
Cas9	CRISPR associated protein 9
CBF	core binding factor
CC	coiled coil
ChIP	chromatin immunoprecipitation
CHT	caudal hematopoietic tissue
CML	chronic myelogenous leukemia
CMML	chronic Myelomonocytic leukemia
CMP	common myeloid progenitor
CN-AML	cytogenetically normal AML
COG	children's oncology group
COX	cyclo-oxygenase
CRE	*cis*-regulatory element
CR	complete remission
CRISPR	clustered regularly interspersed short palindromic repeats
CRISPRi	CRISPR interference
DA	dorsal aorta
dCas9	dead Cas9/catalytically inactive Cas9
DIG	digoxogenin
DNA	deoxyribonucleic acid
DNMT	DNA (cytosine-5)-methyltransferase
DOT1L	Disruptor of telomeric silencing 1-like
E. coli	*Escherichia coli*
EMP	erythromyeloid progenitor
ENU	ethylnitrosourea
[E(z)]	enhancer of zeste
FAB	French-American-British
FACS	fluorescently activated cell sorting
FBS	fetal bovine serum
FDR	false discovery rate
FG	phenylalanine-glycine
FISH	fluorescent *in situ* hybridization
FITC	fluorescein isothiocyanate
FL	fetal liver
FLT3-ITD	fms-like tyrosine kinase 3 with and internal tandem duplication
FSC	forward scatter
gDNA	genomic DNA
GFP	green fluorescent protein
GLFG	glycine-leucine-phenylalanine-glycine
GMP	granulocyte monocyte progenitor
GvHD	graft-versus-host disease
HAT	histone acetyltransferase
HD	homeodomain

HDAC	histone deacetylase
HMG	high-mobility group
HMT	histone methyltransferase
HOX	homeobox
HSC	hematopoietic stem cell
Hyb+	hybridization buffer
Hyb-	pre-hybridization buffer
ICM	intermediate cell mass
indel	insertion/deletion
ITD	internal tandem duplication
JMML	juvenile myelomonocytic leukemia
KIT	KIT proto-oncogene receptor tyrosine kinase
KO	knockout
LB	Lysogeny/luria bertani
LBD	ligand binding domain
LSC	leukemic stem cell
MAB	maleic acid buffer
MAB-T	maleic acid buffer with tween
MDS	myelodysplastic syndrome
MeOH	methanol
MEP	megakaryocyte erythroid progenitor
MPN	myeloproliferative neoplasm
MPP	multi-potent progenitor
mRNA	messenger RNA
MUD	matched unrelated donor
NaCl	sodium chloride
NaOH	sodium hydroxide
NE	nuclear envelope
NFH_2O	nuclease-free water
NHA9	NUP98-HOXA9
NID	nuclear receptor-interaction domain
NND1	NUP98-NSD1
NPC	nuclear pore complex
NRAs	neuroblastoma RAS viral oncogene homolog
NSD1	nuclear receptor binding set domain protein 1
NUP	nucleoporin protein/nucleoporin
NUP98	nucleoporin 98
OM-LZ	octapeptide motif-leucine-zipper
p3e	3’ plasmid element
p5e	5’ plasmid element
PB	peripheral blood
PBI	posterior blood island
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with tween
PCA	principal components analysis
PCR	polymerase chain reaction
PFA	paraformaldehyde

PHD	plant homeodomain
polyA	polyadenylation
pMe	middle element plasmid
ProK	proteinase K
PVFD	polyvinylidene fluoride
qPCR	quantitative PCR
RAPD	random amplified polymorphic DNAs
RARα	retinoic acid receptor alpha
RBC	red blood cell
RIPA	radioimmunoprecipitation
RNA	ribonucleic acid
RNA-seq	RNA sequencing
SET	suppressor of variegation-enhancer of zeste-trithorax
SEM	standard error of the mean
sGFP	superfolder GFP
sgRNA	single guide RNA
SSC	side scatter
SSC-T	saline sodium citrate buffer with tween
SOC	super optimal broth with catabolite repression
[Su(var)]	suppressor of variegation
T7 endo	T7 endonuclease
T-ALL	t-cell ALL
tAML	treatment-related AML
TBS-T	tris-buffered saline with tween
tdPCR	touchdown PCR
Tm	annealing temperature
ubi	ubiquitin-C
WHO	world health organization
WISH	whole-mount *in situ* hybridization
WKM	whole kidney marrow
WT	wildtype
WT1	Wilms tumour 1
YS	yolk sac
ZFN	zinc finger nucleases

ACKNOWLEDGEMENTS

I would like to first acknowledge my co-supervisors, Drs. Jason Berman and Graham Dellaire, for their guidance throughout this project. I especially want to thank Jason for allowing me to be creative in my research goals and for supporting me when I took this project in seemingly weird directions. I would also like to thank my graduate committee, Drs. Robert Liwski, Clinton Campbell, Chantal Pambrun (past member), and Craig McCormick for their contributions.

I want to acknowledge Eileen Kaiser: an invaluable asset to the Pathology department; someone who has helped to support the success of every graduate student; and who has cheerfully reminded me about all of the paperwork I had no idea even existed. You made graduate school much more manageable, and I'm grateful for all of our chats.

I want to thank all of the members of the Berman lab for their support and enthusiasm. I want to give a special acknowledgement to Vinothkumar Rajan and Sergey Prykhozhij for taking me under their wings and teaching me more than I could ever understand about genetics and troubleshooting; without you both this project would not be where it is today, and neither would I.

Thank you to all of the graduate students I met along the way – you encouraged me and taught me a lot about science and myself. I'm glad to have taken part of this journey with you all.

Lastly, thank you to my friends and family who listened to my struggles, and struggled to understand why I was so excited that my fish were finally green – your support was unwavering, thank you for believing in me.

This project was funded by the Leukemia Lymphoma Society of Canada (LLSC), the Nova Scotia Health Research Foundation (NSHRF), and the Beatrice Hunter Cancer Research Institute (BHCRI) Cancer Research Training Program (CRTP).

Chapter 1: Introduction

1.1: Acute Myeloid Leukemia (AML)

1.1.1 AML Diagnosis and Epidemiology

Acute myeloid leukemia is an aggressive blood cancer that results from the accumulation of genetic and epigenetic changes in the hematopoietic stem and progenitor cells. These mutations cause cells to undergo aberrant proliferation and lack of differentiation, leading to a buildup of immature myeloblasts in the bone marrow (Estey & Döhner 2006). AML patients typically present with symptoms of anemia (low red blood cell count), neutropenia (low neutrophil count), and thrombocytopenia (low platelet count) due to the replacement of the normal bone marrow with the immature myeloblasts (Rubnitz et al. 2010). Because of this pancytopenia, patients often have symptoms of fatigue, infections, fever, bleeding, and bone pain (Rubnitz et al. 2010). AML is diagnosed when a blast criterion of >20% is met, meaning that over 20% of blood cells in the bone marrow are the immature myeloblasts. However, the blast criterion has changed over the years, suggesting this number to be somewhat arbitrary and that covariates such as cytogenetics, age, and *de novo* vs. secondary AML must be taken into consideration upon diagnosis.

There are two peaks in incidence of AML, one in childhood and one later in adulthood (Deschler & Lübbert 2006). Patients newly diagnosed with AML have a median age of diagnosis of 65 years, and it is uncommon for a patient to be diagnosed before 40 years of age (Deschler & Lübbert 2006). The overall incidence of AML worldwide is 2.5 per 100,000 persons, but in the United States where AML is more common the incidence increases to 3.4 per 100,000 persons (Weir et al. 2003). The incidence for AML in children and adolescence aged 0-19 years is estimated at 5-7 cases per million people per year (Xie et al. 2003). AML diagnoses in adults are slightly skewed towards males, and although most ethnicities have comparable incidence rates for AML, Hispanic/Latino children are known to have higher rates of AML (Deschler & Lübbert 2006; Puumala et al. 2013).

1.1.2 AML Etiology

Although AML is the most common adult leukemia, the causes of AML are poorly understood. Some established risk factors for adult AML include obesity, tobacco use, other toxin exposures like chemotherapeutic agents, and previous hematologic conditions such as myelodysplastic syndrome (MDS) (Finn et al. 2015). AML resulting from some of these risk factors is classified as secondary AML, which is defined as the development of AML after cytotoxic exposure, like chemotherapeutic agents, or with a previous hematologic condition (Cheson et al. 2003). By contrast, the term '*de novo*' AML refers to an AML diagnosis without prior cytotoxic exposure or any clinical history of hematologic malignancies (Cheson et al. 2003). In children, AML more commonly presents as *de novo* AML (Puumala et al. 2013). Pediatric AML risk factors include the presence of trisomy 21, increased maternal age, and exposure to ionizing radiation *in utero* (Puumala et al. 2013). Although AML is heterogeneous, meaning that there are many different types of mutations that can lead to AML, the causes of these mutations in the first place are not clear. These numerous aberrations have variable clinical characteristics and outcomes, and are the main predictors of treatment outcome as well as the sensitivity or resistance to treatment (Gilliland & Tallman 2002; Estey 2012).

It has been hypothesized that the formation of AML requires at least two different types of genetic events, and that these mutations can be broadly divided into two categories: type I and type II aberrations (Gary Gilliland & Griffin 2002). Type I aberrations lead to uncontrolled cell proliferation and/or survival of leukemic cells, and include oncogenes like *fms-like tyrosine kinase III* with an internal tandem duplication (*FLT3*-ITD), *KIT proto-oncogene receptor tyrosine kinase* (*KIT*), and *Neuroblastoma RAS viral oncogene homolog* (*NRAS*) (Balgobind et al. 2011). In contrast, Type II aberrations lead to impaired cellular differentiation of leukemic cells, are often chromosomal rearrangements, and include gene fusions like *Promyelocytic Leukemia/Retinoic Acid Receptor Alpha (PML-RARα)*, and *Nucleoporin 98/Nuclear receptor-binding SET domain protein 1* (*NUP98-NSD1*) (Balgobind et al. 2011).

1.1.3 AML Classification and Treatment

Classification systems for AML are different than solid tumours, and there are two systems used to classify AML. The French-American-British (FAB) classification system was the original classification system and it focuses on the morphology of the cells; the subtypes are based on the cell of origin and the maturity of the cell. FAB classification is informative for treatment options. For example, the FAB subtype M3 indicates acute promyelocytic leukemia (APL), which has a treatment regimen different than other types of leukemia as it includes all-trans retinoic acid (ATRA) in the presence of the translocation t(15;17) *(PML-RARα)* (Huang et al. 1988). The World Health Organization (WHO) classification system is newer than FAB and is more widely used; it focuses on cytogenetics for recurrent translocations and other genetic abnormalities. The WHO classification is also useful when deciding treatment options, as cytogenetics can be one of the best predictors of outcome.

Standard AML treatment starts with induction therapy followed by consolidation therapy, which may include stem cell transplantation. Induction therapy involves doses of the chemotherapeutic agent cytarabine (araC), a nucleoside analog, followed by an anthracycline such as daunorubicin or idarubicin (Estey 2012). These drugs are given with a goal of inducing remission in a patient (only a small number of remaining blast cells). After induction therapy, consolidation therapy is given. This consists of more chemotherapy and may include an allogeneic stem cell transplant, meaning that the transplanted cells are coming from a donor that can be a family member such as a sibling, or a matched unrelated donor (MUD). The hope is that the new, leukemia-free hematopoietic stem cells (HSCs) from the donor will repopulate the bone marrow. These stem cell transplants are only given to high-risk patients due to the possible complications that can occur from the transplant itself, like graft-versus-host disease (GvHD) and treatment related mortality (Niewerth et al. 2010). If no stem cell transplant is given, consolidation therapy consists of araC. There remains the issue of minimal residual disease (MRD), or subclinical levels of leukemia cells in a patient after treatment that are capable of recapitulating the disease, prompting ongoing research for how best to monitor MRD in patients (Grimwade & Freeman 2014). The long-term survival of pediatric AML patients has improved due to factors such as the increased intensity of chemotherapy regimes, better supportive care at relapse, judicious stem cell

transplants, and better stratification of patients into cytogenetic risk groups (Perel et al. 2005; Abrahamsson et al. 2011; Entz-Werle et al. 2005; Tsukimoto et al. 2009).

1.1.4 Pediatric versus Adult AML

Although AML is generally considered to be a disease of older adults, with a mean age of diagnosis of 65 years, AML makes up almost 25% of pediatric leukemias (Tarlock & Meshinchi 2014; National Cancer Institute 2014). Although acute lymphoblastic leukemia (ALL) is much more common in children, with ~75% of acute leukemias in children diagnosed as ALL, AML accounts for a disproportionate amount of deaths from acute leukemia, at 30% (Rubnitz et al. 2010). The age at diagnosis affects both the overall health of the patient and the type of AML. It is clear that with elderly patients there is a higher frequency of comorbidities such as cardiovascular disease, decreased kidney function, and overall frailty (Weinstein & Anderson 2010; Clegg et al. 2013); these additional health problems result in an inferior response to treatment overall (Appelbaum et al. 2006). Older patients are also more likely to have unfavourable cytogenetics compared to pediatric patients; have their AML preceded by a myelodysplastic phase, which is the production of abnormal, non-functional, immature white blood cells; and to more commonly express multi-drug resistance genes impacting the efficacy of treatment (Appelbaum et al. 2006; Leith et al. 1997). For pediatric patients, approximately 30% will experience disease relapse, and approximately 5-10% of patients will die due to either disease complications or treatment side effects (de Rooij et al. 2015). Treating pediatric patients also comes with different challenges such as the long-term side effects of treatment toxicities; children, with a longer lifespan that adults, will have to manage treatment side-effects for a longer period of time. The cytotoxic chemotherapy given to pediatric patients can affect reproductive ability, heart and lung function, and have other long-term medical and social health consequences; for example, survivors of childhood cancer are less likely to have close interpersonal relationships, and more likely to have learning disabilities (Schwartz 1999; Barrera et al. 2005). The use of targeted therapies, likely in combination with chemotherapy, may reduce overall treatment toxicities and improve the efficacy of these therapies. Unlike chemotherapy, which targets all rapidly dividing cells, a targeted therapy is designed to act on a specific pathway, molecule, or feature of

a cancer cell, and so does not target healthy tissues. However, these targeted therapies can have off-target effects. For example, drugs designed specifically to target tyrosine kinases are known to have cardiotoxic effects (Force et al. 2007). As we increase our understanding of the genetic and epigenetic mutations that cause AML, this opens up more possibilities for the development of targeted therapies (Napper & Watson 2013).

Until recently, genomic events in pediatric AML were detected by extrapolating genomic events from adult AML, leading to the incorrect assumption that there was a lack of genomic lesions in pediatric AML (Appelbaum et al. 2006). Furthermore, the profile of genetic mutations in pediatric/adolescent and young adult (AYA) patients is complex and distinct, with the genetic signatures in younger children differing from that of AYA patients (Tarlock & Meshinchi 2014). There are four cytogenetic categories that most pediatric patients fit into: 25% of patients have core-binding factor (CBF) AML (t(8;21) or inv(16;16)); 12% have APL (t(15;17)); 20% have rearrangements involving the *KMT2A* (formerly the mixed lineage leukemia, *MLL*) gene; and 20% have a seemingly normal karyotype with potential for underlying cryptic chromosomal abnormalities (Tarlock & Meshinchi 2014). Cryptic abnormalities are not detected with conventional karyotyping methods and require other methods of detection such as polymerase chain reaction (PCR) or fluorescent *in situ* hybridization (FISH). Previously undetected cryptic abnormalities, like the translocation *NUP98-NSD1*, are now emerging as significant contributors to AML pathogenesis in children, with a higher incidence in young children and a lower incidence in adults (Tarlock & Meshinchi 2014).

1.1.5 NUP98 Fusions in AML and Other Hematologic Malignancies

Gene fusions are known to be strong driver mutations in cancer, providing insight into mechanisms for disease pathogenesis. Identifying different gene fusions can help with patient stratification and the potential for targeted treatment (Mertens et al. 2015). The *nucleoporin 98* (*NUP98*) gene is located on chromosome 11p15.5, and NUP98 normally functions as a member of the nuclear pore complex (NPC), involved in mediating nuclear-cytoplasmic transport of protein and RNA (Lam & Aplan 2001; Slape & Aplan 2004). Chromosomal rearrangements involving the *NUP98* gene with different partner genes are present in many different types of hematologic malignancies,

including AML, chronic myeloid leukemia in blast crisis (CML-bc), myelodysplastic syndrome (MDS), acute lymphoblastic leukemia (ALL), and bi-lineage/bi-phenotypic leukemia (Gough et al. 2011). *NUP98* was first found to be partnered with the homeobox (HOX) gene *HOXA9* in AML, forming the fusion gene *NUP98-HOXA9* (*NHA9*) (Nakamura et al. 1996; Borrow et al. 1996). Since then, *NUP98* has been found to partner with 32 different genes, listed in Table 1.1.5.1 below. These partner genes can be broadly divided into two groups – homeodomain (HD) proteins, and non-HD proteins. The 12 HD proteins include the HOX- proteins, as well as PMX1, PMX2, HHEX, POU1F1, and GSX2 (Gough et al. 2011). With all of these HD proteins, the glycine-leucine-phenylalanine-glycine (GLFG) repeats on the N-terminus of NUP98 are fused to the C-terminal DNA-binding domains of the HD proteins. The remaining 20 fusions are non-HD proteins, and most but not all of these also have a DNA-binding domain fused to the NUP98 GLFG repeats (Gough et al. 2011). Many of the non-HD proteins also contain a coiled coil (CC) domain which is thought to cause the oligomerization of proteins (Hussey & Dobrovic 2002). Another theme amongst the non-HD fusion partners is the presence of plant homeodomain (PHD) fingers which are chromatin recognition domains that have been shown to be essential for leukemogenesis (Wang et al. 2009). Three of the non-HD fusion genes (*NSD1, NSD3, MLL*) in Table 1.1.5.1 below contain a suppressor of variegation-enhancer of zeste-trithorax (SET) domain, which have a histone methyltransferase function that provides a specific mechanism for leukemogenesis and will be discussed in section 1.2.3.

Table 1.1.5.1: NUP98 fusion partner genes (Adapted from Gough et al. 2011).

Partner Gene	Chromosome	Disease(s)	Relevant Domain(s)	Reference(s)
HOXA9	7p15	MDS, AML, CML, CMML	HD	(Nakamura et al. 1996; Borrow et al. 1996)
HOXA11	7p15	MDS, AML, CML, JMML	HD	(Fujino et al. 2002)
HOXA13	7p15	AML	HD	(Taketani, Taki, Ono, et al. 2002)
HOXC11	12q13	AML	HD	(Taketani, Taki & Shibuya 2002a)
HOXC13	12q13	AML	HD	(Panagopoulos et al. 2003)
HOXD11	2q31	AML	HD	(Taketani, Taki & Shibuya 2002b)
HOXD13	2q31	MDS, AML, CML	HD	(Raza-Egilmez et al. 1998)
PMX1	1q23	AML, CML	HD	(Nakamura et al. 1999)
PMX2	9q34	AML	HD	(Gervais et al. 2005)
HHEX	10q23	AML	HD	(Gurevich et al. 2004)
POU1F1	3p11	tAML	HD	(Lisboa et al. 2013)
GSX2	4q11	AML	HD	(Soler et al. 2013)
PHF23	17p13	AML	PHD, CC	(Reader et al. 2007)
JARID1A	12p13	AML	PHD, CC	(van Zutven et al. 2006)
NSD1	5q35	MDS, AML, T-ALL	PHD, CC, SET	(Jaju et al. 2001)
NSD3	8p11	MDS, AML	PHD, CC, SET	(Rosati et al. 2002)
MLL	11q23	AML	PHD, CC, SET	(Kaltenbach et al. 2010)
SETBP1	18q12	T-ALL	CC	(Panagopoulos et al. 2006)
LEDGF	9p22	AML, CML	CC	(Ahuja et al. 2000)
CCDC28	6q24	AML, T-ALL	CC	(Tosi et al. 2005)
HMGB3	Xq28	AML	CC	(Petit et al. 2010)
IQCG	3q29	T-ALL	CC	(Pan et al. 2008)
RAP1GDS1	4q21	AML, T-ALL	CC	(Hussey et al. 1999)
ADD3	10q25	AML, T-ALL	CC	(Lahortiga et al. 2003)
DDX10	11q22	MDS, AML, CML, CMML	CC	(Arai et al. 1997)
TOP1	20q11	MDS, AML	CC	(Ahuja et al. 1999)
TOP2B	3p24	AML	CC	(Nebral et al. 2005)
LNP1	3q12	AML	CC	(Gorello et al. 2008)
RARG	12q13	AML	CC	(Such et al. 2011)
ANKRD28	3p25	AML	Ankyrin	(Ishikawa et al. 2007)
AF10	10p12	MDS	OM-LZ, Q-rich region	(Soler et al. 2013)
FN1	2q31	AML	Unknown	(Arai et al. 2000)

1.2: *NUP98-NSD1* in AML

1.2.1 Discovery and Characterization of the Patient Group Defined by the NUP98-NSD1 Translocation

In 2001, Jaju and colleagues published the first report of the gene *nuclear receptor-binding SET domain protein 1* (*NSD1*) involved in a human hematologic malignancy. They characterized the *NUP98-NSD1* fusion from *de novo* childhood AML patient samples after discovering the translocation a few years earlier (Jaju et al. 1999; Jaju et al. 2001). It was found that the *NSD1* gene was translocated and fused in frame to *NUP98* (t(5;11)(q35;p15.5)); this translocation is cytogenetically cryptic and therefore cannot be detected by normal karyotyping (Brown et al. 2002). *NUP98-NSD1* was originally detected by FISH, and rapid amplification of cDNA ends (RACE)-PCR was used to detect the *NSD1* gene in the fusion (Jaju et al. 2001). Researchers Jaju and Brown originally detected the reciprocal translocation *NSD1-NUP98 via* PCR, and so initially both the *NUP98-NSD1* and *NSD1-NUP98* transcripts were thought to potentially play a role in leukemogenesis (Jaju et al. 2001; Brown et al. 2002). However, further research by Cerveira and colleagues showed that in their patient samples, only *NUP98-NSD1* could be detected and not *NSD1-NUP98*, giving evidence to the theory that *NUP98-NSD1* likely encoded the biologically relevant fusion protein (Cerveira et al. 2003). *NUP98-NSD1* was later detected in a panel of 20 unselected patient samples, which gave the first indication about frequency of *NUP98-NSD1* in AML patients at approximately 5% (Cerveira et al. 2003). Due to the small number of patient samples with *NUP98-NSD1*, prognostic significance could only state generally that *NUP98-NSD1* indicated poor prognosis (Jaju et al. 2001; Brown et al. 2002; Panarello et al. 2002; Cerveira et al. 2003). A few years later, a mechanism of leukemogenesis was proposed for *NUP98-NSD1*, and this is discussed in section 1.2.3.

It is now known that the presence of *NUP98-NSD1* characterizes an AML patient group with poor prognostic outcome (Hollink et al. 2011; Shiba et al. 2013). Hollink and colleagues performed the first systematic analysis of *NUP98-NSD1* in a large AML cohort. They screened over 1000 cytogenetically normal (CN)-AML patient samples from children and adults and found that *NUP98-NSD1* was detected in 16.1% of pediatric samples and 2.3% of adult samples (Hollink et al. 2011). FAB-M4/M5 monocytic morphology was more frequent in *NUP98-NSD1* positive adult samples,

although other studies have found no association of *NUP98-NSD1* with an FAB subtype (Hollink et al. 2011; Thol et al. 2013). *NUP98-NSD1* is associated with the first report of a *NUP98* fusion being associated with higher blast counts (Fasan et al. 2013). *NUP98-NSD1* was an independent predictor of prognosis with a 4-year event-free survival of less than 10% for both pediatric and adult patients (Hollink et al. 2011). Complete remission (CR) rates at the end of induction therapy are also impacted, with *NUP98-NSD1*-positive patients having a 43% CR rate compared to 77% for *NUP98-NSD1*-negative patients (Thol et al. 2013). When secondary mutations were analyzed, it was found that in *NUP98-NSD1*-positive samples, 82% harboured an internal tandem duplication (ITD) in the *FLT3* gene (*FLT3*-ITD), 45% had a *Wilms tumour 1 (WT1*) mutation, and one had an *NRAS* mutation; all of these secondary mutations are classified as type I aberrations in leukemia, and *NUP98-NSD1* was mutually exclusive with all type II aberrations (Hollink et al. 2011).

1.2.2 NSD1 in Health and Disease

To understand the consequences of the *NUP98-NSD1* fusion gene, it is important to first recognize the normal and aberrant functions of *NSD1* alone. NSD1 was originally discovered by Huang and colleagues using a yeast two-hybrid screen, where NSD1 was baited with the ligand-binding domain (LBD) of retinoic acid receptor alpha (RARα) (Huang et al. 1998). NSD1 contains two nuclear receptor-interacting domains (NIDs) that have characteristics of both co-activators and co-repressors (Huang et al. 1998). NSD1 also contains a SET domain that is part of the SET2 family; the histone methyltransferase (HMT) activity of NSD1, like other HMTs in the SET2 family, is specific for histone position H3K36, meaning that it methylates the lysine (K) at position 36 of the histone H3 tail (Schneider et al. 2002; Strahl et al. 2002; Kouzarides 2007). NSD1 is required for normal development: mouse embryos homozygous mutant for *Nsd1*-/- display high levels of apoptosis and die during gastrulation (Rayasam et al. 2003). Human diseases involving *NSD1* include Weaver syndrome and Sotos syndrome, where a haploinsufficiency of *NSD1* leads to these overgrowth syndromes (Douglas et al. 2003; Turkmen et al. 2003; Rio et al. 2003; Kurotaki et al. 2002). Both Weaver and Sotos syndrome also cause an increased risk of tumourigenesis (Lapunzina 2005). In terms of cancer, NSD1 is known to be involved in a few different types of malignancies.

In neuroblastoma, it was found that the hypermethylation of *NSD1* CpG islands decreased expression of *NSD1*, which was a predictor of poor outcome in high-risk neuroblastoma (Berdasco et al. 2009). Furthermore, restoring the expression of NSD1 in a neuroblastoma cell line reduced colony forming and inhibited cellular growth (Berdasco et al. 2009). As previously discussed in section 1.2.1, *NSD1* is also implicated in AML through the *NUP98-NSD1* fusion gene (Jaju et al. 2001; Cerveira et al. 2003).

1.2.3 Aberrant Expression of NUP98-NSD1 and the Effect on HOX Gene Expression

The first characterization of the molecular mechanisms and transforming properties of *NUP98-NSD1* came from Wang and colleagues in 2007. They were able to specifically link deregulated H3K36 methylation to tumourigenesis, and NSD1 to the transcriptional regulation of the *Hox-A* locus. In an *in vitro* myeloid progenitor transformation assay it was found that marrow-derived progenitors retrovirally transduced to express NUP98-NSD1 proliferated indefinitely as undifferentiated progenitor cells, specifically of myeloblast morphology, whereas cells transduced with an empty vector were able to proliferate transiently and differentiate into macrophages, neutrophils, and mast cells (Wang et al. 2007). Therefore, the principal transforming property of NUP98-NSD1 is blocking cellular differentiation and enforcing progenitor self-renewal. This established *NUP98-NSD1* as a type II aberration: an inhibitor of cellular differentiation. For *in vivo* analysis, bone marrow samples from Balb/c mice were sorted for cells expressing no mature lineage markers (Lin-), and Lin- cells selected for drug resistance were used (Wang et al. 2007). When analyzing the effects of *NUP98-NSD1* in *vivo*, only the transplanted Lin- progenitor cells expressing NUP98-NSD1 induced AML in a mouse model, while mice transplanted with Lin- cells expressing an empty vector remained healthy (Wang et al. 2007).

To explore a mechanism causing differentiation arrest in the myeloid progenitors, a microarray was performed to analyze gene expression profiles between the cells expressing NUP98-NSD1 or empty vector. The six genes that were increased in cells expressing NUP98-NSD1 were *HoxA5, HoxA7, HoxA9, HoxA10, Meis1,* and *Rab38* (Wang et al. 2007). Using chromatin immunoprecipitation (ChIP), it was found that NUP98-NSD1 binds the *HoxA7* and *HoxA9* transcriptional start sites using the PHD fingers of NSD1 (Wang et al. 2007). There are at least two methods by which NUP98-

NSD1 can modify chromatin: first, NUP98 recruits CBP/p300 acetyltransferases *via* the GLFG repeats, and second, NSD1 trimethylates H3K36 *via* the SET domain (Figure 1.2.3.1). The CPB and p300 acetyltransferases are transcriptional co-activators that add acetylation onto the histone, creating an open chromatin formation and increased gene transcription (Ogryzko et al. 1996; Das et al. 2009). The addition of the trimethylation at position H3K36 by NSD1 is activating, and also causes an open chromatin formation leading to increased gene transcription (Wang et al. 2007). Attenuation of functioning NUP98, the NSD1 SET domain, or the PHD fingers of NSD1 prevented progenitor self-renewal, confirming the methylation of H3K36 as essential to the overexpression of *HoxA*- genes (Wang et al. 2007).

After a leukemogenic mechanism for *NUP98-NSD1* was reported, pediatric AML patient samples with and without *NUP98-NSD1* were further explored to molecularly and mechanistically define this subgroup of patients. More recent studies of large pediatric AML samples show that patients harbouring *NUP98-NSD1* make up almost 5% of pediatric AML patients and have a poor prognosis (Shiba et al. 2013; Hollink et al. 2011). Expression data from AML patients shows that patients with *NUP98-NSD1* have a distinct *HOX-A* and *–B* expression pattern, and are separate from other types of AML such as those with *KMT2A* rearrangements (Hollink et al. 2011; Shiba et al. 2013). It is known that certain *HOXA* and *–B* genes are leukemogenic, although more is known about the *HOXA* genes, leaving room for the exploration of the role of *HOXB* gene overexpression in *NUP98-NSD1*-induced AML (Argiropoulos & Humphries 2007; Hollink et al. 2011; Shiba et al. 2013).

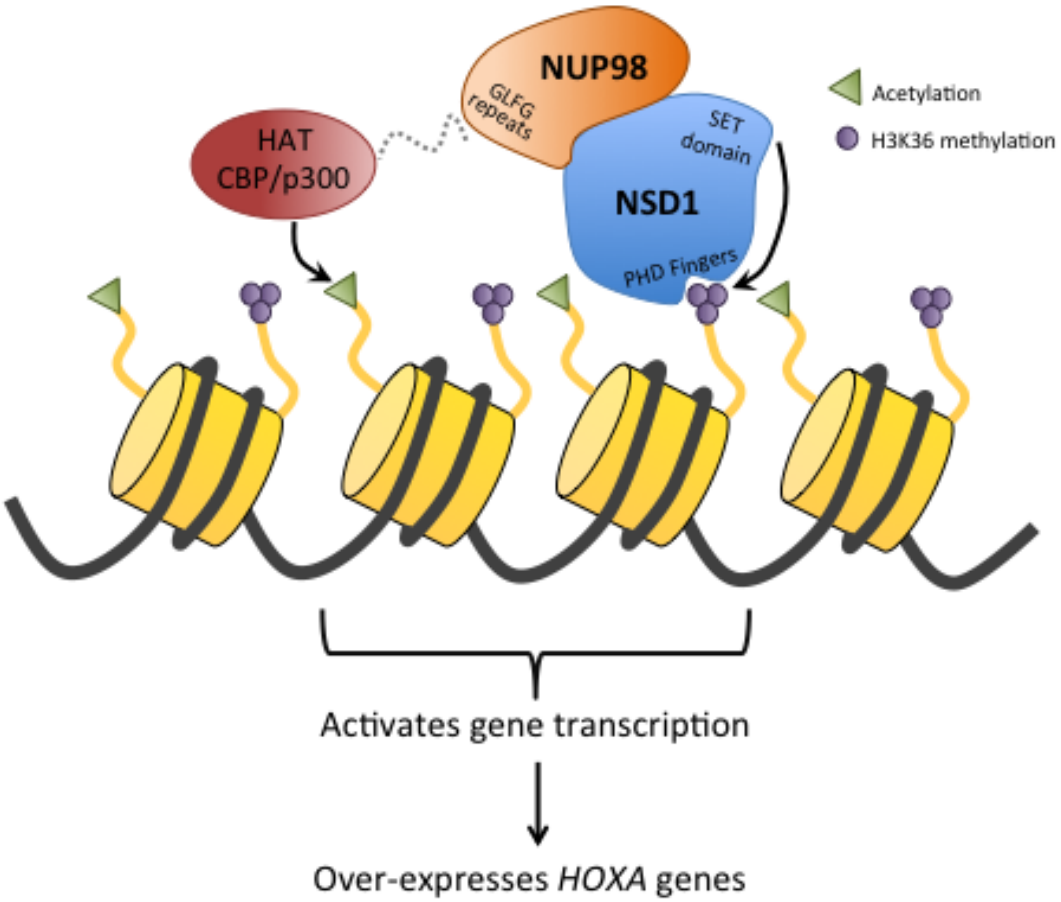


Figure 1.2.3.1: Both parts of the NUP98-NSD1 fusion protein work together to activate transcription and over-express the *HOXA* genes. The NSD1 portion of the fusion protein binds chromatin using the PHD fingers present on the protein. As a histone methyltransferase, NSD1 trimethylates the histone tail position H3K36 with the SET domain part of the protein. This trimethylation on H3K36 promotes gene transcription. The NUP98 portion of the protein attracts proteins such as CBP/p300 *via* the GLFG repeats in the protein. CBP/p300 are histone acetyltransferases (HATs) that add acetylation onto the histone tails and in this way also promote gene transcription. Both of these mechanisms (trimethylation and acetylation) work together to cause an open chromatin formation and gene transcription, which leads to the overexpression of genes like the *HOXA* genes.

