APOBEC3G INFLUENCES HEMATOPOIETIC LINEAGE COMMITMENT OF HEMATOPOIETIC PROGENITORS DERIVED FROM HUMAN EMBRYONIC STEM CELLS

A thesis submitted in partial fulfillment of the requirements
For the degree of Master of Science in Biology
By
Yesai Sevak Fstkchyan

May 2014
The thesis of Yesai Sevak Fstkhyan is approved:

_________________________________________  ______________
Dr. Jerome A. Zack                          Date

_________________________________________  ______________
Dr. Zoran Galic                            Date

_________________________________________  ______________
Dr. Mary-Patricia Stein                    Date

_________________________________________  ______________
Dr. Cindy S. Malone, Chair                 Date

California State University, Northridge
Dedication

Most graduate students regard themselves as lucky if they have one good mentor throughout their academic career, however, I have been blessed to have had five amazing mentors: Dr. Stein, Dr. Malone, Dr. Nixon, Dr. Galic, and Dr. Zack. Without their guidance and support, I would not be where I am today. The amount of training I received as a CIRM Bridges trainee and during my time in the Zack lab is invaluable. These experiences built a solid foundation that will help me achieve my ultimate career goals.

I would not have had the honor of being in this program if it were not for Dr. Stein’s encouragement. She showed me that science is fun and exciting and that if I worked hard enough, there is not anything I cannot accomplish. She also inspired me to explore a career in academia and has been an excellent mentor. Her enthusiasm for teaching and biology is contagious and I often found myself seeking her energy when difficulties arose. Taking multiple classes with her has been extremely rewarding and enjoyable and I owe my current path to her. I am also grateful for Dr. Malone’s continuous involvement as I progress from one phase of my career to another. Her dedication to the Bridges program and her students never ceases to amaze me; she is a tremendous role model for her students. The CSUN community is truly lucky to have these faculty members. I had no desire or confidence in myself to pursue a PhD after completing my graduate studies at CSUN, however, this all changed after I completed my Bridges internship. This program is unique and has been an amazing resource for CSUN students to understand where they can apply their strengths to make a difference in the world.
This study would not have been possible without Dr. Nixon, who spearheaded the preliminary studies and has worked by my side since day one. This study is based on data collected from Dr. Nixon’s dissertation; we would not have known about any of this if it were not for his diligent work. He has been a great mentor and friend. Dr. Galic has been a key member of this project and has taught me almost everything I know about embryonic stem cells. He is the perfect testament to the power of superior work ethic enabling one to achieve anything. His discipline and passion are traits I hope to adopt. Lastly, I would like to thank Dr. Zack for inviting me to join his lab and giving me my first job in the field. His guidance and support for the past two years have been indispensable and have allowed me to develop into a researcher, equipped with the tools and knowledge to succeed. I am forever grateful to have had these amazing professors and scientists teach, train, and direct me over the last two years.

Aside from professors who have been imperative to my professional development, I would not have been able to continue my educational endeavors without the aid and encouragement from my family and friends. I would like to thank, in no particular order, Torrie Goodson, David Piña, Sharene Kassab, and Jessie Strobel. These individuals have been my rock as a graduate student. Without their love and support, I would not have had the drive to venture on and be the best that I can be. I would also like to thank my parents for allowing me to follow my dreams and standing by my side as I tried to discover my calling.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature Page</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vi</td>
</tr>
<tr>
<td>Abstract</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter I: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II: Materials &amp; Methods</td>
<td>15</td>
</tr>
<tr>
<td>Chapter III: Results</td>
<td>23</td>
</tr>
<tr>
<td>Chapter IV: Discussion</td>
<td>37</td>
</tr>
<tr>
<td>Conclusion</td>
<td>46</td>
</tr>
<tr>
<td>References</td>
<td>47</td>
</tr>
<tr>
<td>Appendix: Figures</td>
<td>51</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1: Vector schematic 50
Figure 2: Knockdown of APOBEC3G in hESCs 51
Figure 3: Pluripotency of transduced hESCs 52
Figure 4: Differentiation protocol 1 54
Figure 5: Hematopoietic potential of EB harvest 54
Figure 6: Colony counts 55
Figure 7: Differentiation protocol 2 56
Figure 8: Developmental stage-specific up-regulation of APOBEC3G 57
Figure 9: Emergence of hematopoietic progenitors 57
Figure 10: LNGFR retention of progenitors 58
Figure 11: Hematopoietic potential of EB harvest 58
Figure 12: Hematopoietic potential of CD43+ population 59
Figure 13: Colony counts 60
ABSTRACT

APOBEC3G INFLUENCES HEMATOPOIETIC LINEAGE COMMITMENT OF HEMATOPOEITIC PROGENITORS DERIVED FROM HUMAN EMBRYONIC STEM CELLS

By

Yesai Sevak Fstkchyan

Master of Science in Biology

Apolipoprotein B mRNA editing enzyme catalytic polypeptide like-3 G (APOBEC3G) is a member of a family of cytidine deaminases that includes activation-induced deaminase (AID), the enzyme responsible for somatic hypermutation in B cells. While APOBEC3G (A3G) is best known for its activity in restricting both retroviral infection and the mobility of genomic retrotransposable elements, there is mounting evidence that A3G possesses a far more involved role in the cell. A3G has been localized to P-bodies, subcellular foci of mRNA editing and microRNA activity, and has been found to directly influence microRNA activity. The Broad Institute Differentiation Map (dMap) is a collection of microarray data that quantifies levels of gene expression in hematopoietic cells throughout their development. According to dMap, levels of A3G expression fluctuate tremendously throughout the process of hematopoiesis. This is inconsistent with a purely antiretroviral factor. We chose to determine whether A3G may be functionally important in hematopoietic development.

Analysis by qRT-PCR of A3G expression at the mRNA level reveals that A3G is expressed in human embryonic stem cells (hESC). Moreover, there is a significant rise in A3G expression when hESCs are induced to differentiate into early hematopoietic
progenitors. Together, these data suggest a role for A3G in the earliest stages of human hematopoiesis. While unpublished data from our lab provides evidence that suggests A3G has a role in intermediate hematopoietic lineage commitment, it is unclear how A3G might influence early hematopoietic development. We, therefore, asked whether A3G plays a role in the development of early hematopoietic progenitors derived from hESCs.

