Insights Into Inherited Thrombocytopenia Resulting From Mutations In Etv6 Or Runx1 Using A Human Pluripotent Stem Cell Model

Sara Borst
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Cell Biology Commons, and the Molecular Biology Commons

Recommended Citation
Borst, Sara, "Insights Into Inherited Thrombocytopenia Resulting From Mutations In Etv6 Or Runx1 Using A Human Pluripotent Stem Cell Model" (2021). Publicly Accessible Penn Dissertations. 3838. https://repository.upenn.edu/edissertations/3838

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/3838
For more information, please contact repository@pobox.upenn.edu.
Insights Into Inherited Thrombocytopenia Resulting From Mutations In Etv6 Or Runx1 Using A Human Pluripotent Stem Cell Model

Abstract
Inherited thrombocytopenia results in low platelet counts and increased bleeding. Subsets of these patients have monoallelic germline mutations in either ETV6 or RUNX1 and thus a heightened risk of developing hematologic malignancies. Patients with mutations in either of these transcription factors display the same phenotype of small megakaryocytes that give rise to fewer, less-functional platelets. Utilizing CRISPR/Cas9 technology, we compared and contrasted the in vitro phenotype of hematopoietic progenitor cells and megakaryocytes derived from induced pluripotent stem cell (iPSC) lines harboring mutations in either ETV6 or RUNX1. Both mutant lines display phenotypes consistent with a platelet-related bleeding disorder. Surprisingly, these cellular phenotypes were distinct, suggesting that the mechanisms driving the thrombocytopenia are different. The iPSCs harboring a mutation in ETV6 yield significantly more hematopoietic progenitor cells and megakaryocytes, but the megakaryocytes are immature and less responsive to agonist stimulation. On the contrary, iPSCs with a heterozygous mutation in RUNX1 yield significantly fewer hematopoietic progenitor cells and megakaryocytes, but the megakaryocytes are more responsive to agonist stimulation, though both mutant-MK populations have deficient proplatelet formation. Our work highlights that while patients harboring germline ETV6 or RUNX1 mutations have similar clinical phenotypes, the mechanisms by which these occurs are distinct. This work emphasizes the importance of defining the exact nature of a mutation in patients with a phenotypically similar disorder, as the disease pathology and therapeutic interventions may be different.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Cell & Molecular Biology

First Advisor
Wei Tong

Keywords
ETV6, iPSC, megakaryocyte, RUNX1, thrombocytopenia

Subject Categories
Cell Biology | Molecular Biology

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/3838
INSIGHTS INTO INHERITED THROMBOCYTOPENIA RESULTING FROM MUTATIONS IN ETV6 OR RUNX1 USING A HUMAN PLURIPOTENT STEM CELL MODEL

Sara Borst

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2021

Supervisor of Dissertation:

________________________________________

Paul Gadue, Associate Professor of Pathology and Laboratory Medicine

Graduate Group Chairperson:

________________________________________

Daniel Kessler, Associate Professor of Cell and Developmental Biology

Dissertation Committee:

Christopher Lengner, Associate Professor of Cell and Developmental Biology

Mortimer Poncz, Professor of Pediatrics

Nancy Speck, Professor of Cell and Developmental Biology

Wei Tong, Professor of Pediatrics, Committee Chair
INSIGHTS INTO INHERITED THROMBOCYTOPENIA RESULTING FROM MUTATIONS IN ETV6 OR RUNX1 USING A HUMAN PLURIPOTENT STEM CELL MODEL

COPYRIGHT

2021

Sara Borst
DEDICATION

I dedicate this degree of perseverance to my parents. You have been the driving force behind my success, and for that, I am forever grateful.
ACKNOWLEDGEMENTS

First and foremost, I must thank my thesis advisor, Paul, and my unofficial advisor, Debbie, for all their support and guidance during my time in the lab. You both strived to create and encourage a fun and loving work environment, whether it be through a potluck, happy hour or even axe throwing. I can honestly say that without the two of you, and the former and current French-Gadue lab members, I would not be writing this thesis today, tomorrow, or even five years from now. Thank you all for being there when I needed you most, science related or not. And of course, a special shout out to Somdutta, Alyssa and Jean Ann – you are amazing women, and I am so very grateful for the fun times we have had over the years. I am not only leaving the lab with my PhD, but also with some lifelong friends.

To my thesis committee members, thank you for asking tough questions and encouraging me to think about my project in different ways; your support has shaped my thesis to be what it is. I must give a special thank you to Morty – since meeting you at the Platelet Symposium in Boston, you have been an integral part of my time at Penn, even if it took months of me saying “Hi Morty!” for you to remember me. I’ve enjoyed our conversations about science and life, and I am especially grateful for your help in editing my manuscript.

In addition to my lab family, I found a home in DSRB. A special thank you to Steve, Mary, Meagan and Dan for making this program what it is – undoubtedly, the best CAMB subgroup to be a part of. From the lunches to journal club, we have always had good laughs. To my ride or die DSRBers Hannah, Chris and Sid, and our honorary member, Emmanuelle – I have so many amazing memories with you that I will forever cherish; from the Poconos to holiday parties and even art installations, I couldn’t have asked for better classmates or friends to be by my side on this bumpy six-year journey.
Family is what we make it, even if they have four legs and are covered in fur. To my Murphy and Olive – you have been my stress relievers (...and inducers) since I adopted you two years ago. Thank you both for all the cuddles, licks, laughs and love, especially this year. You always knew exactly what I needed, and I always knew exactly what you wanted – food. To my fur sister, Abby, thank you for being by my side these last 16 years; you will forever be my favorite pup.

Lastly, but most importantly, I thank my human family. Even though none of you understand the vast majority of what I do, I appreciate you asking questions and acting interested. To my parents, your unconditional love and support has undeniably made this journey a little easier. Thank you for always pushing me to see my potential, even when I lost sight. To Aunt Diane, Grandma and Grandpa – thank you for all your support throughout my schooling, all the way back to elementary school; you make me feel special and loved and that means a lot to me. To my sister, Theresa – I cannot imagine going through life without you, and this degree is no different. Thank you for always being there for me to vent. You are my best friend and I love you to pieces. To my brother-in-law Nick, thank you for always being so positive and encouraging. I appreciate all the conversations we’ve had over the years; I am always learning something new from you. And finally – Theresa and Nick, I must thank you for creating my two perfect nieces, Josephina and Giulia. Josephina, you have been the light of my life since the day you started growing in mommy’s belly, and your infectious smile and beautiful soul keep me going. You are the best big sister already and I cannot wait to see the girl you grow to be. Giulia, you may be new, but you are oh so loved; you were the star of 2020 and I know you’ll shine in all the years to come. Although I have the title of “Dr.” now, “Aunt” is still my favorite. <3
ABSTRACT

INSIGHTS INTO INHERITED THROMBOCYTOPENIA RESULTING FROM MUTATIONS IN ETV6 OR RUNX1 USING A HUMAN PLURIPOTENT STEM CELL MODEL

Sara Borst
Paul Gadue

Inherited thrombocytopenia results in low platelet counts and increased bleeding. Subsets of these patients have monoallelic germline mutations in either ETV6 or RUNX1 and thus a heightened risk of developing hematologic malignancies. Patients with mutations in either of these transcription factors display the same phenotype of small megakaryocytes that give rise to fewer, less-functional platelets. Utilizing CRISPR/Cas9 technology, we compared and contrasted the in vitro phenotype of hematopoietic progenitor cells and megakaryocytes derived from induced pluripotent stem cell (iPSC) lines harboring mutations in either ETV6 or RUNX1. Both mutant lines display phenotypes consistent with a platelet-related bleeding disorder. Surprisingly, these cellular phenotypes were distinct, suggesting that the mechanisms driving the thrombocytopenia are different. The iPSCs harboring a mutation in ETV6 yield significantly more hematopoietic progenitor cells and megakaryocytes, but the megakaryocytes are immature and less responsive to agonist stimulation. On the contrary, iPSCs with a heterozygous mutation in RUNX1 yield significantly fewer hematopoietic progenitor cells and megakaryocytes, but the megakaryocytes are more responsive to agonist stimulation, though both mutant-MK populations have deficient proplatelet formation. Our work highlights that while patients harboring germline ETV6 or RUNX1 mutations have similar clinical phenotypes, the mechanisms by which these occurs are distinct. This work emphasizes the importance of
defining the exact nature of a mutation in patients with a phenotypically similar disorder, as the disease pathology and therapeutic interventions may be different.
TABLE OF CONTENTS

DEDICATION ................................................................................................................... iii

ACKNOWLEDGEMENTS ................................................................................................. iv

ABSTRACT ..................................................................................................................... vi

TABLE OF CONTENTS ................................................................................................. viii

LIST OF TABLES ........................................................................................................... xii

LIST OF ILLUSTRATIONS .............................................................................................. xiii

LIST OF ABBREVIATIONS ............................................................................................... xiv

CHAPTER 1: Introduction and Overview ........................................................................ 1

1.1 Overview of Hematopoiesis ....................................................................................... 1

   1.1.1 Classical Hierarchy Model of Hematopoiesis ...................................................... 1

   1.1.2 Primitive and Definitive Hematopoiesis ........................................................... 5

   1.1.3 Primitive vs. Definitive MKs ........................................................................... 5

1.2 Definitive Megakaryopoiesis and Thrombopoiesis ................................................... 6

   1.2.1 Megakaryopoiesis and Cytokines .................................................................. 6

   1.2.2 Regulation of Megakaryopoiesis via Transcription Factors ......................... 7

   1.2.3 MK Maturation during Megakaryopoiesis ...................................................... 8

   1.2.4 Thrombopoiesis and Sites of Platelet Release ............................................. 9

1.3 Inherited Thrombocytopenia and Hematologic Malignancies ................................ 13

   1.3.1 Inherited Thrombocytopenia .......................................................................... 13

   1.3.2 Leukemia (ALL, AML, CLL and CML) .......................................................... 14

   1.3.3 Lymphoma ..................................................................................................... 16

   1.3.4 Myelodysplastic Syndrome ........................................................................... 17

   1.3.5 Multiple Myeloma .......................................................................................... 18

   1.3.6 Chromosomal Translocations ....................................................................... 18
1.4 ANKRD26-Related Thrombocytopenia

1.4.1 The Gene ANKRD26

1.4.2 Thrombocytopenia 2: MASTL, ACBD5 or ANKRD26?

1.4.3 Clinical Phenotype

1.4.4 Mouse Models

1.5 RUNX1-Related Thrombocytopenia

1.5.1 The Gene RUNX1

1.5.2 Clinical Phenotype

1.5.3 Mouse Models

1.6 ETV6-Related Thrombocytopenia

1.6.1 The Gene ETV6

1.6.2 Clinical Phenotype

1.6.3 Mouse Models

1.7 Pluripotent Stem Cells as a Model System for Hematopoietic Disorders

1.7.1 Human Pluripotent Stem Cells

1.7.2 Directed Differentiation of iPSCs to MKs

1.7.3 Pluripotent Stem Cells and their Applications in Hematologic Disease Modeling

1.8 Overview of Research Goals

CHAPTER 2: Materials and Methods

2.1 Patient-derived iPSC Line Generation

2.2 CHOPWT6 iPSC Line Generation

2.3 iPSC Culture

2.4 Differentiation of iPSCs into HPCs, MKs, Erythrocytes and Myeloid Cells

2.5 Western Blot

ix
CHAPTER 3: Insights into Inherited Thrombocytopenia Resulting from Mutations in ETV6 or RUNX1 using a Human Pluripotent Stem Cell Model

3.1 Introduction

3.2 Results

3.2.1 Generation and Characterization of Isogenic iPSC Lines

3.2.2 ETV6 and RUNX1 Mutations Reveal Disparate Effects on Blood Differentiation

3.2.3 Contrasting MK Phenotypes in ETV6- and RUNX1-Mutant iPSC Lines

3.2.4 ETV6-Mutant MKs are Less Responsive to Agonists than RUNX1-Mutant MKs

3.2.5 RNA-Sequencing Analysis of Purified MKs

3.2.6 Common Proplatelet Formation Defects in ETV6- and RUNX1-Mutant MKs
3.3 Discussion........................................................................................................73

CHAPTER 4: Summary and Future Directions.........................................................80

4.1 Defects in Platelet Generation and Function..................................................80

4.2 Mechanistic Differences between *ETV6*- and *RUNX1*-Related IT...............82

4.3 Link between Thrombocytopenia, Inflammation and Cancer.........................87

4.4 *ETV6*- and *RUNX1*-Mutant iPSCs for Drug Screening...............................90

BIBLIOGRAPHY ....................................................................................................92
LIST OF TABLES

CHAPTER 2

Table 2.1: CRISPR/Cas9 Reagents – gRNAs, ssDNA oligos, PCR and Sequencing Primers

Table 2.2: iPSC Lines Generated from Patient Material

Table 2.3: iPSC Lines Generated in the CHOPWT6 Genetic Background

Table 2.4: Antibodies used for Western Blot

Table 2.5: Antibodies used for Flow Cytometry

Table 2.6: qPCR Primers

Table 2.7: Antibody used for Immunofluorescence Microscopy

Table 2.8: Software

CHAPTER 3

Table 3.1: Genes Up- and Down-Regulated in Platelet Activation Pathway
LIST OF ILLUSTRATIONS

CHAPTER 1

Figure 1.1: Classical Hierarchy Model of Hematopoiesis
Figure 1.2: Revised Model of Hematopoiesis
Figure 1.3: Megakaryopoiesis
Figure 1.4: RUNX1 Transcript with Pathogenic Germline Mutations
Figure 1.5: ETV6 Transcript with Pathogenic Germline Mutations
Figure 1.6: Schematic of Adherent Primitive Hematopoietic Differentiation Protocol

CHAPTER 3

Figure 3.1: Characterization of CHOPWT6-Generated iPSC Lines
Figure 3.2: ETV6 and RUNX1 Mutations Reveal Disparate Lineage Potentials in CHOPWT6 Isogenic iPSC Lines
Figure 3.3: Disparate MK Phenotypes in ETV6- and RUNX1-Mutant CHOPWT6 iPSC Lines
Figure 3.4: ETV6-Mutant MKs are Less Responsive than RUNX1-Mutant MKs
Figure 3.5: RNA-Sequencing of Purified CHOPWT6 Isogenic MKs
Figure 3.6: Common Proplatelet Formation Defects in ETV6- and RUNX1-Mutant MKs
Figure 3.7: Model
Figure 3.8: Patient iPSC Line Characterization
Figure 3.9: ETV6 and RUNX1 Mutations Reveal Dis disparate Lineage Potentials in Patient-Derived iPSC Lines
Figure 3.10: ETV6-Mutant MKs are Fewer in Number and Less Responsive to Agonists than RUNX1-Mutant MKs in Patient-Derived iPSC Lines

CHAPTER 4

Figure 4.1: MAPK Modulation during MK Differentiation of CHOPWT6 Isogenic iPSC Lines
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACBD5</td>
<td>Acyl-CoA binding domain containing 5</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>AML</td>
<td>Acute myeloid leukemia</td>
</tr>
<tr>
<td>AML1</td>
<td>Acute myeloid leukemia 1</td>
</tr>
<tr>
<td>ANKRD26</td>
<td>Ankyrin repeat domain 26</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BCL-XL</td>
<td>B-cell lymphoma extra-large</td>
</tr>
<tr>
<td>bFGF</td>
<td>Basic fibroblast growth factor</td>
</tr>
<tr>
<td>BRG1</td>
<td>Brahma-related gene 1</td>
</tr>
<tr>
<td>CAS9</td>
<td>CRISPR-associated protein 9</td>
</tr>
<tr>
<td>CBF</td>
<td>Core-binding factor</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB-binding protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic lymphocytic leukemia</td>
</tr>
<tr>
<td>CLP</td>
<td>Common lymphoid progenitor</td>
</tr>
<tr>
<td>CML</td>
<td>Chronic myeloid leukemia</td>
</tr>
<tr>
<td>CMP</td>
<td>Common myeloid progenitor</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EPO</td>
<td>Erythropoietin</td>
</tr>
<tr>
<td>ESC</td>
<td>Embryonic stem cell</td>
</tr>
<tr>
<td>ETS</td>
<td>E26 transformation-specific</td>
</tr>
<tr>
<td>ETV6</td>
<td>ETS variant transcription factor 6</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence activated cell sorting</td>
</tr>
<tr>
<td>FLI1</td>
<td>Friend leukemia virus integration 1</td>
</tr>
<tr>
<td>FSC-A</td>
<td>Forward scatter</td>
</tr>
<tr>
<td>FV</td>
<td>Factor V</td>
</tr>
<tr>
<td>GATA1</td>
<td>GATA-binding protein 1</td>
</tr>
<tr>
<td>GATA2</td>
<td>GATA-binding protein 2</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>Granulocyte-macrophage colony-stimulating factor</td>
</tr>
<tr>
<td>GMP</td>
<td>Granulocyte-macrophage progenitor</td>
</tr>
<tr>
<td>gRNA</td>
<td>Guide RNA</td>
</tr>
<tr>
<td>hESC</td>
<td>Human embryonic stem cell</td>
</tr>
<tr>
<td>HL</td>
<td>Hodgkin lymphoma</td>
</tr>
<tr>
<td>HPC</td>
<td>Hematopoietic progenitor cell</td>
</tr>
<tr>
<td>HSC</td>
<td>Hematopoietic stem cell</td>
</tr>
<tr>
<td>IMS</td>
<td>Invaginated membrane system</td>
</tr>
<tr>
<td>ID</td>
<td>Inhibitory domain</td>
</tr>
<tr>
<td>iPSC</td>
<td>Induced pluripotent stem cell</td>
</tr>
<tr>
<td>IGF2BP3</td>
<td>Insulin-like growth factor 2 mRNA-binding protein 3</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>INI1</td>
<td>Integrase interactor 1</td>
</tr>
<tr>
<td>IT</td>
<td>Inherited thrombocytopenia</td>
</tr>
<tr>
<td>JAK2</td>
<td>Janus kinase 2</td>
</tr>
<tr>
<td>KLF1</td>
<td>Krüppel-like factor 1</td>
</tr>
<tr>
<td>LMPP</td>
<td>Lymphoid-primed multi-potential progenitor</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharides</td>
</tr>
</tbody>
</table>
MAPK Mitogen activated protein kinase
MASTL Microtubule associated serine/threonine kinase like
M-CSF Macrophage colony-stimulating factor
MDS Myelodysplastic syndrome
MEF Mouse embryonic fibroblast
MEP Megakaryocyte-erythroid progenitor
MIP-1α Macrophage inflammatory protein-1 alpha
miR Micro RNA
MK Megakaryocyte
MM Multiple myeloma
MOZ Monocytic leukemia zinc-finger protein
MPL Myeloproliferative leukemia protein
MPP Multipotent progenitor
MMP3 Matrix metalloproteinase-3
mRNA Messenger RNA
MYB Myb proto-oncogene protein
NET Neutrophil extracellular traps
NFE2 Nuclear factor erythroid 2
NF-κB Nuclear factor κB
NHL Non-hodgkin lymphoma
PCR Polymerase chain reaction
PF4 Platelet factor 4
PI3K Phosphoinositide 3-kinase
PNT Pointed domain
POTE Prostate, ovary, testis and placenta expressed
PSC Pluripotent stem cell
PU.1 Spi-1 proto-oncogene
qPCR Quantitative polymerase chain reaction
PBMC Peripheral blood mononuclear cells
RHD Runt-homology domain
RNA Ribonucleic acid
ROCK Rho-associated protein kinase
RUNX1 Runt-related transcription factor 1
S1P Sphingosine-1-phosphate
SCF Stem cell factor
SDF-1 Stromal derived factor 1
SFD Serum-free defined
SNP Single nucleotide polymorphism
STAT Signaling transducer and activator of transcription
SWI/SNF Switch/sucrose non-fermentable
TAD Transactivation domain
TAL1 T-cell acute lymphoblastic leukemia 1
TALEN Transcription activator-like effector nucleases
TEL1 Translocation-ETS-leukemia
TGFβ Transforming growth factor beta
THC2 Thrombocytopenia 2
THC5 Thrombocytopenia 5
TLR Toll-like receptor
TNFα Tumor necrosis factor-α
TPO Thrombopoietin
UTR Untranslated region
| **VEGF** | Vascular endothelial growth factor |
| **vWF**  | Von Willebrand factor |
| **WT**   | Wild type |
CHAPTER 1
Introduction and Overview

1.1 Overview of Hematopoiesis

Hematopoiesis is the lifelong process of continuously replenishing the body of blood cells necessary to meet every day demands. This process is dynamic and responds to changes in the body, such as during times of injury or infection.