1.2.4 Previous Studies Involving Models of NUP98-NSD1-Induced Leukemia

Much knowledge has been gained from previous studies investigating the role of *NUP98-NSD1* in leukemogenesis. Wang and colleagues linked the methylation of H3K36 by NUP98-NSD1 to *HOXA* overexpression using a combination of *in vitro* work and *in vivo* models, specifically by transplanting Lin- bone marrow cells from Balb/c mice that expressed *NUP98-NSD1* into mice, which resulted in AML (Wang et al. 2007). Thanasopoulou and colleagues used a similar *ex vivo* approach of transducing the same type of cells to express either *NUP98-NSD1*, *FLT3*-ITD, or both, to discover that only cells expressing *NUP98-NSD1* with *FLT3*-ITD resulted in AML in the mice (Thanasopoulou et al. 2014). A study that echoed the work of Thanasopoulou and colleagues looked at drug sensitivity profiling in primary patient cells transduced to express *NUP98-NSD1*, *FLT3*-ITD, or both, and found that cells expressing both oncogenes were sensitive to FLT3 inhibitors, the JAK-inhibitor ruxolitinib, and the BCL2-inhibitor navitoclax (Kivioja et al. 2014). There was also a distinct drug response pattern between the same cells *in vitro* and *in vivo*, suggesting a possible role for the microenvironment and further highlighting the importance of *in vivo* models of *NUP98-NSD1* AML. Other drug response studies include *in vitro* work to determine effective treatments for *NUP98-NSD1*, where it was found that cells once again transduced to express *NUP98-NSD1* were sensitive to small-molecule inhibitors like DOT1L (Disruptor of telomeric silencing 1-like), which is a histone methyltransferase responsible for H3K79 methylation that causes active gene transcription (Deshpande et al. 2014).

1.3: NUP98 and Nucleoporins in Health and Disease

1.3.1 Nucleoporins and the Nuclear Pore Complex (NPC)

Nucleoporin proteins, collectively called NUPs, are vital cellular components that mediate nuclear-cytoplasmic transport. This group of approximately 30 proteins comes together to make up the nuclear pore complex (NPC), which is evolutionarily conserved amongst different species in terms of structure, and is intrinsically linked to eukaryotic gene expression (Raices & D'Angelo 2012; Kohler & Hurt 2010). The NPC forms a channel across the nuclear envelope (NE) allowing the selective passage of ions and other molecules to and from the nucleus (Walde & Kehlenbach 2010). NUPs are organized into categories based on structure. First, approximately one third of NUPs

contain phenylalanine-glycine (FG)-rich repeat sequences that interact with molecules to shuttle them across the NPC (Peters 2009). Second, NUPs without FG-repeats are thought to provide the framework for NPC structure, and are responsible for things like anchoring the NPC to the NE, and forming different parts of the pore itself (Kohler & Hurt 2010). Although NUPs have a predominantly structural role in general, certain NUPs are also involved in many fundamental cellular processes such as differentiation, gene expression, and chromatin organization (Walde & Kehlenbach 2010; Liang et al. 2013; Raices & D'Angelo 2012). For example, NUP98 and a subset of other NUPs containing FG-repeats were found to interact with developmentally regulated genes undergoing transcription induction (Capelson et al. 2010; Kalverda et al. 2010). Due to the alternate functions of NUPs it is unsurprising, yet intriguing, that genetic mutations affecting *NUPs* are linked to many different diseases such as neurological diseases, autoimmune dysfunctions, and cancer (Nofrini et al. 2016).

1.3.2 De-regulation of NUP98 and Disease Consequence in AML

A wide variety of nucleoporins are involved in different diseases, and *NUP98* in particular corresponds with complex levels of pathogenesis. As a nucleoporin, NUP98 is part of the NPC, specifically involved in transcription-dependent mobility, and can also interact with chromatin away from the nuclear envelope (Nofrini et al. 2016; Franks & Hetzer 2013). In terms of disease, NUP98 is most widely known for the promiscuous fusions with genes causing a plethora of different translocations involved in AML (Gough et al. 2011). The current model proposes that the NUP98 fusion partner, either an HD or non-HD protein, is recruited to aberrant loci and may modify the loci in some way (Gough et al. 2011; Wang et al. 2007). What this paradigm does not include is the potential loss of wild type (WT) NUP98 (referring to normal or endogenous levels of expression) in the context of the *NUP98-NSD1* translocation and how this may contribute to disease pathogenesis (Franks & Hetzer 2013).

An important aspect of *NUP98* is that it is bicistronic, meaning that it encodes both *NUP98* and *NUP96* by making a NUP98-NUP96 fusion protein that is cleaved by autoproteolysis to form NUP98 and NUP96 (Fontoura et al. 1999). This is especially interesting as NUP98 is a peripheral NUP, not confined to the NPC, whereas NUP96 is a scaffold NUP; yet these proteins are produced together in stoichiometric amounts. It is

worth noting that there is a splice variant that translates only NUP98, but this transcript is seen in much smaller amounts (Fontoura et al. 1999). As NUP98 and NUP96 are made in relatively similar amounts, deregulation of *NUP98* has consequences for both NUP98 and NUP96. Protein levels of NUP96 are closely regulated during mitosis: levels of NUP96 are decreased approximately 50% by ubiquitination and subsequent degradation at the onset of mitosis, and are increased again during interphase (Chakraborty et al. 2008). It is currently unknown why NUP96 levels are reduced during mitosis, but overexpression of NUP96 causes delays in the cell cycle (Chakraborty et al. 2008)

NUP98 is capable of acting as a transcription factor as it can interact with histone acetyltransferases (HATs) and histone deacetylases (HDACs) through the GLFG repeats, which is also required for the onset of AML in the fusions (Kasper et al. 1999; Bai et al. 2006; Wang et al. 2007). It is not known how NUP98 may activate transcription outside of the context of a fusion protein, as NUP98 does not have an authentic DNA binding domain like some of the fusion protein partners (Franks & Hetzer 2013; Wang et al. 2007; Wang et al. 2009). However, not all of the NUP98 fusion protein partners have a DNA binding domain, and so the question of how some of these fusion proteins cause leukemia is still unanswered. What also requires further exploration is how a broad range of functionally distinct *NUP98* fusions can lead to similar AML phenotypes (Franks & Hetzer 2013; Kohler & Hurt 2010). One similarity that many of these fusions have is the potential for decreased levels of WT NUP98 caused by the translocations. It has been suggested that under normal conditions NUP98 may be a potent transcriptional regulator (Franks & Hetzer 2013). Therefore, a *NUP98* fusion scenario as in AML could improperly regulate transcription in the following three ways: the decrease in NUP98 caused by the fusion may reduce the ability of NUP98 to regulate its target genes; the fusion partner may contribute to misregulation if it is a transcriptional regulator itself (like NSD1); and the fusion protein can also sequester remaining WT NUP98 through the GLFG repeats that are also a self-aggregation domain (Franks & Hetzer 2013). When this hypothesis is combined with what we know of NUP96 as a cell-cycle regulator and how WT NUP98 decreases could perturb nuclear cytoplasmic transport and transcription, we get a very complex picture of what may be occurring in the cell. Further studies need to examine the role of WT

NUP98 in hematopoiesis, as this could greatly increase our understanding of how NUP98 fusions cause leukemia.

There have been a handful of *in vivo* studies examining the role of disrupted *Nup98*. It was found in models of *Xenopus* and mice that defective *Nup98* blocked NPC assembly and RNA nuclear/cytoplasmic transport (Powers et al. 1997; Wu et al. 2001). Of note, mouse models of *Nup98-/-* were embryonic lethal and *Nup98+/-* caused severe developmental delays (Wu et al. 2001). One group has looked at disrupted *nup98* in the zebrafish using a morpholino knockdown, with the goal of examining the role of *nup98* in zebrafish development (Fung et al. 2010). It was found that the protein sequence of nup98 had 65% homology with NUP98, and ectopic expression of *nup98* mRNA was able to rescue a *NUP98* knockdown in HeLa cells (Fung et al. 2010). Fung and colleagues found that *nup98* was expressed diffusely in the eyes and developing brain of the fish starting at 18 hours post-fertilization (hpf). The knockdown was also found to be embryonic lethal, with notable cranial hemorrhage by 48 hpf. In terms of blood development, it was shown by quantitative PCR (qPCR) that fish with a *nup98* knockdown had up-regulated *pu.1* (early myeloid) and *scl* (HSC) expression, with no changes to genes involved in erythropoiesis (Fung et al. 2010). More *in vivo* studies are needed to elucidate the role of *NUP98* in hematopoiesis, and the zebrafish model presents an excellent opportunity to further explore the effects of decreased *nup98* on embryonic and hematopoietic development.

1.4: Zebrafish Blood Development and the Use of Zebrafish as a Model Organism

1.4.1 Overview of the Zebrafish as a Model Organism

The zebrafish, *Danio rerio*, has emerged as an excellent animal model for studying many different malignancies, most notably leukemia. Although zebrafish do not have the evolutionary proximity to humans of a mammalian animal model like rodents, their status as vertebrates infer developmental similarities in many areas, like blood development, making studies in zebrafish relevant to humans (Lieschke & Currie 2007). From a practical standpoint, zebrafish husbandry has a much lower cost per animal compared to a mouse or rat, and requires less infrastructure (Lieschke & Currie 2007). The lifespan of a zebrafish, with rapid development during embryogenesis, means that experiments can be completed in a shorter time frame, also reducing costs.

In terms of the execution of experiments, zebrafish produce large amounts of embryos that develop externally, meaning many fish can be used per experiment with ready access to all developmental stages (Lieschke & Currie 2007). This is further aided by the creation of zebrafish models that lack pigment and are therefore transparent into adulthood, allowing for examination of internal organs and easier fluorescent imaging (White et al. 2008). Zebrafish also have external fertilization of embryos, facilitating genetic manipulation of embryos (Patton & Zon 2001). One of the biggest advantages of the zebrafish is that it lends itself easily to high-throughput drug screens, combining the large screening numbers comparable to *in vitro* cell culture experiments and the whole-organism physiological responses unique to an *in vivo* model (Chakraborty et al. 2009). Another big advantage of the zebrafish model is the ease in which imaging (fluorescence, confocal, bright field) can be undertaken, often in real-time and at a single-cell level in the live organism (Lieschke & Currie 2007; Traver et al. 2003; Lawson & Weinstein 2002). Limitations of the zebrafish model centre around the lack of zebrafish-specific tools for different experiments, like cell surface markers for flow cytometry or antibodies for Western blots, which will hopefully be overcome as demand for zebrafish-specific tools increases (Konantz et al. 2012).

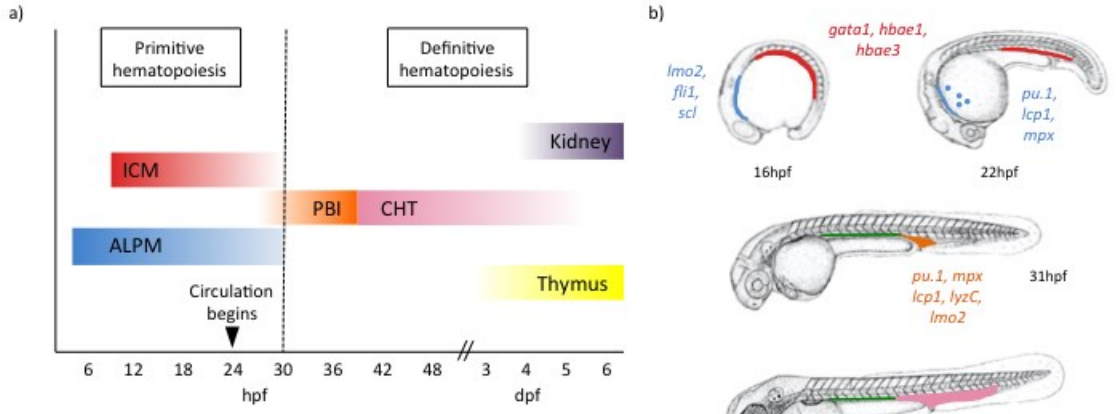
1.4.2 Overview of Hematopoiesis in the Zebrafish

The zebrafish is an excellent model for studying hematopoiesis, and over the past decade there has been an accumulation of knowledge not just about intricacies of zebrafish hematopoiesis, but also how this relates to human hematologic malignancies. All of this translational application is possible because of the similarities of blood development in vertebrates. Despite the differences in zebrafish and mammalian hematopoiesis that are described below, zebrafish and other vertebrates share genetic programs that regulate hematopoiesis, allowing the knowledge gained from zebrafish hematopoiesis to be applied to humans (Paik & Zon 2010).

Zebrafish, like all other vertebrate organisms, experience waves of hematopoiesis (Galloway & Zon 2003). The first hematopoietic wave is primitive, with the main function being to support the embryo through this period of rapid growth and development (Orkin & Zon 2008). In zebrafish, the first primitive wave occurs from ~8-30 hpf in two parts: primitive macrophages arise from the anterior lateral plate

mesoderm (ALPM) and migrate to the yolk sac; and primitive erythrocytes arise from the intermediate cell mass (ICM). In mammals, this primitive hematopoiesis with both unipotent cell types occurs in the extraembryonic yolk sac (also a mesoderm-derived tissue) at an early developmental stage (Cumano & Godin 2007) (Figure 1.4.2.1). The role of the primitive erythrocytes is to provide oxygen to the growing embryo, and the role of the primitive macrophages is still an area of active research; however there is some suggestion that they may help remodel developing tissue through the removal of apoptotic cells (Bertrand & Traver 2009). This primitive wave is transient and gives way to the definitive wave of hematopoiesis, which occurs later on in development at varying time points depending on the vertebrate (Jagannathan-Bogdan & Zon 2013) (Figure 1.4.2.1).

The goal of definitive hematopoiesis is ultimately the production of hematopoietic stem cells (HSCs) that will give rise to all other blood cells. Similar to the primitive wave, the definitive wave also has two parts. In zebrafish, the definitive wave starts around 30 hpf and is first initiated through committed erythromyeloid progenitors (EMPs) found in the posterior blood island (PBI), that arise independently of HSCs (Bertrand et al. 2007). In mammals, these EMPs are formed in the yolk sac and then migrate to the fetal liver (FL) (Palis et al. 1999; Bertrand et al. 2005) (Figure 1.4.2.1). The EMPs can only differentiate into erythroid or myeloid cells, cannot self-renew when transplanted, and likely function in innate immune protection (Bertrand et al. 2007; Bertrand & Traver 2009). The second part of the definitive wave are the HSCs, which in zebrafish bud out of the ventral wall of the dorsal aorta beginning around 30 hpf and then migrate to the caudal hematopoietic tissue (CHT) around 48 hpf (Burns et al. 2005; Murayama et al. 2006; Jin et al. 2007). At 3 days post-fertilization (dpf) the HSCs will migrate to the zebrafish kidney marrow, which is analogous to the bone marrow in mammals (Murayama et al. 2006; Jin et al. 2007). In mammals the HSCs originate in the aorta-gonad-mesonephros (AGM) and then migrate to the fetal liver for HSC maintenance, expansion, and differentiation, with the bone marrow being the primary site of hematopoiesis later in development (Cumano & Godin 2007) (Figure 1.4.2.1).



c)

Vertebrate	Hematopoietic Site				
	Primitive Wave 1	Primitive Wave 2	Definitive Wave 1	Definitive Wave 2	Final
Zebrafish	ALPM	ICM (AGM)	PBI	DA→CHT	Thymus, Kidney
Mouse	YS	YS	YS→FL	AGM→FL	Thymus, Spleen, Bone marrow
Human	YS	YS	YS→FL	AGM→FL	Thymus, Spleen, Bone marrow

Figure 1.4.2.1: An overview of zebrafish hematopoiesis and similarities to other vertebrates. (a) Timeline of zebrafish primitive and definitive hematopoiesis showing when each major hematopoietic site is active. **(b)** Diagrams showing zebrafish at different developmental time points with the anatomical location of the active hematopoietic sites (diagrams not to scale). Lateral views, anterior to the left. Colours of sites correspond to timeline in (a), with green representing the dorsal aorta (DA). **(c)** Table comparing the major hematopoietic sites of zebrafish, mice, and humans during the primitive and definitive waves in embryonic development, and the final major hematopoietic sites. Abbreviations: Intermediate cell mass (ICM), anterior lateral plate mesoderm (ALPM), posterior blood island (PBI), caudal hematopoietic tissue (CHT), hours post fertilization (hpf), days post fertilization (dpf), aorta-gonad-mesonephros (AGM), yolk sac (YS), fetal liver (FL).

While differences exist in the physical sites of blood development in zebrafish compared to other vertebrates, what is the most similar between all of these species is the genetic regulation of blood development. Primitive hematopoiesis is mainly controlled by two transcription factors, gata1 and pu.1 (also known as spi1), which act to functionally cross-antagonize each other (Cantor & Orkin 2002). Gata1 is a zinc finger transcription factor and is the master regulator of erythroid development, required for the differentiation of mature erythrocytes (Cantor & Orkin 2002). The *gata1* gene not only controls erythroid-specific gene regulation, but it also suppresses myeloid fate. The master regulator of myeloid fate, including granulocytes and macrophages, is the gene *PU.1* (Scott et al. 1994). Gene knockdown experiments of either *gata1* or *pu.1* show a reciprocal negative regulation, where a decrease in one gene leads to an increase in the other, and corresponds to changes in numbers of erythroid vs. myeloid cells (Rhodes et al. 2005). Zebrafish orthologs for nearly all other mammalian blood cell genes have been identified, including hemangioblast (*lmo2, fli1, scl, tal1, gata2*), HSCs (*scl, runx1, cmyb*), monocytes (*lcp1*), neutrophils (*mpx*), macrophages (*lyzC, mpeg1*), erythrocytes (*gata1, hbae1, hbae3, hbbe3, hbbe1.1*) and EMPs (*gata1, lmo2*) (Gering et al. 1998; Kalev-Zylinska et al. 2002; Bertrand et al. 2008; Herbomel et al. 1999; Bennett et al. 2001; Liu & Wen 2002; Bertrand et al. 2007; Detrich et al. 1995; Song et al. 2004; Liao et al. 1998). Some of these markers are highlighted in Figure 1.4.2.1b.

1.5: Transgenic and CRISPR Technology in the Context of the Zebrafish

1.5.1 Forward and Reverse Genetics in the Zebrafish

The zebrafish has proved itself as a useful tool for studying both forward and reverse genetics. Forward genetics describes when a phenotype is found, either at random or by chemical induction, and a gene is discovered for the corresponding phenotype. There have been a few large-scale forward genetic screens in zebrafish using chemical mutagens like ethylnitrosourea (ENU) (Driever et al. 1996; Ransom et al. 1996). In this method, the male of the breeding pair is treated with a chemical to mutate the germ cells, and the resulting offspring are grown up and incrossed to eventually produce F_3 offspring that have a phenotype that can be traced back to a gene by analyzing the genomes (Driever et al. 1996; Patton & Zon 2001). This method of genome mapping uses random amplified polymorphic DNAs (RAPD); small PCR

products are amplified randomly from a large piece of genomic DNA using decamer (10 nucleotides) primers, and the pattern of amplification gives a semi-unique profile (Postlethwait et al. 1994). The advantages of this method are that it is inexpensive and the position of the target gene does not need to be known prior to the experiment; however, RAPD can be difficult to reproduce (Postlethwait et al. 1994). Using this method new genes important for development have been discovered, and mutant zebrafish lines such as *moonshine* (*mon* gene, severe anemia; Ransom et al. 1996), *mindbomb* (*ubiquitin E3 ligase* gene, defects in neural development and definitive hematopoiesis; Itoh et al. 2003), and *spadetail* (*tbx16* gene, defects in mesoderm-derived tissues; Griffin et al. 1998), among others, have been created.

Reverse genetics, by contrast, occurs when a known gene is mutated to see the effects of the specific mutation or knockout. The zebrafish lends itself perfectly to reverse genetics, as the external fertilization and development of zebrafish embryos allows for genetic manipulation starting at the one-cell stage. There are a number of different ways that site-specific mutagenesis can be performed, including TALE nucleases (TALENs), zinc finger nucleases (ZFNs), and the Clustered regularly interspersed short palindromic repeats (CRISPR)/CRISPR associated protein 9 (Cas9) gene editing system (Sander et al. 2012; Hwang, Fu, Reyon, Maeder, Shengdar, et al. 2013). CRISPR/Cas9 has opened up a plethora of possibilities for genome editing in zebrafish: loss-of-function studies, establishment of knockout mutant zebrafish lines, and large-scale reverse genetic screens (Ablain & Zon 2016; Shah et al. 2015). Also encompassed by reverse genetics is the use of zebrafish for transgenesis – the incorporation and expression of a gene from another species in the zebrafish (Udvadia & Linney 2003). Transgenic zebrafish allow for the development of disease models, and many transgenic lines have already been created for different types of cancers (Stoletov & Klemke 2008). Both forward and reverse genetics have provided an abundance of knowledge on zebrafish embryonic development and hematopoiesis, much of which has expanded our knowledge about vertebrate development and disease in general.

1.5.2 Transgenic Zebrafish and Their Use as Models

Transgenesis is one of the useful tools available to zebrafish researchers. The ease of genetic manipulation of zebrafish embryos at the one-cell stage and optical clarity of the embryos are two advantages applicable to the creation of transgenic fish. There are a few different methods available for inserting a foreign piece of DNA into the zebrafish to create the transgenic. The first method, pioneered by Stuart and colleagues, inserted high concentrations of a linearized bacterial plasmid containing the gene of interest into zebrafish embryos at early stages of development (Stuart et al. 1988). This method was discovered to be inefficient, as much of the foreign DNA was degraded by the embryo, with only ~5% of fish injected retaining foreign DNA at low copy numbers per cell (Stuart et al. 1988). The germline transmission frequency and reliability of this technique is also low, although occasionally successful (Stuart et al. 1988; Suster et al. 2009). The next transgenesis technology used retroviral vectors carrying the transgene, with embryos infected at the blastula stage (Lin et al. 1994; Gaiano et al. 1996). While this method was more efficient in terms of transgene insertion and transmission, this method was very labour intensive and only small inserts could be used (Suster et al. 2009). A more efficient and popular method for creating a transgenic is through the use of transposable elements and the *Tol2* system.

The *Tol2* system of transgenesis is a relatively easy and efficient method for creating stable transgenic zebrafish lines (Figure 1.5.2.1). The *Tol2* element is a transposable element first found in the medaka fish, *Oryzias latipes*, and was shown to be capable of excision in zebrafish as it encodes the enzyme 'transposase' (Koga et al. 1996; Kawakami et al. 1998; Kawakami & Shima 1999). In this system a bacterial plasmid containing the foreign gene of interest is flanked by *Tol2* exons, and the plasmid is co-injected with transposase mRNA. The foreign gene can then be cut out of the plasmid *via* transposase and inserted into the zebrafish genome through transposition (Kawakami et al. 2000). This method of insertion is efficient, with about 12.5%-50% of injected fish transmitting the gene to progeny (Kawakami et al. 2000; Kawakami 2004). Fish injected with the *Tol2* vector and transposase are mosaic for the transgene, and therefore the creation of a transgenic line can provide stable expression (Figure 1.5.2.1) (Kawakami 2007). The ease of this system has been furthered by the pairing with other methods such as Gateway® cloning to make the injected plasmids (Villefranc et al.

2007). Gateway® cloning allows different plasmids to be combined quickly and easily, with the option to customize plasmids to the specific project.

Creating stable zebrafish transgenics is much easier, cheaper, and faster than creating transgenic mice, and zebrafish come with other advantages such as optical transparency that allows transgenes to be tagged with fluorescent proteins to visualize expression in the fish (Lieschke & Currie 2007). Dominantly acting disease genes can also be put under tissue specific promoters to not only make the malignant gene expression more comparable to human disease, but also to ensure survival and establishment of transgenic lines (Lieschke & Currie 2007). Transgenic zebrafish have been especially dominant in the field of cancer, with this technique being pioneered by Langenau and colleagues with a transgenic model of T-ALL that overexpresses mouse *c-Myc* tagged with GFP under the zebrafish *rag2* (T-cell) promoter (Langenau et al. 2003). With this zebrafish model that developed T-ALL *in vivo*, GFP+ cells could be tracked with fluorescent microscopy to different areas of the fish, giving information on disease progression (Langenau et al. 2003). This model was then improved by making the transgene inducible using the Cre/Lox and heat shock promoter systems, giving even more control of transgene expression (Langenau et al. 2005; Feng et al. 2007). Since these initial transgenic studies, transgenic zebrafish models exist for many different types of cancer, as well as other diseases (Rajan et al. 2016). For example, the Berman lab has successfully modeled a *NUP98* fusion in zebrafish by creating a transgenic fish that expresses the human *NUP98-HOXA9* oncogene under a myeloid cell promoter, *pu.1* (Forrester et al. 2011). This transgenic fish allowed for the discovery of the dysregulation of methyltransferase *DNA (cytosine-5)-methyltransferase 1* (*dnmt1*) as a novel mechanism and target for the development and treatment of *NHA9*-induced AML (Deveau et al. 2016). This highlights the utility of transgenic zebrafish to model cancer, specifically leukemia, and for the development of targeted therapies.

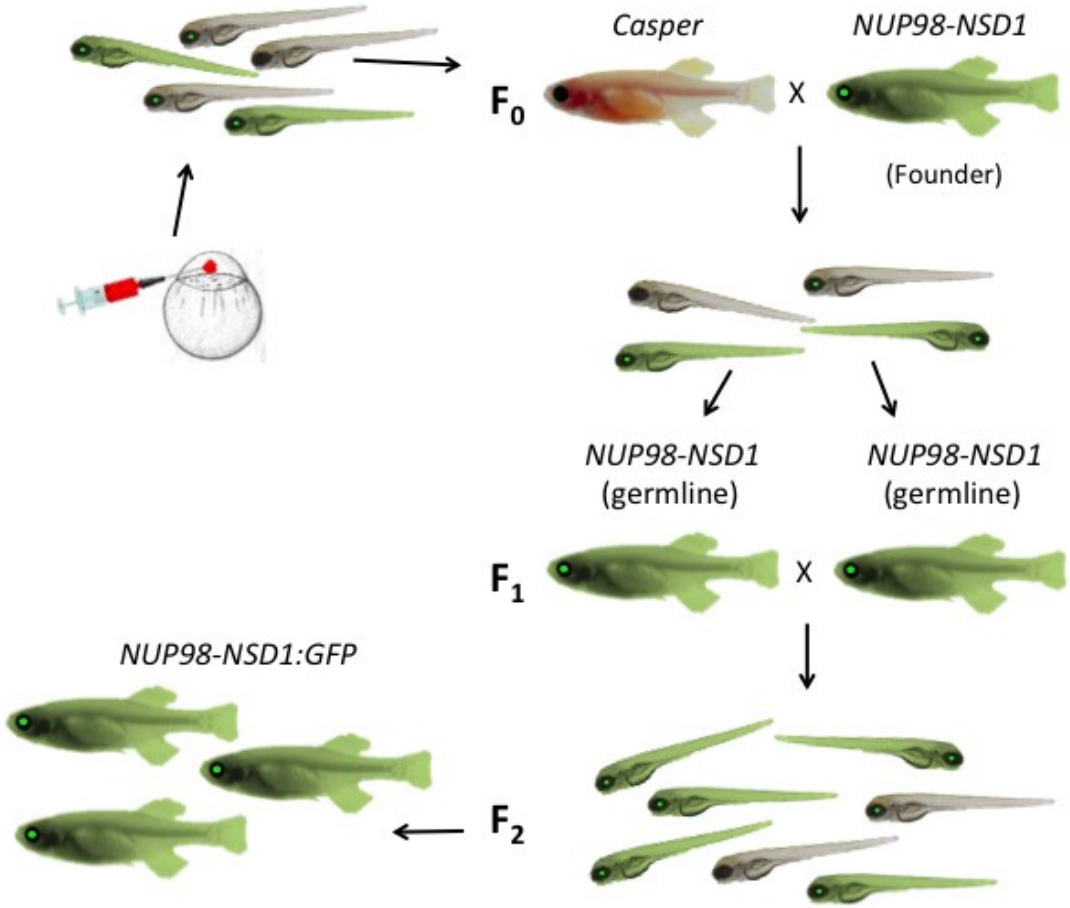


Figure 1.5.2.1: Creation and establishment of a stable transgenic zebrafish line. Zebrafish are first injected at the one-cell stage with *Tol2* mRNA and the bacterial plasmid containing the gene of interest/fluorescent protein. Injected embryos are screened (time point dependent on promoter expression of gene/fluorescent protein) for the fluorescent protein marker, eg. green fluorescent protein (GFP). GFP positive (GFP+) embryos are grown up to breeding age to test for germline incorporation of transgene. Around 3 months of age, injected fish are outcrossed to *casper* fish to test for germline transmission of the gene to F_1 offspring. GFP+ F_1 offspring indicate an F_0 founder fish. F_1 embryos are screened for GFP, and positive embryos grown up for breeding. F_1 embryos are incrossed to obtain F_2 embryos with stronger expression of GFP. Embryos with strong GFP expression can then be used for experiments and/or grown up to adulthood to establish and continue the transgenic zebrafish line.

1.5.3 The Use of CRISPR for Disease Modelling in Zebrafish

The recent advent of the CRISPR/Cas9 gene editing system has drastically changed the capabilities of targeted mutagenesis in *in vivo* systems. What was first found to be a defense mechanism in bacteria and archaea has been harnessed as a powerful tool capable of making specific edits to the genome (Horvath & Barrangou 2010; Wiedenheft et al. 2012; Ran et al. 2013). Briefly, CRISPR technology uses the Cas9 enzyme to make cuts in DNA regions by using synthetic single guide RNAs (sgRNAs) to direct the enzyme to a specific locus. The completion of the sequenced zebrafish genome opened up the possibility that human genes known to cause disease could be verified using gene knockouts in the zebrafish (Howe et al. 2013; Varshney et al. 2015). Prior to the advent of CRISPR, morpholinos - small, targeted oligomers that prevent gene transcription of sequences of RNA and result in decreased gene expression - were widely used for gene knockdown studies. While useful, morpholinos are known to cause off-target effects that can confound the true disease phenotype (Kok et al. 2015; Law & Sargent 2014). Although earlier versions of CRISPR also succumbed to a high frequency of off-target effects, there have been many efforts recently to decrease or eliminate these off-target effects (Fu et al. 2013; Ceasar et al. 2016). The use of zebrafish morphants (morpholino) vs. mutants (CRISPR) is still debated, given evidence to show that compensatory mechanisms are induced by mutations but not gene knockdowns, suggesting that morpholinos are still useful in certain settings (Rossi et al. 2015). However, there is a knockdown approach one can undertake with CRISPR called CRISPR interference (CRISPRi) that uses a catalytically inactive Cas9 ("dead" or dCas9) that represses expression of a target gene, with minimal off-target effects. CRISPR can also be used to transcriptionally activate genes (CRISPR-on), demonstrating the versatility of CRISPR to control gene expression beyond gene knockouts (Qi et al. 2013; Cheng et al. 2013).

The zebrafish has helped to pioneer advances in CRISPR/Cas9 genome editing. Hwang and colleagues first showed that CRISPR could be used *in vivo* to make targeted genetic modifications in zebrafish, with efficiencies comparable to other technologies such as ZFNs and TALENs; these precise mutations made by CRISPR were heritable with an excellent rate of transmission (Hwang, Fu, Reyon, Maeder, Shengdar, et al. 2013; Hwang, Fu, Reyon, Maeder, Kaini, et al. 2013). One of the easier applications of CRISPR

in the zebrafish model is for gene knockouts, where short insertions or deletions (indels) can often disrupt protein coding genes, though larger indels may be required for a full gene knockout (Xiao et al. 2013; Ceasar et al. 2016). From these knockouts, mutant zebrafish lines can be established (Figure 1.5.3.1). Additionally, tissue specific knockouts have been achieved by expressing Cas9 under a tissue specific promoter and sgRNAs ubiquitously, or *vice versa*, by incorporating these components into the zebrafish genome using transgenesis; this highlights the usefulness of transgenic technology in combination with CRISPR (Ablain et al. 2015; Yin et al. 2015; Ablain & Zon 2016). However, gene disruptions in the zebrafish do not have to be a full gene knockout, as the knockdown approach CRISPRi was recently tested in zebrafish. Long and colleagues showed that with multiple sgRNAs, gene expression could be reduced in the zebrafish without eliminating it entirely (Long et al. 2015). Overall, as CRISPR technologies progress, so too will the range and ease of genetic modifications that are possible in the zebrafish.