To address our question, we applied a loss-of-function strategy to investigate whether the absence of A3G might impair or alter early hematopoietic development. To accomplish this, we used hESCs as a model system and knocked down A3G. We then characterized the ability of these genetically altered cells to give rise to early hematopoietic progenitors and their hematopoietic lineage commitment towards terminally differentiated cell types of the blood using flow cytometry and methylcellulose colony forming assays. Results from this study have identified a novel function for A3G in that it influences late progenitor fate decisions, but does not alter early development of hematopoietic progenitors demonstrated by impaired erythropoiesis and increased myelopoiesis. This study has uncovered a novel developmental factor previously not thought to be important in human hematopoietic development.
I
Introduction

Prologue

Producing trillions of new cells on a daily basis to replace aging cells, blood is one of the most regenerative tissues of the body (Doulatov et al. 2012). The immense regenerative ability of this tissue type makes it a key target for the field of regenerative medicine and gene therapy. Cells of the blood are responsible for homeostasis of bodily tissue as well as protection against pathogens; therefore, cells of the blood are unique targets in treating hematologic genetic disorders and engineered immunity (Williams 2013). All blood cells come from a common cell, the hematopoietic stem cell (HSC). To develop into a specific blood cell type, the HSC goes through multiple intermediates that branch further into the multiple lineages of the blood. As these intermediates begin to progress through the process of hematopoiesis, the cell fate decisions of these cells is influenced by various factors. This creates a complex system with many components, such as miRNAs, transcriptional networks, and microenvironments, which are imperative to the development of the terminal cell types of the blood (Novershtern et al. 2011). Although much is known about hematopoiesis, there are still many novel factors that have yet to be uncovered. As interest to use the regenerative ability of this tissue increases, we must uncover factors involved in hematopoietic development to ensure that we do not disrupt the natural processes of this system.

APOBEC3G

Apolipoprotein B mRNA editing enzyme catalytic polypeptide like-3 G (APOBEC3G) is part of the APOBEC family of cytidine deaminases that include
APOBEC1, APOBEC2 and activation-induced deaminase (AID) (Jarmuz et al. 2002). AID is required for somatic hypermutation and terminal differentiation of B-cells (Muramatsu et al. 1999). In Humans, APOBEC3 enzymes consist of 7 subsets: A, D, C, DE, G, F and H (Jarmuz et al. 2002). The existence of APOBEC3 protein subsets is thought to be due to evolutionary expansion unique to the primate genome (Zhang et al. 2004). Mice have only one APOBEC3 gene with no subsets. Since its discovery by Jarmuz et al. (2002), APOBEC3G has been primarily classified as an innate defense factor, notably against HIV-1. Ever increasing interest in APOBEC3G and its antiviral activity has made it a necessity to completely understand its full function in the cell.

In 2002, APOBEC3G was identified as a retroviral restriction factor. Due to the action of vif, a 23kDa viral phosphoprotein, HIV-1 is able to productively infect cells despite the presence of APOBEC3G. Vif targets APOBEC3G for proteosomal degradation, thus evading its antiviral activity. APOBEC3G inhibition of HIV infection is only seen with HIV-1 Δvif. Vif targets and inhibits both the function and synthesis of APOBEC3G protein and mRNA. This interaction is critical for vif to mediate the degradation of APOBEC3G. The interaction causes polyubiquitylation of APOBEC3G and vif via E3 ubiquitin ligase complexes. This causes degradation of both proteins by 26s proteasomes. Evidence reveals vif can interact with mRNA of APOBEC3G and inhibit the translation of the mRNA. It has also been suggested that vif may not only be mediating APOBEC3G degradation, but has the ability to compete with APOBEC3G by binding to the viral nucleocaspid and hindering incorporation into the budding virion. Furthermore, vif can cause APOBEC3G to be excluded from areas where the virion is being assembled. Interestingly enough, when APOBEC3G is over-expressed in cells, it
can overcome the inhibitory effects of vif. With APOBEC3G being a potent HIV-1 inhibitor, there has been great interest in developing drug targets to inhibit this vif-mediated degradation of APOBEC3G (Stopak et al. 2003; Sheehy et al. 2003).

APOBEC3G is a 46-kDa protein consisting of 384 amino acids containing a zinc-binding motif (Jarmuz et al. 2002). It has been found in two forms in T cells and macrophages, a low-molecular mass (LMM, 46 kDa) and high-molecular mass (HMM, 5-15 MDa) form, the LMM being the enzymatically active form. The HMM form of APOBEC3G is a large ribonucleoprotein (RNP) complex, the intramolecular interactions of which inhibit its enzymatic activity. Multiple RNAs are bound to the HMM form of APOBEC3G and, when treated with RNase A, can be converted back to the LMM form (Chui et al. 2006).

The enzymatic activity of APOBEC3G has been defined as deamination of the C4 position of the cytidine base in a hydrolytic manner (Soros et al. 2007). This results in a base conversion of dC to dU in DNA. APOBEC3G contains two catalytic domains and can bind to either DNA or RNA (Harris et al. 2002). It also contains a strong cytosolic retention signal and has not been localized in the nucleus. This cytosolic retention is believed to be important in protecting the cellular genome, especially during cellular division (Mangeat et al. 2003). APOBEC3G has been localized to processing bodies (P-bodies), subcellular foci of mRNA editing and microRNA activity and has been found to directly influence specific microRNA activity (Wichroski et al. 2006). It has also been shown to interact with miRNA processing proteins such as Argonaute 1/2, MOV10, and GW182 (Gallos-Montburn et al. 2007). The APOBEC3G gene can be found in one location on chromosome 22 at q13.2 (Jarmuz et al 2002). APOBEC3G, along with B,
DE, and F, have duplicated cytidine deaminase domains located at each end of the protein while A, C, and H have only one domain. Only the c-terminal cytidine deaminase domain is enzymatically active. Interestingly enough, both cytidine deaminase domains are required to effectively restrict HIV-1 Δvi/ infection (Navarro et al. 2005).