1.1.1 Classical Hierarchy Model of Hematopoiesis

At the apex of the hematopoiesis lineage tree sits the hematopoietic stem cell (HSC). The HSC has two key characteristics: (1) self-renewal capacity and (2) multipotent differentiation potential. Through a series of complex molecular programs driving cell division and cell fate determination, these uncommitted HSCs give rise to more committed hematopoietic progenitor cells (HPCs) that in turn proliferate and further commit to give rise to all fully mature blood lineages (Morrison et al., 1995; Weissman, 2000; Zhu and Emerson, 2002) (Figure 1.1). Specifically, the long-term HSCs differentiate into short-term HSCs, and subsequently differentiate to a multipotent progenitor (MPP) stage where they lose self-renewal capabilities (Morrison and Weissman, 1994). A branch point after the MPP stage yields HPCs with divergent lineage potentials: common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs) (Akashi et al., 2000; Manz et al., 2002). There is a second branch point after CMPs, meaning they have potential to give rise to the bipotent megakaryocyte-erythroid progenitors (MEP) or granulocyte-macrophage progenitors (GMPs); MEPs give rise to megakaryocytes (MKs) and erythrocytes, while GMPs give rise to basophils, eosinophils, macrophages, monocytes and neutrophils (Debili et al., 1996; Pimkin et al., 2014; Tober et al., 2007). In support of the lymphatic
system, CLPs give rise to T-cells, B-cells and natural killer cells (Morrison et al., 1995; Zhu and Emerson, 2002).

Although the classical model of hematopoiesis suggests a well-organized, unidirectional flow of serial differentiation accompanied by progressive loss of pluripotency, recent evidence suggests that this is not necessarily the case (Bao et al., 2019). With advances in single cell technology and genetic model systems comes rebuttals to the classical hematopoietic hierarchy, a model that was largely developed based on cell surface marker expression (Cheng et al., 2020). Studies closely examining the HSCs and their downstream HPCs (MPP, CLP, CMP, GMP and MEP) have found these populations to be largely heterogeneous. In particular, single-cell expression analyses suggest a continuum of HPCs capable of generating myeloid and lymphoid cells, rather than lineage restricted HPCs downstream of the HSC, as the classical model depicts (Cabezas-Wallscheid et al., 2014; Karamitros et al., 2018; Pietras et al., 2015; Rodriguez-Fraticelli et al., 2018) (Figure 1.2). Moreover, many studies provide evidence for a MK-biased HSC with a spatially and functionally distinct bone marrow niche that gives rise to MKs directly (Grinenko et al., 2014; Nishikii et al., 2015; Pinho et al., 2018; Rodriguez-Fraticelli et al., 2018; Sanjuan-Pla et al., 2013; Shin et al., 2014; Yamamoto et al., 2013) (Figure 1.2).

This raises the question: are the MKs derived directly from HSCs functionally and/or molecularly distinct from MKs derived through progenitor intermediates? As technology continues to advance, further studies investigating how these processes are altered in the presence of stress, or in the context of disease, will be crucial to furthering our understanding of HSC and HPC biology in general, but also for potential therapeutic drug discovery.
Figure 1.1: Classical Hierarchy Model of Hematopoiesis

An illustration showing the step-wise hierarchy model of hematopoiesis with the HSC sitting at the top. This multipotent HSC population resides in the bone marrow and contains long-term and short-term HSCs. The quiescent HSCs can self-renew or differentiate into proliferative MPPs. These undergo further differentiation into proliferative CLPs or CMPs. The CLPs give rise to B, T and natural killer cells. The CMP can develop into GMPs and MEPs. GMPs give rise to neutrophils, basophils, eosinophils, monocytes and macrophages. MEPs give rise to erythrocytes and MKs, which produce platelets. Created with BioRender.com.
Figure 1.2 Revised Model of Hematopoiesis

An illustration of a revised model of hematopoiesis in which lineage commitment occurs on a continuum rather than in punctuated stages. MK-biased HSCs give rise to MK progenitors directly. LMPP = lymphoid-primed multi-potential progenitor. Adapted from Bao et. al., EMBO 2019. Created with BioRender.com.
1.1.2 Primitive and Definitive Hematopoiesis

Two distinct hematopoietic programs, designated primitive and definitive, occur during embryonic/fetal development, and both give rise to MKs. These programs can be distinguished by the subtype of globin expressed in erythroid cells, as well as the site of hematopoiesis (Palis, 2014). The first, primitive, wave of hematopoiesis develops from an extraembryonic mesoderm population in the yolk sac (Baron and Fraser, 2005; McGrath and Palis, 2005). These yolk sac progenitors give rise to nucleated erythrocytes expressing fetal ε and ζ globin, macrophages, and primitive MKs (Palis, 2014; Palis et al., 1999; Xu et al., 2001). Toward the end of the primitive wave, definitive erythromyeloid progenitors emerge from the yolk sac and colonize the fetal liver, ultimately acting as the main source of hematopoiesis before the emergence of HSCs (McGrath et al., 2015). Another definitive wave of hematopoiesis occurs in the aorta-gonads-mesonephros (AGM) region and is responsible for the generation of long-term HSCs having multilineage hematopoietic potential, as well as enucleated erythrocytes expressing β and α globin, myeloid cells, lymphocytes, and definitive MKs (Medvinsky and Dzierzak, 1998; Palis, 2014). After specification, these HSCs colonize the fetal liver before transitioning to the bone marrow, where they reside for the remainder of adult life (De Bruijn et al., 2000).

1.1.3 Primitive and Definitive MKs

Although MKs are generated during both the primitive and definitive waves of hematopoiesis, functional differences have been described (Bluteau et al., 2013; Tober et al., 2007b; Xu et al., 2001). Primitive MKs tend to be less proliferative and display a lower ploidy, with each MK releasing a relatively small number of platelets (Mattia et al., 2002; Potts et al., 2014). MKs generated through the definitive hematopoietic program release large numbers of higher functioning platelets ($\leq 10^{3-4}$ platelets per MK in the adult) (Mattia
et al., 2002; Vitrat et al., 1998). Although neonatal MKs are generated from the definitive program of hematopoiesis, the ploidy and functionality of these MKs does not peak until at least a year after birth (Liu and Sola-Visner, 2011). Distinct gene expression signatures between MKs derived from human embryonic stem cells (ESCs), fetal liver, neonate, and adult HPCs supports the finding that MK maturation progresses during ontogeny and results in increased ploidy and platelet release (Bluteau et al., 2013). The exact mechanism of maturation is unknown, but studies suggest insulin-like growth factor 2 mRNA-binding protein 3 (IGF2BP3) as the master switch driving immature fetal/neonatal to mature adult megakaryopoiesis (Elagib et al., 2017). Overall, a better understanding of the regulatory and functional changes that occur during MK ontogeny will be critical for understanding the molecular mechanism of disease disrupting this process.

### 1.2 Definitive Megakaryopoiesis and Thrombopoiesis

Megakaryopoiesis is the process by which HSCs give rise to MKs. Thrombopoiesis is the subsequent release of platelets by these MKs. These processes are tightly regulated by extracellular and intracellular factors that promote the expression or repression of various gene regulatory networks.

#### 1.2.1 Megakaryopoiesis and Cytokines

Despite the low frequency of MKs in the bone marrow (~0.01%), they are critical for proper hemostasis as their main function is platelet generation (Nakeff and Maat, 1974; Pease, 1956). However, secondary roles in the regulation and function of the HSC bone marrow niche have been suggested in recent years (Bruns et al., 2014; Zhao et al., 2014). The derivation of MKs from HSCs in the bone marrow is tightly regulated by several interleukin (IL) family cytokines secreted by innate and adaptive immune cells, such as IL-6 (Lotem
et al., 1989), IL-4 (Catani et al., 2001), IL-10 (Sosman et al., 2000) and IL-1β (Beaulieu et al., 2014). However, the most important cytokine driving megakaryopoiesis is the growth factor thrombopoietin (TPO) and its receptor, myeloproliferative leukemia protein (MPL) (Kaushansky et al., 1995; Lok et al., 1994). TPO is constitutively produced by the liver and predominantly acts to drive the development and early maturation of MKs from HSCs through a feedback loop: low circulating platelet levels lead to higher levels of free circulating TPO to act on the MPL-expressing HSCs, HPCs and MKs in the bone marrow. Overall, this regulatory loop promotes megakaryopoiesis in a MPL-dependent manner (Fielder et al., 1997; Kaushansky et al., 1995; Kuter and Rosenberg, 1995; Lok et al., 1994). Upon binding of TPO to MPL, Janus kinase 2 (JAK2) is phosphorylated (Miyakawa et al., 1995) and functions to trigger a signaling cascade by phosphorylating and activating its downstream target pathways, which include signaling transducer and activator of transcription (STAT) (Miyakawa et al., 1996), mitogen-activated protein kinase (MAPK) (Yamada et al., 1995) and phosphoinositide 3-kinase (PI3K) (Geddis et al., 2001) signal transduction pathways. Through activation of these and other pathways via extrinsic factors, the transcriptional program driving megakaryopoiesis can ensue.

1.2.2 Regulation of Megakaryopoiesis via Transcription Factors

In addition to cytokines and growth factors, transcription factors are necessary for gene regulation driving megakaryopoiesis and other cell fate decisions. Prior to challenging the established dogma of hematopoietic hierarchy, studies aimed to understand the transcription factors and gene expression profiles driving cell fate decisions at bifurcation points (i.e., MEP versus GMP, MK versus erythrocyte). In terms of megakaryopoiesis, antagonistic interplay between transcription factors was found to play a pivotal role in repressing non-MK genes, while simultaneously activating genes driving MK development.
Specifically, the expression levels of GATA-Binding Protein 1 (GATA1) and Spi-1 Proto-Onconege (PU.1) determine the fate of CMPs, as GATA1 is a driver of erythrocytes and MKs, while PU.1 is the master regulator of myeloid development; these transcription factors actively inhibit one another, and thus higher levels of GATA1 drive an MEP fate by repressing PU.1 and its downstream mediators (Arinobu et al., 2007; Rekhtman et al., 1999; Rhodes et al., 2005). This same antagonistic phenomenon was discovered for the subsequent cell fate decision of MK versus erythrocyte. To summarize, Krüppel-Like Factor 1 (KLF1/EKLF) and Friend Leukemia Virus Integration 1 (FLI1) act in opposition, with higher expression levels of FLI1 driving megakaryopoiesis and higher KLF1 levels driving erythropoiesis (Bouilloux et al., 2008).

In addition to these antagonistic transcription factor pairs, the fate of MEPs is also regulated by inhibition of Myb proto-oncogene protein (MYB), a red cell transcription factor, via microRNA (miR)-150 repressive activity (Edelstein et al., 2013). It is important to keep in mind that in addition to these pairs of transcription factors, commitment to the MK lineage is also coordinated by the time- and dose-dependent expression of various hematopoietic transcription factors such as Nuclear Factor Erythroid 2 (NFE2), GATA-Binding Protein 2 (GATA2), T-Cell Acute Lymphoblastic Leukemia 1 (TAL1/SCL) and Runt-Related Transcription Factor 1 (RUNX1) (Tijssen and Ghevaert, 2013).

1.2.3 MK Maturation during Megakaryopoiesis

Upon commitment to the MK lineage, cells begin to upregulate the early MK-specific cell surface marker CD41 (GPIIb) (Mitjavila-Garcia et al., 2002), followed by cell surface expression of CD42a (GPIX) and CD42b (GPIbo), both markers of a more mature MK (Debili et al., 1992). During megakaryopoiesis, dynamic gene expression aids in the
regulation of complex maturation steps necessary for MKs to reach terminal differentiation (Figure 1.3).

This intricate maturation process involves the generation of a polyploid cell (Mattia et al., 2002; Vitrat et al., 1998), the formation of an invaginated membrane system (IMS) (Nakao and Angrist, 1968; Radley and Haller, 1982; Schulze et al., 2006), and the synthesis and accumulation of platelet-specific granules and proteins (Handagama et al., 1987; Heijnen et al., 1998; Youssefian and Cramer, 2000). Although the exact role of acquiring multiple chromosome copies is unknown, the ploidy of MKs during ontogeny increases from 2-4N to up to 128N, suggesting that higher ploidy MKs are correlative with higher function in terms of platelet yield (Mattia et al., 2002). It is thought that through endomitosis – a variation of the cell cycle where chromosomes are replicated but mitosis fails to complete prior to nuclear and cytoplasmic division (Lordier et al., 2008) – MKs drastically increase their cytoplasmic volume to support the large quantities of newly synthesized mRNA and proteins which are packaged in the alpha and dense granules, and ultimately platelets (Heijnen et al., 1998; Youssefian and Cramer, 2000). As the cytoplasmic volume increases, an elaborate membrane system continuous with the plasma membrane is formed; this IMS remains in contact with the extracellular environment and serves as the membrane reservoir for platelets (Nakao and Angrist, 1968; Radley and Haller, 1982; Schulze et al., 2006). After maturation, MKs are equipped to release platelets into the circulation — a process termed thrombopoiesis (Deutsch and Tomer, 2006).

1.2.4 Thrombopoiesis and Sites of Platelet Release

Platelets, also called thrombocytes, are anucleate cellular fragments released from MKs. These cellular fragments play a critical role in hemostasis and thrombus formation, while
also mediating aspects of immunity, inflammation, and angiogenesis (Golebiewska and Poole, 2015; Jenne and Kubes, 2015; Morrell et al., 2014; Semple et al., 2011; Walsh et al., 2015; Ware et al., 2013). With no nucleus, platelets derive most of the RNA and protein required for their activation and function from pre-packaged alpha and dense granules distributed throughout the MK cytoplasm (Roth et al., 1989; Rowley et al., 2012). The exact mechanism of platelet release from MKs, or thrombopoiesis, is not well understood but there are two main models: (1) proplatelet formation and (2) MK rupturing.

The proplatelet formation model can occur in the bone marrow and in the lung capillary beds. In the bone marrow, mature MKs exit the osteoblastic niche and enter the vasculature niche, a process proposed to be dependent on a stromal derived factor 1 (SDF-1) chemotactic gradient. Once in the vasculature niche, MKs extend proplatelet extensions into the bone marrow sinusoids (Junt et al., 2007). Shear stress and high sphingosine-1-phosphate (S1P) levels in the blood are proposed mechanisms supporting the release of platelets into circulation (Zhang et al., 2012). The released pre- and proplatelets are further broken down to yield single platelet fragments. Studies in the mouse suggest that a significant portion of platelet release (~50%) occurs when MKs exit the bone marrow and enter lung capillary beds, where they release platelets into circulation via proplatelet extensions (Howell and Donahue, 1937; Lefrançais et al., 2017; Levine et al., 1993). Although current studies examining platelet release in the lung have suggested this organ as a major site of thrombopoiesis, it is debated whether or not this holds true in the human (Kaufman et al., 1965; Lefrançais et al., 2017; Levine et al., 1993).

Other data suggest MKs can rupture through an IL-1a-dependent pathway. This results in rapid release of many platelets into the bone marrow where they quickly enter the
circulation after acute injury or inflammation (Nishimura et al., 2015). With evidence for both models, it is possible to speculate that thrombopoiesis may occur differently depending on the physiological need of homeostasis versus acute injury. This again raises the possibility that megakaryopoiesis can result in MK subtypes with distinct functions, depending on the stimuli driving MK development.
Figure 1.3 Megakaryopoiesis and Thrombopoiesis

Figure 1.3: Megakaryopoiesis and Thrombopoiesis
Schematic representation of megakaryopoiesis (top) and thrombopoiesis (bottom). (A) As HSCs differentiate into (B) MK-biased HPCs, there are changes in gene and cell surface receptor expression to support normal MK maturation. (C) MKs first express CD41 (GPIIb) on their cell surface. (D) Upon further maturation, MKs begin to express the MK-specific receptors, CD42a (GPIX) and CD42b (GPIbα). (B-E) As MKs mature during megakaryopoiesis, there is an increase in ploidy, size, granules (alpha granules, dense granules) and granular content, as well as the formation of a complex IMS. (E) The fully mature MK migrates down the SDF-1 gradient from the osteoblastic niche to the vascular niche, (F) where it extends proplatelets which fragment to release pre-platelets into the bone marrow sinusoids. Pre-platelet fragments are further shaped by vascular shear forces to become platelets. Intact mature MKs also migrate out of the bone marrow and release platelets in the lung capillary beds through a similar process. (G) Mature MKs can also undergo IL-1α-dependent rupture mechanism to release large numbers of platelets rapidly. Created using BioRender.com.
1.3 Inherited Thrombocytopenia and Hematologic Malignancies

Blood cancers are a diverse group of disorders collectively characterized by abnormal proliferation of a malignant cell in the hematopoietic system. In 2016, a revision to the World Health Organization (WHO) classification system now includes a category of myeloid neoplasms that are attributed to “germline predispositions and pre-existing platelet disorders” (Arber et al., 2016).

1.3.1 Inherited Thrombocytopenia

Thrombocytopenia is a clinical diagnosis of low platelet counts that fall below the normal range of 150,000 – 400,000 platelets/µL (Johnson et al., 2016). This can result from a wide variety of causes such as infection, cancer, medications, and pregnancy (Stasi, 2012). However, some causes are hereditary, in which case, parents pass the genetic mutation leading to defects in megakaryopoiesis and/or thrombopoiesis on to their offspring.

Inherited thrombocytopenia (IT) is a collective group of genetic disorders characterized by abnormally low platelet counts with variability in platelet function, ultimately manifesting as a bleeding diathesis (Balduini et al., 2017). Given the complex regulatory network controlling megakaryopoiesis and thrombopoiesis, it is not surprising that mutations in a wide variety of hematopoietic genes lead to thrombocytopenia. In 1948, the first IT disorder, Bernard-Soulier Syndrome, was described to result from a mutation in the platelet receptor complex, GPIb-IX-V (CD42a/b/c/d) (Lanza, 2006). Since then, high throughput sequencing techniques have aided in the discovery of over 30 novel genes leading to IT (Johnson et al., 2016). Although bleeding is considered the main clinical complication for IT patients, a rare subset of these patients also share a high propensity
to develop hematologic malignancies such as acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS). These patients have been described as having mono-allelic germline mutations in either ankyrin repeat domain 26 (ANKRD26) (discovered in 2010), ETS variant transcription factor 6 (ETV6) (discovered in 2015) or RUNX1 (discovered in 1999) (Noris and Pecci, 2017). Although this subset of IT patients present with a mild to moderate bleeding diathesis, the main clinical concern is hematologic cancer predisposition.