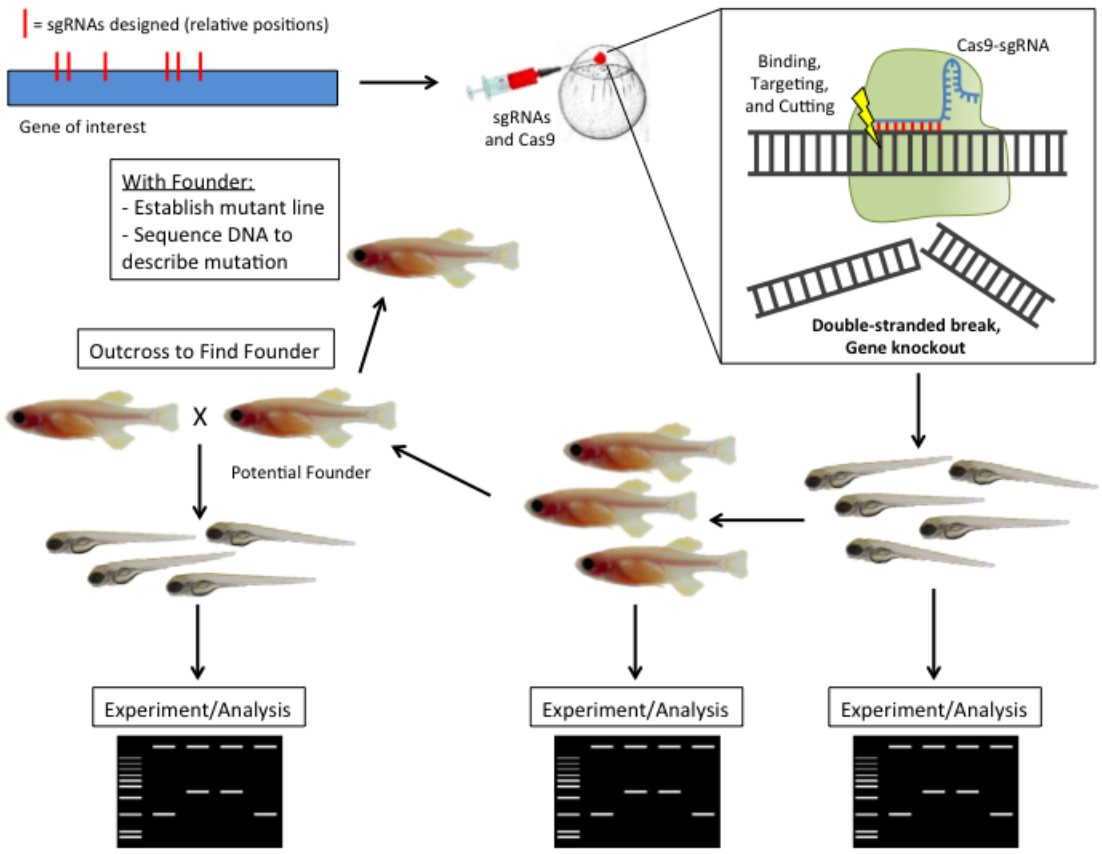


Figure 1.5.3.1: Creating a stable mutant zebrafish line using CRISPR/Cas9 gene knockout. Single guide RNAs (sgRNAs) are designed to target a gene of interest, and then the sgRNAs and the Cas9 enzyme are injected into the zebrafish at the one cell stage, where the gene is targeted and cut. Embryos from the injection can be used for experiments, to analyze the effectiveness of a gene knockout through fin clips or whole embryo analysis with polymerase chain reaction (PCR), a heteroduplex mobility assay (HMA), and/or can be grown up into adult zebrafish. Adults can also be used for experiments or analyzed for the gene knockout through fin clipping and PCR/HMA. To establish a mutant line adult zebrafish can be outcrossed for detection of germline transmission to the embryos. Identified founder fish can then be used to establish a mutant line through the breeding and identification of mutant fish *via* fin clips.

1.6: Rationale

1.6.1 Previous Work Modelling the NUP98-HOXA9 (NHA9) Oncogene in AML Using the Zebrafish

Transgenic zebrafish expressing *NUP98-HOXA9* (*NHA9*) developed in the Berman laboratory demonstrate that the zebrafish is a feasible and robust model for *NUP98* fusions in AML. *NHA9* transgenic fish expressed the transgene under the *pu.1* (*spi1*) promoter, and through *in situ* hybridization experiments the embryonic phenotype showed an increase in early myeloid cells (*pu.1*) at the expense of *gata1*, and to a lesser extent increases in other more differentiated myeloid cells (*mpx, lyzC, lcp1*) (Forrester et al. 2011). At 19-23 months of age, 23% of adult transgenic fish developed a myeloproliferative neoplasm (MPN), which is the overproduction of myeloid-lineage cells such as red blood cells, platelets, or myeloid cells (Forrester et al. 2011). Through microarray analysis it was found that *dnmt1* was overexpressed, providing insight into a mechanism of leukemogenesis. Indeed, it was shown that the knockdowns of *meis1* or *dmnt1*, or epigenetic therapy - specifically with DNMT inhibitors or cyclo-oxygenase (COX) inhibitors - could restore normal hematopoiesis in these embryos (Deveau et al. 2015). DNMT inhibitors also restored normal methylation levels in *NHA9*-treated embryos. In addition, these epigenetic therapies were combined with sub-therapeutic doses of a histone deacetylase inhibitor (valproic acid), and worked synergistically to block the effects of *NHA9* on zebrafish blood development (Deveau et al. 2015). This work highlights the use of the zebrafish to model hematologic malignancies, to discover novel mechanisms of leukemogenesis, and also for use as a platform for drug screens to identify targeted therapies.

1.6.2 Modelling NUP98-NSD1-induced AML and decreased nup98 in the zebrafish

Due to the success of modelling *NUP98-HOXA9* in the zebrafish, it was decided that *NUP98-NSD1* provided an opportunity to model and study a *NUP98* fusion gene that defined a novel patient group with poor prognosis and few treatment options. The *NHA9* model provided invaluable insights into leukemogenic mechanisms and treatment – but because of the transgene insertion method what was missing in this model was changes to levels of *nup98*. *NHA9* and *NND1* were inserted into the zebrafish using the *Tol2* system of transgenesis; the fusion genes were not created using the

zebrafish genome. This has the advantage of humanizing the fish with the human fusion gene and making any mechanistic or drug-related findings relevant to humans; however, it does not model the decrease in endogenous *NUP98* that is seen in human disease. To model both biological consequences of the fusion gene – the oncogene itself and the corresponding decrease in *NUP98* – a transgenic expressing *NUP98-NSD1* as well as the CRISPR/Cas9 gene-editing system for knockout of *nup98* was used.

This project aims to explore the consequences of *NUP98-NSD1* and decreased levels of *nup98*, alone and together, on disruption of hematopoiesis in the zebrafish. I hypothesized that *NUP98-NSD1* would cause a disruption of blood development in zebrafish; specifically, that *NUP98-NSD1* would cause increased myeloid cells, decreased red blood cells, and possibly increased hematopoietic stem cells in zebrafish embryos, similar to the *NHA9* zebrafish model. As *NHA9* is also classified as a type II aberration like *NND1*, and both translocations cause overexpression of *HOXA9*, the zebrafish models were thought to have similar hematopoietic phenotypes. I also hypothesized that decreased levels of *nup98* would not only disrupt blood development - possibly by affecting the myeloid and hematopoietic stem cells as seen by Fung *et al*. - but would synergize the disruption of blood development when combined with *NUP98-NSD1*.

This project combines two methods of genetic manipulation: zebrafish transgenesis of the *NUP98-NSD1* human gene, and CRISPR/Cas9 knockout of the *nup98* gene. With these methods we hope to gain insight into the contributions of each of these genetic aberrations to leukemogenesis *in vivo*. Long-term project goals include the establishment of these zebrafish models to discover downstream pathways that could be targeted with novel therapeutics, and later these models could be used as platforms for drug screening, which will attempt to identify drugs that restore normal blood development in these fish, and suggest potential targeted therapies for use in patients.

Chapter 2: Materials and Methods

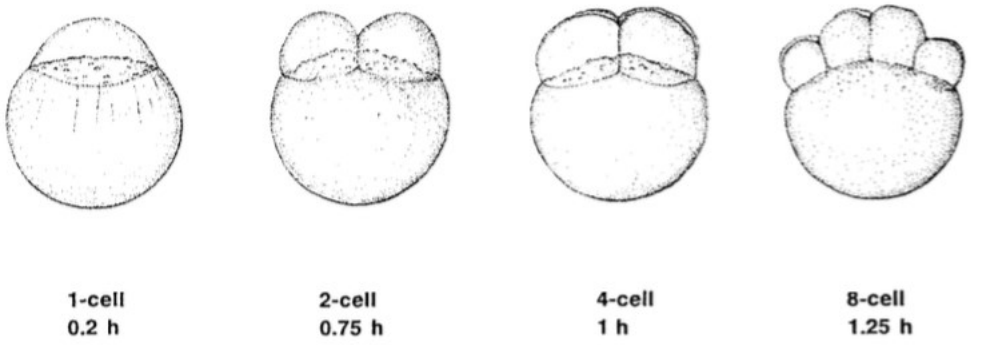
2.1 Zebrafish Husbandry, Embryo Collection, and Embryo Staging

Zebrafish (*Danio rerio*) used in this study were maintained and approved for use by the Dalhousie University Animal Care Committee (protocol #15-127, expires December 1st, 2017, see Appendix A). Fish were kept at 28.5°C on a 14 hour light/10 hour dark cycle. Fish water consisted of reverse osmosis water supplemented with sodium bicarbonate (Aquatic Eco-Systems, Inc., Apopka, FL, USA) and Instant Ocean® sea salt (Spectrum Brands, Inc., Madison WI, USA) to maintain a pH of 7.5. Adult zebrafish were fed once daily Gemma Micro 300 (Skretting, Vervins, France).

The double pigment mutant (*roy orbison -/-, nacre -/-*) translucent *casper* embryos (White et al. 2008) were provided by the Zon Laboratory, Children's Hospital, Boston, MA. The lack of melanocytes, xanthophores, and iridophores in *caspers* permitted real-time analysis of transgene without any auto-fluorescence from pigment that might interfere with image quality. Embryos were collected and grown in E3 embryo medium (5mM NaCl, 0.17mM KCl, 0.4mM $CaCl_2$, and 0.16mM $MgSO_4$ pH 7.5), at 28.5°C that was supplemented with Methylene Blue (1 x 10^{-5}% [v/v]) to inhibit the growth of mold. For early dechorination, embryos were treated with 10 mg/mL of stock Pronase (Roche Applied Science) for 5-15 minutes at 35°C. Embryos were kept in egg water for up to 7 days before being placed on the fish system in adult tanks.

Embryos were developmentally staged according to Figure 2.1.1 below (Kimmel et al. 1995). For experiments requiring embryos at 36 hours post fertilization, embryonic development was delayed by incubation at 22°C instead of 28.5°C, overnight on the second night, for approximately 18-22 hours.

Figure 2.1.1 (*next four pages*): Stages of early development of the zebrafish (from Kimmel *et al*., 1995). Images are camera lucida sketches of zebrafish at different developmental stages; pigmentation is omitted. Animal pole is at the top for early stages ('1-cell' to '90% epiboly'), and anterior is to the top for the later stages ('bud' and beyond). The exception is when the two animal polar (AP) views shown below their side view counterparts for germ-ring and shield gastrulas. Face views are shown during cleavage and blastula stages. After shield stage, the views are of the embryo's left side, but before the shield arises one cannot reliably ascertain which side is which. Arrowheads indicate the early appearance of some key diagnostic features at the following stages: 1 k-cell: yolk syncytial layer (YSL) nuclei. Dome: the doming yolk syncytium. Germ ring: germ ring. Shield: embryonic shield. 75%-epiboly: Brachet's cleft. 90%-epiboly: blastoderm margin closing over the yolk plug. Bud: polster. 3-somite: third somite. 6-somite: eye primordium (upper arrow), Kupffer's vesicle (lower). 10-somite: otic placode. 21-somite: lens primordium. Prim-6: primordium of the posterior lateral line (on the dorsal side), hatching gland (on the yolk ball). Prim-16: heart. High-pec: pectoral fin bud. Scale bar = 250 pn.

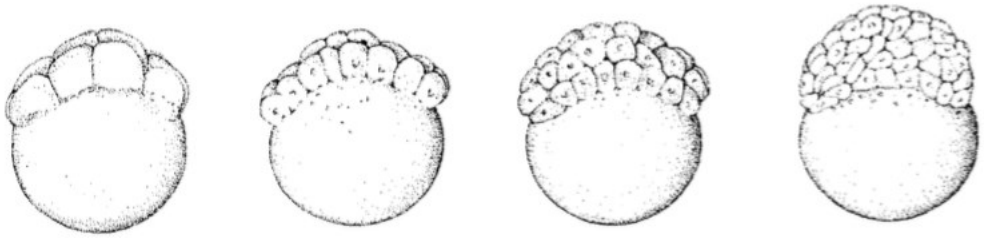


1-cell
0.2 h

2-cell
0.75 h

4-cell
1 h

8-cell
1.25 h



16-cell
1.5 h

32-cell
1.75 h

64-cell
2 h

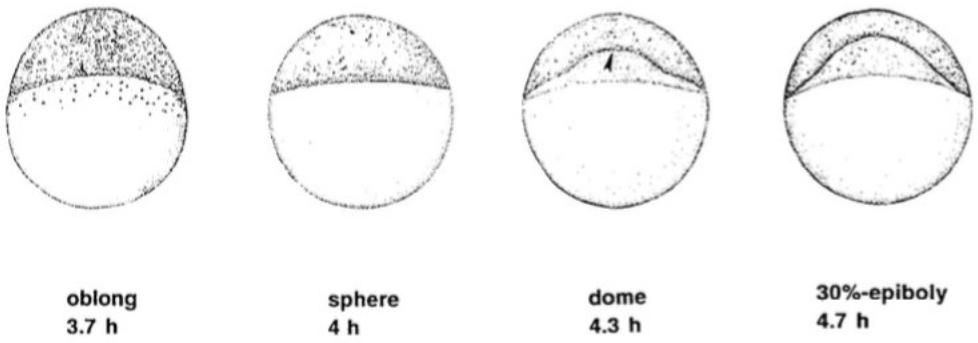
128-cell
2.25 h

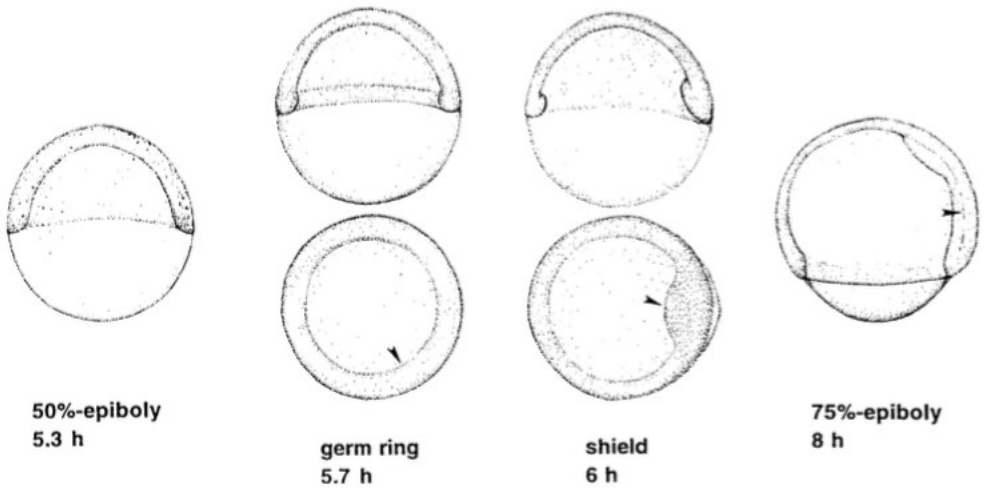
256-cell
2.5 h

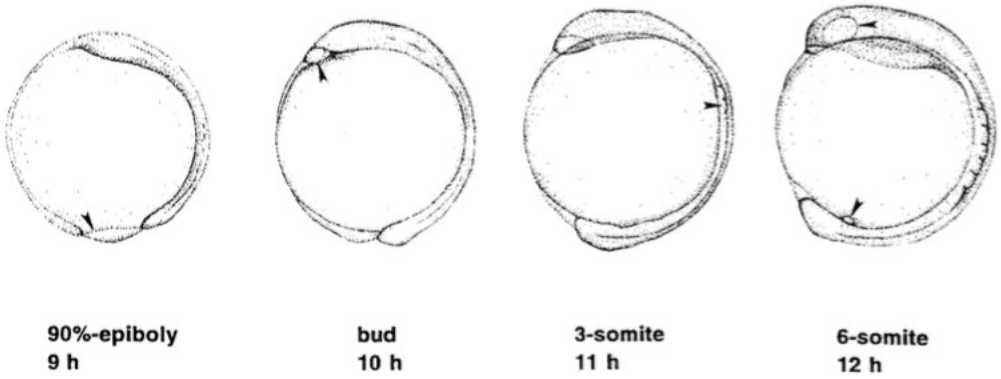
512-cell
2.75 h

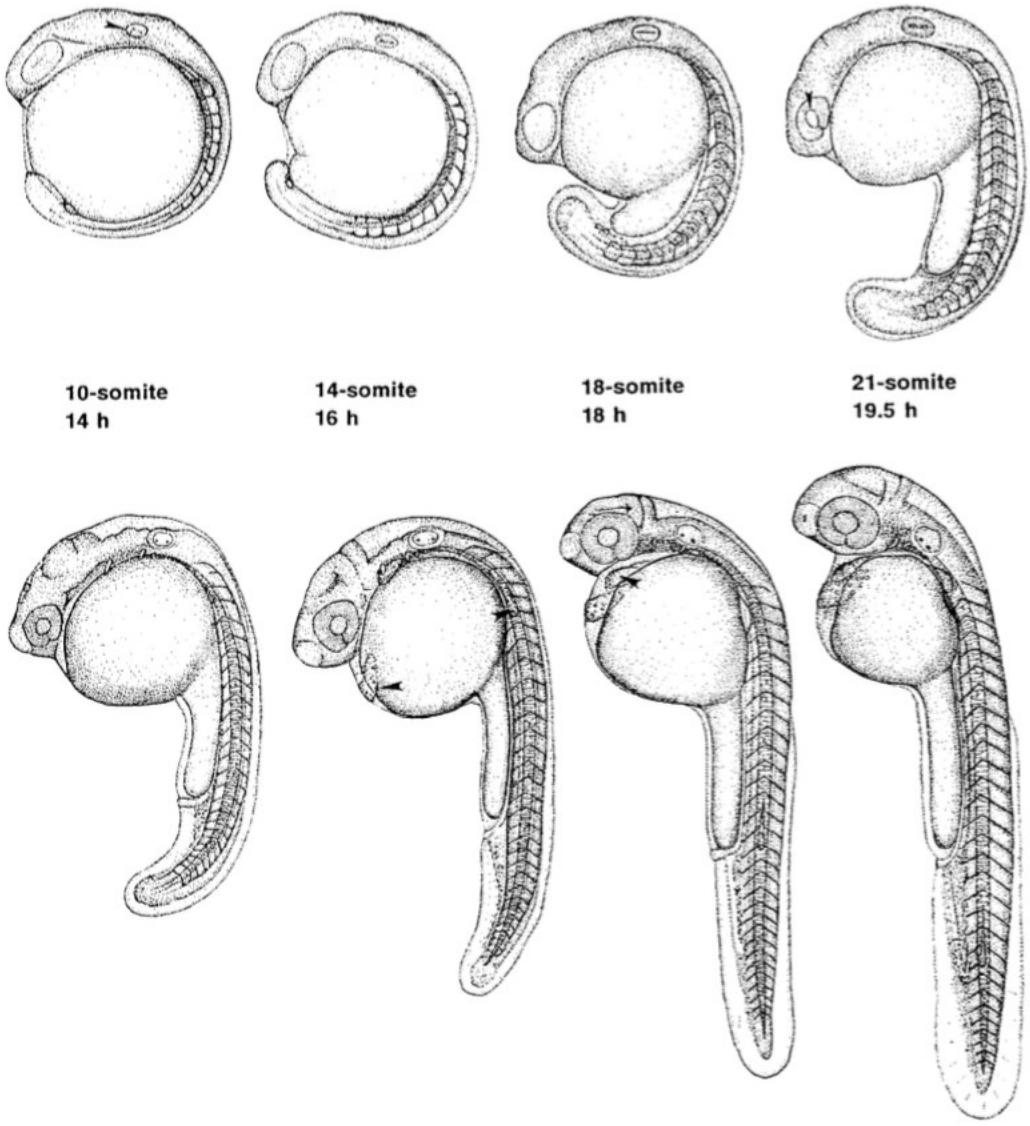
1k-cell
3 h

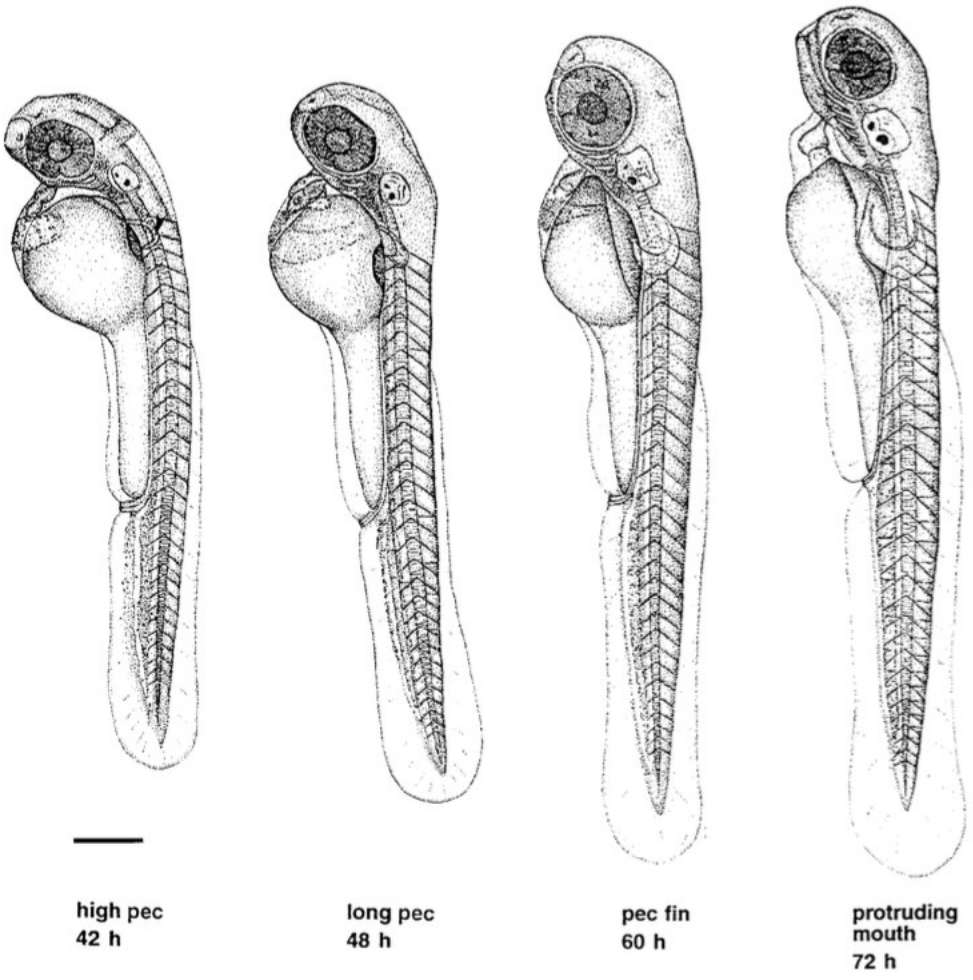
high
3.3 h


oblong
3.7 h
sphere
4 h
dome
4.3 h
30%-epiboly
4.7 h


50%-epiboly
5.3 h
germ ring
5.7 h
shield
6 h
75%-epiboly
8 h


90%-epiboly
9 h
bud
10 h
3-somite
11 h
6-somite
12 h


10-somite
14 h
14-somite
16 h
18-somite
18 h
21-somite
19.5 h
26-somite
22 h
prim-6
25 h
prim-16
31 h
prim-22
35 h


high pec
42 h
long pec
48 h
pec fin
60 h
protruding
mouth
72 h

2.2 Zebrafish Embryo Injections at the One-cell Stage

Adult *casper* zebrafish were bred in the facility, and embryos from these breedings were immediately injected at the one-cell stage. Injection needles were made from capillary tubes (World Precision Instruments, Sarasota, FL, Item no. TW100F-4) using a needle puller (Flaming/Brown Micropipette Puller, Model P-97, Sutter Instrument Co.) to obtain the correct size. Needles were further cut with a scalpel to obtain the optimum width at the tip. Injection solutions were made with nuclease free water (NFH_2O), phenol red (0.1% (w/v)), and specific nucleic acids such as DNA plasmids or RNA components. The injection solution was kept on ice until loaded 1µL at a time into the cut needle. Zebrafish embryos were placed in custom plates made with 2% agar and egg water with troughs to hold the embryos, and injections took place under a dissection scope (Zeiss Discovery.V8 SteREO Scope). The needle was attached to a system of medical air and a micro-injector apparatus to control the pressure and time of injection (Pico-Liter Injector, Warner Instruments PLI-100A). Embryos were injected individually using the needle loaded with the sample, and a bolus amount of ~5% of the total cell size was injected, but was kept consistent between embryo injections. After injections, embryos were kept in an incubator at 28°C for further experimentation at the appropriate time points.

2.3 Creation of a Transgenic Zebrafish Line

2.3.1 Bacterial Cloning

To propagate all plasmids and cloning vectors, DH5α™, Mach1™, or OneShot® TOP10 (Invitrogen Corporation, Carlsbad, CA, USA) strains of chemically competent *Escherichia coli* (*E. coli*) were used for transformation, plating, and culturing. Transformation used 1-5µL of suspended plasmid with 25-50µL of *E. coli* for 10-30 minutes on ice, then transferring to 42°C for 30-45 seconds. Transformed bacteria were then supplied with 250µL Super Optimal broth with Catabolite repression (SOC) (Invitrogen, Cat. #15544-034). Liquid culture was grown at 37°C for 1 hour with agitation, and 50-250µL of culture was plated on Lysogeny/Luria Bertani (LB) agar (Sigma, L2897-1KG) plates (Sarstedt, 83.3902) supplemented with antibiotic (either 50µg/mL kanamycin, or 100µg/mL ampicillin). Plates were inverted and incubated at 37°C overnight. Selected bacterial colonies were cultured in a 12mL conical bottom tube

(Falcon, Ref. 352059) containing 3-5 mL of LB liquid medium supplemented with appropriate antibiotic, and incubated at 37°C overnight with agitation. Liquid cultures were then mini-, midi-, or maxi-prepped using the associated Qiagen Plasmid prep kit according to manufacturer's instructions. The kits used depended on the amount of DNA required and the downstream usage of the material. For example, plasmids for injection were purified with a midi or maxi kit to obtain a large amount of DNA with less contaminant, and a mini prep was used for intermediate steps to check the efficiency of a cloning or ligation reaction. Purified plasmids were stored at 4°C or -20°C.

Purified plasmids were sent for sequencing (Genewiz, Boston, MA, USA), with DNA primers (Integrated DNA Technologies, Coralville, IA, USA) designed using Vector NTI Advance® 11.5.1 software (Invitrogen). Later analysis was carried out with Vector NTI Advance® version 1.6.0.

2.3.2 NUP98-NSD1 Plasmid

The human *NUP98-NSD1* (*NND1*) fusion oncogene was kindly provided by Dr. Soheil Meshinchi (Fred Hutchison Cancer Research Center, Seattle, WA). *NND1* was cloned downstream of either the ubiquitous zebrafish promoter *ubiquitin-C* (*ubi*), or the pan-leukocyte promoter *cd45*. The promoter *cd45* was kindly provided by Dr. David Traver (University of California San Diego, San Diego, CA). Vector NTI (Advance version 1.6.0) was used to find the appropriate restriction enzyme sites for insertion of *NUP98-NSD1* into the plasmid (pMe-MCS #237, Tol2 kit). With this process, *NUP98-NSD1* was cut out of the original vector, isolated by gel extraction (QIAquick® Gel Extraction Kit) and ligated into the plasmid with T4 DNA ligase (New England Biolabs). *NND1* was cloned into a middle element plasmid (pMe) to precede P2A-sGFP, which causes a post-translational cleavage and results in two separate proteins, *NND1* and superfolder green fluorescent protein (sGFP), in order to visualize *NND1* with green fluorescence. The 5' plasmid element (p5e) contained the promoter, either *ubi* or *cd45*, and the 3' plasmid element (p3e) contained the polyadenylation (polyA) sequence. These components were assembled into a PCS2 expression vector using Gateway® Cloning according to manufacturer's instructions (described schematically in Figure 2.3.2.1).

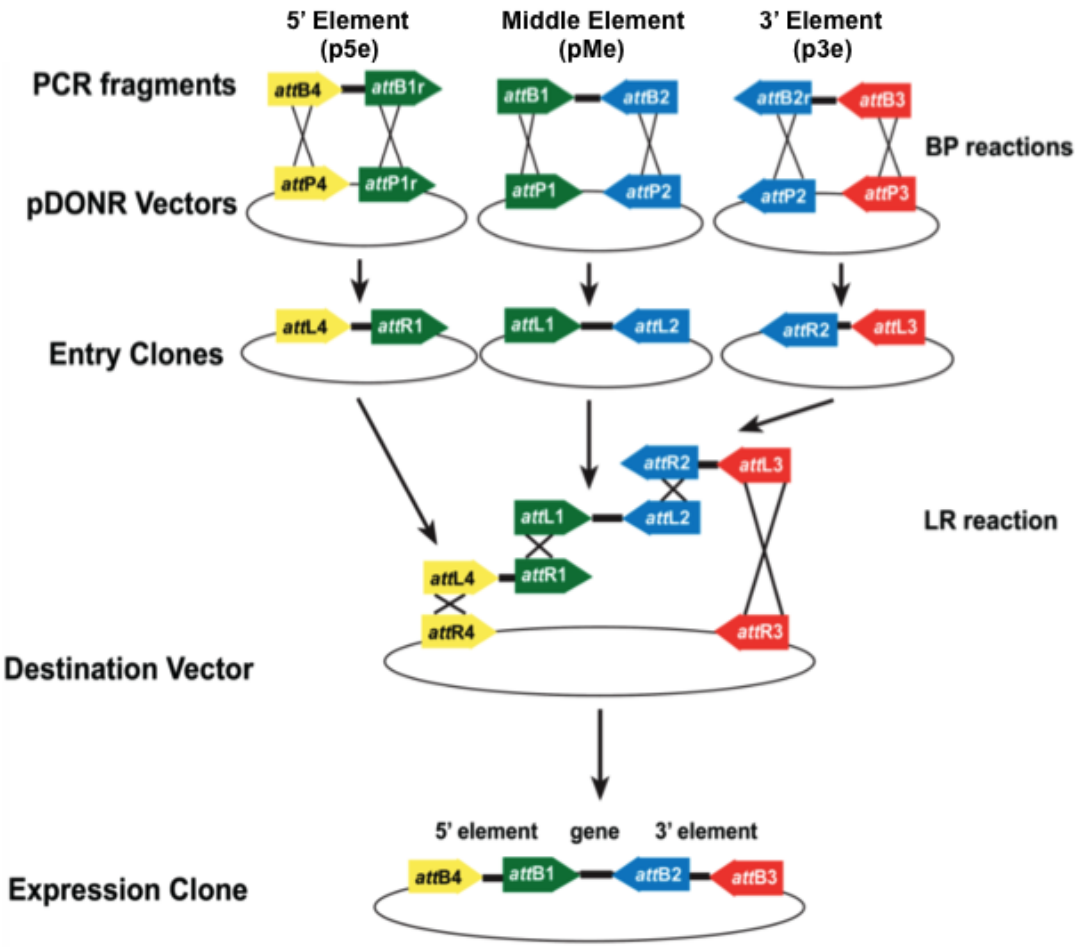


Figure 2.3.2.1: Assembly of Gateway® (Invitrogen) cloning vectors to obtain an expression clone. Gateway® cloning allows for different entry clones (p5e, pMe, and p3e) to be combined quickly and easily into a destination vector. The resulting expression clone contains all parts of the plasmid needed for gene expression – a promoter (normally from the p5e), a gene of interest (normally from the pMe), a fluorescent tag (normally from the pMe or p3e), and a polyadenylation sequence (normally from the p3e). A 'BP' reaction is done to insert the components of each entry clone, and an 'LR' reaction is performed to assemble all of the entry clones into a destination vector in one step. Adapted from (Forrester 2012).

2.3.3 NUP98-NSD1 Zebrafish Line Creation

Purified *ubi-* or *cd45-NND1-P2A-sGFP* was diluted to a working concentration of 125ng/μL, and combined with phenol red (0.1% w/v), transposase (Suster et al. 2011) (175ng/μL), and NFH_2O, and subsequently injected into *casper* embryos at the one-cell stage of development (see section 2.2). Injected embryos were screened for green fluorescence between 24-48 hours post-fertilization (hpf), and embryos expressing GFP were grown up to breeding age. Transgenic founder fish from the injected cohorts were screened for germline transmission of sGFP by outcrossing to *casper* embryos, and screening the F_1 embryos for germline expression to obtain an F_1 founder generation. The F_1 generation was then incrossed to create an F_2 generation and establish the transgenic line (Figure 1.5.2.1).

2.4 Creation of a Mutant Zebrafish Line

2.4.1 nup98 Knockout CRISPR Design

To knock out zebrafish *nup98*, the gene was first taken from Ensembl (http://www.ensembl.org/index.html), and sequences of specific exons for target were put into the website "Sequence Scan for CRISPR" (http://cistrome.org/SSC/). On this webpage, "CRISPR knockout" was selected, along with "19nt" for the length of the spacer. The program then scanned for sgRNAs and gave each sgRNA a score, which was an indication of efficiency with a score of 1.0=100% predicted efficiency. sgRNAs with a positive score over 0.5 were considered, and the sgRNAs with the highest scores were selected for use in the knockout. Exons 3 and 7 on the *nup98* gene were targeted due to their importance in encoding the FG repeats present on nup98, and that these exons contained sequences that could be targeted with sgRNAs that had scores close to 1.0. Three sgRNAs were selected to target each exon, with 6 sgRNAs total, shown in Table 2.4.1.1. sgRNAs were synthesized using the T7 MEGAshortscript™ Kit according to manufacturer's instructions, and aliquots of the sgRNA were stored at -80°C. All 6 sgRNAs were injected into zebrafish embryos at the one cell stage in a mixture of all 6 sgRNAs (~50ng/μL/sgRNA, or ~300ng/μL of sgRNA mix), Cas9 mRNA (1250ng/μL), NFH_2O, and phenol red (0.1% (w/v)). After injections embryos were used in experiments to detect mutations in *nup98*, using assays such as *Taq* LongAmp® polymerase chain reaction (PCR) and the heteroduplex mobility assay (HMA). The

program Vector NTI Express (version 1.6.0) was used to determine placement of the sgRNAs on the gene, which allowed for the development of primers for PCR and HMA, using a combination of Vector NTI and the website Primer3 Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi).