APOBEC3G-mediated restriction of HIV infection can be accomplished in two ways: it can act directly on the viral DNA and cause steric hindrance of viral DNA replication or it acts to hypertmutate the viral genome. APOBEC3G has high affinity for single-stranded DNA, which is the favored enzymatic substrate. The deaminase activity of APOBEC3G does not directly affect nascent viral DNA; instead, APOBEC3G is incorporated into the budding virion from the host cell due to its interaction with the Gag polyprotein in the viral nucleocapsid (Alce & Popik 2004). Once the budding virion has reached the target cells and releases viral RNA, the incorporated APOBEC3G is also released. As the viral RNA is reverse transcribed, APOBEC3G binds to the minus strand of newly synthesized viral DNA and induces a base conversion, dC to dU. This causes an introduction of a premature stop codon in the open reading frame (ORF) of the viral DNA in consequence of the dG to dA mutations in viral dsDNA. The newly synthesized viral DNA with the base conversion can either be degraded as ssDNA due to high amounts of uracil present in DNA, which can be excised by uracil glycosylases then degraded by apurinic-apyrimidinic endonucleases as double-stranded DNA, or integrated into the host genome where it will not produce replication-competent virus due to incorrect viral DNA sequence (Yang et al. 2007). Furthermore, increase of uracil incorporation can also hinder DNA synthesis.

RNase H mediates the activation of APOBEC3G enzymatic activity in the budding
virion resulting from degradation of viral RNA after ssDNA is synthesized, providing APOBEC3G with its favorite substrate. This event also causes the removal of the RNA that is inhibiting APOBEC3G enzymatic activity in HMM RNP complexes. Very few APOBEC3G molecules are needed to sufficiently restrict viral DNA. It has been seen that as few as seven molecules are sufficient (Xu et al. 2007). Most of the cellular APOBEC3G exists in the HMM form and very little LMM can be found within activated T cells. The HMM APOBEC3G incorporated into the budding virions primarily come from newly synthesized enzymes and happen at a fairly rapid rate within 30 minutes. APOBEC3G interaction with viral RNA can also cause inhibition of reverse transcriptase to synthesize viral DNA (Schafer et al. 2004).

Aside from APOBEC3G’s ability to restrict HIV-1 Δvif, it can function as a restriction factor for the mobility of Alu and LINE-1 retrotransposable elements. These retrotransposable elements account for about 17% of the human genome and have been seen to progress many human diseases. Retrotransposons can be responsible for mutations in the genome, which are inserted into the genome via reverse transcription of their mRNA. They may also contain internal promoter sites that can disrupt gene expression. Retrotransposons require the mobility and function of two proteins: ORF1p and ORf2p. Other APOBEC3 family members, such as APOBEC3A and APOBEC3B, have the ability to enter the nucleus where they can interfere with LINE-1 retrotransposition directly, whereas APOBEC3G can only interfere with the RNA of these retrotransposable elements due to being exclusive to the cytosol. It has been shown that APOBEC3G has very little effect in the inhibition of LINE-1 retrotransposition, but can greatly inhibit LINE-1 mediated Alu retrotransposition. It is thought that hY and Alu
RNAs are a natural target for APOBEC3G and do not require its cytidine deaminase activity for retrotranposition restriction. These RNAs are confined to RNA-transporting granules by APOBEC3G in its HMM form. APOBEC3G forms these HMM by recruiting these RNPs into these granules. This in turn takes essential RNAs away from the LINE-1 machinery from entering the nucleus. These granules are Staufen-containing and thought to be the mechanism of APBOEC3G restriction of retrotranpoable elements. RNA granules, which APOBEC3G assembles into its HMM, contain both 40s and 60s ribosome subunits and are not found to assemble in processing bodies. Furthermore, the function of APOBEC3G in human embryonic stem cells on LINE-1 retrotransposition was examined. In the pluripotent state of human embryonic stem cells, endogenous LINE-1 retrotranpostion is increased. It was found that APOBEC3G had no effect on LINE-1 retrotransposition in human embryonic stem cells or in cell lines. However, another APOBEC3 family member, APOBEC3B, was efficient in restricting endogenous LINE-1 retrotransposition in both cell lines and human embryonic stem cells (Wissing et al. 2011).

**Human Hematopoiesis**

Human hematopoiesis, though complex, has been extensively studied in both mice and humans. The increasing importance of studying blood development has been attributed to the fact that many therapies are being developed to treat patients with various diseases and disabilities due to its highly regenerative properties. Hematopoiesis begins with the hematopoietic stem cell (HSC). This highly specialized stem cell has the ability to give rise to all cell types of the blood. It is derived in the bone marrow in adults and can migrate to various locations of the body to mature and commit into specialized
terminally differentiated cells. HSCs give rise to both myeloid and lymphoid lineages of the blood and are characterized by specific cell-surface markers that allow for isolation of these cells. The true HSC is characterized as CD90+CD34+CD38-Lin-. Isolation of this cell population allows for reconstitution of the immune system from its ability to engraft in bone marrow while being able to self-renew and proliferate (Doulatov et al. 2012; Hoang 2004).

The HSC cell population can be broken down further into two types, the long-term (LT-HSC) and short-term (ST-HSC). The LT-HSC has the ability to engraft and self-renew indefinitely 12-weeks after it has been transplanted into the bone marrow. The ST-HSC gives rise to the various cell types of the blood, but cannot self-renew and proliferate 12-weeks post transplantation. Conventionally, HSCs are isolated from human peripheral blood mononuclear cells (PBMC) and bone marrow where they reside. Commitment of HSC towards specialized cell types is achieved through a more committed intermediate known as the multipotent progenitor (MPP). The MPP population is distinguished from HSCs by the loss of CD90 expression. This MPP can then further branch out into various cell types (Doulatov et al. 2012).

Human hematopoietic hierarchy branches into two lineages from the MPP, myeloid and lymphoid. The myeloid lineage includes granulocytes, macrophages, erythrocytes, and megakaryocytes whereas the lymphoid lineage primarily gives rise to T cells, B cells, Dendritic cells, and NK cells. Lymphoid lineages arise from a common lymphoid progenitor (CLP), which can be distinguished by expression of CD34 and CD127. The common myeloid progenitor (CMP) gives rise to those cells of the myeloid lineage and is distinguished by CD34+CD38+CD135+ Lin- profile. The MEP then
branches off into two separate progenitors, the granulocyte-macrophage progenitor (GMP) CD34+CD38+CD135+CD45RA+Lin- and the megakaryocyte/erythroid progenitor (MEP) CD34+CD38+CD135+CD45RA-Lin-. Lineage specific markers, often referred to as Lin+ or Lin-, are expressed on terminally differentiated cell types (Doulatov et al. 2012; Hoang 2004; Laiosa et al. 2006; Mercer et al. 2011). These lineage specific markers include the thrombopoietin receptor (CD 110) and CD45RA (Ziegler et al. 1999).