1.3.2 Leukemia (ALL, AML, CLL and CML)

Leukemia is a blood cancer characterized by abnormal proliferation and clonal expansion of HSCs and HPCs in the bone marrow, ultimately manifesting in the outgrowth of abnormal white blood cells, called leukemia cells. Leukemia is classified into four major categories depending on the type of bone marrow blood cell that initiated the leukemic state, as well as the rate of onset and progression. Myelogenous leukemias develop in cells of the myeloid lineage, whereas lymphocytic leukemia involves bone marrow cells that become lymphocytes. If the leukemia progresses rapidly and is characterized by proliferation of an immature cell population, termed blasts, it is classified as acute. Chronic leukemias on the other hand, are not associated with blasts and progress slowly even without immediate treatment. The four main types of leukemia are (1) acute lymphoblastic leukemia (ALL), (2) acute myeloid leukemia (AML), (3) chronic lymphocytic leukemia (CLL) and (3) chronic myeloid leukemia (CML).

ALL is the most common form of childhood cancer, with 53% of cases occurring in patients under 20 years old (Ma et al., 1997). This leukemia is initiated in immature lymphocytes and is further divided into two subtypes depending on immunotyping of the lymphoblast:
B-cell ALL and T-cell ALL. B-cell ALL is the most common ALL subtype and accounts for ~88% of all childhood cases and ~75% of adult cases (Terwilliger and Abdul-Hay, 2017). T-cell ALL occurs in adults twice as often as in children (Terwilliger and Abdul-Hay, 2017). In approximately 50% of adult ALL cases, lymphoblasts quickly infiltrate lymphatic tissues and result in hepatomegaly (enlarged liver), splenomegaly (enlarged spleen) and/or lymphadenopathy (enlarged lymph nodes) (Davis et al., 2014).

AML accounts for ~80% of acute leukemia cases in the adult population, making it the most common type (Cornell and Palmer, 2012). The risk of AML is increased in people over the age of 65. Most often, there is accumulation of a non- or low-functioning myeloblasts that fail to fully differentiate and mature. These myeloblasts rapidly multiply and overwhelm the bone marrow, thereby reducing the available space for other hematopoietic lineages. Moreover, accumulation of myeloblasts in the bone marrow leads to a concurrent decrease in erythrocytes, platelets and terminally differentiated myeloid cells (i.e., neutrophils, macrophages, eosinophils, etc.), thus altering the composition of peripheral blood. This homeostatic imbalance causes many common symptoms of AML including fatigue, increased incidences of bruising and bleeding, fever and susceptibility to infection (Cornell and Palmer, 2012; Davis et al., 2014).

CLL accounts for nearly one third of all leukemias and is most commonly diagnosed in adults over the age of 55 years (Yee and O’Brien, 2006). CLL is characterized, and thereby diagnosed, by clonal expansion of at least 5,000 B lymphocytes/µL in the peripheral blood (Hallek et al., 2008). Interestingly, ~50% of patients with CLL are incidentally diagnosed when complete blood counts obtained for other reasons show evidence of leukocytosis – a marked increase in the production of white blood cells, typically >100,000 white blood
16 cells/µL (Savage et al., 1997; Yee and O’Brien, 2006). Unsurprisingly, roughly half of all CLL patients are asymptomatic prior to obtaining a diagnosis; this is likely attributed to the slow progression of disease (Yee and O’Brien, 2006).

CML, typically seen in adults, is the consequence of a specific pathogenic translocation event occurring in a single bone marrow myeloid cell. A translocation is an erroneous event where a piece of one chromosome breaks off during mitosis and attaches to a different chromosome; a “fusion gene” is formed if the translocation event joins two genes together, thus forming an abnormal gene with novel function (Lieber et al., 2003). In nearly all cases, CML is caused by a translocation event between chromosomes 9 and 22 (t(9;22)(q34;q11)), leading to the derivation of abnormal chromosome 22, better known as the Philadelphia chromosome, and the fusion gene, BCR-ABL1 (Zhou and Xu, 2015). This oncogene encodes a constitutively active tyrosine kinase that activates multiple signaling pathways causing uncontrolled cellular proliferation and CML progression. Although CML has the propensity to progress from a chronic phase to an accelerated phase to a blast crisis phase, long-term maintenance of the chronic phase is achieved through use of tyrosine kinase inhibitors (Slupianek et al., 2013).

1.3.3 Lymphoma

Lymphoma is a type of blood cancer beginning in the lymphocytes of the lymphatic system – a network of over 500 lymph nodes and organs connected by lymphatic vessels functioning to filter waste and toxins to help fight infection (Cueni and Detmar, 2008). The most common symptom in patients diagnosed with lymphoma is enlargement of one or more lymph nodes, which are biopsied to differentiate between the two main types of lymphoma: Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL).
HL is characterized by the presence of abnormal Reed-Sternberg cells; these large cells, usually derived from B lymphocytes, are often multinucleated and display an unusual morphology with an abnormal immunotype. HL commonly affects people between the ages of 15 and 24, as well as those over the age of 60 (Weber et al., 2003). HL is typically initiated in lymphocytes residing in lymph nodes of the neck, chest or underarms, and progresses in a predictable fashion by spreading to contiguous lymph nodes (Weber et al., 2003). Overall, the prognosis of patients diagnosed with HL is very good, with a five-year survival rate greater than 90%.

In comparison, NHL is characterized by the lack of these abnormal Reed-Sternberg cells and is further classified as one of over 60 subtypes. Unlike HL, NHL is initiated by lymphocytes in lymph nodes located throughout the body, making it more difficult to diagnose at an early stage ultimately leading to a poorer prognosis (Weber et al., 2003).

1.3.4 Myelodysplastic Syndrome

Myelodysplastic syndrome (MDS) refers to a group of clonal stem cell malignancies characterized by inefficient hematopoiesis, cytopenia and dysplasia in one or more myeloid cell lineage (Liew and Owen, 2011; Visconte et al., 2014). In these patients, the bone marrow is overactive and leads to the accumulation of blasts, or immature cells that fail to fully develop into mature blood cells. MDS is a sporadic disease mainly affecting the elderly population, with 70 years being the median age of diagnosis (Liew and Owen, 2011; Visconte et al., 2014). Over time, normal cells in the bone marrow decrease in number while the risk of blast outgrowth increases; MDS patients consequently have an increased risk of developing AML (Kanagal-Shamanna et al., 2017; Liew and Owen, 2011;
Visconte et al., 2014). Further studies following up on the pathogenesis of MDS may aide in the discovery of clinical or therapeutic approaches to prevent AML progression.

1.3.5 Multiple Myeloma

Multiple myeloma (MM), initiated in the bone marrow, is a blood cancer characterized by neoplastic proliferation of plasma cells, called myeloma cells (Kristinsson et al., 2007). Plasma cells are short-lived antibody-producing cells that generate antibodies closely mimicking the receptors of their precursor cell, the B-lymphocyte (Harada et al., 1993). Myeloma cells, like plasma cells, are capable of producing antibodies, but the products are not the same. Rather, myeloma cells generate large quantities of abnormal antibodies called “monoclonal M proteins.” Monoclonal M protein has been found to accumulate in the blood and urine of MM patients and may lead to kidney damage or failure. Although MM is a blood cancer, the primary effect of MM is bone loss and fracture. In MM, myeloma cells congregate to form masses in the bone marrow; these masses are believed to disrupt the structure of surrounding bone by secreting factors which inhibit normal bone repair and growth, thus increasing bone fragility (Hideshima et al., 2007).

1.3.6 Chromosomal Translocations

Chromosomal translocations are frequently associated with a variety of cancers, particularly hematologic malignancies, but the exact mechanisms driving these translocations remains poorly understood. However, mutations in DNA-repair pathways, along with the spatial positions of broken loci, homologous recombination, non-homologous end joining and fragile genomic sites, have been suggested to be factors regulating the production of non-random chromosomal translocations (Meaburn et al., 2007; Mitelman et al., 2007). For oncogenic chromosomal translocations, gene
rearrangements generally lead to cancer by either forming fusion proteins with new or altered activity, as is seen in CML, or by promoting oncogene activation via a new promotor or enhancer (Mitelman et al., 2007).

Double-strand breaks in DNA are necessary for chromosomal translocations to occur. These events are not entirely random and commonly occur in large, evolutionarily conserved genes (Bickmore and Teague, 2002), fragile genomic sites (Burrow et al., 2009), transcriptional start sites (Chiarle et al., 2011) and euchromatic regions (Obe et al., 2002). Given that euchromatin is a lightly packed form of DNA, double-strand breaks and the resultant chromosomal rearrangements can significantly alter normal transcriptional activity of these genes and the resultant proteins. Given that certain genes are more susceptible to translocation events, there are oncogenic chromosomal rearrangements that are frequently seen in various types of cancer.

In B-cell ALL, the most common structural chromosomal abnormality is the translocation t(12;21)(p13;q22) and accounts for ~25% of all pediatric cases (Golub et al., 1995; Romana et al., 1995). This translocation results in an ETV6-RUNX1 chimeric gene by fusing the 5’ region of ETV6 (exons 1-5) with nearly the entire coding region of RUNX1 (Golub et al., 1995; Romana et al., 1995). This fusion gene retains the dimerization and repressive domains of ETV6, along with the DNA-binding and transcriptional activation domains of RUNX1; therefore, it still retains the ability to bind RUNX1 consensus sequences but functions as a histone deacetylase-dependent repressor, thereby causing deregulation of RUNX1 target genes (Golub et al., 1995; Romana et al., 1995). Additionally, the normal transcriptional activity of ETV6 is suggested to be disrupted by this translocation event, as this fusion protein affects the activity of ETV6
heterodimerization partners (De Braekeleer et al., 2012; Zelent et al., 2004). Ultimately, the presence of this fusion gene leads to dysregulation of hematopoiesis by enhancing self-renewal of HPCs, particularly of the B-cell lineage, in mouse (Morrow et al., 2004) and human iPSC models (Böiers et al., 2018).

1.4 ANKRD26-Related Thrombocytopenia

1.4.1 The Gene ANKRD26

ANKRD26 is located on chromosome 10 and encodes a 192-kDa protein that is highly abundant in the hypothalamus, liver, adipose tissue, skeletal muscle and hematopoietic tissue (Raciti et al., 2011). ANKRD26 is an ancestral member of the prostate, ovary, testis and placenta expressed (POTE) gene family and contains N-terminal ankyrin repeats to function in protein-protein interactions, and spectrin repeats which help to coordinate cytoskeletal interactions (Hahn et al., 2006).

1.4.2 Thrombocytopenia 2: MASTL, ACBD5 or ANKRD26?

Thrombocytopenia 2 (THC2) is an autosomal dominant IT disorder characterized by mild to moderate thrombocytopenia, normal in vitro platelet aggregation despite reduced alpha-granules and mean platelet volume, and predisposition to leukemia (Dowton et al., 1985).

The THC2 locus was mapped to chromosome 10p11.2-12 and further characterized by two independent groups. Initially, evidence from these groups suggested that THC2 was caused by missense mutations in two different genes: microtubules associated serine/threonine kinase like (MASTL) and acyl-CoA binding domain containing 5 (ACBD5) (Drachman et al., 2000; Savoia et al., 1999). Following up on this, four unrelated families presenting with THC2 displayed single nucleotide polymorphisms (SNPs) in the 5'
untranslated region (UTR) of a novel gene, \textit{ANKRD26}, but lacked mutations in \textit{MASTL} and \textit{ACBD5} (Noris et al., 2011; Pippucci et al., 2011). Further examination of previously and newly diagnosed THC2 families confirmed that \textit{ANKRD26} SNPs located within a 19 base pair region of the 5’ UTR are pathogenic and lead to THC2 (Noris et al., 2011; Pippucci et al., 2011). Since this discovery, mechanistic studies found these 5’ UTR SNPs to disrupt binding of RUNX1 and FLI1 to the 5’ regulatory region of \textit{ANKRD26}, thereby preventing transcriptional repression of \textit{ANKRD26} (Bluteau et al., 2014). In the unaffected general population, ANKRD26 expression decreases during megakaryopoiesis; in THC2 patients, persistent expression of ANKRD26 induces hyperactivation of MAPK signaling, leading to stalled MK maturation and poor proplatelet formation (Bluteau et al., 2014). These findings highlight the fact that normal megakaryopoiesis depends on the temporal activation and repression of specific signaling pathways via transcription factor regulation.

1.4.3 \textit{Clinical Phenotype}

Patients with germline mono-allelic \textit{ANKRD26} mutations are classified as having the most frequent form of inherited thrombocytopenia termed THC2 (Noris et al., 2013, 2011; Pippucci et al., 2011) The most commonly described pathogenic SNPs and small deletions are located in the 5’ UTR (Noris and Pecci, 2017). These patients typically display increased cellularity of the bone marrow, along with higher numbers of small MKs with hypolobulated nuclei (Tsang et al., 2017). Their platelet counts vary but the thrombocytopenia is generally mild, with few cases of spontaneous bleeds reported (Noris et al., 2011). The mean platelet volume and platelet size is normal, while alpha-granule content and platelet aggregation has been found to vary between patients (Noris et al., 2011; Perez Botero et al., 2016; Zaninetti et al., 2017). One possible reason for the variable platelet aggregation is decreased expression of the collagen receptor, GPIa/IIa.
(CD49b) (Perez Botero et al., 2016). Additionally, a subset of patients display erythrocytosis or leukocytosis, while others have no dysplasia (Perez Botero et al., 2016; Tsang et al., 2017). Approximately 8% of THC2 patients with heterozygous mutations in ANKRD26 go on to develop myeloid malignancies, mainly MDS and AML, with a wide range of age of onset (Noris and Pecci, 2017). Given the heterogeneity in phenotypes relating to platelet structure and function, lineage dysplasia and cancer progression, genetic testing for a mutation in ANKRD26 remains the most reliable diagnosis for THC2. To date, ~45 families have been reported to have ANKRD26-related IT (Tsang et al., 2017).

1.4.4 Mouse Models

In the mouse, homozygous mutations in Ankrd26 led to severe obesity, as well as hyperphagia, insulin resistance and gigantism (Bera et al., 2008). Further studies examining the molecular mechanism provide evidence for enhanced adipogenesis mediated by elevated ERK and mTOR pathway activation (Fei et al., 2011). To date, studies utilizing Ankrd26-mutant mice have not commented on impaired megakaryopoiesis or thrombopoiesis – a common theme in this field, where the observed MK and platelet defects, and risk of myeloid dysplasia, differs between human and mouse.

1.5 RUNX1-Related Thrombocytopenia

1.5.1 The Gene RUNX1

RUNX1 (also known as acute myeloid leukemia 1 (AML1) and core-binding factor alpha (CBFα)), located on chromosome 21, was first identified in 1991 due to its involvement in the t(8;21) translocation in AML. Since then, countless studies have revealed RUNX1 to be the master transcriptional regulator of hematopoiesis. In vertebrates, RUNX1
expression is regulated by two distinct promoters, distal P1 and proximal P2 (Miyoshi et al., 1995). This dual promoter system, in combination with alternative splicing, results in three described isoforms of RUNX1: (1) RUNX1a – P2 promoter, (2) RUNX1b – P2 promoter and (3) RUNX1c – P1 promoter (Miyoshi et al., 1995). Despite being defined in the early 1990’s, the functional role of these various RUNX1 isoforms is highly controversial. However, a few independent studies suggest RUNX1c to be crucial for the specification and emergence of definitive HSCs, with sustained expression of this one isoform through the remainder of HSC life in the bone marrow niche (Bee et al., 2009; Challen and Goodell, 2010; Pozner et al., 2007; Sroczynska et al., 2009).

RUNX1 belongs to a larger RUNX gene family which includes RUNX2 and RUNX3. Although these three genes encode transcription factors with unique spatial-temporal and tissue-specific patterns of expression, they all share conserved motifs thus highlighting the important of these domains in the function of RUNX transcription factors (Levanon et al., 2001; Levanon and Groner, 2004) (Figure 1.4). RUNX proteins contain the conserved C-terminal transactivation domain (TAD) and the C-terminal VWRPY motif, also known as the inhibitory domain (ID) (Levanon et al., 1998). Together, these motifs aide in the transcriptional activation or repression of RUNX-target genes through recruitment of co-activators and co-repressors. However, RUNX proteins, termed the α-subunits, are poor transcriptional activators unless they are bound to the constitutively expressed β-subunit, core-binding factor subunit beta (CBFβ) (Bravo et al., 2001). Interaction of these two proteins is mediated through another highly-conserved motif in the RUNX gene family, the Runt-homology domain (RHD). Formation of the core binding factor (CBF) heterodimer allows for CBFβ to stabilize the interaction of the CBF with DNA, thereby increasing the
transcriptional activity of RUNX proteins (Tang et al., 2000). The RHD of all RUNX transcription factors recognizes the binding motif ‘TGT/cGGT’ (Wang and Speck, 1992).

Despite the weak transcriptional activity of RUNX1 in the absence of CBFβ, RUNX1 is a pioneer transcription factor that drives lineage specific chromatin opening of hematopoietic genes (Hoogenkamp et al., 2009; Lichtinger et al., 2010). Pioneer transcription factors are uniquely characterized by their ability to interact with condensed chromatin to facilitate the opening of closed chromatin sites; these heterochromatic regions of DNA are unable to be transcriptionally regulated without pioneer factors (Mayran et al., 2019; Zaret and Carroll, 2011). RUNX1 transcriptional activity is further regulated through the recruitment of chromatin remodelers characterized by their ability to induce histone modifications. When acting as a transcriptional activator, RUNX1 recruits co-activators, such as p300/CREB-binding protein (CBP) (Kitabayashi et al., 1998) and monocytic leukemia zinc-finger protein (MOZ) (Bristow and Shore, 2003), to the promoters of the RUNX1-target genes macrophage colony-stimulating factor (M-CSF) (Kitabayashi et al., 1998) and macrophage inflammatory protein-1 alpha (MIP-1α) (Bristow and Shore, 2003); these adaptor proteins have inherent histone acetyltransferase activity, thereby promoting chromatin remodeling and gene activation through histone acetylation of lysine residues (Kitabayashi et al., 1998). Additionally, RUNX1 recruits the switch/sucrose non-fermentable (SWI/SNF) chromatin remodeling complex whereby two subunits of this complex, brahma-related gene 1 (BRG1) and integrase interactor 1 (INI1), bind to the promoters of RUNX1-target genes; in an adenosine triphosphate (ATP)-dependent manner, BRG1 and INI1 add epigenetic modifications (H4K16ac and H3K4me2) to the promoters of RUNX1-target genes (IL-3 and GM-SCF) to facilitate transcriptional activation (Bakshi et al., 2010). In contrast, RUNX1 can also recruit co-repressors to aide in downregulation of target genes;
mSin3A, a co-repressor associated with histone deacetylase enzymatic activity, is recruited by RUNX1 to the promoter of p21, ultimately leading to downregulation of p21 (Lutterbach et al., 2000). In some cases, RUNX1 is able to repress or activate specific genes: for example, RUNX1 represses the expression of PU.1 in T-cells and MKs, while it induces expression of this same gene in myeloid and B-cells (Huang et al., 2011).