2.4.2 Taq Polymerase LongAmp® Polymerase Chain Reaction (PCR)

Genomic DNA was extracted from 20-100 pooled embryos using the PureLink™ Genomic DNA Mini Kit (ThermoFisher) according to manufacturer's instructions. The gDNA was subsequently used for PCR reactions with *Taq* LongAmp® Polymerase (New England Biolabs, M0323), using ~100ng of gDNA per reaction. The PCR protocol used was according to the New England Biolabs website, called "PCR Protocol for LongAmp® *Taq* DNA Polymerase", with an annealing temperature (Tm) of 56°C, and an elongation time of 7 minutes and 30 seconds. The forward primer used (5'-GGGGAGGTACAGGAGGATTT-3') started in exon 2 of the *nup98* gene, and the reverse primer (5'-GTCCCAAGAGTGAGCGTAGG-3') started in exon 11 of *nup98*; the expected amplicon size was 9 kilobases (kB).

Table 2.4.1.1: the six small guide RNAs (sgRNAs) used to knockout the *nup98* gene in zebrafish

sgRNA	Components of sgRNA			Exon Targeted
	T7 Sequence	Spacer (unique)	Scaffold Sequence	
1	TAATACGACT CACTATAGG	GGTTTGGTGCAACACCAGG	gttttagagctagaaat agc	3
2	TAATACGACT CACTATAGG	GGACAACAAACAACACAGG	gttttagagctagaaat agc	3
3	TAATACGACT CACTATAGG	TCGGAAGCACCAACACCGG	gttttagagctagaaat agc	4
4	TAATACGACT CACTATAGG	TGTGCCAGCTGCCATGGGG	gttttagagctagaaat agc	7
5	TAATACGACT CACTATAGG	CCAGTACTAGGAGTAGCTG	gttttagagctagaaat agc	7
6	TAATACGACT CACTATAGG	CTGCTGCCTCTCAGGCTGG	gttttagagctagaaat agc	10

2.4.3 Heteroduplex Mobility Assay (HMA)

Genomic DNA was extracted from embryos or fish fin clips prior to this protocol, and used in this assay (see section 2.5, "Fin clipping and Genomic DNA (gDNA) extraction"). First, 2.5μL of gDNA was used in a *Taq* PCR reaction (standard 25μL reaction protocol from New England Biolabs). A touchdown PCR (tdPCR) cycle was used as follows: 94°C for 3 minutes, (94°C for 30 seconds, 61°C for 30 seconds decreasing 1°C every cycle, 72°C for 30 seconds) x 10 cycles, (94°C for 30 seconds, 51°C for 30 seconds, 72°C for 30 seconds) x 25 cycles, 72°C for 5 minutes, Hold at 4°C. Following this tdPCR cycle, a 'T7 endo" cycle was performed with the PCR machine (Eppendorf AG, Hamburg No. 5341): 95°C for 5 minutes, 85°C for 10 seconds, 25°C for 10 seconds, Hold at 4°C. Next, the samples were run on a polyacrylamide gel made according to Table 2.4.3.1 below. The DNA was visualized using SYBR® Green I nucleic acid gel stain (Sigma). A schematic for performing and analyzing the HMA assay is shown in Figure 2.4.3.1. The primers used to analyze the sgRNAs in exon 3 were as follows, with an expected amplicon size of 458 base pairs (bp), and a Tm=51°C: forward, 5'-ATTTATGCAGGCTGCCTTTG-3'; reverse, 5'- TGGATGCTGTGTTGGTTGTT-3'. The primers used to analyze the sgRNAs in exon 7 were as follows, with an expected amplicon size of 433bp, and a Tm=52°C: forward, 5'- GGCTGGGAACCACTGATTTA-3'; reverse, 5'-GGCAAACAAACTTCCTGTGG-3'.

Table 2.4.3.1: Gel mixes for 6, 8% polyacrylamide gels (15mL each)[#]

Acrylamide Concentration	**6%**	**8%**
20X Gel Buffer	0.75mL	0.75mL
Acrylamide stock (30%)	3mL	4mL
H_2O (distilled)	11.13mL	10.14mL
10% AMPS	105µL	105µL
TEMED	15µL	15µL

[#]Mix in order shown and pour into gel cast immediately

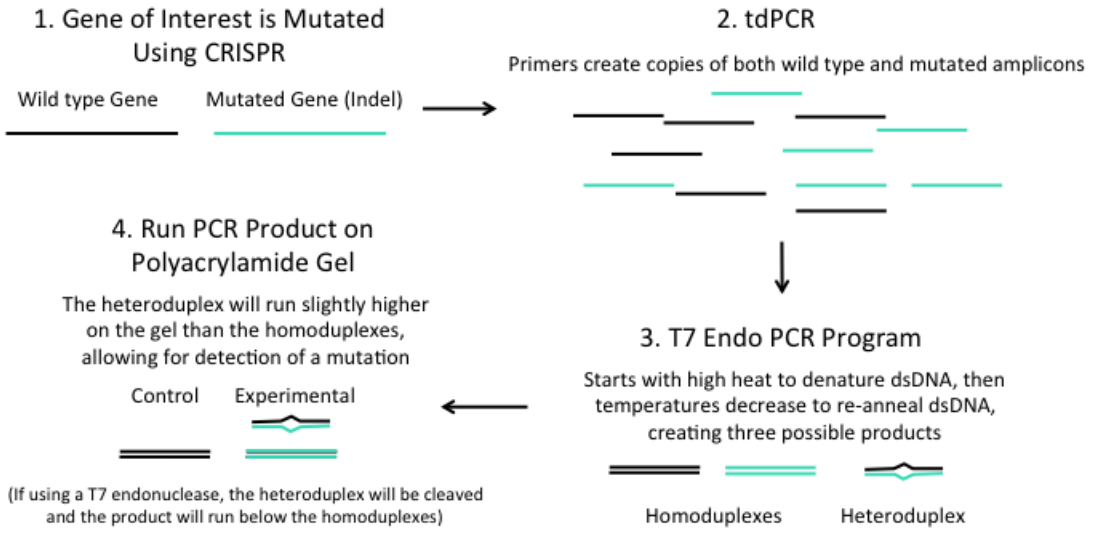


Figure 2.4.3.1: Overview of the process and analysis of a heteroduplex mobility assay (HMA). A gene of interest is first targeted for site-specific mutagenesis using CRISPR/Cas9 gene editing, creating an insertion or deletion (indel). Touchdown polymerase chain reaction (tdPCR) is used to amplify extracted DNA from a sample; if the sample has been mutated, there will be both wildtype and mutated copies of the gene, whereas a control sample would only have wildtype copies. The T7 endonuclease (T7 endo) program causes denaturation and subsequent annealing of the products from tdPCR, and depending on the sample will create either wildtype homoduplexes (control), or three products in a mutated sample: wildtype and mutant homoduplexes, and a heteroduplex containing both wildtype and mutated DNA strands. If the T7 endonuclease enzyme is being used in the assay, it will then cleave the heteroduplexes that form. The products are run on a polyacrylamide gel, where a sample positive for a mutation can be detected by the presence of a heteroduplex that runs higher on the gel than the homoduplexes because of the shape (represented in the figure above). If a T7 endonuclease enzyme was used in the reaction, the endonuclease will have cleaved the heteroduplex, meaning the product will be smaller and will run below the homoduplex on the gel, instead of above as pictured.

2.5 Fin Clipping and Genomic DNA (gDNA) Extraction

Genomic DNA (gDNA) was extracted one of two ways, depending on the downstream applications. The PureLink™ Genomic DNA Mini Kit (ThermoFisher) was used according to manufacturer's instructions to extract gDNA from groups of 10-100 embryos that were euthanized at 24 hpf-5 days post-fertilization (dpf). gDNA acquired from this kit was used in PCR assays where the amplicon expected was more than 1000bp.

Genomic DNA was also extracted using NaOH (sodium hydroxide). This protocol can be used with whole zebrafish embryos or fin clips from adult fish. To obtain fin clips, adult zebrafish were anesthetized in 5% tricaine solution and placed on a dissection plate. Forceps were used to hold the fin, and a cut was made with a scalpel (Lance®, Ref. # L308, no. 4 Fitment B.S. 2982) to remove a small piece of the fin, taking care not to cut past the tail juncture. Fin clips from adult fish, or whole embryos were placed into microcentrifuge or 1.5mL eppendorf tubes containing 20-1000μL NaOH (50mM NaOH (w/v) in H_2O). Samples were heated to 95°C in a PCR machine or heat block for 10 minutes or until tissue was noticeably friable. Tubes were briefly vortexed, and then placed on ice for 2 minutes. Tris (pH 8.3) was added at a ratio of 1:10 (Tris:NaOH) to neutralize the solution. The tubes were briefly centrifuged at low speed to pellet debris. The supernatant was put into a clean tube and 1-5μL of the supernatant was used for PCR assays. gDNA extracted using this protocol was used in PCR assays where the amplicon expected was less than 1000bp.

2.6 Whole-mount *in situ* Hybridization (WISH)

2.6.1 RNA Probe Synthesis

Anti-sense RNA probes for zebrafish *gata1*, *pu.1*, *lcp1*, *mpx*, *runx1*, and *cmyb* were made from plasmids constructed with the cDNA of these genes. Plasmids containing the probes were cloned using DH5α cells, purified using the Qiagen mini-prep kit (QIAprep spin mini prep kit), linearized with the corresponding restriction enzyme (protocol according to New England Biolabs), and then purified (QIAquick PCR purification kit). Anti-sense RNA probes were synthesized using T3, T7, or Sp6 RNA polymerase kits, followed by labeling with digoxogenin (DIG) or fluorescein isothiocyanate (FITC), according to manufacturer's instructions (Roche). A TURBO™

DNase treatment was performed to digest DNA, and probes were purified using NucAway™ Spin Columns (Invitrogen). The concentration of the probes was checked on a nanodrop (NanoDrop One/One^{C}), and the integrity of the RNA probe was checked on an agarose gel alongside the cut DNA vector. Purified probes were stored at -80°C, or at -20°C in a 1:1 ratio of probe to pre-hybridization buffer (Hyb-).

2.6.2 Embryo Preparation

Embryos used for *in situs* were staged at the time point of interest (eg. 24, 28, 36, or 48 hpf, according to Figure 2.1.1). Once staged, embryos were euthanized in tricaine (4mg/mL) and then immediately fixed in 4% paraformaldehyde (PFA) and kept overnight at 4°C. The following day, embryos were washed 3 x 5 minutes in 100% methanol (MeOH) and stored at -20°C overnight. Before staring the WISH experiment, embryos were rehydrated in a series of 5 minute washes in 75% (3:1), 50% (1:1), and 25% (1:3) MeOH:Phosphate-buffered saline with Tween (PBS-T), followed by 2 x 5 minute washes in PBS-T, and permeabilized in Proteinase K (ProK; Roche) according to Table 2.6.2.1 below. After the ProK treatment, embryos were washed in 1X PBS-T for 5 minutes, and re-fixed in 4% PFA for at least 20 minutes. After this step, the WISH protocol commenced.

Table 2.6.2.1: ProK permeabilization times for zebrafish embryos at various stages of development

Embryonic Stage	**Length of treatment (min)**		**ProK Concentration (μg/mL)**	
	Embryos in PFA	Embryos in MeOH	Embryos in PFA	Embryos in MeOH
18 hpf	0	0	N/A	N/A
24 hpf	20	10	1	1
28-30 hpf	5	15	10	1
36 hpf	10	5	10	10
48 hpf	20	15	10	10
60 hpf	10	10	50	50
72 hpf	20	20	50	50

Adapted from Talbot, JC:
https://wiki.zfin.org/display/prot/Triple+Fluorescent+In+Situ

2.6.3 WISH Protocol

WISH assays were adapted from standard protocols (Bennett et al. 2001; Dobson et al. 2008). Each experiment with a specific probe/time point used 20 embryos per group, and two replicates were performed for each probe and time point. Fixed embryos were rinsed and then washed 1 x 5 minutes in 1X PBS-T at room temperature. Embryos were then washed in pre-hybridization buffer (Hyb-; 5X saline sodium citrate (SSC) buffer with 0.1% (v/v) Tween 20 (SSC-T), 50% (v/v) formamide) for 15 minutes at 65°C, then blocked in hybridization buffer (Hyb+; same solution as Hyb- with 5mg/mL torula (yeast) RNA type IV and 50µg/mL heparin added) at 65°C for 1 hour. Embryos were incubated at 65°C overnight in Hyb+ with anti-sense RNA probe added (1:100 or 1:200 dilution). The following day, embryos were washed in the following successive washes (all at 65°C): 1 x 30 minutes and then 1 x 15 minutes in 2X saline sodium citrate buffer with 0.1% (v/v) Tween 20 (SSC-T)/50% (v/v) formamide, 1 x 15 minutes in 2X SSC-T, and 1 x 30 minutes then 1 x 15 minutes in 0.2X SSC-T. Embryos were then washed 3 x 5 minutes in 1X maleic acid buffer with 0.1% (v/v) Tween 20 ((MAB-T); 100mM maleic acid, 150mM NaCl, 10% [w/v] Tris, 0.1% [v/v] Tween-20, pH 7.5) at room temperature. Next, embryos were incubated at room temperature for 1 hour in WISH blocking solution (2% [w/v] blocking reagent [Roche], 10% [v/v] heat-inactivated fetal bovine serum (FBS) in 1X MAB-T). After the addition of blocking solution with anti-digoxogenin-AP antibody (anti-DIG; 1:10,000) embryos were incubated overnight at 4°C. The following day, embryos were washed in blocking solution 1 x 15 minutes, and MAB-T 1 x 15 minutes at room temperature.

For chromogenic development with the BCIP/NBT kit (Vector Laboratories, Inc., Burlingame, CA, USA), embryos were washed for 4 x 5 minutes with 0.1M Tris, pH 9.5. Staining was carried out according to manufacturer's instructions, for 45 minutes – 2 hours (dependent on the probe), in the dark at room temperature. When staining was complete the reaction was stopped with a 1 x 5 minute wash in 1X PBS-T at room temperature. Embryos were then de-stained to remove any background staining in 100% MeOH for 1-10 minutes (dependent on the amount of background staining), and transferred back to 1X PBS-T for imaging. For long-term storage at 4°C embryos were placed in 4% (w/v) PFA. Images were captured with the Zeiss SteREO Discovery.V20 microscope.

2.6.4 Analysis of in situs

The *gata1* probe was analyzed qualitatively using a scoring method according to Figure 2.6.4.1a/b. Three individuals from the lab scored the *gata1* probe separately and were blinded to the group being scored. Each fish in the group was given a score according to the scale in Figure 2.6.4.1a. Responses from each of the three individuals in the lab were tallied and data was presented using 100% stacked bar graphs.

The other *in situ* probes (*pu.1, lcp1, mpx, runx1/cmyb*) were analyzed quantitatively. Individual images were analyzed in FIJI (ImageJ), and the cells in each embryo that stained positive for the respective probe were counted. Counting was done using the plugin "Cell Counter" – each positive cell (stained purple from WISH) was clicked and the number of clicks were counted, which corresponded to the number of positive cells per embryo. The data was compiled in an Excel spreadsheet and statistics were run using the program Stata®. Each probe and time point was analyzed separately. In Stata, the data was first tested for normality using a Shapiro-Wilks test. If the data was normally distributed, a one-way analysis of variance (ANOVA) was run, followed by the *post hoc* Tukey test. If the data was not normally distributed a Kruskal-Wallis test was performed, followed by the *post hoc* Dunn test. Workflow of these statistical tests is shown schematically in Figure 2.6.4.1c.

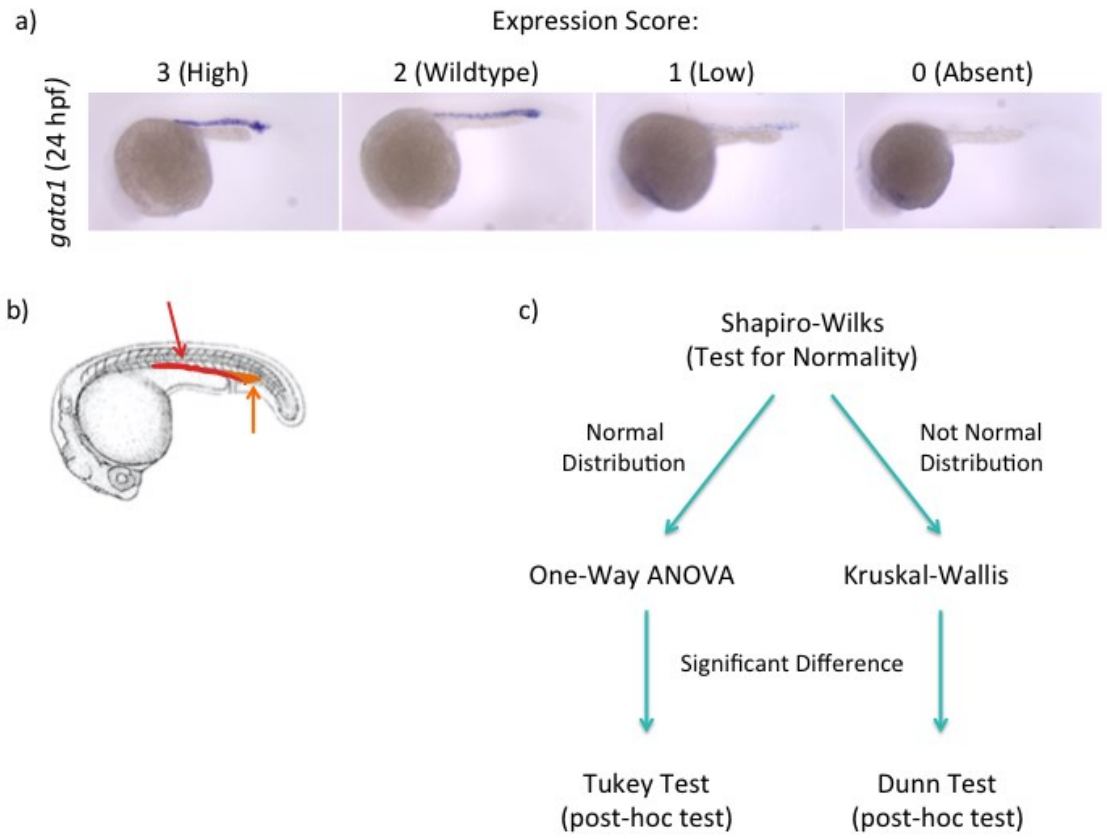


Figure 2.6.4.1: Qualitative and quantitative analysis methods for *in situs.* (a) *gata1* expression with the corresponding expression score used to categorize fish from whole-mount *in situ* hybridization. Embryos at 24 hours post fertilization (hpf) from *gata1 in situs* were scored qualitatively according to expression (purple staining) at two specific anatomical locations, shown in **(b)**. The intermediate cell mass (ICM, red) and posterior blood island (PBI, orange) are two hematopoietic sites with *gata1* expression at 24 hpf. Embryos with normal, or wildtype staining were given a score of 2; embryos with higher staining than normal were given a score of 3; embryos with lower than normal staining were given a score of 1; and embryos with no staining of *gata1* in the ICM or PBI were given a score of 0. **(c)** Workflow of statistical analysis for the cells counted using the ImageJ cell counter plugin. Probes analyzed with this method included *pu.1, lcp1, mpx,* and *runx1/cmyb*. Each data set was first tested for normal distribution of the data. Normally distributed data sets were then analyzed with a one-way analysis of variance (ANOVA) followed by a Tukey *post hoc* test. Data that was not normally distributed was analyzed using the Kruskal-Wallis test followed by the Dunn *post hoc* test.

2.7 Western Blotting

2.7.1 Protein Extraction

Protein was extracted from *casper* embryos (36 hpf), *cd45:NND1:P2A:sGFP* transgenic embryos expressing GFP+ (28 and 36 hpf), and 293T cells transfected with *CMV-NND1-P2A-sGFP* plasmid that would overexpress the *NND1* oncogene and GFP. Embryos (or cells) were dissociated according to the standard protocol (Corkery et al. 2011). After dissociation, lysate was prepared. The cell pellet was resuspended in lysis buffer made of 6:1 RIPA buffer (radioimmunoprecipitation buffer; 0.22% Beta glycerophosphate, 10% Tergitol-NP40, 0.18% Sodium orthovanadate, 5% Sodium deoxycholate, 0.38% EGTA, 1% SDS, 6.1% Tris, 0.29% EDTA, 8.8% Sodium chloride, 1.12% Sodium pyrophosphate decahydrate; pH 7.5) and protease inhibitor (cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Roche). Samples and buffer were incubated on ice for 15 minutes. Cells were sheared by pipet agitation and then incubated on ice for another 15 minutes. Debris was pelleted by centrifugation at 14800 rpm for 15 minutes at 4°C. Supernatant was collected in a new 1.5mL eppendorf tube. Protein was quantified using the Micro BCA™ Protein Assay Kit (ThermoFisher Scientific) according to manufacturer's instructions. Protein was quantified on a nanodrop (NanoDrop One/One^{C}) and stored at -80°C.

2.7.2 Western Blot Protocol

Western blotting was performed on protein extracts. Protein extracts (1-10µg) were added to 5X Laemmli sample buffer (300mM Tris-HCl, pH 6.8, 10% [w/v] SDS, 50% [v/v] glycerol, 25% [v/v] β-mercaptoethanol, with bromophenol blue) and incubated at 95°C for 10 minutes. Samples were run on pre-cast Mini-PROTEAN® TGX Stain-Free™ Precast Gels using the Mini-PROTEAN® Tetra Cell system with 1X running buffer (BioRad, 25mM Tris, 192mM glycine, 0.1% SDS, pH 8.3), run at 200V for 90 minutes. The HiMark™ Pre-stained Protein Standard (ThermoFisher) was used as a marker. Total protein was measured using the ChemiDoc™ Touch Imaging System (BioRad). Protein was transferred to a PVDF (polyvinylidene fluoride) membrane (BioRad, activated in 100% methanol for 1 minute) and run in a Mini Trans-Blot® module (BioRad) submerged in 1X transfer buffer (BioRad, 25mM Tris, 192mM glycine,

pH 8.3) and transferred at 100V for 40-60 minutes. The membrane was rinsed in 1X tris-buffered saline with tween (TBS-T; 20mM Tris, 500mM NaCl, pH 7.6, with 0.1% (v/v) Tween-20). Membrane was blocked in bovine serum albumin (BSA, Bio Basic Inc.; 5% (w/v) in 1X TBS-T) for 1 hour at room temperature, and then washed for 3 x 5 minutes with TBS-T. The primary antibody used was monoclonal rat anti-NUP98 (clone 2H10) IgG2c-κ (Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA; 1:100), and was diluted in 5% (w/v) BSA and TBS-T. The membrane was incubated with the primary antibody overnight at 4°C. Next, the membrane was washed with 3 x 5 minutes TBS-T. The secondary antibody used was goat anti-rat IgG, HRP-linked (Cell Signaling; 1:2000), diluted in 2.5% (w/v) milk powder (Carnation Instant Skim Milk Powder in TBS-T). The membrane was incubated with the secondary antibody for one hour at room temperature, then washed 3 x 5 minutes with TBS-T. The protein was detected using the SuperSignal™ West Dura Extended Duration Substrate kit (ThermoFisher) according to manufacturer's instructions, and then visualized on the ChemiDoc™.

2.8 Zebrafish Peripheral Blood Extraction and Whole Kidney Marrow Dissection

2.8.1 Peripheral Blood (PB) extraction

Fish were first anesthetized using 10% tricaine and brought to a surgical plane of anesthesia, showing no response to stimuli and slow operculum movement. To collect peripheral blood, a 1.5mL eppendorf tube containing 10μL of heparin (Heparin Sodium Injection B.P., 100i.u./mL, Leo Pharma Inc., DIN: 00727520) was placed into a centrifuge. A second eppendorf tube with the bottom cut off was placed into the first eppendorf tube. After the fish was anesthetized, the tail was removed below the caudal fin using a scalpel. Using a pipette, 10μL of heparin was added directly to the cut site to prevent clotting, and the fish was placed tail-first into the cut eppendorf tube. The sample was centrifuged (Thermo Scientific, Heraeus PICO 17 Centrifuge) at 100xg for 5 minutes; this removes the peripheral blood from the fish and collects it in the base eppendorf tube. A schematic representation of this protocol is shown in Figure 2.8.1.1. After centrifugation, the top tube containing the fish is removed, and the fish can now be dissected for the kidney marrow. The blood sample in the bottom collection tube was used for experiments.

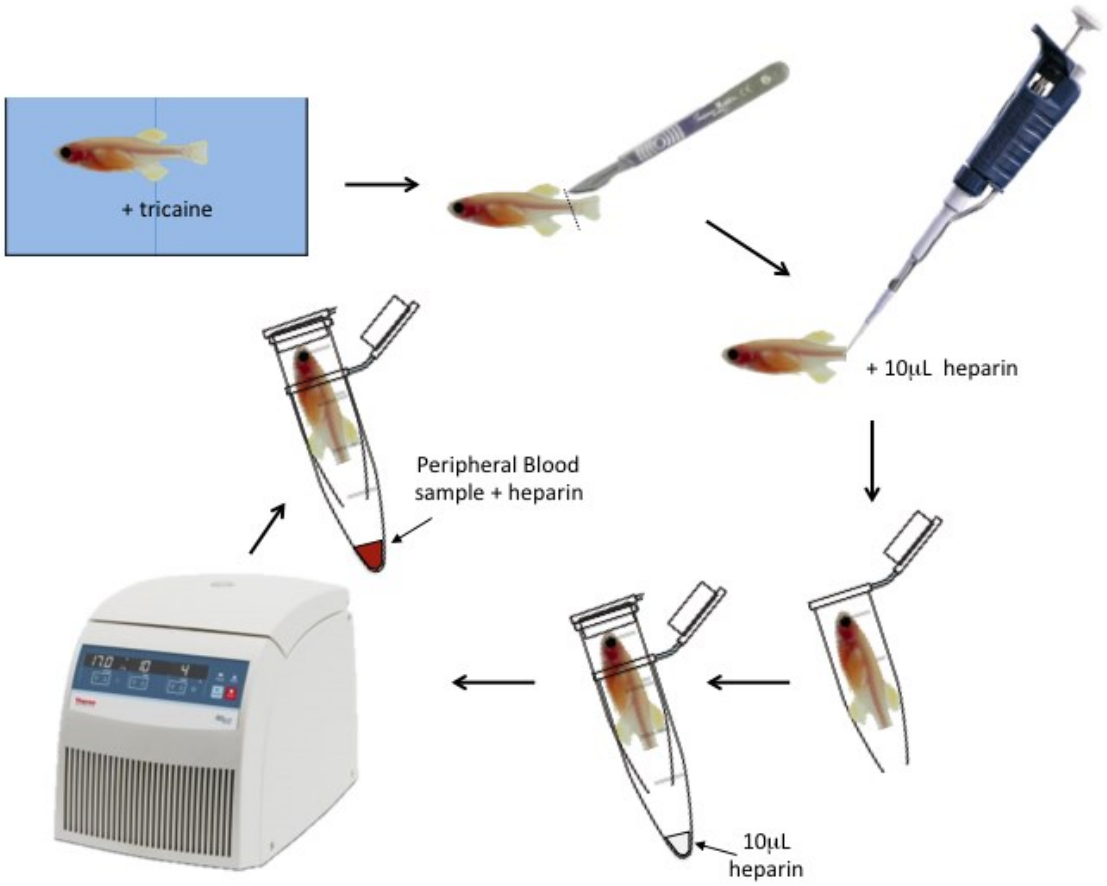


Figure 2.8.1.1: Schematic representation of a zebrafish peripheral blood extraction. First, fish are placed in a tank with 10% tricaine to achieve a surgical plane of anesthesia. Once anesthetized, the tail of the zebrafish is removed below the cloaca using a scalpel, and 10µL of heparin is placed on the cut site. The zebrafish is then placed into a 1.5mL eppendorf tube that has been cut at the bottom. This tube with the fish is placed inside another 1.5mL centrifuge tube that has 10µL of heparin in the bottom for blood collection to prevent coagulation. This assembly of tubes is put into a centrifuge and spun at 100xg for 5 minutes. After centrifugation the peripheral blood has drained into the collection tube and the fish is now exsanguinated. The fish and blood sample can now be used in further experiments.

2.8.2 Whole Kidney Marrow (WKM) Dissection

The exsanguinated fish was placed on a dissection surface, and the head was removed just behind the operculum. Using forceps, the body was secured so that the fish was sitting dorsally against the dissection place, head to the left. A midline incision was made from one end to the other with a scalpel blade. As the kidney lies dorsally, all of the internal organs were carefully removed, revealing the kidney attached to the dorsal cavity. The kidney was removed in pieces using fine forceps, and the kidney pieces were placed in a collection tube containing 2mL of phosphate buffered saline/2% fetal bovine serum (PBS/2%FBS; also known as 'flow buffer'). The kidney sample in flow buffer was used for further experiments.

2.9 Fluorescently Activated Cell Sorting (FACS) and Flow Cytometry

2.9.1 FACS for Western Blotting

Transgenic *cd45:NND1:P2A:sGFP* F_2 embryos were screened for GFP at 28 and 36 hpf. Embryos positive for GFP were euthanized using 10% tricaine, and were placed in a 1.5mL eppendorf tube for dissociation. Embryos were dissociated into a single cell suspension using Tumor dissociation kit-human (Miltenyi Biotech) following the soft tissue dissociation protocol provided by the manufacturer. Of this 5mL dissociation solution, 1mL was placed into the eppendorf tube containing the euthanized embryos. This tube containing the embryos and dissociation solution was heated at 37°C with shaking at 1000rpm for 15 minutes. The tube was removed from the heat once a single cell suspension had been obtained. The solution was then filtered through a 40μM filter, and flow buffer (PBS/2%FBS) was added to the solution to bring the total volume up to 10mL. This solution was centrifuged at 1200rpm for 5 minutes, and the supernatant was then removed. The pellet was resuspended in 2mL of flow buffer and then flow cytometry analysis was performed on the sample using the BD Fortessa SORP equipped with blue (488nm), green (532nm), red (628nm), and violet (402nm) laser. For experiments involving GFP+ cells (Western Blot and Cytospin), cells were sorted for GFP using FACS with the BD FACS Aria III equipped with blue (488nm) and red (633nm) laser.

2.9.2 Flow Cytometry and FACS for Peripheral Blood (PB) and Whole Kidney Marrow (WKM)

Adult *caspers* and the transgenic *cd45:NND1:P2A:sGPF* (>3 months post fertilization) were dissected for peripheral blood and kidney marrow. The peripheral blood and kidney marrow were extracted so that both tissues could be retrieved from one fish (see section 2.8 above). Peripheral blood samples were transferred to a 15mL falcon tube using a pipette tip coated in heparin. The PB sample was diluted to 10mL using flow buffer and filtered with a 40μM filter in preparation for centrifugation. The WKM sample was already in flow buffer, ready for centrifugation. Both samples were centrifuged at 1200rpm for 5 minutes. The PB sample was resuspended in 2mL of flow buffer, ready for flow cytometry or FACS. The WKM was resuspended in 10mL after centrifugation, centrifuged again, and the supernatant removed. The WKM was then dissociated using Tumor dissociation kit-human (Miltenyi Biotech) following the soft tissue dissociation protocol provided by the manufacturer. The single cell suspension was then filtered through a 40μM filter into a falcon tube, and flow buffer was added to bring the volume to 10mL. The sample was centrifuged for 5 minutes at 600xg, the supernatant was removed, and the cell pellet was resuspended in 2mL of flow buffer. Both the PB and WKM samples were then either sorted for GFP+ cells using FACS, with the GFP+ cells collected for the cytospin, or the samples were used for flow cytometry. If being used for flow cytometry the samples were stained for 10 minutes with the dye 7AAD, marking dead cells. The samples were then processed with flow cytometry by gating for live cells (7AAD negative), as well as forward and side scatter to determine the populations of hematopoietic cells.

2.10 RNA-sequencing (RNA-seq)

2.10.1 Preparing Embryos for RNA-seq

Embryos were staged at 36 hpf, euthanized using tricaine, and stored at -80°C until TRIzol™ RNA extraction. The following groups of fish were sent away for sequencing, with each biological replicate containing 50-100 embryos: 1) *casper* control, 2) *casper* injected with *nup98* KO sgRNAs, 3) *cd45:NND1:P2A:sGFP*, 4) *cd45:NND1:P2A:sGFP* injected with *nup98* KO sgRNAs, 5) *casper* injected with *cd45-NND1-P2A-sGFP* plasmid and transposase.

2.10.2 RNA Extraction Using TRIzol™ Reagent

RNA was extracted from whole zebrafish embryos staged at 36 hpf, according to the user guide from TRIzol™ Reagent (cat. #15596026, #15596018). For all extractions, 1mL of TRIzol™ was added before homogenization of whole zebrafish embryos that were stored at -80°C prior to RNA extraction. Extraction was completed using Phasemaker™ tubes (cat. #933249). Other optional steps were not taken. Extracted RNA was tested for purity on a nanodrop (NanoDrop One/One^{C}) and then run on a 2% agarose gel to check for integrity.

2.10.3 RNA-seq Alignment and Analysis

The extracted RNA was sent to the Atlantic Cancer Research Institute (ACRI) in Moncton, New Brunswick, Canada, where the RNA-seq was done in collaboration with Stephen Lewis, Nicolas Crapoulet, Darwin D'souza, and colleagues. At ACRI, the RNA extracted from whole embryos was converted into cDNA fragments; this library was then sequenced using the Ion Torrent Semi-Conductor Sequencer. Raw sequences and base-calling was performed using the Torrent Suite software. Adapter clipping (P1 Ion Torrent adapter) and trimming of trailing low quality bases at the 3' ends of reads were also done by the Ion Torrent Suite. Further removal of low quality bases and reads that were too short was done using *cutadapt*. The Zebrafish genome used for alignment was danRer10, which was downloaded from the UCSC. Reads were mapped initially to the genome using the STAR alignment software. Unmapped reads from the STAR alignment were remapped using the local alignment option of Bowtie2. Samtools was used to merge and sort the mapped reads from both STAR and Bowtie2. Read metrics and alignment metrics were obtained using PICARD. For counting, the number of reads that were mapped to genes was obtained using the featureCounts tool of the Subread software.