Even though hematopoiesis from HSCs is well characterized, hematopoietic development from human embryonic stem cells (hESC) is still vague and does not follow the classic intermediates that will give rise to the terminal cell types of the blood. This is primarily due to the fact that hESCs can either undergo primitive or definitive hematopoiesis. Differentiation of hESCs into hematopoietic progenitors requires both the correct sequential activation of transcriptional factors and microenvironment signaling to support hematopoietic commitment. Since blood is a mesodermal derivative, hESC hematopoietic differentiation requires induction of mesoderm. Primitive and definitive hematopoiesis and can be distinguished by emergence of T lymphocytes from these two populations. Cells programmed for primitive hematopoiesis can be characterized by the coexpression of CD34+ and CD43+, which primarily give rise to erythrocytes, macrophages, and megakaryocytes, whereas definitive hematopoiesis is characterized as CD34+43- population that gives rise to previously mentioned cells types as well as T lymphocytes and granulocytes. Within the primitive hematopoietic population, expression of CD235a and CD41a are indicative of erythroid and myeloid lineage commitment potential, respectively (Kennedy et al. 2012). Furthermore, previous reports
have shown that cell types arising from hematopoietic differentiation of hESC on OP-9 murine stromal layer express a CD34+43+235a+41a+/- phenotype. These cells are known to be early hematopoietic progenitors, which are predetermined to give rise to cells of the myeloid lineage including erythrocytes and megakaryocytes. (Vodyanik et al. 2006).

Although specific cellular phenotypes have been described that shy away from the classical CD34+ phenotype used to isolate HSCs from human umbilical cord blood and fetal liver, there have been reports that show CD34+ cells derived from hESC can give rise to T lymphocytes in vivo (Galic et al. 2009). CD34+ cells were generated in vitro and transplanted to a SCID-hu mouse where they underwent further development. This study offers a basis to which classical hematopoietic development can be achieved both in vitro and in vivo. T lymphocytes arise from the same hematopoietic system as the myeloid lineage cell types; therefore we can assume that isolation of CD34+ cells from hematopoietic differentiation of hESC will also give rise to cell types of the myeloid lineage. However, upon further investigation, it has been found that CD34+ cells derived from hESC are primarily hemogenic endothelium and have low expression of CD45 (Vodyanik et al. 2006). The majority of these cells are incapable of developing into the various cell types of the blood and thus require further development to become true hematopoietic progenitors.

**Human Embryonic Stem Cells**

Human embryonic stem cells (hESC) hold enormous potential in medicine, but there are still many hurdles that we must overcome before providing patients with stem cell-based therapies. Since the isolation of human embryonic stem cells in 1998, the scientific community has been vigorously working with these cells to understand human
development and provide regenerative medicine to those in need (Thomson et al. 1998). The unique characteristic of human embryonic stem cells is their pluripotency, their ability to differentiate into cell types from all three primary germ layers: endoderm, mesoderm, and ectoderm. These cells are able to self-renew and proliferate indefinitely in culture. Pluripotent stem cells have become an ever-increasing interest in the field of medicine and have been a key tool in the field of developmental biology (Chen et al. 2008; Thomson et al. 1998).

Even though we are able to differentiate hES cells into different cells types from all three germ layers in vitro and in vivo, we still do not understand the underlying mechanisms in the differentiation of hESC (Murry and Keller et al. 2008). Therapies that use adult stem cell transplantation, such as hematopoietic stem cell transplants (HSCT) to reconstitute the immune system to treat certain types of cancers have already been used in patients, but these cells are very hard to isolate and not enough cells can be isolated to fully reconstitute the human immune system (Doulatov et al. 2012). Furthermore, finding donors that match patient-specific human leukocyte antigen (HLA) is a very tedious and daunting task (Scheper and Copray 2009). Human ESC hold great potential because they can provide patients with a sufficient number of cells that can be expanded and differentiated in vitro then transplanted into the patients to regenerate tissues and organs. Induced pluripotent stem cells (iPSC) are patient-specific and can potentially circumnavigate the HLA problem as well as help mitigate the graft-versus-host issues encountered with HSC stem cells transplants. Cellular reprogramming to a stem cell-like state is achieved by the introduction of three to four pluripotency factors; Oct-4, Sox-2, klf4, and myc-c (Scheper and Copray 2009). With the discovery and advancement of
induced pluripotent stem cells, the scientific community makes a leap toward personalized regenerative medicine, but first we must understand the underlying mechanisms and elements that influence hESC differentiation and cellular reprogramming.

Human embryonic stem cells can be characterized by their unique expression of pluripotency factors. These factors include both transcriptional factors and cell surface proteins. Oct-3/4 and Sox2 are the transcription factors that are well known to work together and regulate the stem cell like state of hESC. Oct-3/4 (POU5F1) is essential for the self-renewal capacity of hESC. Oct-3/4 expression must be perfectly balanced to regulate self-renewal. Fluctuation of Oct-3/4 expression can cause hESCs to begin to leave the stem cell state and differentiate into specialized cell types. Sox2, along with Oct-3/4, work together with Nanog to transcribe pluripotency genes and regulate the leukemia inhibitory factor (LIF) signaling pathway to control the pluripotency circuitry (Bohler 2009). Specific markers expressed on the cell surface can also confirm pluripotency of stem cells. Human embryonic stem cells highly express stage-specific embryonic antigen (SSEA) 3 and 4 when in the stem cell state. Differentiation reduces the expression of these markers and expression of SSEA 1 increases. High expression of TRA-1-81 and TRA-1-60 also indicates hESC pluripotency (Bohler 2009; Thompson et al. 1998).

**APOBEC3G & Hematopoiesis**

Despite the fact that APOBEC3G has been greatly studied in the context of HIV infection and LINE-1 retrotransposition, unpublished data from our lab suggests that it may also play a bigger role due to differential expression levels of APOBEC3G in the
development and differentiation of early and intermediate hematopoietic progenitors into the different lineages of cells in the blood. Many groups provide data that shows APOBEC3G plays a bigger role in the cell (Izumi et al. 2013). The Broad Institute differentiation map portal (dMap: http://www.broadinstitute.org/dmap/home; Novershtern et al. 2011) is a database of micro array data that quantifies levels of gene expression in cells through their development. Data on APOBEC3G expression levels shows that it is differentially expressed in certain cell types of the blood, in other words, as hematopoiesis progresses, A3G is either up-regulated or down-regulated in cell intermediates, progenitors, as well as terminally differentiated cell types. In agreement with the dMap database, experiments show that the long term and short hematopoietic HSC do not express APOBEC3G. As the hematopoietic stem cell begins to restrict itself into to the MPP, APOBEC3G levels still do not fluctuate. As the MPP begins to commit to the different lineages of the hematopoietic pathway, APOBEC3G levels begin to dramatically fluctuate. APOBEC3G is highly expressed in the CMP and MEP cell populations but not in the GMP. Moreover, A3G has been colocalized with multiple P-body proteins in hematopoietic progenitor cells by confocal microscopy. This differential expression of APOBEC3G in progenitors in the hematopoietic pathway leads us to believe that APOBEC3G is involved in hematopoietic differentiation and lineage commitment.