The wide range of transcriptional activity exerted by RUNX1 raises the question: how do acetyltransferases selectively bind to RUNX1 on specific genetic loci, and can the enzymatic activity of the RUNX1-acetyltransferase complex be regulated under certain conditions? One possible explanation is post-translational modifications. Phosphorylation, acetylation and methylation of RUNX1 can influence RUNX1 activity, the affinity of RUNX1 for DNA motifs and the stability of RUNX1/co-factor complexes (Wang et al., 2009). For example, phosphorylation of RUNX1 has been shown to enhance the transactivation activity of RUNX1 (Tanaka et al., 1996) and reduces its interaction with co-repressors, such as Sin3a (Imai et al., 2004) and histone deacetylases (Guo and Friedman, 2011). Moreover, p300-mediated acetylation of RUNX1 is associated with increased DNA binding and transcriptional activity (Yamaguchi et al., 2004), whereas methylation can influence RUNX1 partner interactions to either induce (Zhao et al., 2008) or repress (Vu et al., 2013) transcriptional activity. Collectively, these modifications regulating the expression and activity of RUNX1, in combination with the extensive transcriptional regulation of genes involved in cell differentiation, proliferation and lineage commitment imparted by RUNX1, highlight why mutations in this gene confer cancer susceptibility.
1.5.2 Clinical Phenotype

In 1999, mono-allelic germline mutations in *RUNX1* were discovered to cause an IT with predisposition for myeloid malignancy through linkage analysis to a region on chromosome 21q22; *RUNX1*-related IT has since been classified as a Familial Platelet Disorder with propensity for Acute Myeloid Leukemia (FPD/AML) (Song et al., 1999). The most common germline mutations in *RUNX1* involve the RHD domain, however mutations within the TAD have also been described. Most of the inherited *RUNX1* mutations are unique to each affected pedigree and frequently result from point mutations or small indels leading to missense, nonsense or frameshift changes in the protein, while few cases of large intragenic deletions and duplications have also been described (Jongmans et al., 2010; Latger-Cannard et al., 2016; Preudhomme et al., 2009; Song et al., 1999) (Figure 1.4). Depending on the nature of the mutation, the affected *RUNX1* allele confers a loss-of-function or dominant-negative protein; large deletions commonly result in RUNX1 haploinsufficiency due to loss-of-function, whereas missense mutations in the RHD, and C-terminal nonsense mutations, can have a dominant-negative effect (Michaud et al., 2002). Although secondary somatic mutations must be acquired for leukemogenesis, the risk of AML is shown to be correlative with total RUNX1 protein levels, thereby conferring a poorer prognosis in patients with dominant-negative mutations. *De novo* germline mutations in *RUNX1* have been reported in patients presenting with thrombocytopenia, but no family history of malignancy (Schmit et al., 2015).

FPD/AML is characterized as an autosomal dominant disorder with qualitative and quantitative platelet defects, and an increased risk of AML and other hematologic malignancies. However, clinical presentation in affected families shows high variability, even within the same pedigree. Patients typically display mild to moderate bleeding
tendencies from childhood, while others do not present with thrombocytopenia until the development of MDS/AML or the occurrence of a "triggering-event" such as trauma or surgery (Jongmans et al., 2010; Latger-Cannard et al., 2016; Yoshimi et al., 2016). In general, the platelet size, mean platelet volume and morphology is normal, though some platelets display a slight gray appearance due to reduced alpha-granule content (Schlegelberger and Heller, 2017). Additionally, most pedigrees display platelet dysfunction as seen by deficiency in dense granules, moderate alpha-granule deficiency, impaired activation of the fibrinogen receptor (GPIIb-IIIa) and defective platelet spreading (Latger-Cannard et al., 2016; Sun et al., 2004). Platelet aggregation is decreased in the presence of the common agonists: collagen, adenosine diphosphate (ADP) and epinephrine. The wide range of platelet abnormalities in FPD/AML patients highlights the critical role of RUNX1 in megakaryopoiesis and thrombopoiesis.

Furthermore, peripheral blood smears in a subset of FPD/AML patients presenting with myeloid neoplasm revealed abnormalities in platelet granulation, as well as red cell macrocytosis and cytoplasmic hypogranulation with abnormal nuclear segmentation in granulocytes (Kanagal-Shamanna et al., 2017). In FPD/AML patients, bone marrow biopsies typically reveal hypocellularity, though patients with myeloid neoplasms have a higher tendency of hypercellularity (Kanagal-Shamanna et al., 2017; Tsang et al., 2017). MK numbers in the bone marrow are heterogeneous amongst patients. Dysmorphic MKs however, are a consistent finding and make up ~10% of MKs in the majority of cases, prior to MDS or leukemic transformation (Kanagal-Shamanna et al., 2017; Schlegelberger and Heller, 2017). In patients with myeloid neoplasms, additional cytopenias are common, as is megakaryocytic dysplasia, dyserythropoiesis, and dysgranulopoiesis (Kanagal-Shamanna et al., 2017). These patients have a roughly 40% chance of developing
MDS/AML with the median age of onset being 33 years (Liew and Owen, 2011). Despite the classification of \textit{RUNX1}-related ITs as being “FPD/AML”, patients have been reported to develop other hematologic malignancies including T-cell ALL, hairy cell leukemia and chronic myelomonocytic leukemia (Galera et al., 2019).

1.5.3 \textit{Mouse Models}

Following the discovery of \textit{RUNX1} in 1991, genetic mouse models were quickly developed to gain insight into the role of this transcription factor in hematopoiesis and malignancy. \textit{RUNX1} is now known to be critical for the initiation of definitive hematopoiesis and HSC emergence. Unsurprisingly, Runx1-null mice are not viable, displaying embryonic lethality at E11.5-12.5 due to central nervous system hemorrhage, fetal liver anemia and failure to develop definitive hematopoiesis (Okuda et al., 1996; Wang et al., 1996). Although FPD/AML patients have heterozygous mutations, \textit{Runx1}^{+/-} mice do not serve as a good genetic model for this disorder as they only display slight thrombocytopenia with failure to develop leukemia, even with secondary mutations introduced (Sun and Downing, 2004).

As an alternative approach to avoid embryonic lethality, \textit{Runx1} was conditionally deleted in adult mice via \textit{Mx1-Cre} (Growney et al., 2005; Ichikawa et al., 2004). These mice displayed slightly higher numbers of HSCs in the bone marrow and transplantation studies showed successful reconstitution of all hematopoietic lineages upon transplantation of \textit{Runx1}-deficient HSCs into irradiated recipient mice. However, competitive repopulation studies indicate a significant reduction in the ability of \textit{Runx1}-deficient HSCs to compete with wild type HSCs during repopulation. Conditional \textit{Runx1} deletion had lineage-specific effects on B- and T-cell maturation with marked inhibition of CLP development; these findings providing evidence for the role of Runx1 during multiple stages of lymphocyte
differentiation and maturation. Mice with excised \textit{Runx1} also displayed reduced peripheral blood platelet counts credited to inefficient MK maturation, again highlighting the role of \textit{Runx1} in megakaryopoiesis and thrombopoiesis. Conditional \textit{Runx1} deletion had no apparent effect on the erythroid lineage. In contrast to the pronounced inhibition of the lymphoid lineage, excision of \textit{Runx1} did not inhibit maturation of the myeloid lineage. Conditional deletion of \textit{Runx1} did not lead to the development of leukemia, even upon the addition of secondary driver mutations. Overall, the current genetic mouse models for the mechanistic studies of FPD/AML are not sufficient.
Figure 1.4 RUNX1 Transcript with Pathogenic Germline Mutations

RUNX1 transcript with germline mutations highlighted. Frameshift mutations are highlighted in blue. Missense mutations are highlighted in green. Nonsense mutations are highlighted in red. The functional domains, RHD, TAD and ID, are shown. Adapted from Schlegelberger et al., *Semin. Hematol.* 2017. Created using BioRender.com.
1.6 *ETV6*-Related Thrombocytopenia

1.6.1 *The Gene ETV6*

*ETV6* (also known as translocation-ETS-leukemia (TEL1)) maps to chromosome 12p13 and encodes a 57kDa protein that mainly functions as a transcriptional repressor. *ETV6* is a member of the ETS family of transcription factors, characterized by the conserved C-terminal DNA-binding domain (ETS) that recognizes the core sequence ‘GGAA/T’ (Green et al., 2010) (Figure 1.5). The DNA-binding capacity of *ETV6* is strongly regulated by autoinhibitory mechanisms (Green et al., 2010). *ETV6* also contains a non-conserved N-terminal pointed domain (PNT) responsible for homodimerization and heterodimerization with other ETS family transcription factors, such as *ETV7* and FLI1 (De Braekeleer et al., 2012) (Figure 1.5). Therefore, mutations in *ETV6* that reduce DNA binding and nuclear localization will result in a dominant-negative effect of mutant *ETV6* over wild type *ETV6*.

In a luciferase reporter assay, deletion of the PNT domain caused a complete loss-of-function, suggesting that this domain is necessary for transcriptional repression (Lopez et al., 1999).

To exert transcriptional repression, *ETV6* must first localize to the nucleus. *ETV6* does not contain a consensus nuclear localization sequence, however experiments deleting various domains of *ETV6* suggest a requirement for residues 332-452 (Park et al., 2006). Upon localization to the nucleus, *ETV6* interacts with other proteins, such as Sin3A and NCOR, to function as co-repressor complexes with histone deacetylases (Chakrabarti and Nucifora, 1999; Guidez et al., 2000; Wang and Hiebert, 2001). This complex facilitates histone condensation and transcriptional repression by inhibiting transcription factor access to target gene promoters.
Although few target genes have been described for ETV6, strong evidence suggests a tumor suppressive role through direct regulation of at least two genes, matrix metalloproteinase-3 \((MMP3)\) and B-cell lymphoma extra-large \((BCL-XL)\): \(MMP3\) aids in cellular migration through remodeling of the extracellular matrix, while \(BCL-XL\) plays a crucial role in inhibiting apoptosis. In one study (Fenrick et al., 2000), \(ETV6\) overexpression inhibited the colony formation and growth of Ras-transformed 3T3 cells; additionally, when these 3T3 cells were injected into nude mice, Ras-transformed cells overexpressing \(ETV6\) prevented metastasis to the surround muscle tissue. Further analysis unveiled significant downregulation of \(MMP3\), along with occupancy of ETV6 on the \(MMP3\) promoter. Moreover, through chromatin immunoprecipitation assays, ETV6 expression was shown to deacetylate histone H3 on the \(MMP\) promoter. Following the discovery of \(MMP\) regulation via ETV6, another group sought to find apoptotic genes regulated by ETS family transcription factors (Irvin et al., 2003). Interestingly, \(BCL-XL\) is an anti-apoptotic gene with multiple ETS-factor binding sites within its promoter. Overexpression of ETV6 in 3T3 cells led to repression of a \(BCL-XL\) promoter-linked reporter gene and also decreased expression of endogenous \(BCL-XL\) mRNA and protein. Collectively, these studies provide strong evidence for ETV6 acting as a tumor suppressor. Although a few large microarray studies have found additional genes thought to be regulated by ETV6, the results are difficult to interpret due to the aberrant expression of ETV6 and the use of immortalized cell lines, such as HeLa cells (Boily et al., 2007). Further studies utilizing a more physiologic system, such as the use of human pluripotent stem cells, will be insightful in terms of discovering novel ETV6 target genes that may be misregulated in a disease setting.
1.6.2 Clinical Phenotype

Although ETV6 translocations are one of the most common gene rearrangements in AML, ALL, MDS, myeloproliferative neoplasms and T-cell lymphoma, with over 30 different described translocation partners, very few families harbor an inheritable ETV6 mutation (~22 pedigrees) (De Braekeleer et al., 2012). These patients with a mono-allelic germline mutation in ETV6 have a rare form of IT that is further categorized as Thrombocytopenia 5 (THC5). For the most part, germline mutations are clustered within the ETS domain (Figure 1.5).

THC5 is broadly considered a non-syndromic autosomal dominant form of IT with increased predisposition for hematologic malignancy. In the bone marrow, patients display increased frequency of small, hypolobulated MKs. Patients usually present with mild to moderate thrombocytopenia and display minimal overt bleeding phenotypes (Feurstein and Godley, 2017). Platelets are normal in size, although some clinicians report a reduced mean platelet volume while others indicate increased platelet diameter (Melazzini et al., 2016; Poggi et al., 2017). Additionally, platelet aggregation, platelet activation and cell surface expression of glycoproteins appear normal in the majority of patients, although some studies have suggested an impaired ability to spread on fibrinogen and decreased responsiveness to ADP (Melazzini et al., 2016). In most patients, hemoglobin and red cell mean corpuscular volume are within normal range, but mild dyserythropoiesis is observed (Melazzini et al., 2016). Patients also commonly display mild myelodysplasia with nuclear hypolobulation and hypogranulation of myeloid cells (Noetzli et al., 2015; Poggi et al., 2017; Topka et al., 2015).
THC5 patients with heterozygous germline ETV6 mutations have an incredibly high predisposition for hematologic malignancies, likely correlated with multilineage dysplasia. To summarize, ~20% of THC5 patients present with B-cell ALL during childhood, and another 30% are diagnosed with other hematologic malignancies – mainly MDS and AML – at a median age of 33 (Feurstein and Godley, 2017). Additionally, patients have been reported to develop non-hematologic cancers including colorectal adenocarcinoma, duodenal adenocarcinoma, skin cancers and renal cell carcinoma (Melazzini et al., 2016; Topka et al., 2015; Zhang et al., 2015).

1.6.3 Mouse Models

In addition to the proposed tumor suppressive role, ETV6 undoubtedly plays a critical role in the development and maintenance of the hematopoietic system. As evidence for this, Etv6–/– mice are not viable and die between embryonic day 10.5 and 11.5 due to defective yolk sac angiogenesis (Wang et al., 1997). Additional studies suggest a requirement for Etv6 in HSC survival, as Etv6–/– HSCs fail to repopulate transplanted wild type mice (Hock et al., 2004). Conditional deletion of Etv6 in B-cells or T-cells through a Cre-lox system does not affect the morphology or yield of these mature hematopoietic lineages, suggesting Etv6 is not required for the maintenance of all hematopoietic lineages (Hock et al., 2004). However, when Etv6 is conditionally deleted from erythrocytes and MKs via GATA1-Cre, the erythroid lineage is unaffected while MK yields are increased and platelet levels are decreased; this suggests a role for Etv6 in megakaryopoiesis, specifically MK maturation (Hock et al., 2004). Although mouse models have improved our understanding of Etv6 in normal hematopoiesis, Etv6–/– mice do not display a hematopoietic phenotype, and conditional deletion of Etv6 does not cause thrombocytopenia or leukemia. Therefore, current mouse models are not adequate to study the pathogenesis of ETV6-related IT.
Figure 1.5 ETV6 Transcript with Pathogenic Germline Mutations

ETV6 transcript with positions of germline mutations. The oligomerization PNT domain and DNA-binding ETS domain are shown. Created using BioRender.com.
1.7 Pluripotent Stem Cells as a Model System for Hematopoietic Disorders

1.7.1 Human Pluripotent Stem Cells

Human pluripotent stem cells (PSCs) are defined as undifferentiated cells with the ability to self-renew indefinitely and differentiate into derivatives of all three germ layers: endoderm, mesoderm and ectoderm. The first human ESC line, described in 1998, was derived from the inner cell mass of a blastocyst. For a long time, the scientific community believed that cell differentiation was a one-way street with irreversible commitment to a specific lineage upon reaching a certain point in the differentiation pathway. However, in 2006, the Yamanaka group proved successful reprogramming of a fully differentiated mouse fibroblast back to its pluripotent state via retroviral induction with four reprogramming factors: Oct3/4, Sox2, Klf4 and c-Myc (Takahashi and Yamanaka, 2006). These four transcription factors caused endogenous pluripotency genes to be re-expressed thus converting the cell back to its pluripotent state. In 2007, Yamanaka showed that this same reprogramming method can be applied to human fibroblasts, thus creating human iPSCs (Takahashi et al., 2007). Similar to human ESCs, human iPSCs have the ability to self-renew and give rise to all somatic cell types.

Although the discovery of human iPSCs has revolutionized the stem cell field, previous studies on somatic cell nuclear transfer and myosin-D provided scientific discoveries that together contributed to the development of iPSC technology. In 1962, the nucleus from an intestinal epithelial cell in a tadpole was transferred to an unfertilized egg and resulted in successful development into a tadpole. This study, which details the first reprogramming event in somatic cells, suggested that the nuclear status of a differentiated cell can be reverted back to a totipotent state by cytoplasmic factors in an oocyte (Gurdon, 1962). In 1989, expression of the transcription factor myosin-D was shown to convert the fate of
multiple cell lines to muscle cells, alluding to the possibility of cell fate changes upon aberrant expression of a transcription factor (Weintraub et al., 1989).

The discovery of human iPSCs has transformed the field of science as it provides potential for studying pluripotency without the ethical concerns and regulations associated with the derivation and use of human ESCs. Moreover, iPSC technology allows for a human based model system where patient material, such as fibroblasts or peripheral blood, can be reprogrammed to generate patient specific stem cells. In addition to disease modeling, iPSC technology has opened the doors for more scientific breakthroughs in regenerative medicine, cell replacement therapy, drug testing and targeted gene repair strategies for genetic disorders. Overall, human iPSCs are the ideal model system for scientific studies of human development.

1.7.2 Directed Differentiation of iPSCs to MKs

The use of primary human MKs to study megakaryopoiesis and thrombopoiesis is difficult because these cells represent only 0.01% of the nucleated cells in the bone marrow. Human stem cell sources, including ESCs and iPSCs, have been used successfully for the in vitro generation of MKs (Figure 1.6). However, an ongoing challenge is identifying conditions to optimize the quantity and quality of in vitro-generated MKs and released platelets. Directed differentiation protocols have been developed for generating HPCs from ESCs/iPSCs with the overall design mimicking embryogenesis (Murry and Keller, 2008). Primitive streak formation is induced through BMP4 and WNT pathway activation, followed by mesoderm specification (Mills et al., 2014; Murry and Keller, 2008). Mesoderm is further specified to hematoendothelial mesoderm, through a still not fully understood mechanism, which ultimately gives rise to HPCs. HPCs expand upon the addition of
hematopoietic cytokines, such as SCF, Flt3L, and TPO (Gori et al., 2015; Mills et al., 2014; Zambidis et al., 2008). The primitive HPCs co-express typical adult MK and erythrocyte markers, CD41 and CD235, respectively, and give rise to primitive erythrocytes, primitive MKs, and myeloid cells (Mills et al., 2014; Paluru et al., 2014). Other widely used differentiation systems use co-culture with stromal cell lines, such as OP9, to generate HPCs (Vodyanik et al., 2006, 2005). These static in vitro cultures generate ~7 platelets/MK.

In recent years, differentiation protocols driving definitive hematopoiesis have been developed through timed modulation of the transforming growth factor beta (TGF-β) and WNT pathways (Ditadi et al., 2015; Ng et al., 2016; Sturgeon et al., 2014). Hemogenic endothelium can be purified from 3-dimensional embryoid body suspension cultures via cell sorting CD34+ cells that are CD73-CD184- (Ditadi et al., 2015; Sturgeon et al., 2014). This population possesses the ability to give rise to hematopoietic progenitors with lymphoid potential, confirming its definitive nature (Kennedy et al., 2012). With the identification of TPO, HPCs can be differentiated in vitro to the MK lineage (Lok et al., 1994). Studies comparing definitive, or adult-like, MKs derived from bone marrow CD34+ stem cells to primitive MKs derived from iPSCs show that the iPSC-derived primitive MKs are smaller, of lower ploidy, and release fewer platelets with a short half-life when infused into mice (Bluteau et al., 2013; Potts et al., 2014). With the newer protocols driving definitive hematopoiesis, further studies will define the characteristics of MKs derived from these more mature progenitor populations.
Figure 1.6: Schematic of Adherent Primitive Hematopoietic Differentiation Protocol

An illustration of the hematopoietic differentiation protocol. iPSCs are first differentiated towards mesoderm through use of regulatory cytokines, BMP4 and VEGF. Upon mesoderm induction, cultures are feed with hematopoietic specific cytokines, SCF and Flt3, to promote development and release of HPCs from the adherent layer into the medium. These HPCs can be differentiated to the MK, erythroid or myeloid lineages through addition of lineage-specific cytokines. Created using BioRender.com.
1.7.3 Pluripotent Stem Cells and their Applications in Hematopoietic Disease Modeling

Although animal models are invaluable for elucidating the underlying mechanism of human diseases, there are instances where they do not recapitulate the human disease phenotype. Human iPSC technology is an invaluable tool for the modeling of human diseases, especially in terms of rare disease where access to patient material is limited, as it allows for the unlimited supply of undifferentiated patient-derived iPSCs to be frozen down and stored long term. Detailed below are a set of MK and platelet disorders that have been studied using human iPSCs.