For statistical analysis, also performed at ACRI, multivariate analysis/visualization to assess transcriptome-wide differences among samples were performed in R statistical software. Principal component analysis (PCA) was performed to see if replicates had a high correlation. Reads were normalized by size factors estimated by the median ratio method of the DESeq tool, which uses a negative binomial distribution model and the Wald test to perform differential expression. Pairwise tests

were performed between each type of condition using DESeq. P-values were corrected using the false discovery rate (FDR) method. Genes were considered differentially expressed if the corrected p-value was less than 0.05 and Log Fold Change >(<)1(-1). To analyze all the groups concurrently and comprehensively to determine if a gene varies in expression across all the groups, a one-way ANOVA test was performed.

Chapter 3: Results

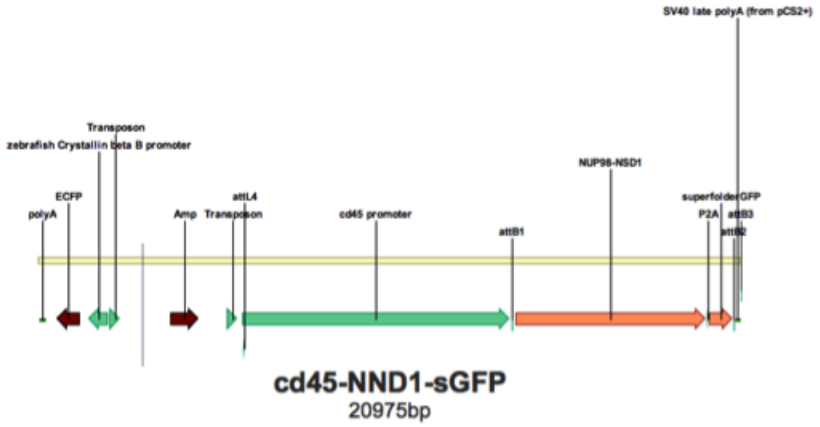
3.1: Transgenic *cd45:NND1:P2A:sGFP* Embryos and Adults Display Green Fluorescence in Hematopoietic Sites and Tissues

3.1.1 Transgenic cd45:NND1:P2A:sGFP Zebrafish Embryos show Green Fluorescence in Sites of Hematopoiesis, and Express the NUP98-NSD1 Protein

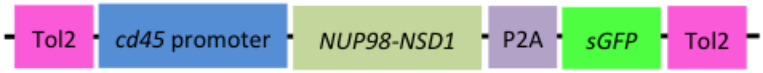
To create the transgenic fish *cd45:NND1:P2A:sGFP* the Gateway® cloning system was used to generate the plasmids for insertion using the *Tol2* system (Figure 3.1.1.1a/b). Originally both a ubiquitous and tissue specific promoter (*ubiquitin-C* [*ubi*] and *cd45*, respectively) were chosen to create the transgenic lines, but as the *ubi* promoter caused embryonic lethality in F_1 fish, we proceeded only with the *cd45* promoter. Only one founder fish was discovered out of over 150 outcrossed potential founders, and was used to establish the transgenic line *cd45:NND1:P2A:sGFP* according to Figure 1.5.2.1. Our transgenic fish *cd45:NND1:P2A:sGFP* displays green fluorescence at sites of hematopoiesis, most notably in the PBI at 24 hpf and the AGM and CHT at 48 hpf (Figure 3.1.1.1c). While the presence of GFP fluorescence should indicate *NND1* gene expression because of the construction of the plasmid, a Western blot was performed to ensure that NUP98-NSD1 was being expressed in the transgenic fish. An anti-NUP98 primary antibody was used, and the NUP98-NSD1 protein was detected in GFP+ cells that were isolated through fluorescence activated cell sorting (FACS) of whole transgenic embryos, but could not be detected in whole transgenic embryos without FACS. The NUP98-NSD1 protein was not detected in control *casper* embryos, as expected (Figure 3.1.1.1d).

Figure 3.1.1.1 (*next page*): Transgenic *cd45:NUP98-NSD1:P2A:sGFP* embryos display green fluorescence at sites of hematopoiesis and express the NUP98-NSD1 protein. A linearized complex **(a)** and simplified **(b)** version of the plasmid used to make the transgenic *cd45:NUP98-NSD1:P2A:sGFP*. **(c)** Transgenic embryos at 24 and 48 hpf display green fluorescence in sites of hematopoiesis (blue arrow: ICM [24 hpf], AGM [48 hpf]; red arrow: PBI [24 hpf], CHT [48 hpf]), and also have off-target expression in the hindbrain and tail (white arrows). **(d)** The NUP98-NSD1 protein is present in the GFP+ cells of transgenic embryos. Lane 1: Hi-Mark ladder, lane 2: protein from whole *casper* fish (negative control), lane 3: protein from whole transgenic fish, lane 4: protein from GFP+ cells sorted from transgenic embryos by FACS, lane 5: protein from cells transfected to express the plasmid *CMV-NUP98-NSD1-P2A-sGFP* (positive control). Abbreviations: green fluorescent protein (GFP), hours post fertilization (hpf), intermediate cell mass (ICM), aorta-gonad-mesonephros (AGM), posterior blood island (PBI), caudal hematopoietic tissue (CHT), fluorescently activated cell sorting (FACS).

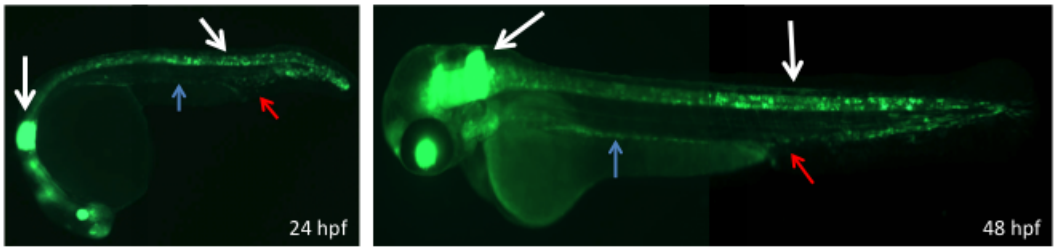
a)


SV40 late polyA (from pCS2+)
Transposon
zebrafish Crystallin beta B promoter
ECFP
polyA
Amp
attL4
Transposon
cd45 promoter
attB1
NUP98-NSD1
superfolderGFP
P2A
attB3
attB2
cd45-NND1-sGFP
20975bp

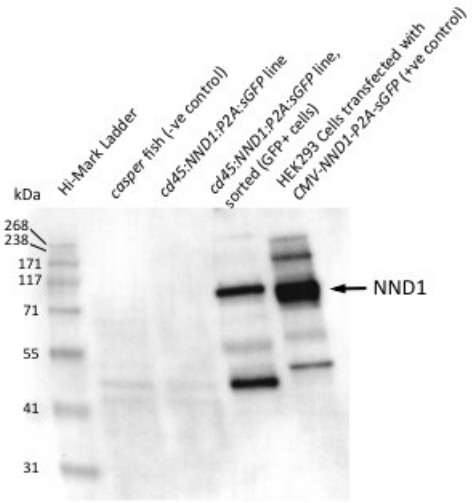
b)


Tol2
cd45 promoter
NUP98-NSD1
P2A
sGFP
Tol2

c)


24 hpf
48 hpf

d)


kDa
Hi-Mark Ladder
casper fish (-ve control)
cd45:NND1-P2A-sGFP line
cd45:NND1-P2A-sGFP line, sorted (GFP+ cells)
HEK293 Cells transfected with CMV-NND1-P2A-sGFP (+ve control)
268
238
171
117
71
55
41
31
NND1

3.1.2 Adult cd45:NND1:P2A:sGFP Fish have GFP+ Cells Present in the Peripheral Blood and Whole Kidney Marrow

Adult transgenic tissues were analyzed for GFP expression. Peripheral blood (PB) and whole kidney marrow (WKM) were extracted from adult *cd45:NND1:P2A:sGFP* fish that were approximately 8 months of age. When FACS analysis was performed on the PB and WKM of transgenic fish, it was found that there was a distinct population of GFP+ cells present in these tissues, showing that the GFP expression extends well into zebrafish adulthood (Figure 3.1.2.1). In the PB it was found that 32.2% of cells were GFP+, and in the WKM it was found that 20.4% of cells were GFP+.

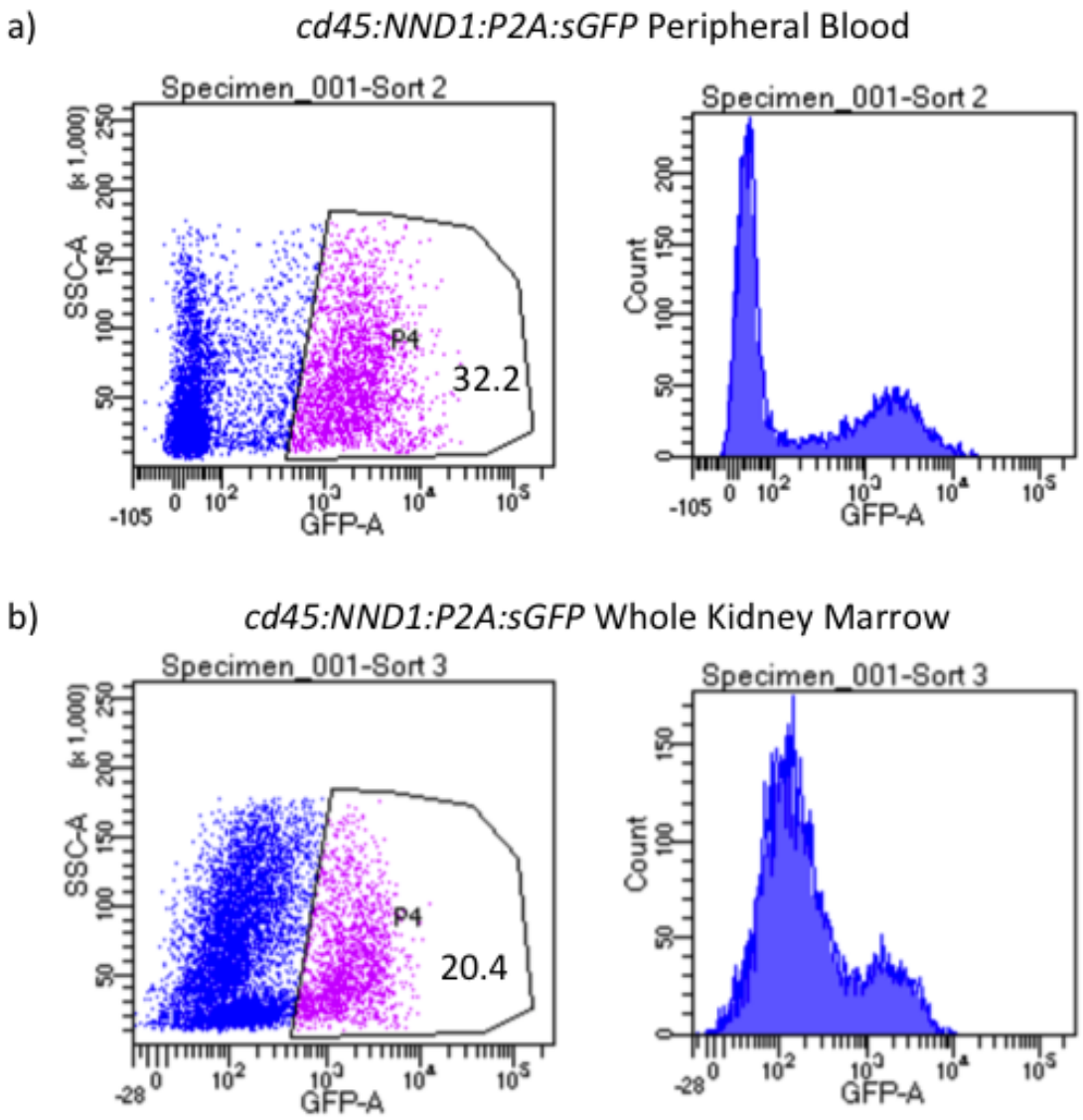


Figure 3.1.2.1: Adult *cd45:NND1:P2A:sGFP* fish have cells positive for green fluorescence in the peripheral blood (PB) (a) and whole kidney marrow (WKM) (b). Samples of PB and WKM were extracted from three adult transgenic fish at 8 months of age, and fluorescence activated cell sorting (FACS) was performed on the samples to sort out the cells positive for green fluorescent protein (GFP). In the PB, 32.2% of cells were GFP+, and in the WKM 20.4% of cells were GFP+. Cells were FACS sorted using BD FACS Aria III equipped with blue (488nm) and red (633nm) laser.

3.2: Successful Knockout of the *nup98* Gene using CRISPR/Cas9

3.2.1 The nup98 Gene was Efficiently Knocked out in Casper Embryos using CRISPR/Cas9 Gene Editing

Six sgRNAs were designed to target the zebrafish *nup98* gene, concentrating on exons 3 and 7, which encode the FG repeats important for protein function (Table 2.4.1.1, Figure 3.2.1.1a). To knockout the *nup98* gene, the 6 sgRNAs were injected with Cas9 mRNA into *casper* zebrafish embryos at the one-cell stage. To test the efficiency of the gene knockout in injected embryos, DNA was extracted from groups of injected embryos and analyzed at 24 hpf. A *Taq* Long-Amp® polymerase assay designed to amplify the 9kb region encompassing all of the sgRNAs was used to analyze the presence of a knockout in injected embryos. Analysis with gel electrophoresis showed a less intense 9kB band in injected embryos compared to un-injected *casper* control embryos, with the presence of smaller bands (indicating sgRNA cuts) present in the lanes of injected embryos (Figure 3.2.1.1b). After confirmation that the *nup98* gene was being targeted for knockout in injected embryos, groups of embryos were injected either for experiments carried out within 2 days, or grown up to breeding age to establish a mutant fish line. Founder fish that can be used to establish *nup98* knockout mutant lines were identified. Injected fish were grown up to breeding age, outcrossed to an un-injected *casper* fish, and the resulting F_1 embryos were pooled and analyzed using a heteroduplex mobility assay (HMA) (refer to Figure 2.4.3.1 in Chapter 2 for the overview of HMA experimental steps and analysis). Positive embryos were identified based on the presence of bands above the wildtype control band (Figure 3.2.1.1c). Currently, 3 founder fish with a mutation in *nup98* have been found using the HMA assay.

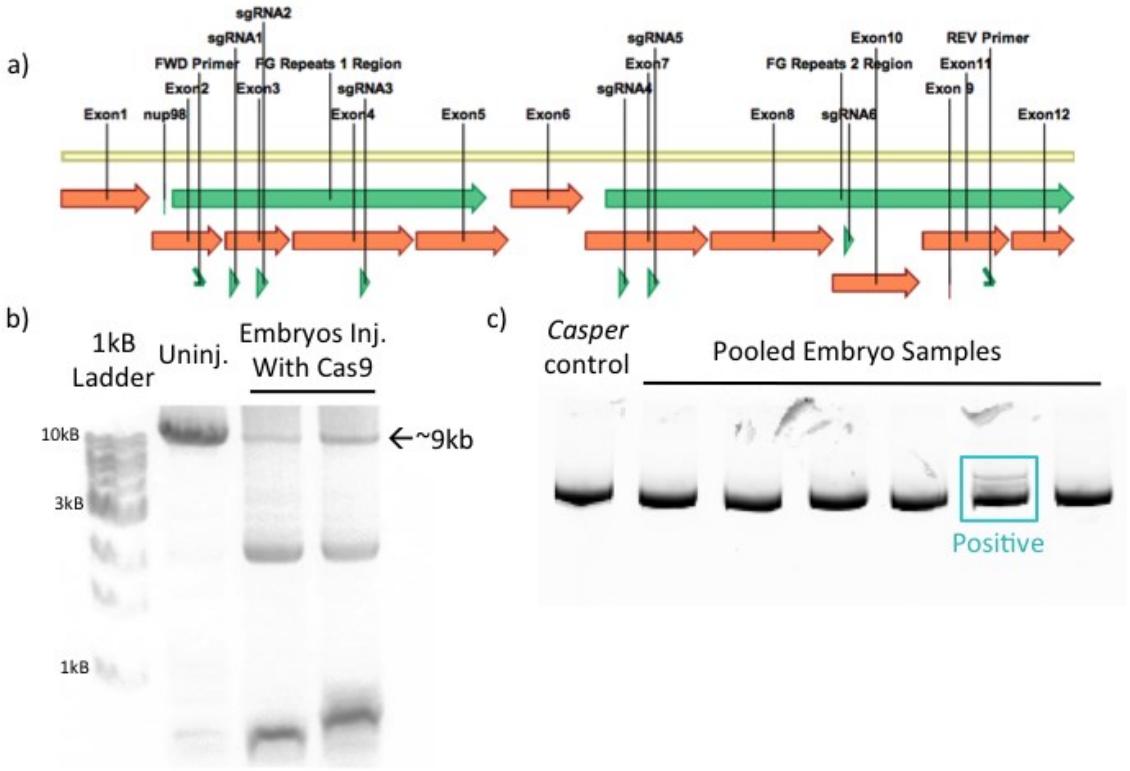


Figure 3.2.1.1: Assays demonstrating the efficiency of *nup98* knockout (KO). (a) A picture from the program Vector NTI showing the design of short guide RNA (sgRNA) targets on the *nup98* gene, as well as exons, primers, and important domains. A CRISPR multi-targeted approach with 6 sgRNAs was used to target exons 3 and 7. The *nup98* gene is displayed without introns for simplicity. **(b)** *Taq* Long-Amp® PCR shows primary injected embryos have cuts in *nup98* (smaller bands) compared to un-injected control. A 9kb band amplified the region containing all the sgRNAs. The forward and reverse primers used to amplify this region are shown in (a), located in exons 2 and 11, respectively. Lane 1: 1kB ladder (New England Biolabs), lane 2: un-injected control, lanes 3,4: *casper* embryos injected with *Cas9* and sgRNAs. **(c)** Heteroduplex mobility assay (HMA) on pooled embryos from outcrossed injected fish show a positive for a *nup98* mutation in lane 6. Lane 1: *casper* control, lanes 2-7: pooled embryo samples 1—6.

3.3: Transgenic Embryos with and without *nup98* Knockout, and Adult Transgenic Fish Display Disrupted Blood Development Compared to *Casper* Controls

3.3.1 The NUP98-NSD1 Fusion and the nup98 Knockout (KO) Contribute to Disrupted Blood Development in Zebrafish Embryos

Expression of *NUP98-NSD1* and/or a reduction in *nup98* expression have an effect on blood development in zebrafish embryos. Figure 3.3.1.1 shows graphs with the change in phenotype score (a) or cell numbers (c-j), accompanied with representative images for each *in situ* probe. In transgenic embryos with and without *nup98* KO there is a decrease in *gata1*, marking red blood cells (RBCs), at 24 hours post-fertilization (hpf) (Figure 3.3.1.1a). In transgenic embryos with and without *nup98* KO, *pu.1*, which marks early myeloid cells, was significantly decreased at both 24 and 28 hpf (Figure 3.3.1.1c, d). Ectopic expression of *pu.1* was also seen (Figure 3.3.1.1c/d, arrows), but only in the transgenic embryos with and without *nup98* KO at both 24 and 28 hpf. In addition, *mpx*, which marks neutrophils, was significantly decreased at 28 and 36 hpf in transgenic embryos with and without *nup98* KO (Figure 3.3.1.1e, f). A significant decrease was also seen in *lcp1*, marking monocytes, but was only observed in transgenic fish with *nup98* KO; a similar decrease in monocytes was also seen in control fish in which *nup98* was knocked out (28 and 36 hpf) (Figure 3.3.1.1g, h). Interestingly, hematopoietic stem cells (HSCs) marked by *runx1* and *cmyb* were increased at 36 hpf, but only in the transgenic embryos injected with the *nup98* KO, and the increase seen at 48 hpf is in the *casper* embryos injected with *nup98* KO (Figure 3.3.1.1i, j). These results are summarized in Table 3.3.1.1.

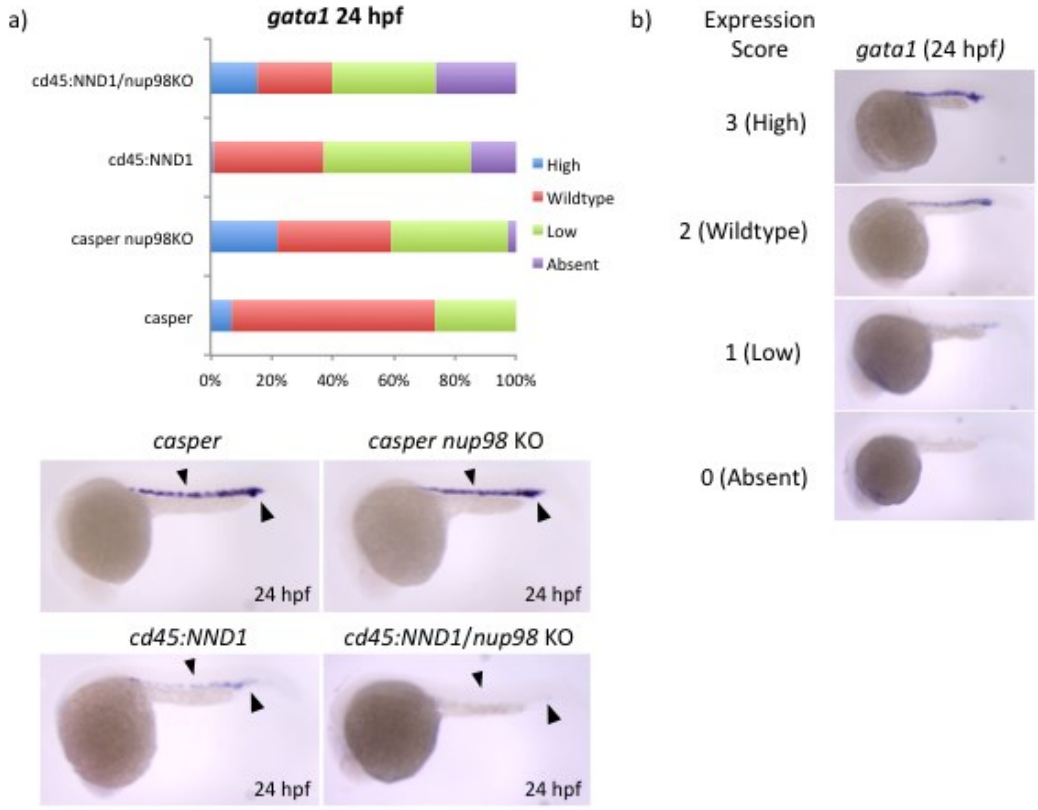
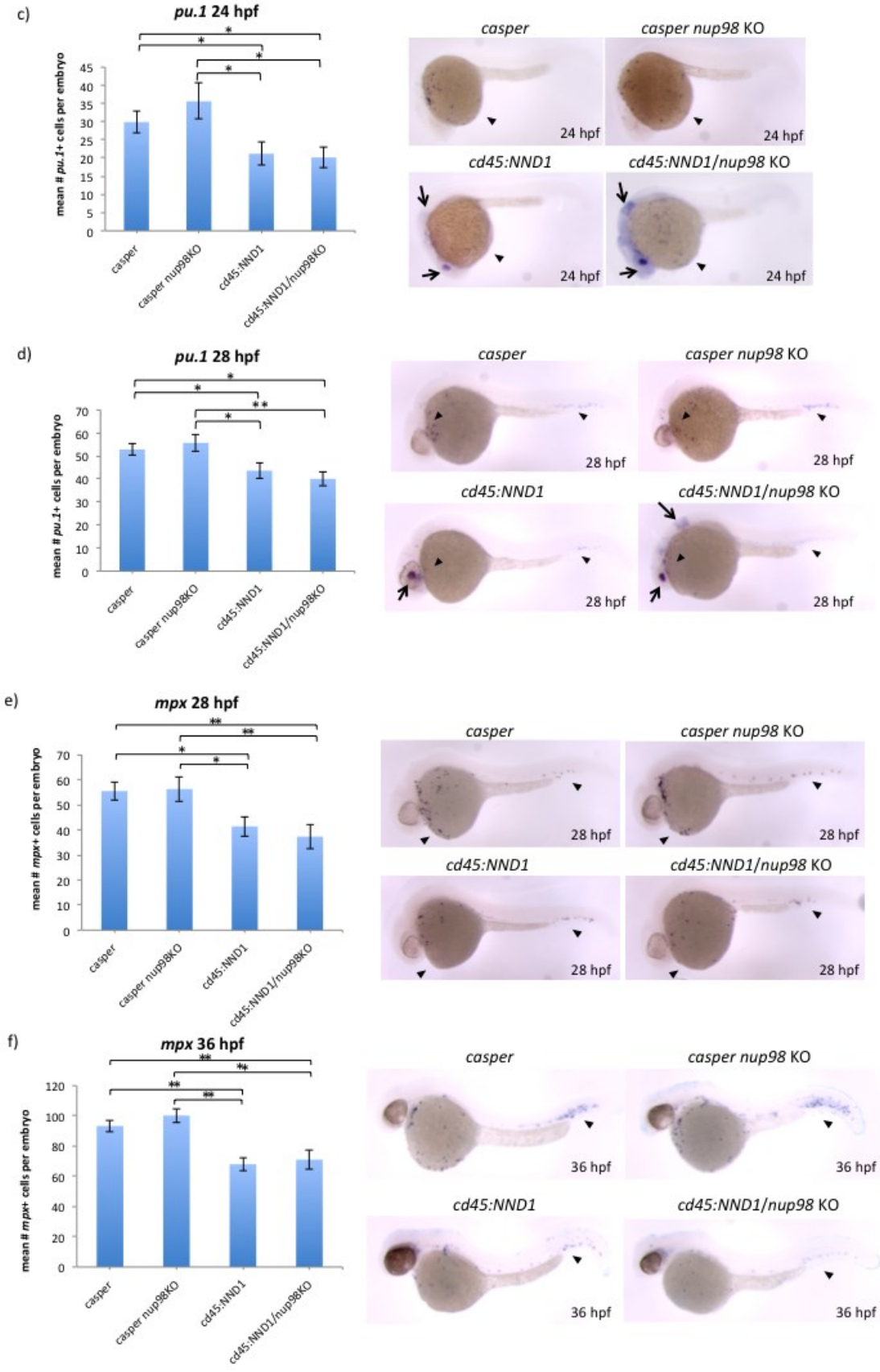


Figure 3.3.1.1 (*continued on next 2 pages*): whole-mount *in situ* hybridization (WISH) assays show that blood development is disrupted in the transgenic *cd45:NND1:P2A:sGFP* and in fish injected with *nup98* knockout (KO). (a) The red blood cell (RBC) marker *gata1* was decreased in transgenic fish with and without the *nup98* KO, according to a blinded qualitative scoring analysis of each group. Fish were scored for *gata1* expression according to the scale shown in **(b)**. Quantitative analysis was performed using the ImageJ cell counter plugin for probes *pu.1* (early myeloid cells), *lcp1* (monocytes), *mpx* (neutrophils), and *runx1/cmyb* (hematopoietic stem cells [HSCs]). Data analysis was performed using the program Stata®. Data presented are mean number of positive cells per embryo (± standard error of the mean (SEM)). The myeloid marker *pu.1* was decreased in transgenic fish with and without *nup98* KO at 24 **(c)** and 28 **(d)** hours post fertilization (hpf). Neutrophils marked by *mpx* were decreased in the transgenic with and without *nup98* KO at 28 **(e)** and 36 hpf **(f)**. The monocyte marker *lcp1* was decreased during *nup98* KO in both *casper* and transgenic fish at 28 **(g)** and 36 **(h)** hpf. HSCs were increased only in the transgenic with *nup98* KO at 36 hpf **(i)**, and in caspers with *nup98* KO at 48 hpf **(j)**. Arrowheads highlight regions containing positive cells, and arrows point to ectopic expression of the gene analyzed. Two replicates were performed for each probe/time point, with 20 embryos per group in each replicate. * = $p<0.05$, ** = $p<0.001$.


c)
pu.1 24 hpf
mean # pu.1+ cells per embryo
casper
casper nup98KO
cd45:NND1
cd45:NND1/nup98KO
casper
casper nup98 KO
cd45:NND1
cd45:NND1/nup98 KO
24 hpf
d)
pu.1 28 hpf
mean # pu.1+ cells per embryo
28 hpf
e)
mpx 28 hpf
mean # mpx+ cells per embryo
f)
mpx 36 hpf
mean # mpx+ cells per embryo
36 hpf

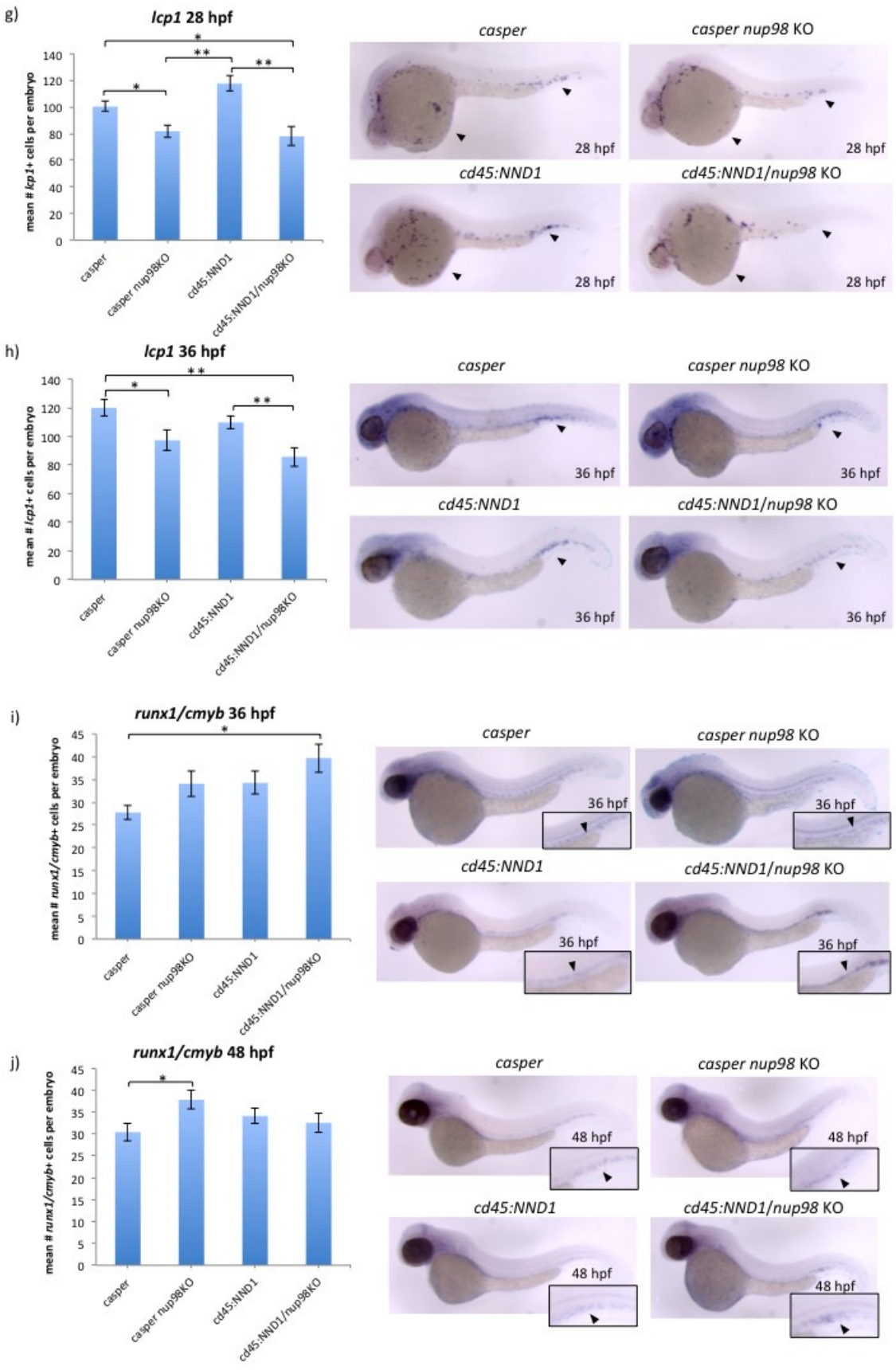

g)
lcp1 28 hpf
mean # lcp1+ cells per embryo
casper
casper nup98KO
cd45:NND1
cd45:NND1/nup98KO
casper
casper nup98 KO
cd45:NND1
cd45:NND1/nup98 KO
28 hpf
h)
lcp1 36 hpf
36 hpf
i)
runx1/cmyb 36 hpf
mean # runx1/cmyb+ cells per embryo
j)
runx1/cmyb 48 hpf
48 hpf

Table 3.3.1.1: Summary of whole-mount *in situ* hybridization results analyzing changes in blood development of *cd45:NND1:P2A:sGFP* transgenic embryos and *casper* embryos with *nup98* knockout, alone and together, compared to the *casper* control

	Probe and Time Point (hours post-fertilization)								
	gata1	*pu.1*		*mpx*		*lcp1*		*runx1/cmyb*	
Group	24	24	28	28	36	28	36	36	48
casper with *nup98* KO	-	-	-	-	-	↓*	↓*	-	↑*
cd45:NND1	↓	↓*	↓*	↓*	↓**	-	-	-	-
cd45:NND1 with *nup98* KO	↓	↓*	↓*	↓**	↓**	↓*	↓**	↑*	-

↑ = increase in expression compared to *casper* embryos (staining/cell number)
↓ = decrease in expression compared to *casper* embryos (staining/cell number)
- = no change in expression compared to *casper* embryos (staining/cell number)
* = $p<0.05$
** = $p<0.001$

3.3.2 Adult cd45:NND1:P2A:sGFP Transgenic Fish show Changes to the Proportions of Blood Cell Populations Present in the Peripheral Blood (PB) and Whole Kidney Marrow (WKM)

The disruption in blood development in *cd45:NND1:P2A:sGFP* transgenic fish is extends beyond the embryonic phenotype. When the peripheral blood (PB) of adult *casper* and *cd45:NND1:P2A:sGFP* F_2 fish were compared, transgenic fish were found to have different proportions within blood cell populations than *casper* controls (Figure 3.3.2.1). Flow cytometry revealed that in the PB of the transgenic fish there were more than double the number of monocytes (49.9% compared to 22.1% in *caspers*), and less than half the number of lymphocytes (8.71% compared to 22.4% in *caspers*). There was also slightly higher numbers of granulocytes in the transgenic versus the *casper* fish (1.83% versus 0.35%). Similar trends were also seen in the WKM: in the transgenic fish there were slightly more monocytes (16% versus 14% in *casper*) and less lymphocytes (3.88% compared to 9.52% in *casper*) (Figure 3.3.2.2). There was also an increase in precursor cells in the transgenic fish (6.58% compared to 2.75% in *casper*), and a decrease in red blood cells (55.5% compared to 62.5% in *casper*). Overall, the adult transgenic fish had more monocytes/granulocytes and precursor cells, and less lymphocytes and red blood cells than the adult *casper* fish.