The involvement of APOBEC3G in influencing hematopoietic commitment and hematopoietic cell fate was further confirmed by functional analysis experiments in HSCs. In these experiments, APOBEC3G expression was diminished using shRNA. Human hematopoietic stem cells are transduced with lentiviral vectors containing these
shRNA and differentiated into terminal blood cell types. Examination of colony-forming units (CFU) derived from these HSCs containing APOBEC3G, shRNA-containing vectors revealed an apparent pattern in the lineage commitment of these progenitors. Specifically, in cells with the APOBEC3G knockdown vectors, erythroid and megakaryocyte colonies were reduced by 50%. Moreover, myeloid colonies containing a mixture of granulocytes and macrophages increased. This is indicative of APOBEC3G influencing the lineage commitment of the common myeloid progenitor. When APOBEC3G was over-expressed in CD34+ cells, no colonies were formed. Knockdown of APOBEC3G in sorted progenitors such as the HSC, CMP, GMP and MEP showed similar impairment of erythroid colony formation. Further analysis of gene expression using DNA microarrays when APOBEC3G was knocked-down in the CMP population revealed that there was severe impairment of important hematopoietic genes, of note the erythropoietin receptor, erythropoietin receptor, IL-6 receptor, and KLF-4. These genes are imperative in the development of the common myeloid progenitor in its commitment toward erythrocytes, and megakaryocytes (Feinberg et al. 2007; Sato et al. 2000; Tenen et al. 1997).

Together, these data suggest a role for APOBEC3G in the earliest stages of human hematopoiesis. While there is evidence that suggests APOBEC3G has a role in intermediate hematopoietic lineage commitment, it is unclear how APOBEC3G might influence the development of hematopoietic progenitors from human embryonic stem cells, however, it has been shown that APOBEC3G is expressed and does not mediate the restriction of retrotranspositional elements in hESCs; knockdown studies have shown that APOBEC3G is not an important regulator of hESC pluripotency, but were incomplete
(Wissing et al. 2011). We therefore asked whether APOBEC3G plays a role in the early development of hESC into hematopoietic progenitors.
II

Materials and Methods

Cell Culture

H1 Human embryonic stem cells (hESC) were cultured on gelatin (0.01% gelatin) coated plates that were coated overnight in a 37°C 5% CO\textsubscript{2} incubator. Feeder cells, CF-1\(\gamma\)-irradiated mouse embryonic fibroblast (MEF), were seeded at a density of 1.8x10\(^6\) cells per 6-well plate. Cells were cultured in human embryonic stem cell (HES) medium consisting of Dulbecco’s modified eagle’s medium with Ham’s F-12 (DMEM/F-12, Gibco). The medium was supplemented with 20% knockout replacement serum (KSR, Invitrogen), 0.01mM \(\beta\)-mercaptoethanol (Sigma), 100µM MEM non-essential amino acid (Gibco), 2mM Glutamax (Gibco), and 10ng/ml basic fibroblast growth factor (Invitrogen). Cells were passaged either once a week or when cells became too dense to remain pluripotent using collagenase IV (Gibco) at a concentration of 1mg/ml in DMEM/F-12 to dissociate colonies. Cells were then scraped off plates using 10ml glass pipettes (Fisher). Cells were washed twice with HES media and spun down for 5 mins at 1000x RPM. Cell suspension were then put through a 40 -µm cell strainer and seeded on MEF plates. Cells were then expanded for transduction and subsequent embryoid body formation.

Transduction and Derivation of hES Cell Lines

Pseudotyped lentiviral vectors for hESC transduction were prepared by co-transfection of the HEK293FT cell line with Lipofectamine 2000 (Invitrogen) following manufacturer protocol using VSV-G envelope, Δ8.2 packaging vector, and one of the three vectors designed to manipulate or control for A3G expression. Supernatant was
collected 48 hours later and concentrated to a titer of $1 \times 10^8$ infectious units per ml. Infection of hESCs were done on day 6 since last passage using one 6-well plate of confluent cells. Cells were enzymatically dissociated using collagenase IV. Cell suspensions were spun down and washed with media twice. The suspension was then put through a 100µm cell strainer followed by a 40µm filter. Top of the 40µm cell strainer was washed off with media to give us colonies of optimal size for infection; colonies that are between 100 µm and 40µm in size. Filtered cell suspensions were then spun down and re-suspended in 300µl of media. Using a 1:1 ratio yielding a total volume of 300µl, 150µl cell suspension and 150 µl of tittered virus, lentivirus was mixed in and put on a rocker in a 37°C 5% CO$_2$ incubator for 2 hours. Cell suspensions with virus were then plated on irradiated puromycin resistant DR4 MEF’s at density of $1 \times 10^6$ cells/plate. To select for transduced hESC colonies, HES medium was supplemented with puromycin (Sigma) at a concentration of 1µg/ml. Colonies were maintained in drug selection for 2 weeks then individual colonies were selected mechanically. These individual colonies were expanded and maintained in HES media supplemented with puromycin for an additional 2 weeks. Transduced cells were then plated and expanded on CF-1 MEFs.