FPD/AML is a rare genetic disorder characterized by pathogenic $RUNX1$ mutations, thrombocytopenia and cancer predisposition. Since it was first described in 1999, multiple groups have leveraged iPSC technology to successfully model this disease (Breton-Gorius et al., 1995; Connelly et al., 2014; Iizuka et al., 2015; Sakurai et al., 2014). iPSCs were derived from FPD/AML patients and differentiated towards the MK lineage. In all cases, a severe decrease in MK yield was described, consistent with the patient phenotype. Some studies provided evidence for a deficit in HPC generation (Connelly et al., 2014; Iizuka et al., 2015; Sakurai et al., 2014). Lentiviral expression of wild type $RUNX1$, or gene correction via transcription activator-like effector nucleases (TALEN) gene editing, rescued the MK phenotype, suggestive of these mutations leading to $RUNX1$ haploinsufficiency.

Paris-Trousseau syndrome is caused by a large heterozygous deletion in chromosome 11q resulting in hemizygous expression of $FLI1$ — a critical transcription factor for megakaryopoiesis (Breton-Gorius et al., 1995). Patients present with macrothrombocytopenia, a condition of large platelets circulating at lower numbers. $FLI1$
mutations have recently been ascribed to cause thrombocytopenia, but the link between FLI1 mutations and Paris-Trousseau syndrome had not been confirmed (Stevenson et al., 2015; Stockley et al., 2013). Patient-derived iPSCs were differentiated to the MK lineage, where phenotypes were consistent with clinical presentation of Paris-Trousseau syndrome. Specifically, the FLI1-mutant iPSCs displayed impaired megakaryopoiesis resulting in decreased MK and platelet yields compared to wild type iPSCs. Upon transfusion of these MKs into mice, few FLI1-mutant platelets were released into the circulation, where they were functionally defective and had a reduced half-life (Vo et al., 2017).

Glanzmann thrombasthenia is a rare autosomal recessive disease caused by mutations in ITGA2B and ITGB3 that result in lack of GPIIb/IIIa (CD41/CD61) (Coller and Shattil, 2008). Although patients with Glanzmann thrombasthenia have normal platelet counts, they present with a bleeding diathesis due to the role of αIIbβ3 in fibrinogen binding and platelet activation (Massberg et al., 2005). The iPSC lines generated from different Glanzmann thrombasthenia patients demonstrated a consistent physiological phenotype (Hu et al., 2017; Orban et al., 2015; Sullivan et al., 2014). iPSC-derived MKs did not express GPIIb/IIIa because of defective αIIb/CD41, but did express a mature MK marker, CD42b (GPIbα). Through expression of wild type ITGA2B under the GP1BA promoter, αIIbβ3 receptor levels were restored to normal, evident by normal PAC-1 binding after agonist stimulation.

1.8 Overview of Research Goals

Although ETV6 and RUNX1 mutations have been discovered to be causative for a rare subset of IT patients with severe hematologic malignancy predisposition, the
mechanism(s) underlying these disorders, including common pathways effected, is not well understood. With that, key questions in terms of how ETV6 and RUNX1 mutations relate to MK and platelet defects, as it pertains to human biology, remains. To approach these questions and gain a deeper understanding of ETV6- and RUNX1-related IT, we utilized the human iPSC model system and previously established hematopoietic differentiation protocols. We derived iPSCs from two patients, one with a germline ETV6 mutation and one with a RUNX1 mutation. We gene edited these patient-derived iPSCs to establish isogenic control iPSC lines. Additionally, we gene edited a wild type control iPSC line to individually introduce the pathogenic mutation in ETV6 or RUNX1, mirroring that in the patient-derived iPSCs. Given that mono-allelic germline mutations in ETV6 or RUNX1 present with the similar clinical phenotype of thrombocytopenia and myeloid malignancy predisposition, I hypothesized that these two transcription factors act in similar pathways, therefore disrupting normal megakaryopoiesis. The scope of this thesis includes the creation of relevant iPSCs to study ETV6- and RUNX1-related IT as well as detailed in vitro characterization of the HPCs and MKs derived from them. Chapter 2 outlines the generation of iPSCs and the methods utilized to study megakaryopoiesis in vitro. Chapter 3 details the phenotypic differences in HPCs and MKs seen in ETV6-mutant and RUNX1-mutant iPSCs. Finally, Chapter 4 discusses the relevant findings and potential future directions from these studies.
2.1 Patient-derived iPSC Line Generation

Two patient samples were used to generate iPSC lines: the ETV6 iPSC line was generated from a patient harboring a monoallelic germline mutation in ETV6 (Arg369Gln) and the RUNX1 iPSC line was generated from a previously-described patient harboring a monoallelic germline mutation in RUNX1 (splice acceptor site mutation)(Pedigree 3, W.-J. Song et al., 1999). These two iPSC lines were generated from patient peripheral blood mononuclear cells (PBMCs) approved by the CHOP Institutional Review Board (IRB), protocol 09-007042. Cells were infected with integration-free Sendai virus containing four reprogramming genes (OCT4/SOX2/KLF4/MYC) (Fusaki et al., 2009; Tokusumi et al., 2002). CRISPR/Cas9 was utilized to correct the patient mutations, thus generating two isogenic control iPSC lines with normal karyotype (Figures 3.8A – 3.8D), as previously described (Maguire et al., 2019). All guide (g) RNAs, single-stranded DNA oligonucleotides, polymerase chain reaction (PCR) primers, and sequencing primers are listed in Table 2.1. The iPSC lines generated in these patient genetic backgrounds are listed in Table 2.2. PluriTest (Coriell Institute), tri-lineage differentiation (StemXVivo, R&D Systems), flow cytometry and gene expression studies confirmed the stem cell features including pluripotency of the patient-derived iPSC lines (Figure 3.8D). The antibodies used to confirm pluripotency marker expression are detailed in Table 2.4. This independent set of isogenic iPSC lines was utilized in a confirmatory subset of studies (Figures 3.8 – 3.10).
2.2 CHOPWT6 iPSC Line Generation

A wildtype iPSC line (CHOPWT6, referred below as WT), generated by lentiviral infection of four reprogramming genes (OCT4/SOX2/KLF4/MYC), was described previously (Somers et al., 2010). Teratoma formation, flow cytometry and gene expression studies confirmed the stem cell features including pluripotency (Sullivan et al., 2014). Two distinct patient-specific mutations, described above, were introduced via CRISPR/Cas9 technology, as previously described (Maguire et al., 2019), to generate a set of isogenic lines with normal karyotypes (Cell Line Genetics) (Figures 3.1A – 3.1C). The iPSC lines generated in the CHOPWT6 background are described in Table 2.3. This set of isogenic iPSC lines is the focus of the presented studies.

2.3 iPSC Culture

iPSC lines were cultured at 37°C in an environment of 5% CO₂, 5% O₂ and 90% N₂. Mouse embryonic fibroblasts (MEFs) were isolated from CF1 mouse strain (Envigo) embryos, expanded and irradiated as described (Jiang et al., 2016). iPSCs were cultured on 0.1% gelatin and MEF feeder cells in hESC medium: DMEM/F12 (Corning) supplemented with 20% knockout serum replacement (KOSR) (Gibco), 2 mM glutamine (Corning), 1x non-essential amino acids (Gibco), 1x penicillin/streptomycin (Corning), 0.1 mM β-mercaptoethanol (Gibco) and 10 ng/mL basic fibroblast growth factor (bFGF) (R&D). Cells were passaged at 80% confluence using TrypLE (Gibco). In all cultures, 5 μM Rho-associated protein kinase (ROCK) inhibitor Y-27632 (Selleck Chemicals) was added for ~18 hours during passaging or thawing to enhance survival (Li et al., 2009; Watanabe et al., 2007).
2.4 Differentiation of iPSCs into HPCs, MKs, Erythrocytes and Myeloid cells

The iPSCs were differentiated into primitive HPCs as previously described (Mills et al., 2014, 2013) (Figure 1.6). The HPCs were isolated as single cells that were differentiated to MKs by culturing in serum-free defined (SFD) medium supplemented with TPO (50 ng/mL) (R&D) and stem cell factor (SCF) (25 ng/mL) (R&D) for 5-6 days. HPCs were cultured in SFD medium supplemented with SCF (50 ng/mL) and erythropoietin (EPO) (2 U/mL) (Janssen) for 5 days to promote erythroid differentiation. HPCs were cultured in SCF (50 ng/ml), interleukin-3 (IL-3) (5 ng/mL) (R&D) and granulocyte-macrophage colony-stimulating factor (GM-CSF) (10 ng/mL (R&D) for 5 days to promote myeloid differentiation. Media was added on top every other day. SFD media is defined as: 750 mL Iscove’s modified Dulbecco’s medium (IMDM) (Sigma), 250 mL HAMS/F12 (Corning), 5 mL N2 supplement (Gibco), 10 mL B27 supplement (Gibco), 5 mL 10% bovine serum albumin (BSA) (Sigma) in phosphate-buffered saline (PBS) (Corning), 2 mM glutamine and 1x penicillin/streptomycin.

2.5 Western Blot

Protein was quantified using the Pierce™ BCA Protein Assay Kit (Thermo Fischer Scientific (TFS)) and 20 μg was loaded onto a 4-12% Bis-Tris SDS-polyacrylamide gel (Invitrogen). Samples were transferred onto a 0.45 μM pore size nitrocellulose membrane (TFS) that was blocked using 5% nonfat dry milk. Primary antibody incubation was overnight at 4°C and secondary antibody incubation was for 1 hour at room temperature. The membrane was washed five times, 5 minutes each, with 1x tris buffered saline (TBS) (Biorad) and was incubated with Pierce™ enhanced chemiluminescent (ECL) substrates (TFS) before exposure to Hyblot CL autoradiography film (Denville Scientific). All antibodies are described in Table 2.4.
2.6 Flow Cytometry

For adherent cells, single cell suspensions were prepared by treating with 0.25% Trypsin/ethylenediaminetetraacetic acid (EDTA) (Gibco) for 3 to 5 minutes. Cells were stained in FACs buffer, defined as: Dulbecco’s PBS (DPBS) (Corning) containing 0.5% BSA and 0.05% sodium azide (Sigma). Cells were incubated with antibodies for 15 minutes at room temperature then washed and suspended in FACs buffer. All cells were analyzed on a Cytoflex flow cytometer (Becton Dickinson) using a FlowJo (Treestart) software program. All antibodies are described in Table 2.5. All software utilized is described in Table 2.8.

2.7 FACs Analysis of MKs

MKs were stained with CD41, CD42a, and CD42b in IMDM (Corning) for 30 minutes on ice. Cells were washed twice with IMDM, resuspended in 300 μL IMDM containing 1% DNAse and filtered using a snap cap cell strainer and put on ice. The CD41<sup>+</sup>CD42a<sup>+</sup>CD42b<sup>+</sup> MKs were sorted for RNA isolation using a FACS Aria II (Becton Dickinson). All antibodies are described in Table 2.5.

2.8 RNA Isolation and Reverse Transcription and Quantitative PCR (qPCR)

Cells were lysed in a buffer provided in the RNAeasy micro kit (Invitrogen) and stored at -80°C. Total RNA was extracted according to the manufacturer’s instructions. cDNA was prepared using the SuperScript<sup>™</sup> III Reverse Transcriptase kit (Invitrogen). qPCR reactions were performed in technical triplicates using SYBR-GreenER qPCR Master Mix (Roche) on a LightCycler 480 II (Roche). The TATA box-binding protein (TBP) gene was used as a housekeeping gene to determine relative gene expression (Veazey and Golding, 2011). The primers used for qPCR are listed in Table 2.6.
2.9 RNA Sequencing and Rosalind Analysis

Three biological replicates were used for RNA sequencing. Library prep was done using the SMARTer® Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian kit (Takara Bio) and run on HiSeq 2500 (Illumina) by the CHOP Center for Applied Genomics. FASTq files were uploaded to Rosalind. Data were analyzed by Rosalind (https://rosalind.onramp.bio), with a HyperScale architecture developed by OnRamp BioInformatics. Reads were trimmed using cutadapt (Martin, 2011). Quality scores were assessed using FastQC. Reads were aligned to the Homo sapiens genome build hg19 using STAR (Dobin et al., 2013). Individual sample reads were quantified using HTseq(Anders et al., 2015) and normalized via Relative Log Expression (RLE) using DESeq2 R library (Love et al., 2014). Read Distribution percentages, violin plots, identity heatmaps, and sample MDS plots were generated as part of the QC step using RSeQC (Wang et al., 2012). DEseq2 was also used to calculate fold changes, p-values and to perform optional covariate correction. Clustering of genes for the final heatmap of differentially expressed genes was done using the PAM (Partitioning Around Medoids) method using the fpc R library (https://cran.r-project.org/web/packages/fpc/).

Hypergeometric distribution was used to analyze the enrichment of pathways, gene ontology, domain structure, and other ontologies. The topGO R library (https://bioconductor.org/packages/release/bioc/html/topGO.html) was used to determine local similarities and dependencies between gene ontology (GO) terms in order to perform Elim pruning correction. Several database sources were referenced for enrichment analysis, including Interpro (Mitchell et al., 2019), NCBI (Geer et al., 2009), MSigDB (Liberzon et al., 2011; Subramanian et al., 2005), REACTOME (Fabregat et al., 2018), WikiPathways (Slenter et al., 2018). Enrichment was calculated relative to a set of
background genes relevant for the experiment. Functional enrichment analysis of pathways, GO, domain structure and other ontologies were performed using HOMER (Heinz et al., 2010). All differentially expressed genes in mutant-MKs that pass a threshold of +/- 1.2-fold change and have a p-value < 0.05 are listed in the associated excel spreadsheet. Additional gene enrichment is available from Advaita (http://www.advaitabio.com/ipathwayguide) (Donato et al., 2013; Draghici et al., 2007). All software utilized is described in Table 2.8.

2.10 MK Activation

MKs were suspended in Tyrode’s salt solution (Sigma) with 0.1% BSA to a final concentration of 1x10^6 cells/mL. Cells were stained with anti-CD42a, anti-CD42b, PAC-1 antibodies (see Table 2.5) and stimulated with thrombin (0.01 U - 0.1 U final) (Sigma) or adenosine 5’-diphosphase sodium salt (ADP) (1 μM – 10 μM final) (Sigma) in a total volume of 100 μL for 10 minutes at room temperature followed by incubation on ice, in the dark.

2.11 Pulse Labeling of MKs with Coagulation Factor V (FV)

MK cultures were washed with RPMI. MKs were resuspended in RPMI and pulse-labeled with FV by incubating with 200 nM of a previously described FV variant tagged with Alexa-488 for 1 hour at 37℃ (Ivanciu et al., 2014). For the last 20 minutes, anti-CD42b antibody was added to the MK cultures. MKs were washed twice with fresh RPMI. MKs were resuspended in FACS buffer before flow cytometry analyses.
2.12 Proplatelet Formation Assay

HPCs were cultured in StemSpan™ Serum-Free Expansion Media II (SFEM II) (Stem Cell Technologies) with TPO (50 ng/mL) and SCF (25 ng/mL) for 5-7 days at 37°C with 5% CO₂ to promote MK differentiation. In 12 well tissue culture-treated plates, sterile glass coverslips (EMS) were washed with PBS and coated with 100 µg/mL fibrinogen (Millipore) in PBS for 2 hours. Coverslips were blocked with 1% BSA for 1 hour, removed 15 minutes prior to MK seeding. MKs were resuspended in seeding media (SFEM II with TPO (50 ng/mL)) at a concentration of 2,000 MKs/µL and 1x10⁵ MKs were seeded on the center of each coverslip. MKs were left to adhere for 1 hour at 37°C before flooding with 500 µL of seeding media per well and incubating for an additional 24 hours. Coverslips were fixed with 4% paraformaldehyde (PFA) (Thermo Scientific) in PBS for 20 minutes at room temperature, washed 3x with PBS and left in 1 mL PBS for imaging. Images were taken at 20x in random locations and proplatelet forming MKs were counted using FIJI software.

2.13 Immunofluorescence Microscopy

Proplatelet forming MKs were fixed on coverslips with 4% PFA in PBS for 20 minutes followed by permeabilization with 0.3% Triton X-100 (Sigma) in PBS for 10 minutes at room temperature. After 3 washes with PBS, coverslips were blocked with 3% BSA in PBS for 1 hour, followed by conjugated anti-α-tubulin antibody (see Table 2.7) diluted in 3% BSA for 2 hours at room temperature. After PBS washes, coverslips were mounted with VECTASHIELD mounting medium with DAPI (Vector Laboratories). Digital images were taken on a Leica DMI 4000B microscope with Leica Application Suite X software (see Table 2.8).
2.14 Hematopoietic and MK Colony Assays

HPCs (1x10^4) were plated in a methylcellulose-based medium, Methocult™ H4435 Enriched (STEMCELL technologies), according to the manufacturer's instructions. Colonies were counted 12-14 days after plating. For MK colonies, HPCs (5-10x10^3) were plated in a collagen-based medium, Megacult™-C (Invitrogen) according to manufacturer's instructions. MK colonies were counted after 12-14 days of incubation.

2.15 Mitogen-activated Protein Kinase (MAPK) Inhibition using PD98059

HPCs were collected on day 8 of differentiation and each sample was split in two. Half of the HPCs were treated with DMSO as a control, and the other half of HPCs were treated with 10µM PD98059 (MEK1 and MEK2 inhibitor) (Calbiochem) for 5 days during HPC to MK expansion and differentiation.