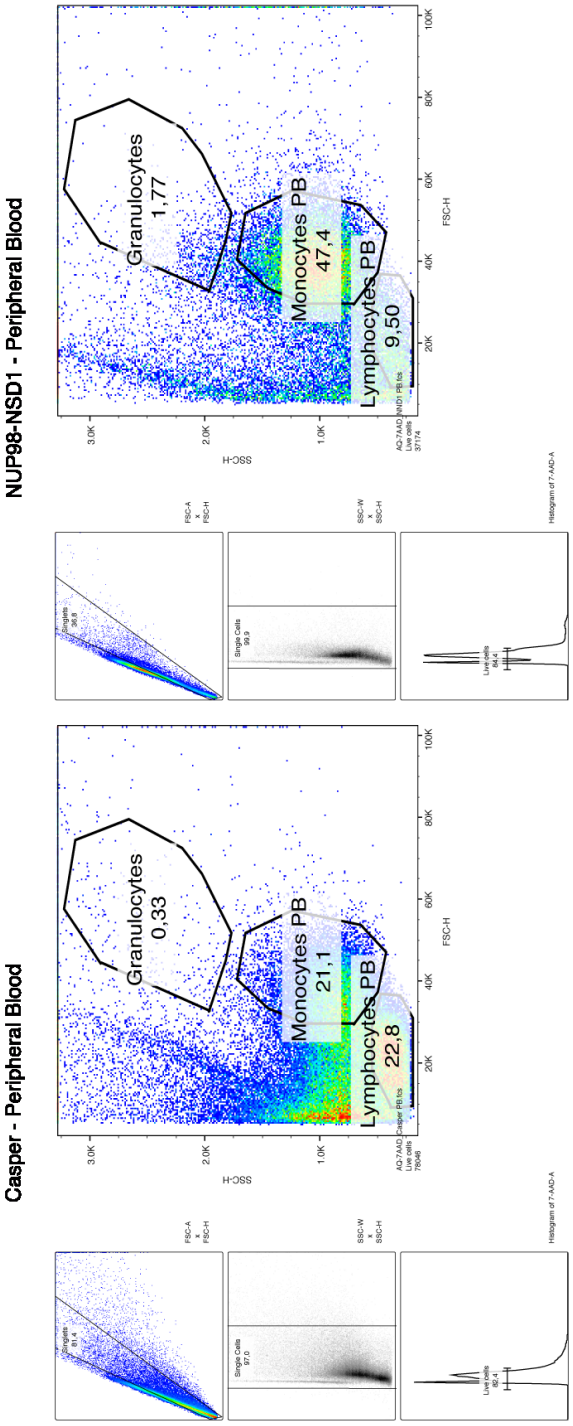


Figure 3.3.2.1: Peripheral blood from adult *casper* and *cd45:NND1:P2A:sGFP* zebrafish show different amounts of blood cell populations. Transgenic fish have more monocytes and granulocytes, and less lymphocytes compared to *casper* fish. Cells were analyzed by selecting only for single cells, and live cells (7AAD negative). Three adult fish were used per group, and this experiment represents one replicate.

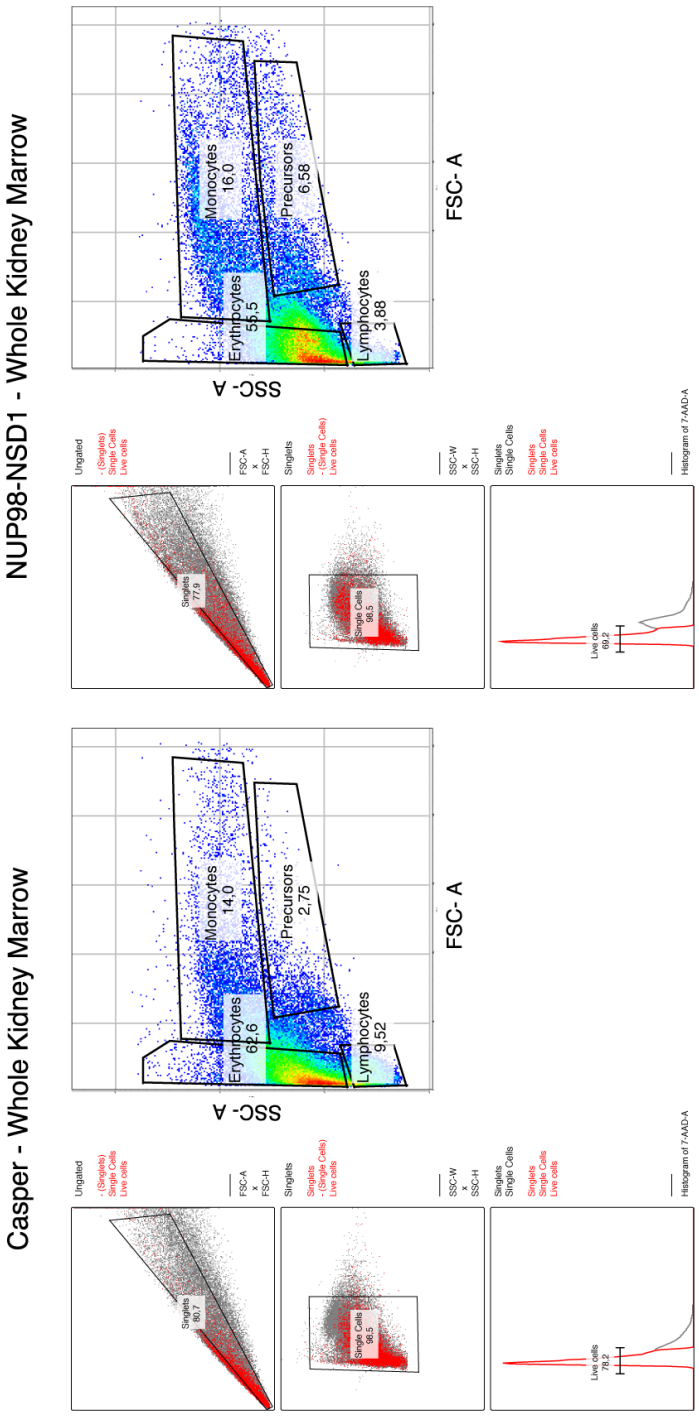


Figure 3.3.2.2: Whole kidney marrow from adult *casper* and *cd45:NND1:P2A:sGFP* zebrafish show different amounts of blood cell populations. Transgenic fish have more monocytes and precursor cells, and less lymphocytes and red blood cells compared to *casper* fish. Cells were analyzed by selecting only for single cells, and live cells (7AAD negative). Three adult fish were used per group, and this experiment represents one replicate.

3.4: RNA-sequencing shows Changes in Gene Expression of Transgenic Embryos Compared to *Casper* Embryos

3.4.1 RNA Sequencing Results show a Developmental Delay in Transgenic Zebrafish Embryos

Five groups of embryos were analyzed using RNA sequencing (RNA-seq): *casper* control embryos, transgenic *cd45:NND1:P2A:sGFP* embryos, *casper* embryos injected with *nup98* KO, transgenic embryos injected with *nup98 KO*, and *casper* embryos injected with the plasmid *cd45-NND1-P2A-sGFP*. A heat map showing the top differentially expressed genes between these groups is shown in Figure 3.4.1.1. Red indicates increased gene expression and green indicates decreased gene expression. The largest difference in gene expression is seen between the *casper* control and the *cd45:NND1:P2A:sGFP* transgenic line. Most of the downregulated genes are associated with muscle development, and most of the upregulated genes are associated with ribosomal proteins.

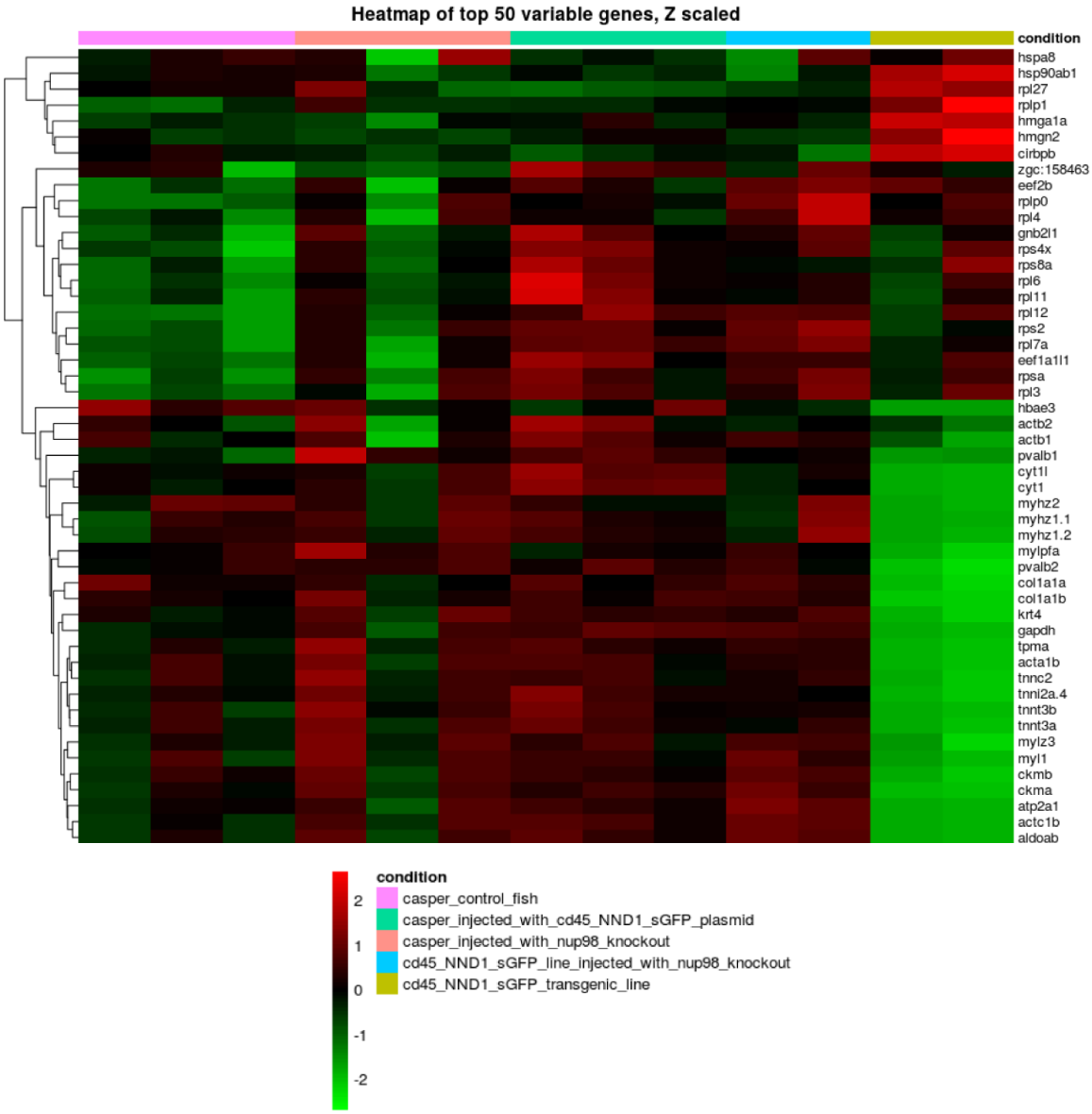


Figure 3.4.1.1: Heat map showing the top genes with increased (red) and decreased (green) expression from RNA sequencing analysis. Five groups of fish were analyzed (left to right): *casper* control fish (pink), *casper* injected with *nup98* knockout (orange), *casper* injected with the plasmid *cd45-NND1-sGFP* (green), the transgenic line *cd45:NND1:P2A:sGFP* injected with *nup98* knockout (blue), and the transgenic line *cd45:NND1:P2A:sGFP* (gold). All groups analyzed contained RNA extracted from 50-100 embryos staged at 36 hours post-fertilization. RNA was extracted using TRIzol™ Reagent.

Chapter 4: Discussion

4.1 *cd45:NND1:P2A:sGFP* embryos display GFP in sites of hematopoiesis and have off-target expression

4.1.1 Choice of Promoter Matters for Expression of NUP98-NSD1 in a Transgenic Line

When making a transgenic, we chose to use both a ubiquitous and tissue-specific promoter to express the transgene, as little was known about the consequences of expression of *NUP98-NSD1* in zebrafish; the transgene may have caused embryonic lethality, infertility, or no effect depending on the amount of expression. In our case, the ubiquitous zebrafish promoter *ubi* caused embryonic lethality in F_1 fish, and therefore the fish line *ubi:NND1:P2A:sGFP* was not successfully grown to maturity. F_0 embryos that were transiently injected with the plasmid *ubi-NND1-P2A-sGFP* did not show embryonic lethality, probably because plasmid incorporation was mosaic and the expression was not immediate. True ubiquitous expression only occurred in F_1 fish when a founder with plasmid incorporation in germline cells was able to pass down the transgene to the next generation where it could be truly ubiquitously expressed and was embryonic lethal.

The *cd45* promoter was also chosen to express *NUP98-NSD1* because *cd45* allows for selective expression in white blood cells, which are central to the disease progression of AML. The *cd45* promoter (named *ptprc* in zebrafish), was previously discovered to be a pan-leukocyte promoter that reliably marked all myeloid cells (Bertrand et al. 2008; Wittamer et al. 2011). In the transgenic fish *cd45:dsRed,* expression of *cd45* started around 24 hpf in the intermediate cell mass (ICM) and posterior blood island (PBI), and was also seen in the aorta-gonad-mesonephros (AGM) around 30 hpf (Bertrand et al. 2008). Wittamer and colleagues further explored *cd45* expression in adult zebrafish, and it was found that *cd45* was not observed in erythrocytes or B cells, and that almost all myeloid cells and T cells were marked by *cd45* (Wittamer et al. 2011). Our transgenic *cd45:NND1:P2A:sGFP* also shows expression in the ICM and PBI starting around 24 hpf, as well as in the AGM around 30 hpf (Figure 3.1.1.1c) and has fluorescence into adulthood (Figure 3.1.2.1). However, our transgenic also has green fluorescence in parts of the rhombomere, a developmental segment of

the hindbrain, as well as throughout the tail (Figure 3.1.1.1c). This type of expression was not expected and is considered off-target.

Having off-target expression can be common when making transgenic zebrafish. For example, the transgenic *gata1:GFP* was found to express GFP in blood cells, as expected, but also showed GFP expression in neurons along the tail; the *pu.1:NHA9:GFP* transgenic displayed expression in myeloid cells along with off-target CNS expression (Long et al. 1997; Forrester et al. 2011). One of the reasons why we may be seeing off-target expression in *cd45:NND1:P2A:sGFP* fish is that the *cd45* promoter could be 'leaky', meaning that it is being expressed in different locations and/or time points than expected. There are a few different factors than can cause a promoter to be leaky: the influence of enhancers near the integration site of the transgene, the loss of *cis*-regulatory elements (CREs) when the promoter was cloned, and changes in chromatin-mediated regulation in transgenics (Hernandez-Garcia & Finer 2014). As other zebrafish transgenics that have used the *cd45* promoter have not reported off-target expression, the potential 'leakiness' of the *cd45* promoter in our model is not likely to be due to the loss of CREs or changes in chromatin-mediated regulation, as these factors would likely affect the *cd45* promoter in any transgenic fish. It is possible that the leakiness could be due to a "position effect", where the insertion of our construct could have occurred near endogenous enhancers in the zebrafish genome that are causing the aberrant expression. As our transgenic most likely has many insertions of the transgene it would be difficult to map exactly where all the insertions are located in the zebrafish genome. An easy way to determine if the off-target expression is due to a positional effect would be to identify another *cd45:NND1:P2A:sGFP* founder fish and start another transgenic line, as the chances of having the exact same positional effect in two different founders is quite low; therefore, if a *cd45:NND1:P2A:sGFP* transgenic line from another founder fish had the same off-target expression in the rhombomeres and tail, it is not likely a positional effect. Unfortunately, the integration rate of the *cd45-NND1-P2A-sGFP* plasmid into zebrafish germline cells was very low for this construct, with over 150 injected embryos grown up to breeding age and outcrossed before 1 founder was identified, and so it would not be practical to screen for more founders. This percentage of founders is much lower than the normal 50% for the *Tol2* system, and is most likely due to the large size of the *cd45-NND1-P2A-sGFP* construct (Kawakami 2004).

4.1.2 NUP98-NSD1 was Detected in the GFP+ Cells of Transgenic Embryos

Although GFP should be an indicator of *NUP98-NSD1* translation, we wanted to ensure that we could detect this protein in transgenic fish. To do so, the anti-NUP98 antibody was used, which was specific to human NUP98 and did not cross-react with zebrafish nup98 (Forrester et al. 2011). Originally we could not detect this protein when whole transgenic embryos were dissociated into a single-cell suspension with subsequent protein extraction. This was most likely a titer issue due to relatively low expression of NUP98-NSD1 under the *cd45* promoter compared to all of the other zebrafish cells not expressing the transgene. However, after dissociating transgenic embryos, using FACS to extract all of the GFP+ cells, and performing a protein extraction solely on the GFP+ cells, the protein could be detected (Figure 3.1.1.1d). As no zebrafish positive control expressing NND1 was available, HEK293 cells were transfected with the plasmid *CMV-NND1-P2A-sGFP* to overexpress NND1 under a ubiquitous mammalian promoter. In Figure 3.1.1.1d it is apparent that there are protein bands above and below the identified *NND1* band. The bands below the NND1 band have been previously described in another assay detecting NUP98-HOXA9 in transgenic zebrafish, and were expected in this assay as well (Forrester et al. 2011). The bands above the NND1 band may be due to post-translational modifications of NUP98-NSD1, or may also be the NND1 protein not yet separated from GFP *via* the P2A site. Regardless, the NND1 band is present in the positive control and the transgenic fish sorted for GFP+ cells, indicating that NND1 is being expressed in the transgenic fish.

4.2 CRISPR/Cas9 was used to knockout *nup98* in zebrafish embryos

4.2.1 Design of the nup98 Knockout and Taq LongAmp® PCR Assay

CRISPR/Cas9 is a novel gene editing technique that can be efficiently used on zebrafish embryos. One of the most common applications of CRISPR/Cas9 in zebrafish is for gene knockouts, and a multi-targeted approach with multiple single guide RNAs (sgRNAs) can be used to effectively knock out a gene (Prykhozhij et al. 2015). Six sgRNAs were used to knockout the *nup98* gene, and targeted exons 3 and 7 because of their importance in encoding the FG repeats that are crucial for protein function (Figure 3.2.1.1a). The targeting of such a large region of DNA required the use of the *Taq* LongAmp® assay for analysis of cutting efficiency. This was not ideal due to the

financial and time costs of running these experiments. Also, while this assay could demonstrate that the sgRNAs were guiding the Cas9 to cut the DNA (Figure 3.2.1.1b), there was no information available about which sgRNAs were cutting, or if this knockout impedes the translation of nup98. As there is currently no antibody for zebrafish nup98, one way to demonstrate that no protein is being translated is to create a mutant zebrafish line and characterize the disruption of the DNA. For example, a deletion that causes a frameshift mutation would likely prevent translation of a functional nup98 protein. Additionally, qPCR should be used to detect and quantify the efficiency of a *nup98* knockout in zebrafish embryos, and could be a proxy measurement for a decrease in nup98 protein as well.

4.2.2 The Use and Limitations of the HMA Assay to Identify Founder Fish with Mutated nup98

One of the first steps to establishing a mutant zebrafish line is to identify founder fish and characterize their mutation (Prykhozhij et al. 2017). The HMA assay has allowed for the identification of 3 founder fish that could potentially be used to make a *nup98 +/-* or *-/-* mutant line, depending on the lethality of the *nup98* knockout (a positive result indicating a founder is shown in Figure 3.2.1.1c, lane 6). The HMA assay is used to detect mutations encompassed in a region containing sgRNAs, and requires an amplicon of 350-450bp. Because the sgRNAs designed for *nup98* were in different exons, two amplicons for the HMA assay (one for each exon) were needed. In addition, only 2/3 sgRNAs in each location could be analyzed with the HMA assay because of the requirement for amplicon size and the placement of the sgRNAs (refer to Figure 3.2.1.1a). Therefore, for exon 3 only sgRNAs 1 and 2 were captured by the amplicon, and for exon 7 only sgRNAs 4 and 5 were captured by the amplicon. In addition, it was found that the region amplifying exon 3 was polymorphic, and therefore could not be used for the HMA assay, due to the creation of heteroduplexes in the control sample. The 3 founder fish that were identified used only the amplicon analyzing sgRNAs 4 and 5 in exon 7 to detect founder fish positive for a mutation (Figure 3.2.1.1a/c). The HMA assay for *nup98* would benefit from the placement of sgRNAs within a region that could be amplified by one amplicon, and this would require the targeting of one exon, not two, with sgRNAs placed closer together.

4.3 Hematopoiesis is disrupted in embryos that express *NUP98-NSD1* and/or have a *nup98* knockout

4.3.1 Both NUP98-NSD1 Expression and nup98 Knockout Contribute to the Disruption of Blood Development in Zebrafish Embryos

It is known that cells expressing the *NUP98-NSD1* fusion through viral transduction of cell lines or primary patient samples exhibit self-renewal and differentiation arrest, and that when these cells are put into a mouse model acute myeloid leukemia can develop (Wang et al. 2007; Thanasopoulou et al. 2014). However, there is currently no transgenic animal model of this disease, which would provide new opportunities for studying disease progression *in vivo*. We aimed to fill this gap in knowledge by creating a transgenic zebrafish that displayed disrupted blood development akin to the *in vitro* and *ex vivo* models that have already laid the groundwork for the analysis of *NUP98-NSD1*-induced AML.

In transgenic embryos with and without *nup98* KO it was found that red blood cells were decreased, and that early myeloid cells and neutrophils were significantly decreased (Figure 3.3.1.1a-f). Monocytes were also found to be significantly decreased, but only in *casper* and transgenic fish with *nup98* KO (Figure 3.3.1.1g/h). Intriguingly, HSCs were found to be significantly increased but only in the transgenic embryos with *nup98* KO at 36 hpf and in *casper* embryos with *nup98* KO at 48 hpf (Figure 3.3.1.1i/j). Taken together, these results show that *NUP98-NSD1* and *nup98* KO have independent effects on blood development. However, from this data, it is difficult to say if the effects of *NUP98-NSD1* and *nup98* KO are additive or synergistic. With *mpx* at 28 hpf, the transgenic with *nup98* KO has a more significant decrease in neutrophil number than the transgenic alone, in reference to *casper* embryos (Figure 3.3.1.1e). With *lcp1* at 36 hpf, there is a more significant decrease in the transgenic with *nup98* KO than in *casper* with *nup98* KO, in reference to the *casper* embryos (Figure 3.3.1.1h). These two trends of a more significant decrease ($p=0.001$ instead of 0.05) when both genetic aberrations are present are only seen with these two markers and these time points, and are not consistent between time points or markers. More replicates may help to determine if these trends are present in other markers or time points. However, there are instances where it is clear that *NUP98-NSD1* is causing a change in blood cell number compared to *caspers* (i.e. with *pu.1* and *mpx* at both time points, Figure 3.3.1.1c-f), and where *nup98*

is causing this decrease compared to *caspers* (i.e. with *lcp1* at all time points, Figure 3.3.1.1g/h). Therefore, based on our results we can see that *NUP98-NSD1* and *nup98* KO are both having effects on blood development, but the additive or synergistic effects of these aberrations together remains to be determined fully.

NUP98-NSD1 and *nup98* KO seem to not only have independent effects on hematopoiesis, but also affect different waves of hematopoiesis. The transgenic expressing *NUP98-NSD1* has decreased RBCs (*gata1*) at 24 hpf and decreased early myeloid cells (*pu.1*) at 24 and 28 hpf – the markers at these time points correspond to the two primitive waves of hematopoiesis. Neutrophils (marked by *mpx*) are also decreased in the transgenic, but it is unclear if this is because *NUP98-NSD1* may also be having an effect during the definitive waves of hematopoiesis, or because early myeloid cells - that should be differentiating into neutrophils – are decreased. By contrast, the embryos with *nup98* KO show effects in both primitive and definitive hematopoiesis, with decreased monocytes at 28 hpf (primitive hematopoiesis) and 36 hpf (definitive hematopoiesis), and increased HSCs at 48 hpf (definitive hematopoiesis). The transgenic embryos with *nup98* KO show disruptions in both primitive and definitive hematopoiesis, with effects on blood development seen in all of the types of blood cells analyzed at almost every time point (Table 3.3.1.1).

It is known that *NUP98-NSD1* causes overexpression of *Hox* genes such as *HoxA7*, *HoxA9*, and *HoxA10*, and that overexpression of *HOX* genes causes cells to gain self-renewal properties; it has been hypothesized that *NUP98* fusions causing overexpression of *HOX* genes is necessary but not sufficient for the development of leukemia (Wang et al. 2007; Argiropoulos & Humphries 2007). In our transgenic model, overexpression of these *hox* genes could be what's causing the embryonic phenotype. If hematopoietic cells are gaining self-renewal properties this could lead to impaired differentiation and the pancytopenia that is occurring. In terms of the *nup98* knockout embryos, a suspected mechanism of blood disruption could be the deregulation of genes that are normally regulated by nup98. NUP98 acts as a transcriptional regulator in addition to being part of the nuclear pore complex, and decreased amounts of nup98 could be causing decreased monocytes and increased HSCs possibly due to the misregulation of *nup98*-targeted genes, which are currently unknown (Franks & Hetzer 2013; Capelson et al. 2010; Kalverda et al. 2010). This effect could potentially be

exacerbated when combined with *NUP98-NSD1* as the GLFG repeats on the NUP98-NSD1 fusion protein can also sequester nup98, potentially leading to even more severe misregulation and mis-expression of genes due to even lower amounts of available nup98 (Xu & Powers 2010). Additionally, decreased *nup98* could also lead to decreased nup96, which could potentially be one of the mechanisms behind the increase in HSCs that is seen in the embryos as Nup96 is known to be a cell cycle regulator (Chakraborty et al. 2008). As of now these proposed mechanisms need to be explored. RNA-seq can shed light on what genes may be over or under expressed in transgenic and/or *nup98* KO embryos; our RNA-seq data (discussed in section 4.5) is unfortunately inconclusive and requires further experimentation.

4.3.2 The Embryonic Phenotype of NUP98-NSD1 Transgenic Embryos Differs from NUP98-HOXA9 Transgenic Embryos

When taking into account previous research in transgenic zebrafish modelling *NUP98* fusions in leukemia, the phenotype seen in the *NUP98-NSD1* transgenic embryos is slightly unexpected. In *NUP98-HOXA9* transgenic zebrafish embryos there was a decrease in red blood cells, an increase in myeloid cells (*pu.1*, *lcp1*, *mpx*, *lyzC*), and an increase in HSCs (*runx1*/*cmyb*) (Forrester et al. 2011; Deveau et al. 2015). This phenotype has been referred to as a "pre-leukemic state", and adult *NHA9* transgenic fish developed a myeloproliferative neoplasm, although not overt leukemia (Forrester et al. 2011; Deveau et al. 2015). A similar embryonic phenotype was hypothesized in the fish expressing *NUP98-NSD1* because of the fact that *NUP98-NSD1* causes overexpression of *HOXA9* similar to *NUP98-HOXA9*, is a type II aberration like *NHA9*, and so may display a phenotype similar to *NHA9* transgenic embryos (Wang et al. 2007; Kroon et al. 2001; Balgobind et al. 2011). However, the same embryonic phenotype as the *NHA9* transgenic embryos is not seen in the *NND1* transgenic embryos. This may be due to the fact that *NHA9* and *NND1* likely have different mechanisms of leukemogenesis, despite leading to similar effects on gene expression, and so more research is needed into the genes and pathways involved in *NND1* leukemogenesis. This was one of the reasons we performed RNA-sequencing analysis (discussed in section 4.5), but as of now the results are not conclusive. In addition, *NUP98-NSD1* is most often found with type I aberrations like *FLT3*-ITD, *WT1*, and *NRAS*, unlike *NHA9*, and so *NND1*

may require the effects of other oncogenes to see a pre-leukemic state in embryos (Thanasopoulou et al. 2014; Kroon et al. 2001).

4.3.3 The Embryonic Phenotype of nup98 KO Induced by CRISPR/Cas9 Differs from nup98 Knocked Down with Morpholino

Only one other published study exists that looked at decreased *nup98* in zebrafish. This study used morpholino to knock down *nup98*, and showed a disruption in blood development using qPCR to demonstrate that there was increased *pu.1* (early myeloid marker) and *scl* (early HSC marker), and no changes to erythroid genes (Fung et al. 2010). This study showed that decreases in *nup98* have consequences for blood development; this could be especially relevant in the context of a *NUP98* fusion. Our zebrafish embryos injected with CRISPR/Cas9 and *nup98* sgRNAs show decreased monocytes (*lcp1*), and an increase in HSCs (*runx1/cmyb*) (Figure 3.3.1.1g-j). This is not entirely consistent with the findings of Fung and colleagues, which may be due to a few reasons.

First, the main difference in approach is morpholino versus CRISPR/Cas9 KO. Morpholinos are known to have off-target effects, but are less likely to have genetic compensation occur (Rossi et al. 2015). The differences in blood phenotype between the knockdown by Fung *et al*. and our *nup98* knockout could therefore be due to either off-target effects in the model by Fung and colleagues, or genetic compensation occurring in our model. As we are still seeing an effect on blood development in our model and not the absence of any phenotype, this may suggest genetic compensation is not occurring. Also, if a certain gene was compensating for the decreased amount of *nup98* it is reasonable that another *nup* gene should be upregulated in fish that have the *nup98* knockout; we did not detect any upregulated *nup* genes in our RNA-seq analysis. In addition, *NUP98* is the only vertebrate NUP with a GLFG domain, whereas other NUPs only have an FG domain, further supporting that *nup98* is not being compensated by other NUPs, as NUP98 is more unique than other NUPs (Fahrenkrog 2014).

Another reason we may have seen a different blood phenotype compared to the morphant embryos is that the embryos we used for analysis are injected at the one-cell stage and do not come from a mutant *nup98* +/- or -/- line. Although we have shown the *nup98* KO assay to be efficient (Figure 3.2.1.1b), *nup98* KOs in primary injected embryos

would be mosaic, and it is reasonable that not every zebrafish used in the WISH assay had a robust *nup98* KO, as there is no way to screen for this mutation in embryos before use in the WISH assay. Once a *nup98* knockout mutant line is established, WISH could be repeated in germline mutant embryos to see if the same effect on blood development is seen. Furthermore, the study by Fung and colleagues used qPCR to detect changes in genes involved in blood development, whereas our study did not have this truly quantitative approach. While valuable information on blood cell number was gleaned from the WISH experiments, qPCR should be done in the future to validate the increases or decreases seen in our WISH experiment. Lastly, Fung and colleagues focused on earlier hematopoietic time points, between 18-24 hpf, whereas our study wanted to capture both primitive and definitive hematopoiesis in embryos, and analyzed probes from 24-48 hpf. As speculated above, it seems that the *nup98* KO in our embryos is affecting primitive and definitive hematopoiesis. Definitive hematopoiesis was not analyzed by Fung and colleagues because of the earlier time points used in their study. Although different probes and time points were used, there is still a lack of similarity between the overall increases or decreases in types of hematopoietic cells, requiring further investigation into the true blood phenotype of zebrafish embryos with decreased levels of *nup98*.

4.3.4 cd45:NND1:P2A:sGFP Embryos with nup98 KO may be Displaying an MDS Phenotype

Myelodysplastic syndromes (MDSs) are a group of disorders that have a high risk of developing into AML, with approximately one third of MDS patients progressing to AML (Walter et al. 2012; Greenberg et al. 1997). MDS is characterized by the disordered growth and differentiation of cells during hematopoiesis, and can be caused by many types of mutations (Papaemmanuil et al. 2013). The relative amounts of different cell populations in healthy patients and those with low or high risk MDS is illustrated in Figure 4.3.4.1. Although *NUP98-NSD1* has not previously been associated with patients that have MDS, other *NUP98* fusion genes have such as *HOXA9*, *HOXD13*, *TOP1*, *AF10*, and *NSD3* (refer to Table 1.1.5.1). The cytopenia seen in Figure 4.3.4.1b is similar to the cytopenia with HSC expansion that is seen our transgenic zebrafish model with *nup98* KO: a decrease in *gata1* (RBCs), *pu.1* (GMPs), *lcp1* (monocytes), *mpx* (neutrophils), and an increase in HSCs (*runx1/cmyb*). This makes sense in the context of

our fish model, as *NUP98-NSD1* is known to be a type II aberration, causing impairment in cellular differentiation. It is important to note that the transgenic embryos alone or the *nup98* KO embryos alone do not recapitulate all of the aspects of MDS as per Figure 4.3.4.1b, as *nup98* KO embryos do not have a decrease in RBCs, early myeloid cells, or neutrophils, and transgenic embryos do not have a decrease in monocytes, or an increase in HSCs. This highlights the importance of modelling both aspects of the *NUP98-NSD1* fusion in zebrafish: the human transgene and the endogenous decrease in *nup98*, as this is where we are seeing a phenotype most similar to MDS and the largest effect on blood development.