**Differentiation and EB formation of hESC**

Embryoid bodies were formed using the scrap method. Cells were treated with dispase (Gibco) at a concentration of 0.05mg/ml in DMEM/F-12 for 30 minutes, then colonies were gently lifted from the plate using a glass pipette. Cells were then washed twice with HES media and spun for 2 minutes at 100xRPM using EB media which consists of Iscove’s Modified Dulbecco’s Medium (Gibco) supplemented with 15% FBS (Gemcell), 0.01mM β-mercaptoethanol (Sigma), 100µM MEM non-essential amino acid
(Gibco), 2mM Glutamax (Gibco), 1% 100x pen/strep (Gibco). Lifted colonies were plated on 6-well ultra-low attachment plates (Corning) and the medium was changed every other day. Cell culture plates were changed depending on whether the EBs have attached to the bottom of the plate, typically the first 3 media changes. On day 3 of EB formation 10ng/ml Bone morphogenic protein-4 (BMP-4) (Invitrogen), 50 ng/ml of stem cell factor (SCF) (Invitrogen), 20ng/ml of flt3 ligand (flt3l) (Invitrogen), and 30ng/ml of vascular endothelial growth factor (VEGF) (Invitrogen). BMP-4 was removed from culture on day 12. After 14 days of differentiation EB cultures were harvested on day 15. Cultures were washed 3 times with 1x phosphate-buffered saline (PBS), and then digested using 0.25% trypsin-EDTA (Gibco) with 2% chicken serum. EBs were digested for 30 minutes in 37°C water bath while swirling the mixture every 5 minutes. Harvests were then mechanically broken and washed twice with EB media then put through a 40 - µm mesh filter. The resulting suspension was spun down for 2 minutes at 1000x rpm and re-suspended in 2ml of EB media.

For CD43 early hematopoietic progenitor differentiation of EBs 10ng/ml bone morphogenic protein-4 (BMP-4) (Invitrogen), 5ng/ml basic fibroblast growth factor (bFGF)(Invitrogen), 0.3ng/ml Activin A (Invitrogen), 15ng/ml of vascular endothelial growth factor (VEGF) (Invitrogen), 50 ng/ml of stem cell factor (SCF) (Invitrogen), 20ng/ml of flt3 ligand (flt3l) (Invitrogen), 30ng/ml thrombopoetin (TPO)(Invitrogen), and 2U/ml erythropoietin (EPO)(Invitrogen) were supplemented into the differentiation media as described above. Figure 7 shows the time line of cytokine additions starting from day 0, the day EBs were formed, until day 11 when EBs were harvested. EB harvest and media changes methods remained the same as it is described above.
Methylcellulose Colony Forming Assay

Methylcellulose colony forming assays were carried out in two different ways; EBs that were differentiated in the protocol illustration by figure 4 were harvested and plated at a density of 100,000 cells/plate for each conditions in triplicate. EBs that were differentiated in the protocol illustrated in figure 7, were sort based on expression on early hematopoietic markers, CD43. Cells were sorted based on the expression of CD43 using magnetic-activated cell sorting (MACS). CD43 direct beads (Miltenyi) and LS columns (Miltenyi) were used to sort CD43+ cells. Both positive and negative fractions were plated in complete methylcellulose (Methocult-H4435, Stem Cell Technologies). Colony forming assays in methylcellulose were used to evaluate the hematopoietic potential of vector-transduced cells. For each condition, 10,000 cells were seeded in each of three 100mm grid plates. Plates were maintained in a 37°C 5% CO₂ incubator for two weeks in a humidified environment. To score colonies, a Nikon microscope was used at a magnification of 400x and counted in a blinded fashion.

Western Blot Analysis

hES cells were collected from wells using collagenase IV and washed once with hES media. Cells were pelleted and lysed using 2x SDS lysis buffer. Bradford assays (Biorad) were used to determine protein concentrations. Cell extracts were mixed with loading buffer and 50 µg of protein. The mix was boiled at 100°C for 3-4 mins. Extracts were loaded on 8% polyacrylamide loading gel and run on 10% polyacrylamide resolving gel next to the 10 µl Benchmark™ protein ladder (Invitrogen). Gels ran at 100 volts for 15 mins then voltage was increased to 150 volts for 1.5 hours. Proteins were then transferred onto nitrocellulose membranes. Membranes were blocked using PBS
supplemented with 5% what does bovine serum albumin (BSA) and incubated with primary antibodies, purified mouse anti-Oct 3/4 (human) (BD Pharmingen) and purified mouse anti-sox 2 (BD Pharmigen), with a 1:1000 dilution overnight at 4°C. Membranes were washed and secondary antibodies, goat-anti mouse IgG Texas Red conjugate (Santa Cruz Biotechnology) were added the next day at the same dilution for an one hour incubation. Membranes were imaged using ECL™ plus and Typhoon phosphoimager scanner (GE Healthcare Biosciences).

**RNA isolation and qRT-PCR analysis to detect APOBEC3G mRNA levels**

Quantitative real-time PCR analysis was carried out in two steps: first cDNA was synthesized from isolated RNA and secondly cDNA was amplified to detected mRNA quantities using real-time amplification measurements. RNA was isolated from hESCs and time-point EBs using RNeasy Mini kit (Qiagen) per manufacturer protocol, RNA was eluted using nuclease free water and quantified by a NanoDrop 2000c spectrophotometer. To generate cDNA, a High Capacity cDNA Reverse Transcription Kit by Applied Biosystems (Invitrogen) was used; 20µl reactions were set up containing 2.0µl of 10x Reverse Transcription Buffer, 3.2µl of Nuclease-free H₂O, 0.8 µl of 100mM 25x dNTPs, 2.0µl of 10x random primers, 1.0µl of RNase inhibitor, 1.0µl of Multiscribe™ Reverse Transcriptase, and 10µl of RNA template per reaction. Reverse transcription of template RNA to synthesis cDNA was conducted using an Eppendorf Mastercycler gradient in three thermocycler steps: first 10 minutes at 25°C, then 30 minutes at 48°C, and finally 5 minutes at 95°C. Once synthesis of cDNA was completed, the products were quantified using a NanoDrop 2000c spectrophotometer and dilutions were made to yield a working concentration of 100ng of cDNA per 5µl. qRT-PCR
reactions were set-up in 20µl reactions containing 10µl of 2x Taqman® RT buffer by Applied Biosystems (Invitrogen), 1µl of validated primer/probe mix from Applied Biosystems (Invitrogen), and 4µl of Nuclease-free H2O per reaction. Reactions were then seeded in a 96-well TempPlate Semi-Skirt 0.2mL PCR plate (USA Scientific) with 15µl of the previously described master mix and 5µl of cDNA template. cDNA from each sample was analyzed in triplicate on a BioRad IQ5 cycler consisting of 45 cycles: a 30 minute cycle at 50˚ C, then a 10 minute cycle at 55˚ C, then a 15 minute cycle at 95˚ C, then 40 cycles of 15 seconds at 95˚ C, then a 30 second cycle at 55˚ C, and finally a 1 minute cycle at 60˚ C. Primers and probes for APOBEC3G mRNA are as follows: APOBEC3G forward primer: CGCAGCCTGTGTCAGAAAAG, APOPBEC3G reverse primer: CCAACAGTGCTGAAATTCGTCATA, APOBEC3G probe: 6-FAM-CCGGTGCCACCATGAAGATCA-BHQ. Copy numbers were normalized to either the amount of cDNA per reaction or to GAPDH standards.