2.16 Statistical Analysis

Statistical analyses were performed using GraphPad Prism 6 software. The results are represented as mean ± standard error of the mean (SEM). Ordinary one-way ANOVA with multiple comparisons were used with correction for multiple comparisons using statistical hypothesis testing performed using Tukey. In figures, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>gRNA – sense</strong></td>
<td></td>
<td>GATAGTGGATCCCAACGGAC</td>
</tr>
<tr>
<td><strong>Single stranded DNA oligo – WT</strong></td>
<td></td>
<td>CGTCTATCAGTTGCTTTCTGACAGCCGGTACG AAAACTTCCATCGAGAGGGAGAAGAAAAGATCT AGATCCCAACGGGCT AGCTCGAAGCTGTGGGGAAACCATAGGttaaagg gcagcagatagtctctcat</td>
</tr>
<tr>
<td><strong>Single stranded DNA oligo – mutant</strong></td>
<td></td>
<td>CGTCTATCAGTTGCTTTCTGACAGCCGGTACG AAAACTTCCATCGAGAGGGAGAAGAAAAGATCT AGATCCCAACGGGCT AGCTCGAAGCTGTGGGGAAACCATAGGttaaagg gcagcagatagtctctcat</td>
</tr>
<tr>
<td><strong>PCR primer – Forward intron 5</strong></td>
<td></td>
<td>5’ – CGGTGGGTCAAAAGGTAC – 3’</td>
</tr>
<tr>
<td><strong>PCR primer – Reverse, intron 6</strong></td>
<td></td>
<td>5’ – GTCAAGGGAGTGGAAACAA – 3’</td>
</tr>
<tr>
<td><strong>Sequencing primer – forward</strong></td>
<td></td>
<td>5’ – CAAGCTAGGCAGAAGCAG – 3’</td>
</tr>
<tr>
<td><strong>gRNA – antisense</strong></td>
<td></td>
<td>GCCATCTGGGAACATCCCTA</td>
</tr>
<tr>
<td><strong>Single stranded DNA oligo – WT</strong></td>
<td></td>
<td>atcactacacaatatccctaaagaattgatttataacacccctgttgt gcattggcctttgtgacggtgatttagGTGTGGCCCTAGG CGACGTCCACAGTGACTGTGCTACTCTGTAG GTGGCTGGCAATGATGAAAACTACTCGGC</td>
</tr>
<tr>
<td><strong>Single stranded DNA oligo – mutant</strong></td>
<td></td>
<td>atcactacacaatatccctaaagaattgatttataacacccctgttgt gcattggcctttgtgacggtgatttagGTGTGGCCCTAGG CGACGTCCACAGTGACTGTGCTACTCTGTAG GTGGCTGGCAATGATGAAAACTACTCGGC</td>
</tr>
<tr>
<td><strong>PCR primer – forward intron 3</strong></td>
<td></td>
<td>5’ – TCTTGTCCGACTGTGCTATGT – 3’</td>
</tr>
<tr>
<td><strong>PCR primer – reverse intron 4</strong></td>
<td></td>
<td>5’ – TCTTACTACGGTGGATCTGTC – 3’</td>
</tr>
<tr>
<td><strong>Sequencing primer – forward</strong></td>
<td></td>
<td>5’ – TCATTGCTATTCTCTGCAACCT – 3’</td>
</tr>
</tbody>
</table>

For single stranded DNA oligos: uppercase letters denote exons; lowercase letters denote introns; red letters denote silent mutations introduced; blue letters denote patient mutation; underlined letters denote silent restriction site introduced after successful CRISPR/Cas9 targeting.
### Table 2.2: iPSC Lines Generated from Patient Material

<table>
<thead>
<tr>
<th>Name</th>
<th>Background</th>
<th>Allele #1</th>
<th>Allele #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETV6\textsuperscript{mut/mut}</td>
<td>ETV6 Patient</td>
<td>Point Mutation</td>
<td>Point Mutation</td>
</tr>
<tr>
<td>ETV6\textsuperscript{+/mut}</td>
<td>ETV6 Patient</td>
<td>Wild Type</td>
<td>Point Mutation</td>
</tr>
<tr>
<td>ETV6\textsuperscript{+/+}</td>
<td>ETV6 Patient</td>
<td>Wild Type</td>
<td>Wild Type</td>
</tr>
<tr>
<td>RUNX1\textsuperscript{+/mut}</td>
<td>RUNX1 Patient</td>
<td>Wild Type</td>
<td>Splice Mutation</td>
</tr>
<tr>
<td>RUNX1\textsuperscript{+/+}</td>
<td>RUNX1 Patient</td>
<td>Wild Type</td>
<td>Wild Type</td>
</tr>
</tbody>
</table>

The colors used for each iPSC line in this table will be consistent throughout all remaining figures.

### Table 2.3: iPSC Lines Generated in the CHOPWT6 Genetic Background

<table>
<thead>
<tr>
<th>Name</th>
<th>Background</th>
<th>Allele #1</th>
<th>Allele #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHOPWT6</td>
<td>CHOPWT6</td>
<td>Wild Type</td>
<td>Wild Type</td>
</tr>
<tr>
<td>CHOPWT6/ETV6\textsuperscript{+/mut}</td>
<td>CHOPWT6</td>
<td>Wild Type</td>
<td>Point Mutation</td>
</tr>
<tr>
<td>CHOPWT6/ETV6\textsuperscript{mut/mut}</td>
<td>CHOPWT6</td>
<td>Point Mutation</td>
<td>Point Mutation</td>
</tr>
<tr>
<td>CHOPWT6/RUNX1\textsuperscript{+/mut}</td>
<td>CHOPWT6</td>
<td>Wild Type</td>
<td>Splice Mutation</td>
</tr>
</tbody>
</table>

The colors used for each iPSC line in this table will be consistent throughout all remaining figures.

### Table 2.4: Antibodies used for Western Blot

<table>
<thead>
<tr>
<th>Fluorophore</th>
<th>Antibody</th>
<th>Dilution</th>
<th>Company</th>
<th>Identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconjugated</td>
<td>Actin</td>
<td>1:2000</td>
<td>Sigma</td>
<td>A1978</td>
</tr>
<tr>
<td></td>
<td>ETV6</td>
<td>1:1000</td>
<td>Thermo Fischer Scientific</td>
<td>PA5-79224</td>
</tr>
</tbody>
</table>
### Table 2.5: Antibodies used for Flow Cytometry

<table>
<thead>
<tr>
<th>Fluorophore</th>
<th>Antibody</th>
<th>Dilution</th>
<th>Company</th>
<th>Identifier</th>
<th>Cell type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexa Fluor 488</td>
<td>SSEA3</td>
<td>1:50</td>
<td>BioLegend</td>
<td>330306</td>
<td>iPSC</td>
</tr>
<tr>
<td></td>
<td>TRA-1-60</td>
<td>1:50</td>
<td>BioLegend</td>
<td>330614</td>
<td>iPSC</td>
</tr>
<tr>
<td>Alexa Fluor 647</td>
<td>SSEA4</td>
<td>1:400</td>
<td>BioLegend</td>
<td>3304008</td>
<td>iPSC</td>
</tr>
<tr>
<td></td>
<td>TRA-1-81</td>
<td>1:50</td>
<td>BioLegend</td>
<td>330706</td>
<td>iPSC</td>
</tr>
<tr>
<td>Allophycocyanin (APC)</td>
<td>CD18</td>
<td>1:20</td>
<td>BD</td>
<td>551060</td>
<td>Myeloid</td>
</tr>
<tr>
<td></td>
<td>CD42b</td>
<td>1:20</td>
<td>BioLegend</td>
<td>303912</td>
<td>MK</td>
</tr>
<tr>
<td></td>
<td>CD235</td>
<td>1:1000</td>
<td>BD</td>
<td>551336</td>
<td>HPC &amp; Erythroid</td>
</tr>
<tr>
<td>Allophycocyanin (APC)</td>
<td>eFluor450</td>
<td>1:50</td>
<td>eBioscience</td>
<td>45-0428-42</td>
<td>MK</td>
</tr>
<tr>
<td>eFluorescein Isothiocyanate (FITC)</td>
<td>CD41</td>
<td>1:50</td>
<td>BioLegend</td>
<td>303704</td>
<td>MK</td>
</tr>
<tr>
<td></td>
<td>CD43</td>
<td>1:50</td>
<td>BioLegend</td>
<td>315204</td>
<td>HPC</td>
</tr>
<tr>
<td></td>
<td>PAC1</td>
<td>1:10</td>
<td>BD</td>
<td>340507</td>
<td>MK</td>
</tr>
<tr>
<td>Pacific Blue (PB)</td>
<td>CD45</td>
<td>1:100</td>
<td>BioLegend</td>
<td>304022</td>
<td>Myeloid</td>
</tr>
<tr>
<td>Phycoerythrin (PE)</td>
<td>CD41</td>
<td>1:50</td>
<td>BD</td>
<td>555467</td>
<td>HPC &amp; Erythroid</td>
</tr>
</tbody>
</table>

### Table 2.6: qPCR Primers

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETV6</td>
<td>Forward: 5’ – TATCAGCACAAATCACCGGCTTCTC – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – GGTAGGACTCCTGGTGTGGTTCTC – 3’</td>
</tr>
<tr>
<td>GP9</td>
<td>Forward: 5’ – CTGGCCAAACAGACCTTCAGT – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – CCAGCCAGGGCGAGATAG – 3’</td>
</tr>
<tr>
<td>ITGA2B</td>
<td>Forward: 5’ – CCAGAGCTTTGTGCACCTGCAAG – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – ATAGAAGGTGAGCTGCACTGGGT – 3’</td>
</tr>
<tr>
<td>MYL9</td>
<td>Forward: 5’ – GCCACATCCATGTTCTCGCAATGT – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – TGGTCAGGTCTGACCAGTCTGTT – 3’</td>
</tr>
<tr>
<td>PTGS1</td>
<td>Forward: 5’ – GTACTGGAATCCGAGCACATTGG – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – GGAAGAAACCTAGGAGAGCTCCTT – 3’</td>
</tr>
<tr>
<td>SMOX</td>
<td>Forward: 5’ – AGATCTGCGGTGTGGTCTCTCAC – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – TGCTCGTCTCATCACTTCTCCAT – 3’</td>
</tr>
<tr>
<td>RUNX1</td>
<td>Forward: 5’ – GCAGCATGGTGAGTGCTG – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – AAAGCGATGCGGCGGTTCTTGT – 3’</td>
</tr>
<tr>
<td>TBP</td>
<td>Forward: 5’ – TTGGAGAGAGTGAGCACAGTG – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – CGTAAGGTGGCAGGCTGTTGT – 3’</td>
</tr>
<tr>
<td>TUBB1</td>
<td>Forward: 5’ – GCAGCTGCTTTGTGGAGGTGATT – 3’</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ – AGATCTCTGGATGGCCGTTGT – 3’</td>
</tr>
</tbody>
</table>
Table 2.7: Antibody used for Immunofluorescence Microscopy

<table>
<thead>
<tr>
<th>Fluorophore</th>
<th>Antibody</th>
<th>Dilution</th>
<th>Company</th>
<th>Identifier</th>
<th>Cell Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluorescein Isothiocyanate (FITC)</td>
<td>α-Tubulin</td>
<td>1:50</td>
<td>Sigma</td>
<td>F2168</td>
<td>MK</td>
</tr>
</tbody>
</table>

Table 2.8: Software

<table>
<thead>
<tr>
<th>Name of software</th>
<th>Used for:</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biorender</td>
<td>Graphics/schematics</td>
<td><a href="http://www.biorender.com">www.biorender.com</a></td>
</tr>
<tr>
<td>GraphPad Prism</td>
<td>Figure generation</td>
<td><a href="http://www.graphpad.com">www.graphpad.com</a></td>
</tr>
<tr>
<td>Flowjo</td>
<td>Flow cytometry analysis</td>
<td><a href="http://www.flowjo.com">www.flowjo.com</a></td>
</tr>
<tr>
<td>ImageJ – Fiji</td>
<td>Proplatelet formation cell counting</td>
<td><a href="http://www.imagej.net/Fiji">www.imagej.net/Fiji</a></td>
</tr>
<tr>
<td>Leica Application Suite X</td>
<td>Immunofluorescence microscopy</td>
<td><a href="http://www.leica-microsystems.com">www.leica-microsystems.com</a></td>
</tr>
<tr>
<td>Rosalind</td>
<td>RNA-sequencing analysis and related figures</td>
<td><a href="http://www.onramp.bio">www.onramp.bio</a></td>
</tr>
</tbody>
</table>
CHAPTER 3

Insights into Inherited Thrombocytopenia Resulting from Mutations in ETV6 or RUNX1 using a Human Pluripotent Stem Cell Model

3.1 Introduction

IT caused by mutations in a variety of genes, is a genetic bleeding disorder resulting in low blood platelet (Dowton et al., 1985; Savoia et al., 2017). A subset of these patients harbor monoallelic germline mutations in the transcription factors ETV6 or RUNX1, or the gene ANKRD26 (Savoia et al., 2017). Patients with monoallelic germline mutations in any of these three genes phenocopy one another in terms of platelet defects and heightened hematologic malignancy predisposition; however, the disease pathology is not well understood (Dowton et al., 1985). One reason for this lack of understanding is the need for appropriate model systems. Mouse models have been utilized to study IT with risk of malignancy, but they fail to fully recapitulate the human disease phenotype (De Braekeleer et al., 2012; Sakurai et al., 2014; Sood et al., 2010; Zhang et al., 2015). Specifically, mice heterozygous for ETV6 have unperturbed hematopoiesis, whereas RUNX1-haploinsufficient mice have minimal thrombocytopenia (~15% reduction in platelet count) and do not develop leukemia, even with secondary hits (De Braekeleer et al., 2012; Hock et al., 2004; Sakurai et al., 2014; Sood et al., 2010).

iPSCs have proven to be a suitable model system to study developmental genetic disorders such as ETV6- and RUNX1-related thrombocytopenia (Cherry and Daley, 2013; Yamanaka, 2012). As knowledge of the signaling pathways driving developmental processes is gained, in vitro differentiation protocols are continually developed and improved to enable the generation of various cell types from iPSCs. This model system
allows for the investigation of pathogenic events that may be transient, and otherwise missed, during hematopoiesis. iPSC technology also allows for the generation of unlimited quantities of cells, which is important when studying disorders with limited patient population size and/or key samples. Additionally, patient-derived iPSCs can be genetically modified via CRISPR/Cas9 to yield isogenic control iPSC lines, which is critical when using human samples due to the large variation found in the human population. Genome engineered iPSCs allow for detailed analysis of pathogenic mutations on the development of a disease relevant cell type, such as the MK (Musunuru, 2013; Ran et al., 2013).

Although the first case of IT with heightened risk of AML was described in 1978, heterozygous germline RUNX1 mutations were not confirmed to be the cause until 1999 (Luddy et al., 1978; Song et al., 1999). RUNX1 has since been studied extensively in megakaryopoiesis. In recent years, iPSC technology has been utilized to characterize the in vitro phenotype of iPSCs harboring RUNX1 mutations (Antony-Debré et al., 2015; Connelly et al., 2014; Iizuka et al., 2015; Sakurai et al., 2014). These studies consistently highlight a defect in MK differentiation from HPCs. However, an HPC phenotype is not always noted. More recently, mutations effecting ETV6 have also been found to result in IT with predisposition for hematologic malignancy, mainly B-cell acute lymphoblastic leukemia (Zhang et al., 2015). Since this discovery in 2015, studies in HeLa and primary human CD34+ cells show a defect in yielding large, proplatelet-forming ETV6-haploinsufficient MKs (Noetzli et al., 2015; Poggi et al., 2017; Zhang et al., 2015). To date, no ETV6-mutant patient-derived iPSC studies have been reported.

In this study, we investigated the mechanism of IT caused by monoallelic mutations in the transcription factors ETV6 or RUNX1. Although this disease has two parts -
thrombocytopenia and hematologic malignancy - we focused our studies on the thrombocytopenia aspect as a more tractable endpoint. Here, we generated iPSCs from two IT patients: one harboring a mutation in *ETV6* (Arg369Gln) and the other with a mutation in *RUNX1* (splice acceptor site mutation) (Song et al., 1999) and then established isogenic controls for each line. We also introduced these patient mutations into a control iPSC line so that there was one isogenic control common for both mutant lines. We found completely disparate phenotypes between the *ETV6*-mutant and *RUNX1*-mutant iPSC lines. The *ETV6*-mutant iPSCs yielded higher numbers of HPCs that gave rise to more MKs when compared to the isogenic control, whereas the *RUNX1*-mutant iPSCs yielded fewer HPCs that gave rise to fewer MKs. Upon further characterization of the MKs, we demonstrated that *ETV6*-mutant MKs were less mature as they responded poorly to agonist stimulation and endocytosed less FV, which we have previously shown is a sign of MK maturation (Sim et al., 2017). In contrast, *RUNX1*-mutant MKs were more responsive to agonist stimulation and endocytosed more FV compared to an isogenic control. Genome-wide gene-expression analysis of MKs derived from the various genotypes further supported the notion that the mechanism of thrombocytopenia is different between *ETV6*- and *RUNX1*-mutants. These studies highlight that pathogenic mutations may lead to the same phenotype in very different ways, which can affect therapeutic treatment options.

### 3.2 Results

#### 3.2.1 Generation and Characterization of Isogenic iPSC Lines

To examine the role of *ETV6* and *RUNX1* in HPC and MK development, two different sets of iPSC lines were used: (1) patient-derived iPSCs harboring monoallelic germline *ETV6* or *RUNX1* mutations and (2) a WT iPSC line with the *ETV6* or *RUNX1* patient mutations
introduced. The latter system allowed for investigation of the pathogenic effects of these mutations in a common genetic background (CHOPWT6) (Somers et al., 2010). The ETV6 patient mutation affects DNA binding, whereas dimerization is unaffected (Zhang et al., 2015). The RUNX1 patient was previously described as having a splice-acceptor site mutation that results in haploinsufficiency (Pedigree 3, W.-J. Song et al., 1999). To introduce the patient mutations in the WT iPSC line and correct the mutation in the patient lines, CRISPR/Cas9 technology was used (Maguire et al., 2019). The gRNAs were designed near the patient mutations of interest (Figures 3.1A and 3.8A) and karyotypically normal clones (Figures 3.1B and 3.8B) expressing either heterozygous or homozygous mutations were confirmed by sequencing (Figures 3.1C and 3.8C). The WT isogenic set of iPSC lines are the main focus of these studies, while the patient-derived iPSC lines and isogenic corrected lines are used for validation in a second genetic background.

To establish the kinetics of ETV6 or RUNX1 expression during blood cell development, mRNA and protein expression were examined during the differentiation of iPSCs to HPCs and MKs. ETV6 mRNA is expressed at low levels in undifferentiated iPSCs, whereas RUNX1 mRNA is undetectable. Transcript levels of both ETV6 and RUNX1 increased significantly at the HPC and MK stages of differentiation (Figure 3.1D). Protein expression of ETV6 was analyzed by western blot. The ETV6mut/mut line had significantly lower levels of ETV6 protein when compared to the WT and ETV6+mut lines (Kirkpatrick et al., 2015; Zhang et al., 2015), suggesting that this homozygous mutation also affected protein production and/or stability (Figure 3.1E). The RUNX1 splice defect was confirmed using reverse transcription-PCR (RT-PCR), demonstrating aberrant splicing at a downstream cryptic splice acceptor site (Song et al., 1999) (Figure 3.1F).
Figure 3.1: Characterization of CHOPWT6-Generated iPSC Lines. (A) Schematic of ETV6 and RUNX1 genes with arrows denoting patient mutations and bars denoting gRNA. (B) Karyotypes of lines after CRISPR/Cas9 gene editing. (C) Sanger sequencing of CRISPR/Cas9 edited lines, showing introduction of patient-specific mutations. (D) qPCR analyses during hematopoietic differentiation. (E) Western blot for ETV6 and actin. (F) RT-PCR for RUNX1 showing WT allele at 149 bp and RUNX1-mutant spliced allele at 136 bp.
3.2.2 ETV6 and RUNX1 Mutations Reveal Disparate Effects on Blood Differentiation

To determine the effect of these mutations on blood cell development, the iPSC lines were differentiated into HPCs using a previously described protocol (Mills et al., 2014) (Figure 1.6). The HPCs were isolated as single cells on day 8 of differentiation and analyzed by flow cytometry for CD34 and CD43 expression: CD34 is an early HPC marker, whereas CD43 is a pan-hematopoietic cell surface marker. We observed similar HPC flow profiles for all genotypes, but there was significantly fewer CD34⁺CD43⁺ HPCs in the RUNX1⁺/mut iPSC line compared to the WT, consistent with prior publications (Antony-Debré et al., 2015; Sakurai et al., 2014) (Figures 3.2A and 3.2B). In contrast, the yield of ETV6⁺/mut CD34⁺CD43⁺ HPCs was not significantly different when compared to the WT, while the ETV6mut/mut line generated significantly more HPCs (Figure 3.2B).