4.3.5 Ectopic Expression of pu.1 Seen in Transgenic Embryos

WISH is not a quantitative assay, but we can get information on where blood cells are found or where a gene is being expressed. One of the surprising findings from the WISH data is that *pu.1* is ectopically expressed in the lens and hindbrain of transgenic embryos (Figure 3.3.1.1c/d, arrows). This is the only probe showing ectopic expression, and it is in areas of reporter fluorescence (lens) and off-target fluorescence (hindbrain) in the embryo (Figure 3.1.1.1a/c, white arrows). As no other probes are being ectopically expressed in this way, this may indicate a connection between leaky expression of *NUP98-NSD1* and *pu.1*. However, in other zebrafish models showing off-target expression, it was found that the areas of off-target expression only expressed GFP and not the transgene itself (Long et al. 1997). In addition, the transgene *NND1* should not be expressed in the lens, as this lens reporter fluorescence is a part of the construct used to measure integration rates of the plasmid separate from transgene expression. One of the first steps to discerning why *pu.1* is being expressed in these areas of our transgenic fish would be to isolate the GFP+ cells from the lens, hindbrain, and tail and perform immunofluorescence to see if NUP98-NSD1 is present in these areas. If NUP98-NSD1 is being expressed in these areas, then there could be some common transcription factor or otherwise that is enhancing the expression of both the transgene and *pu.1*. At this time, it is unclear what is causing the ectopic expression of *pu.1* in transgenic zebrafish.

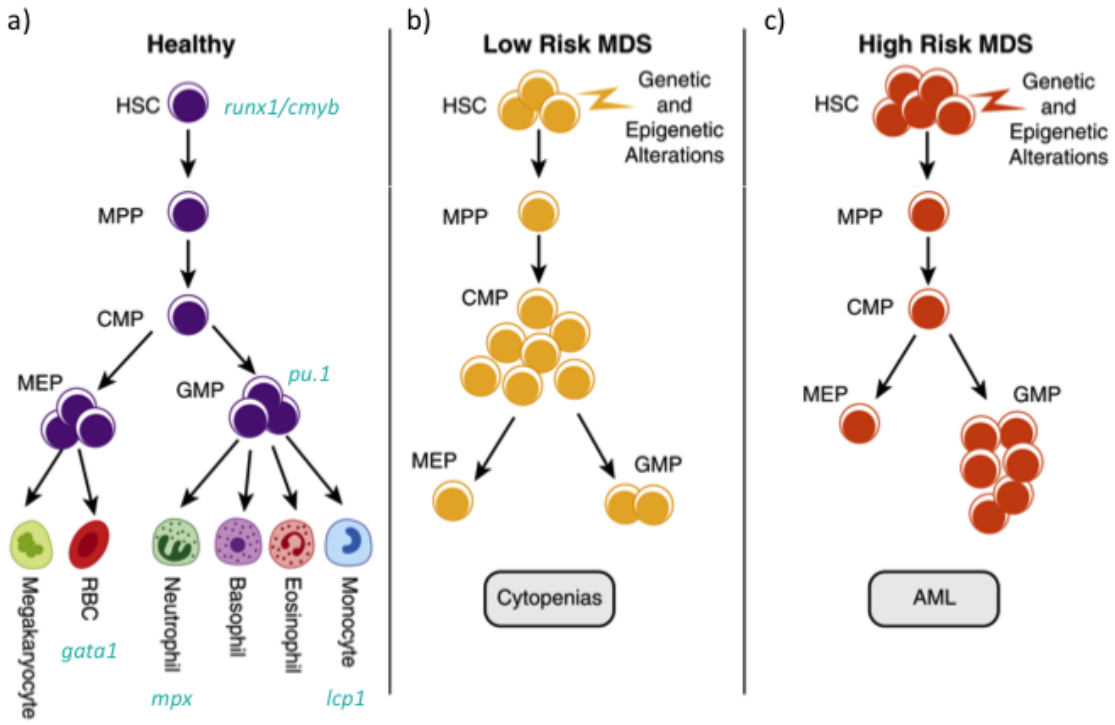


Figure 4.3.4.1: Hematopoietic stem and progenitor cell populations in a healthy individual (a) versus an individual with a low (b) or high (c) risk myelodysplastic syndrome (MDS). (a) In normal hematopoiesis, hematopoietic stem cells (HSCs) develop into multi-potent progenitors (MPPs), which in turn develop into common myeloid progenitor cells (CMPs). These CMPs have two different fates: CMPs can differentiate into megakaryocyte erythroid progenitors (MEPs) which further differentiate into red blood cells (RBCs) or megakaryocytes; or CMPs can differentiate into a granulocyte monocyte progenitor (GMP) which can further differentiate into neutrophils, basophils, eosinophils, or monocytes. The corresponding zebrafish *in situ* probes for different cell types are shown in blue. **(b)** In low risk MDS, genetic and epigenetic mutations cause the expansion of HSCs and CMPs, leading to cytopenia as the progenitor cells are not differentiating properly. **(c)** In high risk MDS there is a large expansion of HSCs and GMPs, leading to an accumulation of myeloblasts in the bone marrow, which can lead to AML. Adapted from (Shastri et al. 2017).

4.4 Hematopoiesis is Disrupted in Adult Transgenic Fish

4.4.1 cd45:NND1:P2A:sGFP Adult Fish have Different Peripheral Blood and Kidney Marrow Blood Cell Populations Compared to Caspers

A zebrafish embryonic phenotype showing disrupted blood development can give insight into leukemogenesis in a transgenic fish, but pathologic phenotypes in adult zebrafish can be more representative of disease. Embryos do not have a fully developed kidney marrow, which is analogous to the bone marrow in humans, and overt leukemia or other hematologic malignancies may not have immediate onset detectable in embryos as it may take time to develop in the adult zebrafish. In addition, zebrafish do not have a functioning adaptive immune system until approximately one month of life, and so analysis in adult zebrafish gives a more complex picture as we have additional types of mature blood cells present, more akin to the environment of human leukemogenesis (Trede et al. 2004). We wanted to compare the whole kidney marrow (WKM) and peripheral blood (PB) of adult transgenic fish to *casper* control fish to see if the transgene had any effect on long-term hematopoiesis. Flow cytometry uses forward scatter (FSC) and side scatter (SSC) to measure the size and complexity/granularity of the cells, respectively, and therefore can differentiate between different cell types present in a sample from hematopoietic tissue. In our analysis with flow cytometry it was found that overall, transgenic fish have more monocytes, granulocytes, and precursor cells, and less lymphocytes and RBCs than *casper* fish (Figures 3.3.2.1 and 3.3.2.2). This experiment needs to be repeated with more replicates so that statistical analysis may determine the degree to which blood cell populations have changed. This shift from lymphocytes to precursor cells/myelocytes in transgenic fish could indicate that the WKM and PB are being occupied by blast cells that are invading these tissues and preventing the proper formation of other blood cells, similar to what happens in AML. This scenario in adult transgenic fish is also similar to the schematic for high-risk MDS in Figure 4.3.4.1c, where there is an abundance of myeloid cells and precursor cells. From the data we have now it is not possible to distinguish whether this phenotype in adult transgenic fish is more similar to MDS or AML; for a proper diagnosis, sections of the WKM of transgenic fish should be fixed on a slide and analyzed with a stain used for hematopoietic cells, such as the Giemsa stain; if more than ~20%

of cells are blasts this would indicate an AML-like phenotype in the adult fish, whereas <20% blast cells would indicate an MDS-like phenotype.

This analysis in adult fish is only analogous to the WISH data from transgenic embryos without the *nup98* KO, as these adult fish express *NUP98-NSD1* but also express endogenous nup98. Future experiments with flow cytometry and histology should be performed on fish generated from crossing the transgenic fish with the mutant fish line *nup98* +/- or -/- (once established), to see if decreased levels of *nup98* have an effect on blood cell populations in the WKM and PB. Data from the embryonic WISH samples suggest that there would be an effect, possibly that a more severe phenotype would emerge in transgenic adult fish with *nup98* +/-. For example, if our transgenic adult fish were found to develop MDS as adults, transgenic adult fish with *nup98* KO may develop overt AML. To see the effects of *nup98* +/- on other *NUP98* fusions, a *nup98* mutant line could also be crossed with the *NUP98-HOXA9* transgenic fish already present in the Berman lab. These *NHA9* transgenic fish were shown to develop a MPN at 23 months of age, and *NUP98-HOXA9* adults with *nup98* +/- could possibly show a phenotype different than an MPN, like AML. However, analysis of *NUP98-NSD1* transgenic fish, even with *nup98* KO, may not show overt AML, and may require a second genetic mutation, such as *FLT3*-ITD, before developing more severe disease.

4.5 RNA-sequencing Found Many Genes Downregulated in Transgenic Zebrafish Embryos Compared to *Casper* Embryos

4.5.1 RNA-seq Data Shows a Developmental Delay in cd45:NND1:P2A:sGFP Transgenic Embryos

Zebrafish from five different experimental groups were analyzed to look at changes in gene expression that may help to explain the disruption in hematopoiesis seen in transgenic embryos and/or embryos with *nup98* KO. A 36 hpf time point was chosen for analysis as this should capture potential changes in genes involved in HSC development, and is on the cusp of primitive and definitive hematopoiesis. Overall, the differentially expressed genes show a developmental delay in the transgenic fish, with no obvious association to genes involved in hematopoiesis except for the downregulation of the gene *hbae3*, which is involved in red blood cell development. The

biggest difference in gene expression is seen between the *casper* control group and the transgenic embryos, and is illustrated in the heat map of Figure 3.4.1.1. Most of the genes that are downregulated in the transgenic compared to the *casper* embryos are genes involved in muscle development, indicating that the transgenic fish are developmentally delayed. This is despite stringent staging of all zebrafish embryos at the 36 hpf time point. This developmental delay shown by downregulation of muscle genes may have confounded results, and could explain why we are not seeing a decrease in the myeloid genes that were decreased in WISH. Most of the genes that are upregulated in the transgenic fish compared to the *casper* groups are genes that encode ribosomal proteins. This may hint at a ribosomopathy, which have been found to underlie bone marrow failure syndromes such as Diamond-Blackfan anemia and Shwachman-Diamond syndrome that predispose to the development of AML, however these ribosomopathies are caused by a decrease in ribosomal gene expression, not an increase (Nakhoul et al. 2014). Furthermore, this increase in ribosomal gene expression is not consistent within the replicates of *cd45:NND1:P2A:sGFP* embryos (Figure 3.4.1.1), and therefore it could be that the ribosomal genes have very variable expression, meaning that the increase we're seeing in the transgenic fish is not a true finding. It was thought that the *hoxa* genes would be increased in the transgenic embryos as NUP98-NSD1 is known to cause overexpression of *HOXA* genes; one reason this may not have been detected in our RNA-seq is that whole-embryo analysis may have overpowered changes in gene expression seen in cells expressing GFP, which are a minority in the embryos due to the expression under the *cd45* promoter. Another possibility is that 36 hpf is too early of a time point to detect changes in *hoxa* gene expression, and that looking at 48 or even 72 hpf may yield different results. Tissue specific analysis at a later time point, for example the kidney marrow of adult fish, may help to eliminate these variables. Furthermore, performing qPCR to detect changes in *hoxa* genes can be done to confirm the changes in gene expression in embryos.

Other genes of interest from the RNA-seq include the *hmga1a* and *hmgn2* genes, which are high-mobility group (HMG) proteins associated with transcriptionally active chromatin and the regulation of gene transcription, and may be genes of interest to further explore to determine their effect on leukemogenesis due to their increased gene expression in transgenic embryos. The HMGN family of proteins facilitates DNA repair,

where HMGN2 is specifically a nucleosome-binding protein that modulates global genome repair (Reeves 2015; Subramanian et al. 2009). The *HMGA1* group of genes are known to be involved in transcriptional regulation, embryogenesis, transformation, differentiation and DNA repair (reviewed by (Resar 2010)). In previous studies focusing on leukemia, *HMGA1* overexpression was correlated with relapse in pediatric B-cell ALL, and *HMGN2* was found to have an age-dependent expression in AML patients where there was higher expression in younger patients and lower expression in older patients (Roy et al. 2013; Homme et al. 2010). The differential expression of the genes mentioned from our RNA-seq analysis should be validated with qPCR in embryos, and if the trend holds these genes should then be explored for their potential role in *NUP98-NSD1*-induced leukemogenesis.

4.5.2 Limitations of the RNA-seq and Suggestions for Future Experiments

Overall, the RNA-seq data has some confounding factors that make it difficult to glean any information related to blood development in the embryos analyzed. The data set needs to be simplified to obtain clearer results. Suggestions for achieving this goal include future RNA-seq experiments where the analysis focuses on two groups of embryos before doing analysis on multiple groups at once. Sending a higher number, such as 6 replicates each, for only *casper* control embryos and *NUP98-NSD1* transgenic embryos may allow for less noise in the data set to reveal trends not seen when 5 groups were analyzed together. In addition, analyzing a mutant line with *nup98* knocked out instead of primary injected embryos would allow for better consistency in terms of *nup98* gene expression, and potentially less variability in gene expression amongst embryos. We know through this RNA-seq analysis that there is a lot of background noise in the samples, highlighted by the fact that transgenic embryos are showing a developmental delay despite staging. To eliminate this effect specific tissues could be analyzed in lieu of whole embryos. Expression analysis in zebrafish does not necessarily have to be tissue specific to obtain valuable results, as showcased by the microarray that was performed on whole *NHA9* transgenic embryos that revealed the overexpression of *dnmt1* as a mechanism of leukemogenesis (Deveau et al. 2015); however, potentially due to factors such as the off-target expression seen in the *NUP98-NSD1* transgenic embryos, further expression analysis may benefit from focusing on

certain tissues, for example the kidney marrow of adult fish. Now that a phenotype has been established in both embryos and adult transgenic fish, it may be more appropriate to focus on the blood cells instead of the whole fish at once. All of these suggestions may help to eliminate noise in the data set in future experiments.

4.6 Limitations

4.6.1 General Limitations of the Zebrafish Model

Zebrafish are useful models to study human disease because their status as vertebrates allows for similarities in developmental programs, such as hematopoiesis (Avagyan & Zon 2016). Zebrafish have many advantages over mammalian models such as lower experiment cost, larger sample sizes, ease of genetic manipulation, and optical clarity for imaging (Lieschke & Currie 2007). However, there are limitations of the zebrafish model in hematological research. There are small differences in morphology of blood cells, such as the nucleation of red blood cells, which researchers need to be aware of, although this may not affect experiments (Carradice & Lieschke 2008). In terms of experiment design, there are a lack of antibodies and markers for zebrafish cells, making flow cytometry experiments more difficult in zebrafish (Carradice & Lieschke 2008). While zebrafish possess many genes also important to human hematopoiesis, zebrafish have many duplicated genes, which can require more complex gene knockout experiments, and the mutation of zebrafish genes does not always correspond to human disease (Carradice & Lieschke 2008). Human NUP98 and zebrafish nup98 share 65% identity (Fung et al. 2010), and human NSD1 and zebrafish nsd1 share 30% identity, meaning that expression of NUP98-NSD1 in zebrafish may not exactly recapitulate downstream transcriptional effects due to the difference in species. Additionally, although zebrafish drug screens are valuable due to the whole-organism feedback for bioavailability and toxicity, zebrafish cannot replace drug testing in mammals, where the safety of a drug can be better evaluated for use in humans (Carradice & Lieschke 2008; Zon & Peterson 2005).

4.6.2 Many Transgenic Approaches Can Be Taken to Model Cancer

Transgenic zebrafish have been widely used to model cancers, especially different types of leukemia. The versatility of the zebrafish and ease of genetic

manipulation leaves many options for how a cancer is modeled. Our transgenic *cd45:NND1:P2A:sGFP* expresses the oncogene *NUP98-NSD1* under the pan-leukocyte promoter *cd45*, with the expression visualized by GFP. A popular option for modelling cancers is an inducible zebrafish model, which has a few distinct advantages. For example, if a gene is found to be lethal under a certain promoter, such as *NND1* expressed with *ubi*, a model in which one can control the timing of gene expression with methods such as heat shock or the addition of tetracycline can allow for the development of a transgenic line as fish will live to breeding age, as well as expression of a transgene at specific developmental time points (Suster et al. 2009). In addition to the control of expression, this scenario can more closely model how a cancer like leukemia arises, which is generally from somatic mutations (Abdel-Wahab & Levine 2013). A zebrafish transgenic line with germline mutations may not model a cancer as well as a transgenic fish with an inducible transgene that acts more akin to a somatic mutation. Although the transgenic *cd45:NND1:P2A:sGFP* is not embryonic lethal, an inducible model for *cd45:NND1:P2A:sGFP* or *ubi:NND1:sGFP* could be considered in the future.

4.6.3 Off-target Expression May Confound Disease Phenotype

The transgenic *cd45:NND1:P2A:sGFP* showed effects on blood development in both embryos and adults, but also has off-target expression of GFP in the hindbrain and tail. Other transgenic fish with off-target expression demonstrated that the transgene was not expressed in these off-target areas, but that has not been verified in our model (Long et al. 1997). If our transgenic was found to express *NUP98-NSD1* in the hindbrain and tail, this could have consequences for development that may affect hematopoiesis distinct from the transgene being expressed in white blood cells. This may have even confounded the results of the RNA-seq, requiring a tissue specific approach to analyze effects on hematopoiesis, and preventing whole-embryo analysis. This fact was also relevant in the WISH experiments, where the *pu.1* probe was ectopically expressed in areas of off-target fluorescence, but has unknown significance at this time. This off-target expression may be eliminated through the creation of a new transgenic fish line under a different promoter, such as *pu.1*, or by the detection of another *cd45:NND1:P2A:sGFP* founder fish which may not have off-target expression.

4.7 Future Directions

4.7.1 Patients with a NUP98-NSD1 Translocation Often Harbour a FLT3-ITD Mutation

The association of *NUP98-NSD1* with *fms-like tyrosine kinase 3* with an internal tandem duplication (*FLT3*-ITD) has not only been found when examining AML patient samples, but this association extends further to define a prognostically relevant subgroup that has implications for survival and treatment (Akiki et al. 2013; Ostronoff et al. 2014). A large study done by Ostronoff and colleagues from the Children's Oncology Group (COG) showed that 82% of patients with *NUP98-NSD1* also harboured *FLT3*-ITD, and the presence of *FLT3*-ITD reduced the complete remission rate from 69% to 27% compared with expression of *NUP98-NSD1* alone, with the corresponding 3-year overall survival decreasing from 48% to 31% (Ostronoff et al. 2014). Thanasopoulou and colleagues looked at the biological implications of *NUP98-NSD1* and *FLT3*-ITD, alone and together, in a mouse model transplanted with cells that were retrovirally transduced to express NUP98-NSD1, FLT3-ITD, or both. Expression of both mutations increased cell proliferation and self-renewal *in vitro* and resulted in AML in the mice after a short latency, whereas cells expressing either NUP98-NSD1 or FLT3-ITD alone were unable to initiate AML in the mice (Thanasopoulou et al. 2014). Taking all of this into consideration, it is paramount that the transgenic fish expressing *NUP98-NSD1* are eventually either bred with a transgenic expressing *FLT3*-ITD, or are injected with a construct containing *FLT3*-ITD.

Lu and colleagues have already explored the effects of *FLT3*-ITD in a transgenic fish, and have published the model *spi1:FLT3ITD:eGFP* (Lu et al. 2016). In this model, they showed that zebrafish expressing *FLT3*-ITD under the *spi1* (*pu.1*) promoter had a greater number of myeloid cells and blast cells in adult fish, compared with a wild type control. Another group has also published that ubiquitous expression of *FLT3*-ITD in zebrafish embryos caused myeloid cell expansion and clustering (He et al. 2014). Given the clinical relevance of *NUP98-NSD1* and *FLT3*-ITD together in patients, our transgenic model should include *FLT3*-ITD, so that the effect on embryonic blood development and adult WKM and PB can be analyzed. As *FLT3*-ITD is classified as a type I aberration, causing cell proliferation, and *NUP98-NSD1* is classified as a type II aberration, impairing cellular differentiation, the combination of these genetic aberrations is likely

to lead to AML, and would be able to give us a lot of information on leukemogenesis and disease progression in patients, as well as provide a preclinical screening tool.

4.7.2 Downstream Use of the Transgenic cd45:NND1:P2A:sGFP as a Platform for Drug Discovery

Finding a blood cell phenotype in transgenic zebrafish embryos that differs from normal zebrafish is essential to the eventual use of a transgenic line for drug screening. Drugs that restore normal hematopoiesis in the transgenic embryos can only be found if there is an initial disruption in normal hematopoiesis. Our transgenic embryos display disruption of hematopoiesis through decreased red blood cells and myeloid cells, and increased HSCs when combined with *nup98* KO. The overarching goal of this project is to apply what we learn about *NUP98-NSD1* leukemogenesis to finding novel therapies for patients. The zebrafish provides a unique platform for performing high-throughput drug discovery experiments in an *in vivo* model. Drugs can be put directly in the water if water soluble, and the embryonic blood disruption occurs at early time points (between 24-48 hpf), allowing for dosing and analysis in a short timeframe. Down the road, this zebrafish model could be used to find targeted therapies for patients, most likely to be used in combination therapy with existing chemotherapeutics. As patients with *NUP98-NSD1*-induced AML are often children and adolescence with dismal survival rates, finding new therapy options that are effective and less toxic is paramount in the treatment of this high-risk AML.

4.8 Conclusions

In summary, this project has created a transgenic zebrafish to model high-risk pediatric AML. The transgenic *cd45:NND1:P2A:sGFP* shows disrupted blood development in embryos as well as adult fish, and overall reinforces the role of *NUP98-NSD1* as a type II aberration that impairs cellular differentiation. This project also explored the importance of decreased *nup98* in the context of *NUP98* fusions, and determined that in zebrafish embryos, both *NUP98-NSD1* and *nup98* knockout contributes to disease phenotype. Therefore, modelling a *nup98* KO should be taken into consideration when creating transgenic zebrafish with insertions of human *NUP98* fusion genes. This project highlights the technical feasibility in the zebrafish of applying

different genetic manipulations, such as transgenesis and CRISPR/Cas9 gene editing, as well as the effectiveness of using zebrafish to study hematopoiesis and leukemogenesis. This model will be used in the future to investigate the cellular pathways involved in *NUP98-NSD1*-induced AML, and as a platform to find drugs that are able to restore normal hematopoiesis in embryos, which could be used in patients harbouring the *NUP98-NSD1* transgene.

REFERENCES

Abdel-Wahab, O. & Levine, R.L., 2013. Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia. *Blood*, 121(18), pp.3563–72. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3643757&tool=pmcentrez&rendertype=abstract [Accessed October 29, 2014].

Ablain, J. et al., 2015. Resource A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish Resource A CRISPR / Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish. *Developmental Cell*, 32(6), pp.756–764. Available at: http://dx.doi.org/10.1016/j.devcel.2015.01.032.

Ablain, J. & Zon, L.I., 2016. *Tissue-specific gene targeting using CRISPR/Cas9*, Elsevier Ltd. Available at: http://dx.doi.org/10.1016/bs.mcb.2016.03.004.

Abrahamsson, J. et al., 2011. Response-guided induction therapy in pediatric acute myeloid leukemia with excellent remission rate. *Journal of Clinical Oncology*, 29(3), pp.310–315.

Ahuja, H.G. et al., 2000. t (9; 11)(p22; p15) in acute myeloid leukemia results in a fusion between NUP98 and the gene encoding transcriptional coactivators p52 and p75-lens epithelium-. *Cancer research*, 75(7), pp.6227–6229. Available at: http://cancerres.aacrjournals.org/content/60/22/6227.short.

Ahuja, H.G., Felix, C. a & Aplan, P.D., 1999. The t(11;20)(p15;q11) chromosomal translocation associated with therapy-related myelodysplastic syndrome results in an NUP98-TOP1 fusion. *Blood*, 94(9), pp.3258–3261.

Akiki, S. et al., 2013. NUP98-NSD1 Fusion in Association with FLT3 -ITD Mutation Identifies a Prognostically Relevant Subgroup of Pediatric Acute Myeloid Leukemia Patients Suitable for Monitoring by Real Time Quantitative PCR. *Genes Chromosomes and Cancer*, 52, pp.1053–1064.

Appelbaum, F.R. et al., 2006. Age and acute myeloid leukemia. , 107(9), pp.3481–3486.

Arai, Y. et al., 2000. Heterogeneous fusion transcripts involving the NUP98 gene and HOXD13 gene activation in a case of acute myeloid leukemia with the t (2 ; 11)(q31 ; p15) translocation. , (May), pp.1621–1629.

Arai, Y., Hosoda, F. & Kobayashi, H., 1997. The inv(11)(p15q22) chromosome translocation of de novo and therapy-related myeloid malignancies results in fusion of the nucleoporin gene, NUP98, with the putative RNA helicase gene, DDX10. *Blood*, 89(11), pp.3936–3944. Available at: http://bloodjournal.hematologylibrary.org/content/89/11/3936.short.

Argiropoulos, B. & Humphries, R.K., 2007. Hox genes in hematopoiesis and leukemogenesis. *Oncogene*, 26(47), pp.6766–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17934484 [Accessed May 23, 2013].

Avagyan, S. & Zon, L.I., 2016. Fish to Learn: Insights into Blood Development and Blood Disorders from Zebrafish Hematopoiesis. *Human gene therapy*, 27(4), pp.287–294.

Bai, X. et al., 2006. Trans-Repressive Effect of NUP98-PMX1 on PMX1-Regulated c- FOS Gene through Recruitment of Histone Deacetylase 1 by FG Repeats. *Cancer research*, 66(9), pp.4584–4591.

Balgobind, B. V et al., 2011. Integrative analysis of type-I and type-II aberrations underscores the genetic heterogeneity of pediatric acute myeloid leukemia. , 96(10), pp.1478–1487.

Barrera, M. et al., 2005. Educational and social late effects of childhood cancer and related clinical, personal and familial characteristics. *Cancer*, 104(8), pp.1751–1760.

Bennett, C.M. et al., 2001. Myelopoiesis in the zebrafish, Danio rerio. *Blood*, 98(3), pp.643–651.

Berdasco, M. et al., 2009. Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. *PNAS*, 106(51), pp.21830–21835.

Bertrand, J.Y. et al., 2008. CD41+ cmyb+ precursors colonize the zebrafish pronephros by a novel migration route to initiate adult hematopoiesis. *Development (Cambridge, England)*, 135(10), pp.1853–1862.

Bertrand, J.Y. et al., 2007. Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo. *Development*, 134(23), pp.4147–4156.

Bertrand, J.Y. et al., 2005. Three pathways to mature macrophages in the early mouse yolk sac. *Blood*, 106(9), pp.3004–11.

Bertrand, J.Y. & Traver, D., 2009. Hematopoietic cell development in the zebrafish embryo. *Current opinion in hematology*, 16(4), pp.243–248.

Borrow, J. et al., 1996. The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. *Nat Genet*, 12(2), pp.159–167. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8563754.

Brown, J. et al., 2002. A cryptic t(5;11)(q35;p15.5) in 2 children with acute myeloid leukemia with apparently normal karyotypes , identified by a multiplex fluorescence in situ hybridization telomere assay. *Blood*, 99(7), pp.2526–2531.

Burns, C.E. et al., 2005. Hematopoietic stem cell fate is established by the Notch – Runx pathway. *Genes & Development*, 19, pp.2331–2342.

Cantor, A.B. & Orkin, S.H., 2002. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene*, (21), pp.3368–3377.

Capelson, M. et al., 2010. Chromatin-Bound Nuclear Pore Components Regulate Gene Expression in Higher Eukaryotes. *Cell*, 140(3), pp.372–383. Available at: http://dx.doi.org/10.1016/j.cell.2009.12.054.

Carradice, D. & Lieschke, G.J., 2008. Zebrafish in hematology: sushi or science? *Blood*, 111(7), pp.3331–3343.

Ceasar, S.A. et al., 2016. Insert, remove or replace: A highly advanced genome editing system using CRISPR/Cas9. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1863(9), pp.2334–2344. Available at: http://dx.doi.org/10.1016/j.bbamcr.2016.06.009.

Cerveira, N. et al., 2003. Frequency of NUP98-NSD1 fusion transcript in childhood acute myeloid leukaemia. *Leukemia*, 2244(Figure 1), pp.2244–2247.

Chakraborty, C. et al., 2009. Zebrafish: A Complete Animal Model for In Vivo Drug Discovery and Development. *Current Drug Metabolism*, 10(2), pp.116–124. Available at: http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1389-2002&volume=10&issue=2&spage=116.

Chakraborty, P. et al., 2008. Nucleoporin Levels Regulate Cell Cycle Progression and Phase- Specific Gene Expression. *Dev Cell*, 15(5), pp.657–667.

Cheng, A.W. et al., 2013. Multiplexed activation of endogenous genes by CRISPR-on , an RNA-guided transcriptional activator system. *Nature Publishing Group*, 23(10), pp.1163–1171. Available at: http://dx.doi.org/10.1038/cr.2013.122.

Cheson, B.D. et al., 2003. Revised Recommendations of the International Working Group for diagnosis, standardization of response criteria, treatment outcomes, and reporting standards for therapeutic trials in acute myeloid leukemia. *Journal of Clinical Oncology*, 21(24), pp.4642–4649.

Clegg, A. et al., 2013. Frailty in elderly people. *The Lancet*, 381, pp.752–762.

Corkery, D.P., Dellaire, G. & Berman, J.N., 2011. Leukaemia xenotransplantation in zebrafish - chemotherapy response assay in vivo. *British Journal of Haematology*, 153, pp.786–789.

Cumano, A. & Godin, I., 2007. The Ontogeny of the Hematopoietic System. *Annual Review of Immunology*, 25, pp.745–85.

Das, C. et al., 2009. CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature*, 459(7243), pp.113–117. Available at: http://www.nature.com/doifinder/10.1038/nature07861.

Deschler, B. & Lübbert, M., 2006. Acute myeloid leukemia: Epidemiology and etiology. *Cancer*, 107(9), pp.2099–2107.

Deshpande, A.J. et al., 2014. AF10 Regulates Progressive H3K79 Methylation and HOX Gene Expression in Diverse AML Subtypes. *Cancer Cell*, pp.896–908.

Detrich, H.W. et al., 1995. Intraembryonic hematopoietic cell migration during vertebrate development. *Proceedings of the National Academy of Sciences of the United States of America*, 92(23), pp.10713–7. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40682&tool=pmcentrez&rendertype=abstract.

Deveau, A.P. et al., 2015. Epigenetic therapy restores normal hematopoiesis in a zebrafish model of NUP98–HOXA9-induced myeloid disease. *Leukemia*, 29(10), pp.2086–2097. Available at: http://www.nature.com/doifinder/10.1038/leu.2015.126.

Deveau, A.P., Bentley, V.L. & Berman, J.N., 2016. Using zebrafish models of leukemia to streamline drug screening and discovery. *Experimental Hematology*, 0(0), pp.3–15. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0301472X16306488.

Dobson, J.T. et al., 2008. Carboxypeptidase A5 identifies a novel mast cell lineage in the zebrafish providing new insight into mast cell fate determination. *Blood*, 112(7), pp.2969–2972.

Douglas, J. et al., 2003. NSD1 Mutations Are the Major Cause of Sotos Syndrome and Occur in Some Cases of Weaver Syndrome but Are Rare in Other Overgrowth Phenotypes. , pp.132–143.

Driever, W. et al., 1996. A genetic screen for mutations affecting embryogenesis in zebrafish. *Development*, 123(1), pp.37–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9007227.

Entz-Werle, N. et al., 2005. Results of 58872 and 58921 trials in acute myeloblastic leukemia and relative value of chemotherapy vs allogeneic bone marrow transplantation in first complete remission: the EORTC Children Leukemia Group report. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK*, 19(12), pp.2072–2081. Available at: papers3://publication/doi/10.1038/sj.leu.2403932.

Estey, E. & Döhner, H., 2006. Acute myeloid leukaemia. *Lancet*, 368, pp.1894–1907.

Estey, E.H., 2012. Acute myeloid leukemia: 2012 update on diagnosis, risk stratification, and management. *American Journal of Hematology*, 87(1), pp.89–99.

Fahrenkrog, B., 2014. Nucleoporin Gene Fusions and Hematopoietic Malignancies. *New Journal of Science*, 2014, pp.1–18. Available at: http://www.hindawi.com/journals/njos/2014/468306/.

Fasan, a et al., 2013. A rare but specific subset of adult AML patients can be defined by the cytogenetically cryptic NUP98-NSD1 fusion gene. *Leukemia*, 27(1), pp.245–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22945772 [Accessed November 13, 2014].

Feng, H. et al., 2007. Heat-shock induction of T-cell lymphoma/leukaemia in conditional Cre/lox-regulated transgenic zebrafish. *British Journal of Haematology*, 138, pp.169–175.

Finn, L. et al., 2015. Epidemiology of adult acute myeloid leukemia: Impact of exposures on clinical phenotypes and outcomes after therapy. *Cancer Epidemiology*, 39(6), pp.1084–1092. Available at: http://dx.doi.org/10.1016/j.canep.2015.09.003.

Fontoura, B.M.A., Blobel, G. & Matunis, M.J., 1999. A Conserved Biogenesis Pathway for Nucleoporins : Proteolytic Processing of a 186-Kilodalton Precursor Generates Nup98 and the Novel. , 144(6), pp.1097–1112.

Force, T., Krause, D.S. & Van Etten, R.A., 2007. Molecular mechanisms of cardiotoxicity of tyrosine kinase inhibition. *Nature Reviews Cancer*, 7(5), pp.332–344. Available at: http://www.nature.com/doifinder/10.1038/nrc2106.