Flow Cytometry

Both EB harvest and hESC were stained using the same protocol. EB harvests were stained after being treated with the digestion buffer and removal of cellular debris. hESC were harvest by being treated with actuaase to obtain a single cell suspension. Cells were spun down and re-suspended in 100ul of PBS supplemented with 10% human AB serum. Specific antibodies were added in the amount indicated by the manufacturer for one test. Cells were incubate in human AB serum and antibodies for 30 minutes in 4ºC. Cells were then washed twice with FACS buffer consisting of 1X PBS supplemented with 1% FBS twice and then re-suspended in 4% paraformaldehyde (PFA). Acquisition of data was done using the BD Biosciences LSRFortessa cell analyzer at the Broad Stem
Cell Research Center flow cytometry core. The following antibodies were used: anti-CD271 AF647 (BD Pharmingen), anti-SSEA-4 PE-CF594 (BD Horizon), anti-TRA-1-60 BV421 (BD Horizon), anti-SSEA-1 PE (BD Pharmingen), anti-CD34 PE-CY7 (eBioscience), anti-CD45 AF700 (AbD Serotec), anti-CD41a FITC (BD), anti-CD43 APC (BD Pharmingen), anti-CD235a PE Cy5.5 (BD Pharmingen), and anti-CD271 BV421 (BD Pharmingen). Acquired data were analyzes using FloJO 10.0.6 (Tree Star Inc.).

**Teratoma Assay**

To determine if transduced hESC were able to differentiate into all three germ layer, teratomas were formed and histologically stained. Cells were prepared for injection as follows; 3 confluent wells of hESC colonies (day 6 post passage) were incubated with collagenase IV (1mg/ml in DMEM/F-12) for 10 minutes. Once colonies began to peel off, a 10ml glass pipette was used to scrap cells off the plate. Cells were washed twice in HES media. Media was then aspirated and re-suspended in 50ul of BD Matrigel™ (BD Biosciences). Cells were placed on ice until they were injected. Injections were done in each testicle of severe combined immunodeficient (SCID/Beige) mice. Teratomas were allowed to grow for 6 weeks and then harvested. Teratomas were fixed in 4% PFA for 24 hours in room temperature. PFA was then replaced by 70% ethanol and put into tissue cassettes and submitted to UCLA Translational Pathology Core Laboratory (TPCL) for histological sectioning and H&E staining.
Statistical Analysis

All statistics were calculated using Prism software (GraphPad). Unpaired student t-tests were calculated to assess significance of data. A p-value < 0.05 was considered significant.
Results

Stable Transduction of Human embryonic Stem Cells

To assess the role of APOBEC3G in hESCs and early hematopoietic development we employed a loss-of-function strategy to investigate any variations in cellular phenotypes caused by diminished APOBEC3G expression. Two vectors were prepared, one as a control and another to knockdown APOBEC3G. Both vectors were constructed using an FG12 base and are driven by a CMV promoter for expression following integration into the hESC genome (Fig. 1). The H1 promoter drives one of two constructs, shRNA against A3G (Fig. 1a) or a scramble control (Fig. 1b). The scramble control is a randomized shRNA sequence, which allows for control over vector toxicity, off-target gene silencing of our shRNA sequence, and transduction selection. The ubiquitin C promoter is upstream of a delta low-affinity nerve growth factor receptor (dLNGFR) reporter gene with a truncated cytosolic end, disabling any intracellular signaling, and a puromycin selectable marker, which are separated by an internal ribosomal entry site (IRES) that will allow for translation of two proteins off a single mRNA. These vectors were packaged into a vesicular stomatitis virus G protein envelop (VSV-G) pseudotyped virus for hESC transduction. Human embryonic stem cell colonies were infected then plated on DR-4 puromycin resistant mouse embryonic fibroblasts (MEF). To confirm successful transduction and vector integration of hESC, LNGFR expression was measured by flow cytometry after four weeks of puromycin drug selection. Both scramble control and shRNA transduced hES colonies show high expression of dLNGFR when compared to non-transduced H1 hESCs (Fig. 2), indicating
successful transduction of hESC colonies and stable expression of the vectors.

Transduction of hESCs with the scramble control vectors were 96.5% positive for LNGFR (Fig. 2a) expression and those that were transduced with the APOBEC3G shRNA vector were 98.4% positive (Fig. 2b). To further confirm that APOBEC3G was knocked down in transduced hESCs, qRT-PCR analysis was conducted to measure APOBEC3G mRNA levels; cDNA was synthesized from isolated RNA and then subsequently amplified to measure APOBEC3G mRNA copy number (Fig. 2c). Non-transduced hESCs were used as a control to quantify unaltered endogenous APOBEC3G expression levels and to test whether our scramble control altered APOBEC3G mRNA expression. Messenger RNA copy numbers were normalized to the amount of RNA seeded per well, in quantities of 100ng cDNA per well. Endogenous APOBEC3G is expressed at 0.08 Copies/100ng cDNA in non-transduced hESCs with similar expression levels in hESCs transduced with the scramble control vector, however, hESCs that were transduced with the shRNA vector for APOBEC3G knockdown expressed 0.02 copies of A3G/100ng of cDNA, indicating a 4-fold decrease in APOBEC3G mRNA compared to both non-transduced and scramble control hESCs. Decreases in mRNA levels in the knockdown group were statistically significant (p<0.05). No statistically significant change in APOBEC3G mRNA levels were seen when comparing scramble control hESCs to non-transduced hESCs, which indicates that expression levels were not altered due to transduction and that the scramble sequence does not mediate silencing of APOBEC3G expression. Both expression of our reporter gene, LNGFR, and a significant decrease in APOBEC3G mRNA in hESCs transduced with APOBEC3G shRNA vector is indicative of successful transduction of hESCs and knockdown of APOBEC3G.
Human Embryonic Stem Cells Retain Pluripotency After Transduction