Next, the lineage biases of these HPCs were analyzed by flow cytometry and colony assays. Cells were co-stained for cell surface expression of CD41 and CD235, markers for MK and erythroid commitment, respectively, whereas myeloid precursors were included within the population of double-negative cells within the CD34⁺CD43⁺ HPC population (Paluru et al., 2014; Vodyanik et al., 2005). Both the ETV6⁺/mut and ETV6mut/mut HPCs were biased toward double-negative CD41⁻CD235⁻ cells, whereas the RUNX1⁺/mut HPCs were biased toward CD235⁺ cells when compared to the WT (Figure 3.2C). Colony assays confirmed these findings. The ETV6mut/mut HPCs gave rise to more myeloid colonies and fewer erythroid colonies, whereas the ETV6⁺/mut lines gave rise to fewer erythroid colonies (Figure 3.2D). Although the RUNX1⁺/mut HPCs gave rise to fewer erythroid colonies (Figure 3.2D), they were generally larger in size when compared to the WT. Culturing these HPCs in erythroid or myeloid liquid expansion conditions also supported these data: the ETV6mut/mut HPCs generated more myeloid cells, whereas the
RUNX1\textsuperscript{+\textasciitilde\textasciitilde\textasciitilde} HPCs generated more erythroid cells (Figure 3.2E). The MK potential of the HPCs was analyzed using the megacult colony assay. The ET\texttv6\textsuperscript{mut\textasciitilde\textasciitilde\textasciitilde} HPCs generated more MK colonies, whereas RUNX1\textsuperscript{+\textasciitilde\textasciitilde\textasciitilde} HPCs generated fewer MK colonies (Figure 3.2F). These differences were confirmed using isogenic pairs of iPSC lines in a second genetic background. (Figures 3.9A – 3.9F).
Figure 3.2: **ETV6** and **RUNX1** Mutations Reveal Disparate Lineage Potentials in CHOPWT6 Isogenic iPSC Lines. (A) Representative flow profiles of HPCs. (B) Quantification of fold change in CD34⁺CD43⁺ HPCs per iPSC plated on day -2; normalized to WT. N = 7. (C) Representative flow profiles of CD34⁺CD43⁺ HPC lineage biases: erythroid is CD41⁻CD235⁺, myeloid is CD41⁻CD235⁻. (D) Quantification of number of erythroid (left) and myeloid (right) colonies after 12-14 days in methylcellulose-based medium. N = 9 for erythroid and N = 6 for myeloid. (E) Quantification of erythroid (left) and CD18⁺CD45⁺ myeloid (right) cells after 5 days of culture; normalized to WT. N = 4. (F) Quantification of number of MK colonies after 12 days in collagen-based medium. N = 4. For all statistical analyses, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
3.2.3 Contrasting MK Phenotypes in ETV6- and RUNX1-Mutant iPSC Lines

To analyze the MK phenotype, CD34⁺CD43⁺ HPCs were cultured in medium containing TPO and SCF. By flow cytometry, all of the lines generated populations of CD41⁺CD42a⁺ MKs (Figure 3.3A). However, when calculating MK yield per input CD34⁺CD43⁺ HPC, major differences were observed: the ETV6mut/mut HPCs generated more MKs (Figure 3.2B, left). In contrast, the RUNX1⁺/mut HPCs generated fewer MKs (Figure 3.3B, left). This same phenomenon was observed when calculating MK yield per input iPSC (Figure 3.3B, right).

To determine if these differences were reflective of maturation defects, the expression of various receptors was analyzed during MK maturation using flow cytometry. The expression of CD41 (GPIIb), an early marker of MK generation (Mitjavila-Garcia et al., 2002), was decreased on ETV6⁺/mut and ETV6mut/mut MKs but increased on RUNX1⁺/mut MKs (Figure 3.3C, left). The expression of CD42a (GPIX) and CD42b (GPIba), both later markers of MK maturation, (Nishikii et al., 2015) were decreased on ETV6⁺/mut and ETV6mut/mut MKs, whereas CD42a was increased on RUNX1⁺/mut MKs (Figure 3.3C, middle and right). Despite increased expression on RUNX1⁺/mut MKs, these MKs were slightly smaller in size, suggesting that receptor expression was not correlated with surface area on these cells (Figure 3.3D). ETV6mut/mut MKs were also smaller compared to the WT iPSC line. These data suggest that the ETV6 mutant lines generated increased numbers of MKs with abnormal maturation marker expression, whereas the RUNX1 mutant line generated decreased numbers of MKs with normal maturation marker expression. These findings were consistent in a second genetic background (Figures 3.10A – 3.10C).
Figure 3.3: Disparate MK Phenotypes in ETV6- and RUNX1-Mutant CHOPWT6 iPSC Lines. (A) Representative flow profiles of CD41⁺CD42a⁺ MKs. (B) Quantification of MKs per CD34⁺CD43⁺ HPC (left) and per iPSC plated on day -2 (right) after 5 days of culture; normalized to WT. N = 6. (C) Quantification of mean fluorescence intensity (MFI) of MK markers after 5 days of MK culture; normalized to WT. For CD41 (left): N = 10 for WT, ETV6⁺/mut and ETV6mut/mut; N = 6 for RUNX1⁺/mut. For CD42a (middle): N = 12 for WT, ETV6⁺/mut and ETV6mut/mut; N = 8 for RUNX1⁺/mut. For CD42b (right): N = 14 for WT, ETV6⁺/mut and ETV6mut/mut; N = 10 for RUNX1⁺/mut. (D) Quantification of MK size after 5 days of culture via forward-scatter area (FSC) MFI; normalized to WT. N = 9 for WT, ETV6⁺/mut and ETV6mut/mut; N = 6 for RUNX1⁺/mut. For all statistical analyses, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
3.2.4 ETV6-Mutant MKs are Less Responsive to Agonists than RUNX1-Mutant MKs

To test MK functionality, the responsiveness of MKs to stimulation by various agonists was analyzed. Activation with thrombin or ADP results in a conformational change in surface αIIbβ3 receptors, thus enabling the binding of the conformation-specific monoclonal antibody, PAC1 (Shattil et al., 1987, 1985). All MKs showed no basal PAC1 binding, demonstrating no spontaneous pre-activation of these cells in culture (Figure 3.4A, top). Upon thrombin stimulation, ~30% of WT MKs responded to agonist stimulation (Figure 3.4A, bottom). This percentage was decreased to an average of ~15% for ETV6mut/mut MKs, but significantly increased to ~50% for RUNX1+/mut MKs (Figure 3.4A, bottom). MK responsiveness to ADP was similar to that of thrombin stimulation (Figure 3.4B).

Another assay measuring MK function is uptake of coagulation FV. We previously demonstrated that the uptake of fluorescently-labeled FV positively correlated with MK maturation (Ivanciu et al., 2014; Sim et al., 2017). The uptake of FV by ETV6mut/mut MKs was decreased when compared to the WT, whereas the uptake of FV by RUNX1+/mut MKs was increased (Figure 3.4C). These data suggest that although ETV6mut/mut MKs may be increased in number, they are less mature and responsive, whereas the RUNX1+/mut MKs were decreased in number, but appeared to be fully mature and more responsive when compared to the WT MKs. Both phenotypes are consistent with disease, but suggestive of distinct mechanisms. All of these findings were similar in the second genetic background (Figures 3.10D & 3.10E).
Figure 3.4: *ETV6*-Mutant MKs are Less Responsive than *RUNX1*-Mutant MKs. (A) Representative flow histograms of PAC1 binding in unstimulated MKs (top) and thrombin stimulated MKs (bottom) after 6 days of MK culture. (B) Quantification of PAC1 MFI after stimulation with thrombin (left) or ADP (right) after 6 days of culture; fold change normalized to WT. N = 7. (C) Quantification of FV uptake over background in day 5 MKs; fold change normalized to WT. N = 6. For all statistical analyses, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
3.2.5 RNA- Sequencing Analysis of Purified MKs

To probe the molecular mechanism for these differences, genome-wide gene expression was performed on purified CD41⁺CD42b⁺ MKs from all four genotypes using RNA-sequencing (Figure 3.5A). Comparing the mutant MKs to the WT and representing the data as volcano plots, ETV6⁺/mut MKs had 323 down-regulated genes and 723 up-regulated genes, which increased dramatically in number when the second allele of ETV6 was mutated (1177 and 1434, respectively) (Figures 3.5B and 3.5C). Compared to the WT, RUNX1⁺/mut MKs had 937 down-regulated genes and 686 up-regulated genes (Figures 3.5B and 3.5C). These data fit with ETV6 acting predominantly as a transcriptional repressor, and RUX1 being a transcriptional activator (Noetzli et al., 2015; Rasighaemi et al., 2015; Topka et al., 2015; Zhang et al., 2015).

We compared differences in gene expression between the various genotypes via a meta-analysis that utilizes an unsupervised clustering algorithm to bin genes into distinct subpopulations. When comparing the WT to both the ETV6⁺/mut and ETV6⁰/mut MKs, a trend was observed in up- and down-regulated genes, with a higher degree of dysregulation in the ETV6⁰/mut MKs compared to the ETV6⁺/mut MKs (Figure 3.5D, left). In contrast, when RUNX1⁺/mut and ETV6⁺/mut MKs were compared to the WT, there was minimal overlap in dysregulated genes, with genes up or down in one mutant line but not the other (Figure 3.5D, right). Overall, these data are in line with our other findings that show disparate phenotypes and support the idea that ETV6 and RUNX1 haploinsufficiency leads to MK defects via distinct mechanisms. Despite the differences in gene expression between ETV6- and RUNX1-mutant MKs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed platelet activation and cancer related...
signaling pathways as the top GO terms in both mutant-MK populations (Figures 3.5E and Table 3.1).
Figure 3.5: RNA-sequencing of Purified CHOPWT6 Isogenic MKs. (A) Gating strategy for MK purification sort prior to RNA isolation. (B) Volcano plots showing differential gene expression compared to WT. Each dot represents a gene, with the coloring of the dot reflecting the clustering information for each gene: black dots are genes that do not pass the filter parameters; green dots are genes that are up-regulated compared to WT; purple dots are genes that are down-regulated compared to WT. The number of up- and down-regulated genes are noted in red and blue, respectively. (C) Heat maps of up- and down-regulated genes in CHOPWT6 and ETV6^{+/-mut} (left), ETV6^{mut/mut} (middle) and RUNX1^{+/-mut} (right). (D) Meta-analysis heat map independently comparing WT to ETV6^{+/-mut} and ETV6^{mut/mut} (left), and WT to RUNX1^{+/-mut} and ETV6^{+/-mut} (right). The overall trend in gene expression for each cluster is depicted above the total number of genes in each cluster. (E) Top GO terms from KEGG pathway analysis. For all statistical analyses, \*P<0.05, \**P<0.01, \***P<0.001 and \****P<0.0001.
3.2.6 Common Proplatelet Formation Defects in ETV6- and RUNX1-Mutant MKs

Given that platelet activation was one of the top dysregulated pathways in ETV6- and RUNX1-mutant MKs, we wanted to examine the effects of these mutations on thrombopoiesis. Comparison of our RNA-sequencing data with a platelet transcriptome gene list from a prior publication (Rowley et al., 2011) revealed a similar degree of downregulation in genes crucial for platelet formation and function between ETV6\textsuperscript{mut/mut} and RUNX1\textsuperscript{+/mut} MKs (Figure 3.6A). Lastly, the proplatelet forming abilities of ETV6-mutant and RUNX1\textsuperscript{+/mut} MKs was examined through adhesion to fibrinogen substrate. Interestingly, the efficiency and complexity of the resultant proplatelet structures was significantly reduced in ETV6- and RUNX1-mutant MKs compared to the isogenic control (Figures 3.6 B and 3.6C). These findings suggest that the genes dysregulated as a result of mutations in ETV6 or RUNX1 may share common pathways to ultimately disrupt proplatelet formation and thrombopoiesis, despite the disparate defects on megakaryopoiesis.
Figure 3.6: Common Proplatelet Formation Defects in ETV6- and RUNX1-Mutant MKs. (A) Graphs validating platelet gene expression from MK RNA-sequencing data via qPCR. N = 3. (B) Representative immunofluorescence images of proplatelet forming MKs on glass coverslips coated with fibrinogen. Blue = DAPI; Green = α-tubulin. (C) Graph quantifying the percent proplatelet forming CD41+ MKs out of total adhered CD41+ MKs. N = 4. For all statistical analyses, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
<table>
<thead>
<tr>
<th>Genotype</th>
<th>GO term</th>
<th>Genes UP</th>
<th>Genes DOWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETV6&lt;sup&gt;+&lt;/sup&gt;/mut</td>
<td>Platelet activation</td>
<td>COL1A2, ARHGEF1, PLCB4, PIK3CG, GP5, SRC, ORAI1, GNAI2, ADCY6, RAP1A, PLCB3</td>
<td>PIK3R3, GUCY1A2, ITGA2, MAPK12</td>
</tr>
<tr>
<td>ETV6&lt;sup&gt;mut/mut&lt;/sup&gt;</td>
<td>Platelet activation</td>
<td>ARHGEF1, ORAI1, PLCB4, PIK3CG, PLCB2, PRKCI, COL1A2, GNAI1, VAMP8, PPP1CB, GNAI2, RAP1A</td>
<td>PLA2G4A, PIK3R3, PRKG1, GUCY1A2, ITGA2, RASGRP1, GUCY1A3, MAPK12, ADCY3, RASGRP2, FCGR2A, P2RY12, MAPK13, GUCY1B3</td>
</tr>
<tr>
<td>RUNX1&lt;sup&gt;+&lt;/sup&gt;/mut</td>
<td>Platelet activation</td>
<td>AKT3, GNAI1, GUCY1A3, PRKCI, PPP1CC, MAPK12, MAPK11</td>
<td>MYLK3, ITGA2, ADCY3, MAPK13, VWF, AKT2, APBB1IP, PRKG1, GP9, PTGS1, ITGA2B, COL1A1, GP6, PLCG2, MAPK3, COL1A2, ACTB, PTGIR, ADCY6, SRC, LCP2, F2RL3, TLN2</td>
</tr>
</tbody>
</table>
3.3 Discussion

These studies suggest that, despite patients with monoallelic mutations in *ETV6* or *RUNX1* having similar phenotypes, the mechanism by which these mutations act to disrupt megakaryopoiesis is different (Glembotsky et al., 2014; Kirkpatrick et al., 2015; Song et al., 1999; Topka et al., 2015). We demonstrate that HPCs and MKs generated from iPSCs harboring mutations in *ETV6* or *RUNX1* show defects when compared to isogenic controls (Figure 3.7).

We explored the *in vitro* phenotype of HPCs and MKs derived from iPSCs harboring a mutation in *ETV6*. Currently, *in vitro* studies analyzing the effect of *ETV6* mutations on megakaryopoiesis have only been conducted in primary human CD34+ cells and HeLa cells, with results showing a broad defect in megakaryopoiesis (Noetzli et al., 2015; Zhang et al., 2015). We found that *ETV6*+/mut and *ETV6*mut/mut iPSCs unexpectedly produce higher yields of HPCs, with a greater ability to expand and differentiate into MKs. The MKs generated, however, are less mature and responsive when compared to the isogenic control iPSC line, and display significant defects in proplatelet formation.

Previous studies utilizing iPSCs derived from RUNX1 patients have mixed results in regards to HPC yield, likely due to the fact that total percentages, rather than absolute yield, were reported (Antony-Debré et al., 2015; Connelly et al., 2014; Iizuka et al., 2015; Sakurai et al., 2014). Reassuringly, all papers report a significant defect in MK generation, but none report on MK maturation or functionality. We observed a reduction in HPCs generated from RUNX1+/mut iPSCs, as well as a severe defect in MK differentiation. However, the MKs generated appear to be more mature, as seen by MK receptor
expression and functional studies, such as MK agonist responsiveness and FV uptake – a late marker of MK maturation.

The differences in ETV6 and RUNX1 mutant iPSC-derived MKs became even more apparent after analyzing the RNA-sequencing data generated from purified MKs. The results show that MKs generated from ETV6+/mut or ETV6mut/mut iPSCs and RUNX1+/mut iPSCs had minimal overlap in differential gene expression when compared to the WT isogenic control. However, all mutant MK populations showed dysregulation in pathways related to platelet activation and transcriptional mis-regulation in cancer, though the individual dysregulated genes in these pathways were mostly distinct between the ETV6- and RUNX1-mutants. Interestingly, both ETV6-mutant and RUNX1+/mut MKs had reduced mRNA expression of platelet-specific genes regulating proplatelet formation and platelet function. These mutant MK populations also displayed severe defects in the ability to form proplatelet extensions, and the complexity of these platelet producing structures was significantly reduced in comparison to the isogenic WT control.