Forrester, A.M., 2012. *Zebrafish Tools to Combat Acute Myeloid Leukaemia – the Nucleoporin 98 Kilodalton-Homeobox A9 (NUP98-HOXA9) Fusion Oncogene Drives Myeloproliferation by Upregulating DNA (Cytosine-5-)-Methyltransferase 1 (DNMT1)*. Dalhousie.

Forrester, a. M. et al., 2011. NUP98-HOXA9-transgenic zebrafish develop a myeloproliferative neoplasm and provide new insight into mechanisms of myeloid leukaemogenesis. *British Journal of Haematology*, 155(2), pp.167–181. Available at: http://doi.wiley.com/10.1111/j.1365-2141.2011.08810.x [Accessed November 19, 2014].

Franks, T.M. & Hetzer, M.W., 2013. The role of Nup98 in transcription regulation in healthy and diseased cells. *Trends in Cell Biology*, 23(3), pp.112–117.

Fu, Y. et al., 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, 31(9), pp.822–826. Available at: http://dx.doi.org/10.1038/nbt.2623.

Fujino, T. et al., 2002. Single-translocation and double-chimeric transcripts: Detection of NUP98-HOXA9 in myeloid leukemias with hOXA11 or HOXA13 breaks of the chromosomal translocation t(7;11)(p15;p15). *Blood*, 99(4), pp.1428–1433.

Fung, T.K. et al., 2010. Role of a novel zebrafish nup98 during embryonic development. *Experimental Hematology*, 38(11), p.1014–1021.e2. Available at: http://dx.doi.org/10.1016/j.exphem.2010.07.010.

Gaiano, N. et al., 1996. Highly efficient germ-line transmission of proviral insertions in zebrafish. *Genetics*, 93(July), pp.7777–7782.

Galloway, J.L. & Zon, L., 2003. Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. *Current Topics in Developmental Biology*, 53, pp.537–542.

Gary Gilliland, D. & Griffin, J.D., 2002. The roles of FLT3 in hematopoiesis and leukemia. *Blood*, 100(5), pp.1532–1542.

Gering, M. et al., 1998. The SCL gene specifies haemangioblast development from early mesoderm. *EMBO Journal*, 17(14), pp.4029–4045.

Gervais, C. et al., 2005. A new translocation t(9;11)(q34;p15) fuses NUP98 to a novel homeobox partner gene, PRRX2, in a therapy-related acute myeloid leukemia. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K*, 19, pp.145–148. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15496970.

Gilliland, D.G. & Tallman, M.S., 2002. Focus on acute leukemias. *Cancer Cell*, 1(5), pp.417–420.

Gorello, P. et al., 2008. t(3;11)(q12;p15)/NUP98-LOC348801 fusion transcript in acute myeloid leukemia. *Haematologica*, 93(9), pp.1398–1401.

Gough, S.M., Slape, C.I. & Aplan, P.D., 2011. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. *Blood*, 118(24), pp.6247–57. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3236115&tool=pmcentrez&rendertype=abstract [Accessed November 5, 2014].

Greenberg, P. et al., 1997. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*, 89(6), pp.2079–88. Available at: http://www.bloodjournal.org/content/89/6/2079.abstract.

Griffin, K.J. et al., 1998. Molecular identification of spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes. *Development (Cambridge, England)*, 125, pp.3379–3388.

Grimwade, D. & Freeman, S.D., 2014. Defining minimal residual disease in acute myeloid leukemia: which platforms are ready for "prime time"? *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program*, 2014(1), pp.222–33. Available at: http://www.asheducationbook.org/cgi/doi/10.1182/asheducation-2014.1.222%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/25696859.

Gurevich, R.M. et al., 2004. has potent leukemogenic activities independent of an engineered catalytic site mutation NUP98-Topoisomerase I acute myeloid leukemia – associated fusion gene has potent leukemogenic activities independent of an engineered catalytic site mutation. *Blood*, 104(4), pp.1127–1136.

He, B.-L. et al., 2014. Functions of flt3 in zebra fish hematopoiesis and its relevance to human acute myeloid leukemia. *Blood*, 123(16), pp.2518–2529.

Herbomel, P., Thisse, B. & Thisse, C., 1999. Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development (Cambridge, England)*, 126(17), pp.3735–45.

Hernandez-Garcia, C.M. & Finer, J.J., 2014. Identification and validation of promoters and cis-acting regulatory elements. *Plant Science*, 217–218, pp.109–119. Available at: http://dx.doi.org/10.1016/j.plantsci.2013.12.007.

Hollink, I.H.I.M. et al., 2011. NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. *Blood*, 118(13), pp.3645–56. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21813447 [Accessed November 12, 2014].

Homme, C. et al., 2010. Low SMC1A protein expression predicts poor survival in acute myeloid leukemia. *Oncology reports*, 24, pp.47–56.

Horvath, P. & Barrangou, R., 2010. REVIEW CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*, 327(January), pp.167–171.

Howe, K. et al., 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, 496(April), pp.498–504.

Huang, M. et al., 1988. Use of All-Trans Retionic Acid in the Treatment of Acute Promyelocytic Leukemia. , 72(2), pp.567–572.

Huang, N. et al., 1998. Two distinct nuclear receptor interaction domains in NSD1 , a novel SET protein that exhibits characteristics of both corepressors and coactivators. *The EMBO Journal*, 17(12), pp.3398–3412.

Hussey, D.J. et al., 1999. The (4;11)(q21;p15) translocation fuses the NUP98 and RAP1GDS1 genes and is recurrent in T-cell acute lymphocytic leukemia. *Blood*, 94(6), pp.2072–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10477737.

Hussey, D.J. & Dobrovic, A., 2002. Recurrent coiled-coil motifs in NUP98 fusion partners provide a clue to leukemogenesis. *Blood*, 99(3), pp.1097–1098.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Shengdar, Q., et al., 2013. Efficient In Vivo Genome Editing Using RNA-Guided Nucleases Woong. *Nature biotechnology*, 31(3), pp.227–229.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Kaini, P., et al., 2013. Heritable and Precise Zebrafish Genome Editing Using a CRISPR-Cas System. *PloS one*, 8(7), pp.1–9.

Ishikawa, M. et al., 2007. A novel gene, ANKRD28 on 3p25, is fused with NUP98 on 11p15 in a cryptic 3-way translocation of t(3;5;11)(p25;q35;p15) in an adult patient with myelodysplastic syndrome/acute myelogenous leukemia. *International Journal of Hematology*, 86, pp.238–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17988990.

Itoh, M. et al., 2003. Mind bomb is a ubiquitin ligase that is essential for efficient activation of notch signaling by delta. *Developmental Cell*, 4(1), pp.67–82.

Jagannathan-Bogdan, M. & Zon, L.I., 2013. Hematopoiesis. *Development*, 140(12), pp.2463–2467. Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.083147.

Jaju, R.J. et al., 1999. A new recurrent translocation, t(5;11)(q35;p15.5), associated with del(5q) in childhood acute myeloid leukemia. The UK Cancer Cytogenetics Group (UKCCG). *Blood*, 94(2), pp.773–780.

Jaju, R.J. et al., 2001. Brief report A novel gene , NSD1 , is fused to NUP98 in the t (5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. *Blood*, 98(4), pp.1264–1268.

Jin, H., Xu, J. & Wen, Z., 2007. Migratory path of definitive hematopoietic stem/progenitor cells during zebrafish development. *Blood*, 109(12), pp.5208–5214.

Kalev-Zylinska, M.L. et al., 2002. Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development (Cambridge, England)*, 129(8), pp.2015–30. Available at: http://dev.biologists.org/content/129/8/2015.abstract.

Kaltenbach, S. et al., 2010. NUP98-MLL fusion in human acute myeloblastic leukemia Brief report NUP98-MLL fusion in human acute myeloblastic leukemia. *Blood*, 116(13), pp.2332–2335.

Kalverda, B. et al., 2010. Nucleoporins Directly Stimulate Expression of Developmental and Cell-Cycle Genes Inside the Nucleoplasm. *Cell*, 140(3), pp.360–371. Available at: http://dx.doi.org/10.1016/j.cell.2010.01.011.

Kasper, L.H. et al., 1999. CREB Binding Protein Interacts with Nucleoporin-Specific FG Repeats That Activate Transcription and Mediate NUP98-HOXA9 Oncogenicity. *Molecular and Cellular Biology*, 19(1), pp.764–776.

Kawakami, K. et al., 1998. Excision of the Tol2 transposable element of the medaka fish Oryzias latipes in Xenopus laevis and Xenopus tropicalis. *Gene*, 225, pp.93–98.

Kawakami, K., 2007. Tol2: a versatile gene transfer vector in vertebrates. *Genome Biology*, 8(Suppl 1), p.S7. Available at: http://genomebiology.biomedcentral.com/articles/10.1186/gb-2007-8-s1-s7.

Kawakami, K., 2004. Transgenesis and Gene Trap Methods in Zebrafish by Using the Tol2 Transposable Element. *Methods in Cell Biology*, 77, pp.201–222.

Kawakami, K. & Shima, A., 1999. Identification of the Tol2 transposase of the medaka fish Oryzias latipes that catalyzes excision of a nonautonomous Tol2 element in zebrafish Danio rerio. *Gene*, 240(1), pp.239–244.

Kawakami, K., Shima, A. & Kawakami, N., 2000. Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. *Proceedings of the National Academy of Sciences*, 97(21), pp.11403–11408. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.97.21.11403.

Kimmel, C.B. et al., 1995. Stages of embryonic development of the zebrafish. *Developmental dynamics : an official public*, 203(3), pp.253–310.

Kivioja, J. et al., 2014. Identification of Novel Therapeutic Strategies for NUP98-NSD1-Positive AML By Drug Sensitivity Profiling. *Blood*, 124(21), p.2160 LP-2160. Available at: http://www.bloodjournal.org/content/124/21/2160.abstract.

Koga, A. et al., 1996. Transposable element in fish. *Nature*, 383(6595), pp.30–30. Available at: http://www.nature.com/doifinder/10.1038/383030a0.

Kohler, A. & Hurt, E., 2010. Gene Regulation by Nucleoporins and Links to Cancer. *Molecular Cell*, 38(1), pp.6–15.

Kok, F.O. et al., 2015. Reverse Genetic Screening Reveals Poor Correlation between Morpholino-Induced and Mutant Phenotypes in Zebrafish. *Developmental Cell*, 32(1), pp.97–108. Available at: http://dx.doi.org/10.1016/j.devcel.2014.11.018.

Konantz, M. et al., 2012. Zebrafish xenografts as a tool for in vivo studies on human cancer. *Annals of the New York Academy of Sciences*, 1266(1), pp.124–137.

Kouzarides, T., 2007. Chromatin modifications and their function. *Cell*, 128(4), pp.693–705. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17320507 [Accessed July 9, 2014].

Kroon, E. et al., 2001. NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. *The EMBO journal*, 20(3), pp.350–61. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=133485&tool=pmcentrez&rendertype=abstract.

Kurotaki, N. et al., 2002. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nature genetics*, 30(April), pp.365–366.

Lahortiga, I. et al., 2003. NUP98 is fused to Adducin 3 in a patient with T-cell acute lymphoblastic leukemia and myeloid markers, with a new translocation t(10;11)(q25;p15). *Cancer Research*, 63, pp.3079–3083.

Lam, D.H. & Aplan, P.D., 2001. NUP98 gene fusions in hematologic malignancies. *Leukemia*, 15(11), pp.1689–95. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11681408.

Langenau, D.M. et al., 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), pp.6068–6073.

Langenau, D.M. et al., 2003. Myc-Induced T Cell Leukemia in Transgenic Zebrafish. *Science*, 299(FEBRUARY), pp.887–891.

Lapunzina, P., 2005. Risk of Tumorigenesis in Overgrowth Syndromes : A Comprehensive Review. *American Journal of Medical Genetics Part C*, 71, pp.53–71.

Law, S.H.W. & Sargent, T.D., 2014. The Serine-Threonine Protein Kinase PAK4 Is Dispensable in Zebrafish : Identification of a Morpholino- Generated Pseudophenotype. *PloS one*, 9(6).

Lawson, N.D. & Weinstein, B.M., 2002. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental biology*, 248(2), pp.307–318.

Leith, B.C.P. et al., 1997. Acute Myeloid Leukemia in the Elderly: Assessment of Multidrug Resistance (MDR1) and Cytogenetics Distinguishes Biologic Subgroups With Remarkably Distinct Responses to Standard Chemotherapy. A Southwest Oncology Group Study. *Blood*, 89(9), pp.3323–3329.

Liang, Y. et al., 2013. Dynamic Association of NUP98 with the Human Genome. *PLoS Genetics*, 9(2).

Liao, E.C. et al., 1998. SCL/Tal-1 transcription factor acts dow nstream of cloche to specify hem atopoietic and vascular progenitors in zebrafish. *Research Communication*, 12, pp.621–626.

Lieschke, G.J. & Currie, P.D., 2007. Animal models of human disease: zebrafish swim into view. *Nature reviews. Genetics*, 8(5), pp.353–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17440532 [Accessed July 10, 2014].

Lin, S. et al., 1994. Integration and germ-line transmission of a pseudotyped retroviral vector in zebrafish. *Science*, 265, pp.666–669. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.8036514.

Lisboa, S. et al., 2013. POU1F1 is a novel fusion partner of NUP98 in acute myeloid leukemia with t(3;11)(p11;p15). *Molecular Cancer*, 12(5), pp.1–5.

Liu, F. & Wen, Z., 2002. Cloning and expression pattern of the lysozyme C gene in zebrafish. *Mechanisms of Development*, 113(1), pp.69–72.

Long, L. et al., 2015. Regulation of transcriptionally active genes via the catalytically inactive Cas9 in C. elegans and D. rerio. *Cell Research*, pp.1–4. Available at: http://www.nature.com/doifinder/10.1038/cr.2015.35.

Long, Q. et al., 1997. GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene. *Development*, 124, pp.4105–4111.

Lu, J.-W. et al., 2016. Overexpression of FLT3-ITD driven by spi-1 results in expanded myelopoiesis with leukemic phenotype in zebrafish. *Leukemia*, (May), pp.1–4. Available at: http://www.nature.com/doifinder/10.1038/leu.2016.132.

Mertens, F. et al., 2015. The emerging complexity of gene fusions in cancer. *Nature Reviews Cancer*, 15(6), pp.371–381. Available at: http://dx.doi.org/10.1038/nrc3947%5Cnpapers3://publication/doi/10.1038/nrc3947.

Murayama, E. et al., 2006. Tracing Hematopoietic Precursor Migration to Successive Hematopoietic Organs during Zebrafish Development. *Immunity*, 25(6), pp.963–975.

Nakamura, T. et al., 1996. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nature genetics*, 12(2), pp.154–158.

Nakamura, T. et al., 1999. NUP98 is fused to PMX1 homeobox gene in human acute myelogenous leukemia with chromosome translocation t(1;11)(q23;p15). *Blood*, 94(2), pp.741–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10397741.

Nakhoul, H. et al., 2014. Clinical Medicine Insights : Blood Disorders Ribosomopathies : Mechanisms of Disease. *Clinical Medicine Insights: Blood Disorders*, 7, pp.7–16.

Napper, A.D. & Watson, V.G., 2013. Targeted drug discovery for pediatric leukemia. *Frontiers in oncology*, 3(July), pp.1–13. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3703567&tool=pmcentrez&rendertype=abstract [Accessed November 6, 2014].

National Cancer Institute, 2014. SEER Cancer Statistics Review 1975-2011 National Cancer Institute SEER Cancer.

Nebral, K. et al., 2005. NUP98 is fused to topoisomerase (DNA) IIB 180 kDa (TOP2B) in a patient with acute myeloid leukemia with a new t(3;11)(p24;p15). *Clinical Cancer Research*, 11(18), pp.6489–6494.

Niewerth, D. et al., 2010. A review on allogeneic stem cell transplantation for newly diagnosed pediatric acute myeloid leukemia. *Blood*, 116(13), pp.2205–2214.

Nofrini, V., Giacomo, D. Di & Mecucci, C., 2016. Nucleoporin genes in human diseases. *European Journal of Human Genetics*, 24(10), pp.1388–1395. Available at: http://dx.doi.org/10.1038/ejhg.2016.25.

Ogryzko, V. V. et al., 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*, 87(5), pp.953–959.

Orkin, S.H. & Zon, L.I., 2008. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell*, 132(4), pp.631–644.

Ostronoff, F. et al., 2014. NUP98/NSD1 and FLT3/ITD coexpression is more prevalent in younger AML patients and leads to induction failure: a COG and SWOG report. *Blood*, 124(15), pp.2400–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25145343 [Accessed October 17, 2014].

Paik, E.J. & Zon, L.I., 2010. Hematopoietic development in the zebrafish. *International Journal of Developmental Biology*, 54, pp.1127–1137.

Palis, J. et al., 1999. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development*, 126, pp.5073–5084. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10529424.

Pan, Q. et al., 2008. A new fusion gene NUP98-IQCG identified in an acute T-lymphoid/myeloid leukemia with a t(3;11)(q29q13;p15)del(3)(q29) translocation. *Oncogene*, 27(24), pp.3414–23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18084320.

Panagopoulos, I. et al., 2006. Fusion of NUP98 and the SET binding protein 1 (SETBP1) gene in a paediatric acute T cell lymphoblastic leukaemia with t(11;18)(p15;q12). *British Journal of Haematology*, 136, pp.294–296.

Panagopoulos, I. et al., 2003. Fusion of the NUP98 gene and the homeobox gene HOXC13 in acute myeloid leukemia with t(11;12)(p15;q13). *Genes Chromosomes and Cancer*, 36(1), pp.107–112.

Panarello, C., Rosanda, C. & Morerio, C., 2002. Cryptic translocation t(5;11)(q35;p15.5) with involvement of the NSD1 and NUP98 genes without 5q deletion in childhood acute myeloid leukemia. *Genes Chromosomes and Cancer*, 35(3), pp.277–281.

Papaemmanuil, E. et al., 2013. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*, 122(22), pp.3616–3627.

Patton, E.E. & Zon, L.I., 2001. The art and design of genetic screens: Zebrafish. *Nature reviews. Genetics*, 2, pp.956–966. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11972155.

Perel, Y. et al., 2005. Treatment of childhood acute myeloblastic leukemia: dose intensification improves outcome and maintenance therapy is of no benefit--multicenter studies of the French LAME (Leucémie Aiguë Myéloblastique Enfant) Cooperative Group. *Leukemia*, 19, pp.2082–2089.

Peters, R., 2009. Translocation through the nuclear pore: Kaps pave the way. *BioEssays*, 31, pp.466–477.

Petit, A. et al., 2010. NUP98–HMGB3: a novel oncogenic fusion. *Leukemia*, 24, pp.654–658. Available at: http://www.nature.com/doifinder/10.1038/leu.2009.241.

Postlethwait, J.H. et al., 1994. A genetic linkage map for the zebrafish. *Science*, 264(5159), pp.699–703. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8171321.

Powers, M.A. et al., 1997. The Vertebrate GLFG Nucleoporin, Nup98, Is an Essential Component of Multiple RNA Export Pathways. *The Journal of Cell Biology*, 136(2), pp.241–250.

Prykhozhij, S. V. et al., 2017. A rapid and effective method for screening, sequencing and reporter verification of engineered frameshift mutations in zebrafish. *Disease Models & Mechanisms*, p.dmm.026765. Available at: http://dmm.biologists.org/lookup/doi/10.1242/dmm.026765.

Prykhozhij, S. V. et al., 2015. CRISPR multitargeter: A web tool to find common and unique CRISPR single guide RNA targets in a set of similar sequences. *PLoS ONE*, 10(3), pp.1–18.

Puumala, S.E. et al., 2013. Epidemiology of Childhood Acute Myeloid Leukemia. *Pediatric Blood Cancer*, 60(5), pp.728–733. Available at: http://journals.lww.com/jpho-online/Abstract/1995/05000/Epidemiology_of_Childhood_Acute_Myelogenous.2.aspx.

Qi, L.S. et al., 2013. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. , 152(5), pp.1173–1183.

Raices, M. & D'Angelo, M.A., 2012. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nature Reviews Molecular Cell Biology*, 13(November), pp.687–699.

Rajan, V., Dellaire, G. & Berman, J.N., 2016. *Zebrafish Methods and Protocols (Chapter 12)* Second. K. Kawakami, E. Patton, & M. Orger, eds., Humana Press.

Ran, F. et al., 2013. Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, 8(11), pp.2281–308. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24157548%5Cnhttp://www.nature.com/nprot/journal/v8/n11/abs/nprot.2013.143.html.

Ransom, D.G. et al., 1996. Characterization of zebrafish mutants with defects in embryonic hematopoiesis. *Development*, 123, pp.311–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9007251.

Rayasam, G.V. et al., 2003. NSD1 is essential for early post-implantation development and has a catalytically active SET domain. , 22(12), pp.3153–3163.

Raza-Egilmez, S.Z. et al., 1998. NUP98-HOXD13 Gene Fusion in Therapy-related Acute Myelogenous Leukemia Advances in Brief NUP98-HOXD13 Gene Fusion in Therapy-related Acute Myelogenous Leukemia1. *Cancer Research*, 58, pp.4269–4273.

Reader, J.C. et al., 2007. A novel NUP98-PHF23 fusion resulting from a cryptic translocation t(11;17)(p15;p13) in acute myeloid leukemia. *Leukemia*, 21(4), pp.842–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17287853.

Reeves, R., 2015. High mobility group (HMG) proteins: Modulators of chromatin structure and DNA repair in mammalian cells. *DNA Repair*, 36, pp.122–136.

Resar, L.M.S., 2010. The high mobility group A1 gene: transforming inflammatory signals into cancer? *Cancer Research*, 70(2), pp.436–439.

Rhodes, J. et al., 2005. Interplay of pu.1 and Gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Developmental Cell*, 8(1), pp.97–108.

Rio, M. et al., 2003. Spectrum of NSD1 mutations in Sotos and Weaver syndromes. *Journal of Medical Genetics*, 40, pp.436–440.

de Rooij, J., Zwaan, C. & van den Heuvel-Eibrink, M., 2015. Pediatric AML: From Biology to Clinical Management. *Journal of Clinical Medicine*, 4(1), pp.127–149. Available at: http://www.mdpi.com/2077-0383/4/1/127/.

Rosati, R. et al., 2002. NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). *Blood*, 99(10), pp.3857–3860.

Rossi, A. et al., 2015. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature*, Aug 13(524), pp.230–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26168398.

Roy, S. et al., 2013. HMGA1 overexpression correlates with relapse in childhood B-lineage acute lymphoblastic leukemia. *Leuk Lymphoma*, 54(11), pp.2565–2567. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23472968.

Rubnitz, J.E., Gibson, B. & Smith, F.O., 2010. Acute Myeloid Leukemia. *Hematology/Oncology Clinics of North America*, 24, pp.35–63. Available at: http://dx.doi.org/10.1016/j.hoc.2009.11.008.

Sander, J.D. et al., 2012. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nature Biotechnology*, 29(8), pp.697–698.

Schneider, R., Bannister, A.J. & Kouzarides, T., 2002. Unsafe SETs : histone lysine methyltransferases and cancer. *Trends in biochemical sciences*, 27(8), pp.396–402.

Schwartz, C.L., 1999. Long-term survivors of childhood cancer: the late effects of therapy. *The Oncologist*, 4, pp.45–54.

Scott, E.W. et al., 1994. Requirement of Transcription Factor PU . 1 in the Development of Multiple Hematopoietic Lineages Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2884565 Your use of the JSTOR archive indicates yo. *Science*, 265(5178), pp.1573–1577.

Shah, A.N. et al., 2015. Rapid Reverse Genetic Screening Using CRISPR in Zebrafish. *Nature Methods*, 12(6), pp.535–544. Available at: http://online.liebertpub.com/doi/10.1089/zeb.2015.29000.sha.

Shastri, A. et al., 2017. Stem and progenitor cell alterations in myelodysplastic syndromes. *Blood*, 129(12), pp.1586–1595.

Shiba, N. et al., 2013. NUP98-NSD1 gene fusion and its related gene expression signature are strongly associated with a poor prognosis in pediatric acute myeloid leukemia. *Genes, chromosomes & cancer*, 52(7), pp.683–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23630019 [Accessed November 7, 2014].

Slape, C. & Aplan, P.D., 2004. The Role of NUP98 Gene Fusions in Hematologic Malignancy. *Leukemia & Lymphoma*, 45(7), pp.1341–1350. Available at: http://informahealthcare.com/doi/abs/10.1080/10428190310001659325 [Accessed November 6, 2014].

Soler, G. et al., 2013. Identification of GSX2 and AF10 as NUP98 partner genes in myeloid malignancies. *Blood Cancer Journal*, 3, pp.e124-3. Available at: http://dx.doi.org/10.1038/bcj.2013.20.

Song, H.-D. et al., 2004. Hematopoietic gene expression profile in zebrafish kidney marrow. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), pp.16240–16245.

Stoletov, K. & Klemke, R., 2008. Catch of the day: zebrafish as a human cancer model. *Oncogene*, 27, pp.4509–4520. Available at: http://www.nature.com/doifinder/10.1038/onc.2008.95.

Strahl, B.D. et al., 2002. Set2 Is a Nucleosomal Histone H3-Selective Methyltransferase That Mediates Transcriptional Repression. , 22(5), pp.1298–1306.

Stuart, G.W., McMurray, J. V & Westerfield, M., 1988. Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. *Development*, 103, pp.403–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2852096.

Subramanian, M. et al., 2009. The nucleosome-binding protein HMGN2 modulates global genome repair. *FEBS Journal*, 276(22), pp.6646–6657.

Such, E. et al., 2011. A novel NUP98 / RARG gene fusion in acute myeloid leukemia resembling acute promyelocytic leukemia Brief report A novel NUP98 / RARG gene fusion in acute myeloid leukemia resembling acute promyelocytic leukemia. *Blood*, 117(1), pp.242–245.

Suster, M. et al., 2009. *Transgenesis in Zebrafish with the Tol2 Transposon System (Chapter 3)* vol. 561. E. J. Cartwright, ed., Humana Press.

Suster, M.L. et al., 2011. Transposon-mediated BAC transgenesis in zebrafish. *Nature Protocols*, 6(12), pp.1998–2021. Available at: http://dx.doi.org/10.1038/nprot.2011.416.

Taketani, T., Taki, T., Ono, R., et al., 2002. The chromosome translocation t(7;11)(p15;15) in acute myeloid leukemia results in fusion of the NUP98 gene with a HOXA cluster gene, HOXA13, but not HOXA9. *Genes Chromosomes and Cancer*, 34(4), pp.437–443.

Taketani, T., Taki, T. & Shibuya, N., 2002a. Novel NUP98-HOXC11 Fusion Gene Resulted from a Chromosomal Break within Exon 1 of HOXC11 in Acute Myeloid Leukemia with t (11 ; 12)(p15 ; q13) Advances in Brief Novel NUP98-HOXC11 Fusion Gene Resulted from a Chromosomal Break within. *Cancer Research*, 62, pp.4571–4574.

Taketani, T., Taki, T. & Shibuya, N., 2002b. The HOXD11 Gene Is Fused to the NUP98 Gene in Acute Myeloid Leukemia with t (2 ; 11)(q31 ; p15) Advances in Brief The HOXD11 Gene Is Fused to the NUP98 Gene in Acute Myeloid Leukemia with. *Cancer research*, 62, pp.33–37.

Tarlock, K. & Meshinchi, S., 2014. Pediatric Acute Myeloid Leukemia. *Pediatric Clinics of North America*. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0031395514001862 [Accessed November 13, 2014].

Thanasopoulou, A., Tzankov, A. & Schwaller, J., 2014. Potent co-operation between the NUP98-NSD1 fusion and the FLT3-ITD mutation in acute myeloid leukemia induction. *Haematologica*, 99(9), pp.1465–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24951466 [Accessed November 7, 2014].

Thol, F. et al., 2013. Analysis of NUP98/NSD1 translocations in adult AML and MDS patients. *Leukemia*, 27(3), pp.750–4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22929522 [Accessed November 7, 2014].

Tosi, S. et al., 2005. Characterization of 6q abnormalities in childhood acute myeloid leukemia and identification of a novel t(6;11)(q24.1;p15.5) resulting in a NUP98-C6orf80 fusion in a case of acute megakaryoblastic leukemia. *Genes Chromosomes and Cancer*, 44(3), pp.225–232.

Traver, D. et al., 2003. Transplantation and in vivo imaging of multilineage engraftment in zebrafish bloodless mutants. *Nature immunology*, 4(12), pp.1238–1246.

Trede, N.S. et al., 2004. The Use of Zebrafish to Understand Immunity. *Immunity*, 20(4), pp.367–379.

Tsukimoto, I. et al., 2009. Risk-stratified therapy and the intensive use of cytarabine improves the outcome in childhood acute myeloid leukemia: The AML99 trial from the Japanese childhood AML cooperative study group. *Journal of Clinical Oncology*, 27(24), pp.4007–4013.

Turkmen, S. et al., 2003. Mutations in NSD1 are responsible for Sotos syndrome , but are not a frequent finding in other overgrowth phenotypes. *European Journal of Human Genetics*, 11, pp.858–865.

Udvadia, A.J. & Linney, E., 2003. Windows into development: Historic, current, and future perspectives on transgenic zebrafish. *Developmental Biology*, 256(1), pp.1–17.

Varshney, G.K. et al., 2015. High-throughput gene targeting and phenotyping in zebrafish using CRISPR / Cas9. *Genome Research*, 25, pp.1030–1042.

Villefranc, J.A., Amigo, J. & Lawson, N.D., 2007. Gateway compatible vectors for analysis of gene function in the zebrafish. *Developmental Dynamics*, 236(11), pp.3077–3087.

Walde, S. & Kehlenbach, R.H., 2010. The Part and the Whole: functions of nucleoporins in nucleocytoplasmic transport. *Trends in Cell Biology*, 20(8), pp.461–469. Available at: http://dx.doi.org/10.1016/j.tcb.2010.05.001.

Walter, M.J. et al., 2012. Clonal Architecture of Secondary Acute Myeloid Leukemia. *New England Journal of Medicine*, 366(12), pp.1090–1098. Available at: http://www.nejm.org/doi/abs/10.1056/NEJMoa1106968.

Wang, G.G. et al., 2009. Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature*, 459(7248), pp.847–851.

Wang, G.G. et al., 2007. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nature cell biology*, 9(7), pp.804–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17589499 [Accessed November 7, 2014].

Weinstein, J.R. & Anderson, S., 2010. The aging kidney: physiological changes. *Advances in chronic kidney disease*, 17(4), pp.302–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20610357%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2901622.

Weir, H.K. et al., 2003. Annual Report to the Nation on the Status of Cancer, 1975-2000, Featuring the Uses of Surveillance Data for Cancer Prevention and Control. *Journal of the National Cancer Institute*, 95(17), pp.1276–1299. Available at: https://academic.oup.com/jnci/article-lookup/doi/10.1093/jnci/djg040.

White, R.M. et al., 2008. Transparent Adult Zebrafish as a Tool for In Vivo Transplantation Analysis. *Cell Stem Cell*, 2(2), pp.183–189.

Wiedenheft, B., Sternberg, S.H. & Doudna, J.A., 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, 46(482), pp.331–338.

Wittamer, V. et al., 2011. Characterization of the mononuclear phagocyte system in zebrafish. *Blood*, 117(26), pp.7126–7135.

Wu, X. et al., 2001. Disruption of the FG nucleoporin NUP98 causes selective changes in nuclear pore complex stoichiometry and function. *PNAS*, 98(6), pp.3191–3196.

Xiao, A. et al., 2013. Chromosomal deletions and inversions mediated by TALENs and CRISPR / Cas in zebrafish. *Nucleic acids research*, 41(14), pp.1–11.

Xie, Y. et al., 2003. Trends in leukemia incidence and survival in the United States (1973-1998). *Cancer*, 97(9), pp.2229–2235.

Xu, S. & Powers, M.A., 2010. Nup98-Homeodomain Fusions Interact with Endogenous Nup98 during Interphase and Localize to Kinetochores and Chromosome Arms during Mitosis. *Molecular biology of the cell*, 21(22), pp.4042–4056.

Yin, L. et al., 2015. Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs. *Genetics*, 200(June), pp.431–441.

Zon, L.I. & Peterson, R.T., 2005. In vivo drug discovery in the zebrafish. *Nature Reviews Drug Discovery*, 4(1), pp.35–44. Available at: http://www.nature.com/doifinder/10.1038/nrd1606.

van Zutven, L.J. et al., 2006. Identification of NUP98 Abnormalities in Acute Leukemia: JARID1A (12p13) as a New Partner Gene. *Genes, chromosomes & cancer*, 45, pp.437–446.

APPENDIX A: STUDY APPROVAL FOR *NUP98-NSD1* PROJECT

NOTICE OF PROTOCOL APPROVAL
UNIVERSITY COMMITTEE ON LABORATORY ANIMALS

Protocol Number:	**15-127**
Previous Protocol Number:	**Renewal**
Principal Investigator:	**Dr. Jason Berman**
Start Date/Expiry Expiry Date:	**December 1, 2016/December 1, 2017**
Title of Study:	**Using Zebrafish to Elucidate Novel Mechanisms and Therapeutics in NUP98-NSD1 Pediatric Acute Myeloid Leukemia (AML)**
Species:	**Zebrafish**

Jennifer Wipp
UCLA@dal.ca
University Committee on Laboratory Animals
902-494-1270
WEBSITE: http://www.dal.ca/dept/animal-ethics.html -

In compliance with granting agency and Dalhousie University policy, Dalhousie Research Services is not permitted to release funding instalments into research accounts until documentation of all necessary approvals are submitted (i.e. Human ethics, animal ethics, biohazard and radiation permits).

IMPORTANT FUNDING INFORMATION:

To ensure the research funds related to this protocol are released, fill out the information below and Scan and send to katie.merwin@dal.ca or fax the entire page to Dal Research Services 494-1595

For IWK funded projects fax to IWK Research Office - FAX 470-6767