Overall, the experimental studies and RNA-sequencing data suggest that the molecular mechanism driving at least the thrombocytopenia differs in patients with monoallelic mutations in ETV6 or RUNX1. However, despite differing phenotypes, the end result is the same: for ETV6-mutants, there are more, but less responsive, MKs predicted to give rise to fewer platelets per MK, and for RUNX1 mutants, there are fewer, but more responsive, MKs though proplatelet formation and therefore subsequent platelet release may be reduced as well (Figure 3.7).
Figure 3.7: Model. In comparison to the WT isogenic control, \( ETV6 \)-mutant iPSCs generate more HPCs that yield higher numbers of less responsive MKs with lower expression of CD41, CD42a and CD42b, whereas \( RUNX1^{+/\text{mut}} \) iPSCs generate fewer HPCs that yield fewer numbers of more responsive MKs with higher expression of CD41. Both \( ETV6 \)- and \( RUNX1 \)-mutant MKs display defects in the ability to form proplatelet extensions, suggesting that platelet production and release will be poor in comparison to the WT isogenic control. PPF-MK = proplatelet forming MK. Created using BioRender.com.
Figure 3.8: Patient iPSC Line Characterization. (A) Schematic of ETV6 and RUNX1 genes with arrows denoting patient mutations and bars denoting gRNA. (B) Karyotypes of iPSC lines after reprogramming and CRISPR/Cas9 gene editing. (C) Sanger Sequencing of CRISPR/Cas9 edited lines, showing introduction of patient-specific mutations. (D) Representative flow profiles of the stemness surface markers SSEA3, SSEA4, TRA1-60 and TRA-1-81.
Figure 3.9

A

ETV6mutmut  ETV6mut  ETV6+/+  RUNX1mutmut  RUNX1+/+

CD34

CD34

B

CD34⁺CD45⁺ HPC per iPSC

Normalized to Patient iPSC Line

C

CD41

CD235

D

Erythroid Colony Count

Myeloid Colony Counts

ETV6mutmut  ETV6mut  ETV6+/+  RUNX1mutmut  RUNX1+/+

Number of Colonies

ETV6mutmut  ETV6mut  ETV6+/+  RUNX1mutmut  RUNX1+/+

Number of Colonies

E

Erythroid Expansion (CD41⁺CD235⁺)

Myeloid Expansion (CD18⁺CD45⁺)

Fold Change

ETV6mutmut  ETV6mut  ETV6+/+  RUNX1mutmut  RUNX1+/+

Fold Change

F

Megakaryocyte Colony Count

ETV6mutmut  ETV6mut  ETV6+/+  RUNX1mutmut  RUNX1+/+

Number of Colonies

77
Figure 3.9: ETV6 and RUNX1 Mutations Reveal Disparate Lineage Potentials in Patient-Derived iPSC Lines. (A) Representative flow profiles of HPCs. (B) Quantification of fold change in CD34^+CD43^+ HPCs per iPSC plated on day -2; normalized to respective isogenic controls. N = 10. (C) Representative flow profiles of CD34^+CD43^+ HPC lineage biases; erythroid is CD41^+CD235^-^, myeloid is CD41^-CD235^-^. (D) Quantification of erythroid (left) and myeloid (right) colonies after 12-14 days in methylcellulose-based medium. N = 6. (E) Quantification of erythroid (left) and myeloid (right) cells after 5 days of culture; normalized to respective isogenic controls. For erythroid: N = 7 for ETV6 iPSC lines, N = 10 for RUNX1 iPSC lines. For myeloid: N = 8. (F) Quantification of MK colonies after 12 days in collagen-based medium. N = 5. For all statistical analyses, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
Figure 3.10: *ETV6*-Mutant MKs are Fewer in Number and Less Responsive to Agonists than *RUNX1*-Mutant MKs in Patient-Derived iPSC Lines. (A) Representative flow profiles of CD41⁺CD42a⁺ MKs. (B) Quantification of MKs per CD34⁺CD43⁺ HPC after 5 days of culture; fold change normalized to respective isogenic controls. N = 9 for *ETV6*mut/mut; N = 13 for *ETV6*+/mut and *ETV6*+/+; N = 14 for *RUNX1* iPSC Lines. (C) Quantification of MFI of the MK markers CD41 (left), CD42a (middle) and CD42b (right) after 5 days of MK culture; values normalized to respective isogenic control. N = 6. (D) Quantification of PAC1 MFI after MK stimulation with thrombin (left) or ADP (right); fold change normalized to respective isogenic controls. For thrombin: N = 12 for *ETV6*mut/mut; N = 16 for all other iPSC lines. For ADP: N = 6. (E) Quantification of FV uptake over background in day 5 MKs; fold change normalized to respective isogenic controls. N = 5. For all statistical analyses, *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001.
In conclusion, we have used human iPSCs to model a rare disorder resulting from mutations in \textit{ETV6} or \textit{RUNX1} to gain insight into the disease mechanism. Despite similar clinical features, we found disparate phenotypes in the HPCs and MKs derived from \textit{ETV6}- and \textit{RUNX1}-mutant iPSCs. Mutations in \textit{ETV6}, respective to an isogenic control, result in more HPCs that differentiate toward the MK lineage more efficiently, though the resultant MKs are not as responsive to agonist stimulation and fail to take up FV efficiently. In contrast, \textit{RUNX1} mutations result in low HPC generation with poor yield in MKs; these MKs take up higher levels of FV and respond to agonists better than isogenic controls, but display lower proplatelet forming capabilities. Global gene expression changes on purified MKs due to loss of \textit{ETV6} or \textit{RUNX1} is very different, suggesting that these two transcription factors do not function in the same pathway to regulate megakaryopoiesis. Additional future studies investigating the effects of \textit{ETV6} and \textit{RUNX1} mutations on thrombopoiesis are warranted.

4.1 Defects in Platelet Generation and Function

The studies presented in this thesis uncover differing effects of \textit{ETV6} and \textit{RUNX1} haploinsufficiency on megakaryopoiesis. Mutations in \textit{ETV6} improve MK lineage commitment, but block maturation of MKs; whereas mutations in \textit{RUNX1} hinder MK commitment, but do not impede maturation of MKs, though preliminary thrombopoiesis data suggests defective proplatelet formation. Despite these differences in mechanism, the end result of thrombocytopenia can still be explained. Current platelet yields per MK have been described based on in vitro experimental data, as the in vivo studies required
to provide a concrete answer are not possible in human subjects. However, there is extensive evidence suggesting MK maturity to be correlative with platelet production. Therefore, we originally hypothesized that ETV6-mutant MKs will release fewer platelets compared to isogenic controls, and that RUNX1-mutant MKs will release more platelets, considering the increased markers of maturation. However, proplatelet formation on fibrinogen coated coverslips is decreased in both ETV6- and RUNX1-mutant MKs compared to isogenic control MKs (Figures 3.6B and 3.6C), suggesting that mutations in ETV6 and RUNX1 impede normal thrombopoiesis.

Additional studies to support how mutations in ETV6 and RUNX1 affect thrombopoiesis and platelet functionality would be beneficial. Although iPSC technology provides a good system for disease modeling and mechanistic studies of hematopoietic disorders, it does not provide a robust system for generating platelets. An in vivo system has been developed, whereby iPSC-derived MKs can be infused into mice to subsequently release functional platelets into circulation. To perform these experiments, MKs derived from ETV6 and RUNX1 iPSCs would be labeled with fluorescently tagged FV and injected into non-obese diabetic (NOD)/severe combined immunodeficiency (SCID)/interleukin receptor 2 (G)-deficient (NSG) mice. These mice would be bled for 24 hours post-injection, with more frequent bleeds during the first 2 hours. The collected blood samples would be stained with anti-human CD41 to distinguish the iPSC-derived platelets from native mouse platelets. Flow cytometry analysis of CD41 and FV expression would allow us to determine the effects of ETV6 and RUNX1 mutations on the platelet-generating capability of MKs by seeing how many CD41+FV+ platelets are released into the circulation, as well as for how long they circulate.
In addition to examining the efficiency of which ETV6- and RUNX1-mutant MKs can generate and release platelets, the functionality of the released platelets needs to be examined as well. Through an arteriole laser injury model, we can determine if ETV6- and RUNX1-mutant MKs release platelets that contribute to thrombus formation. NSG mice would first be injected with anti-mouse CD41 Fab fragments to visualize the clot, after which FV-labeled MKs would be injected. Laser injury would be induced in the cremaster arterioles of NSG mice and incorporation of the MK-released platelets into thrombus formation would be visualized through confocal imaging. These studies, combining iPSC technology with mouse models, would provide a more physiologic approach to understanding the effects of ETV6 and RUNX1 mutations on thrombopoiesis.

4.2 Mechanistic Differences between ETV6- and RUNX1-Related IT

Despite ETV6- and RUNX1-related IT having very similar clinical presentations, the studies described here suggest that the mechanism of disease resulting from either mutation is distinct. Mutations in a third gene, ANKRD26, also confer an IT with cancer predisposition. In a previous study investigating the molecular mechanism of ANKRD26-related IT, researchers found MAPK hyperactivation to be responsible for the MK phenotype in vitro (Bluteau et al., 2014). Interestingly, Kyoto encyclopedia of genes and genomes (KEGG) analysis from our RNA-sequencing studies revealed hypoactivation of MAPK signaling in ETV6+/mut MKs and hyperactivation of MAPK mediators in RUNX1+/mut MKs.

MAPK signaling is known to play a role in the transcriptional activity of ETV6 and RUNX1: phosphorylation of ETV6 negatively regulates its activity (Maki et al., 2004), while MAPK-mediated phosphorylation of RUNX1 increases its transcriptional activity (Tanaka et al.,
Given the role of MAPK phosphorylation on the transcriptional regulation of ETV6 and RUNX1, along with our RNA-sequencing data and ANKRD26-related IT mechanistic studies, we analyzed this pathway as a potential molecular mechanism for the disparate phenotypes seen in vitro. Preliminary studies using a small molecule inhibitor of MEK1 and MEK2 (PD98059) during MK differentiation showed no significant, or specific, rescue in the MK phenotypes (Figure 4.1). However, we will continue to follow up on MAPK signaling as a potential mechanism driving the phenotypes displayed in ETV6- and RUNX1-mutant MKs by modulating the concentration and timing of PD98059 treatment.

Alternatively, we can use small molecules inhibiting various MAPK signaling mediators to see if HPC and MK phenotypes are rescued by MAPK pathway modulation. In a recent study, the use of a c-Jun N-terminal kinase (JNK) inhibitor, JNK-IN-8, resulted in a quantitative rescue of MK-biased HPCs from RUNX1+/mut iPSCs (Estevez et al., 2020). The use of a second JNK inhibitor, JNK inhibitor IX, did not show the same rescue in phenotype. This highlights the fact that specific arms of MAPK signaling may need to be inhibited for phenotypic rescue. Therefore, we could use the small molecular SP600125, a potent JNK inhibitor, on ETV6- and RUNX1-mutant MKs to see if there is a rescue in the yield and quality of this resultant cell population.

We can also test another arm of MAPK signaling, specifically targeting extracellular signal-regulated kinases (ERKs). The FDA approved drug Ulixertinib is a potent inhibitor of ERK1 and ERK2 (Sullivan et al., 2018). ERK1/2 signaling upregulates transcription factors such as NF-κB, a protein known to be highly dysregulated in AML (Zhou et al., 2015). In addition to treating MKs with this drug alone, we can also try it in combination with the other MAPK inhibitors. Through modulation of these signaling pathways known to be
dysregulated in *ETV6*- and *RUNX1*-mutant MKs, we may be able to rescue the MK and platelet phenotypes. These studies may also provide insight into the mechanistic differences driving the disparate in vitro phenotypes observed from *ETV6*- and *RUNX1*-related IT iPSCs.
Figure 4.1 MAPK Modulation during MK Differentiation of CHOPWT6 Isogenic iPSC Lines

A

CD41 MFI - MAPK modulation

Normalized to WT

WT + DMSO
WT + PD98059
ETV6 +/mut + DMSO
ETV6 +/mut + PD98059
ETV6 mut/mut + DMSO
ETV6 mut/mut + PD98059
RUNX1 +/mut + DMSO
RUNX1 +/mut + PD98059

CD42a MFI - MAPK modulation

Normalized to WT

WT + DMSO
WT + PD98059
ETV6 +/mut + DMSO
ETV6 +/mut + PD98059
ETV6 mut/mut + DMSO
ETV6 mut/mut + PD98059
RUNX1 +/mut + DMSO
RUNX1 +/mut + PD98059

CD42b MFI - MAPK modulation

Normalized to WT

WT + DMSO
WT + PD98059
ETV6 +/mut + DMSO
ETV6 +/mut + PD98059
ETV6 mut/mut + DMSO
ETV6 mut/mut + PD98059
RUNX1 +/mut + DMSO
RUNX1 +/mut + PD98059

B

FV Uptake - MAPK modulation

Normalized to WT

WT + DMSO
WT + PD98059
ETV6 +/mut + DMSO
ETV6 +/mut + PD98059
ETV6 mut/mut + DMSO
ETV6 mut/mut + PD98059
RUNX1 +/mut + DMSO
RUNX1 +/mut + PD98059
Figure 4.1: MAPK Modulation during MK Differentiation of CHOPWT6 Isogenic iPSC Lines. (A) Surface expression of the MK markers CD41 (top), CD42a (middle) and CD42b (bottom) were quantified via flow cytometry after 5 days of treatment with DMSO alone or PD98059. Data normalized to WT. N = 3 (B) MKs incubated with FV with uptake quantified via flow cytometry. Data normalized to WT. N = 3.
4.3 Link between Thrombocytopenia, Inflammation and Cancer

Germline mutations in \textit{ETV6} or \textit{RUNX1} result in a rare IT disorder characterized by thrombocytopenia and hematologic cancer predisposition. Although ETV6 and RUNX1 are known to play an important role in the transcriptional regulation of hematopoiesis and megakaryopoiesis, the exact mechanism leading to these two blood-related phenotypes is unknown. However, it appears that thrombosis, inflammation and cancer are all interrelated, with circulating platelets being the one common cellular element in each of these processes (Franco et al., 2015; Morrell et al., 2014).

Platelet biology and its role in thrombosis and hemostasis is well established. The role of platelets in inflammatory pathways and cancer however, is less defined and appears to be quite complex. In the mid-1800’s, the link between cancer and inflammation was elucidated: the risk of cellular transformation can be enhanced by cell proliferation in an environment characterized by inflammatory cells, growth factors, activated stroma and factors promoting DNA damage (Balkwill and Mantovani, 2001). Today, the causal relationship between inflammation, innate immunity and cancer development is widely accept, but the role of platelets in this process is still a black box.

Over the years, studies have provided strong evidence in support of platelets acting as immune-like cells. One key observation is the expression of Toll-like receptors (TLRs) on the platelet surface, thus giving the ability to directly engage microbial pathogens, similar to leukocytes (Clemetson, 2010). When platelet TLRs are activated through interaction with a microbial species, the cellular fragment degranulates and releases a myriad of proinflammatory mediators including CD40 ligand – a potent secretory molecule that elicits lymphocyte activation (Phipps, 2000; Stark et al., 2012). Additionally, when platelets bind
von Willebrand factor (vWF) on endothelial cells, the interaction encourages recruitment of leukocytes via tethering and rolling on the endothelial surface (Bernardo et al., 2005). Through platelet-mediated leukocyte migration, platelets aid in the clearance of infection by boosting the recruitment of white blood cells (Heinz et al., 2010; Kuckleburg et al., 2011). In addition to TLRs, there are other receptors on the platelet surface which aid in their interaction with microbes: GPIIb-IIIa (CD41/CD61) and GPIb-IX-V (CD42a/b/c/d) (Kerrigan and Cox, 2010; Tilley et al., 2013). Bacterial adhesion via these platelet receptors leads to platelet activation and subsequent release of secondary mediators, such as PF4, to create a positive feedback loop (Brennan et al., 2009; Kerrigan and Cox, 2010; Tilley et al., 2013).

Platelets have also been found to interact with and activate neutrophils. In vivo, neutrophils clear bacteria through the release of neutrophil extracellular traps (NETs) – a web-like network of chromatin, histones and degradative enzymes – which capture circulating microbes (Zawrotniak and Rapala-Kozik, 2013). Interestingly, in vitro experiments showed that neutrophils only induced NET formation in response to lipopolysaccharides (LPS) stimulation when co-cultured with platelets, suggesting that platelets can directly respond to stimuli and activate the immune system (Ma and Kubes, 2007).

RUNX1 is frequently mutated in myeloid malignancies and has been shown to negatively regulate TLR4 (LPS receptor) signaling through nuclear factor κB (NF-κB) signaling (Tang et al., 2017). Recent evidence suggests a link between RUNX1, TLR signaling, neutrophils and inflammation. Hematopoietic-specific loss of RUNX1 in mice was shown to increase the production of proinflammatory mediators, such as tumor necrosis factor-α (TNFα), by bone marrow neutrophils in response to LPS stimulation (Bellissimo, 2020). However,
when \textit{RUNX1} was deleted specifically in neutrophils, the mice no longer displayed the inflammatory phenotype caused by pan-hematopoietic loss of \textit{RUNX1} (Bellissimo, 2020). These data suggest that neutrophils are hyperactivated in the presence of \textit{RUNX1} loss, but another cell type is causal of this phenotype. This brings the platelet into question.

Patients with \textit{RUNX1}-related IT have approximately a 40\% chance of developing myeloid malignancy with a median age of onset of 33 years old. Their heightened risk for cancer is not fully understood, but recent evidence suggests a role for platelets in inflammation and immunity. Therefore, we are asking the question: Are \textit{RUNX1}-mutant neutrophils hyperactivated by \textit{RUNX1}-mutant platelets, and does this inflammatory state predispose FPD/AML patients to AML and other malignancies?

To answer this question, we could generate neutrophils from the \textit{RUNX1}^{+/mut} iPSC line through a CMP intermediate (Choi et al., 2011). Upon neutrophil differentiation, we could stimulate the cells with LPS at various concentrations and look at cytokine production over time through intracellular flow cytometry. If cytokine production is too low to measure, or if there is no difference between wild type and \textit{RUNX1}-mutants, we could co-culture iPSC-derived neutrophils with MKs to see if this increases cytokine production from the neutrophils. As platelets are derived from MKs, they share many of the same receptors, such as TLRs, so it would be interesting to see if MKs can activate neutrophils as well. Additionally, we could utilize a flow chamber assay in which we can look for NET formation from \textit{RUNX1}-mutant neutrophils in the presence of LPS, with and without MKs. Throughout these differentiations, DNA-damage can be examined in the different cell populations (CMPs and neutrophils) through immunofluorescence microscopy, staining for P53 and γH2AX. Additionally, we could subject the CMPs and neutrophils to ionizing
radiation and oxidative stress to see if the RUNX1-mutant cells have impaired DNA repair mechanisms compared to the isogenic control. These experiments, collectively, would provide insight into the role of MKs and platelets in neutrophil hyperactivity resulting from loss of RUNX1 activity. Completing these experiments with the ETV6+/mut and ETV6mut/mut iPSC lines would be interesting as well as no studies have elucidated the link between ETV6, inflammation and immunity.

4.4 ETV6- and RUNX1-Mutant iPSCs for Drug Screening

Traditionally, human genetic disease models have included the use of immortalized cell lines and gene expression systems. However, both of these systems raise concern for faithful disease modeling given their non-physiologic nature (Maqsood, 2013). In addition to this, the use of animal models for human disease pathology and drug discovery also has notable caveats, mainly due to significant species differences (Uhl, 2015 & Lin, 2008). In any case, primary tissue and cells derived from a disease patient would serve as the best model for human disease studies, but accessing such material is often challenging or prohibited. The development of iPSC technology has since allowed for a nearly unlimited supply of patient-derived material for disease modeling and drug discovery.

Therefore, a potential use for the set of CHOPWT6 isogenic ETV6+/mut, ETV6mut/mut and RUNX1+/mut iPSC lines is in drug discovery. Given that these three lines have specific pathogenic mutations in an otherwise wild type genetic background, they could serve as a useful tool to screen for compounds specifically increasing ETV6 or RUNX1 expression, or activity, to wild type levels. If activity levels of ETV6 or RUNX1 could be increased through use of a drug, perhaps the MK and platelet defects would be abrogated. One study examining the RUNX1+/mut iPSC line found two drugs, JNK-IN-8 and RepSox,
targeting the JNK and TGFβ signaling pathways respectively, able to correct quantitative defects in *RUNX1*-mutant HPC generation (Estevez et al., 2020). This finding supports the possibility of druggable pathways for clinical management of thrombocytopenia in *ETV6-* and *RUNX1*-related IT.

In summary, we have highlighted the use of patient derived iPSC lines to gain insight into the mechanism driving the thrombocytopenia phenotype in IT patients with mono-allelic germline mutations in *ETV6* or *RUNX1*. By generating and utilizing genetically engineered sets of isogenic iPSC lines, we uncovered disparate phenotypes in the HPCs and MKs derived from *ETV6-* and *RUNX1*-related IT patient iPSCs. Pathogenic mutations in *ETV6* result in increased HPC and MK generation, though the yielded MKs are less mature in comparison to the isogenic controls. In contrast, mutations in *RUNX1* result in deficient HPC and MK differentiation, though the resultant MKs appear to be more responsive to agonist stimulation and FV uptake; however, preliminary data suggests a defect in proplatelet formation. Further studies will be important to further define the role of ETV6 and RUNX1 in proplatelet formation, platelet release and platelet functionality. Additionally, determining whether there is a link between platelet levels, inflammation and cancer predisposition in these patients will be enlightening for future therapeutic treatments.


Bera, T.K., Liu, X.-F., Yamada, M., Gavrilova, O., Mezey, E., Tessarollo, L., Anver, M.,


cancer predisposition syndrome. Oncotarget 6, 535–538. https://doi.org/10.1038/nt.3253.Although


Murry, C.E., Keller, G., 2008. Differentiation of embryonic stem cells to clinically relevant


Veazey, K.J., Golding, M.C., 2011. Selection of stable reference genes for quantitative


Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T.,