The uniqueness of human embryonic stem cells is their ability to self-renew and proliferate indefinitely. This is due to the action of two pluripotency genes that are well known to be essential in the complex regulatory network of hESC self-renewal and proliferation, Oct 3/4 and Sox 2. Although Oct 3/4 and Sox 2 are not the only genes involved in hESC pluripotency, it has been shown that down-regulation of these transcription factors alters the pluripotent state of hESCs immensely (Bohler 2009). The expression of these two transcription factors is essential in maintaining the pluripotent state of hESCs. We tested for Oct 3/4 and Sox 2 protein expression in the transduced hESC cell lines to confirm that manipulation of A3G expression does not impair expression of pluripotency markers (Fig. 3a). When assayed by Western blot, expression levels of the Oct3/4 protein, which is 45 kDa in size, were consistent in the scramble and siRNA transduced hES cell lines. The same can be said of Sox 2 expression, which is 35kDa in size. Furthermore, we tested for expression of stage specific embryonic antigen-1 (SSEA-1), this marker indicates the onset of differentiation and deviance from the pluripotent state (Fig 3b), by flow cytometry. There was very little to no expression of SSE-1 when compared to our non-transduced H1 hESC control, however, some expression of SSE-1 can be seen: non-transduced 0.11%, scramble 2.67%, and siRNA 0.43% of cells are SSEA-1 positive. This is due to cell culture conditions; it can be expected to see some differentiated cells within the cultures due to long-term maintenance in a laboratory setting (Baker et al. 2007). Nonetheless, low expression of SSE-1 indicates that transduced hESCs are remaining in an undifferentiated state. Further analysis of additional human embryonic stem cell surface markers revealed that
transduced colonies highly expressed these markers (Fig. 3c). SSEA-4 and TRA-1-60 expression on human embryonic stem cells is indicative of the pluripotent state of these transduced colonies. When compared to non-transduced H1 hESCs, transduced colonies showed similar high expression of both SSEA-4 and TRA-1-80 as seen by flow cytometry, 99% of cells are double positive for these markers for all conditions. Lastly, teratoma assays were used to confirm the ability of transduced hESC to form all three germ layers (Fig. 3d). All three embryonic germ layers are present in teratoma sections that were H&E stained. Neural rosettes can characterize emergence of ectoderm, mesoderm by the formation of immature bone cartilage, and endoderm by the presence of gut endothelium. Retention of pluripotency markers and the ability of transduced hESC to form all three embryonic germ layers indicates that knockdown of APOBEC3G does not impair the pluripotency of hESCs in agreement with previous studies (Wissing et al. 2011).

**Emergence of Hematopoietic Progenitors from Embryoid Body Cultures**

Human embryonic stem cells were differentiated using the embryoid body (EB) method in culture conditions shown in figure 4. To induce hematopoietic progenitor differentiation, multiple cytokines were added to the differentiation media (Galic et al. 2009). Addition of bone morphogenic protein-4 (BMP-4) early on in EB culture allowed for mesodermal differentiation. Media was also supplemented with stem cell factor (SCF), flt3 ligand (flt3l), and vascular endothelial growth factor (VEGF) to push for hematopoietic lineage commitment. After 14 days of differentiation EB’s were harvested and analyzed by flow cytometry (Fig. 5). Whole EB harvests were stained for CD45 and CD34. CD45 and CD34 co-expression are commonly used in conjunction to assess the
hematopoietic potential of progenitors via flow cytometry; cells expressing these markers are known to be hematopoietic in nature. Gates were set at $10^2$ for CD45 expression and $10^3$ for CD34, on average 60,000 events were recorded. Cells expressing both CD34 and CD45 are hematopoietic progenitors that emerge from the EB harvests. Double positive populations are highlighted in red. Non-transduced cells had 2.10% double positive represented by flow cytometry plots, the siRNA and scramble control had 2.71% and 2.07% double positive cells (Fig.5). There were no apparent differences in the ability of hESCs to differentiate into hematopoietic progenitors. When compared to the non-transduced cells, the scramble control cells were able to differentiate similarly due to the emergence of CD34+45+ in similar percentages, indicating that the vector does not have a toxic effect on and does not alter hematopoietic differentiation. When the scramble control and siRNA groups are compared there are no significant differences in the percentages of double positive cells: siRNA 2.71% and scram 2.07%. Overall, these data demonstrate that APOBEC3G is not a critical factor for early development of hematopoietic progenitors as seen by the emergence of early progenitors and no change in the hematopoietic potential of the EB harvests due to no significant differences in CD34 and CD45 co-expression when APOBEC3G expression levels are diminished.

**Colony Forming Assay of Hematopoietic Progenitors**

Whole EB harvests were plated in complete methylcellulose at a density of 100,000 cells per 100mm grid dish in triplicate to assess initial hematopoietic lineage commitment. Colonies were phenotyped as either erythroid, granulocyte, macrophage, or mixed granulocyte/macrophage colonies. These progenitors were allowed to further differentiate into the terminal phenotypes of the blood for 14 days and form colonies
Plates were then scored in a blinded fashion. This allowed us to dismiss any bias in the colony counts between the conditions. First we assessed the ability of these progenitors to form colonies. Total colony counts remained consistent throughout all conditions and no significant changes were seen. The knockdown of APOBEC3G gave rise to colonies similar in numbers to that of the scramble control (Fig. 6a). Each transduced hESC line was mechanically derived from a single colony that survived after two weeks of puromycin selection, and was subsequently cultured for an additional two weeks in the presence of puromycin. This allows for control over vector integration, yielding a more homogenous transduced hESC culture. Through drug selection, transduced lines that were established include: siRNA1, siRNA2, siRNA3, scram1, and scram2. The different lines of APOBEC3G knockdown were assessed by scoring colony formation in methylcellulose; siRNA1, siRNA2, and siRNA 3. The average colonies in the knockdown of APOBEC3G ranged from 22-32 colonies: siRNA1 32±7.2, siRNA2 26±9.5, and siRNA3 21.7±1.5. The scramble control group averaged between 18-26 colonies: scram1 26±7.8 and scram 2 18±5, there were no significant differences in the ability of these progenitors to form colonies. When the colonies were phenotyped (Fig. 6b), there were slight differences in the phenotype distribution throughout all experimental conditions; however, none yielded statistically significant changes due to poor ability of progenitors to give rise to erythroid colonies, even in control conditions. Erythroid colony counts are represented as a percentage of total colonies (Fig. 2c). The siRNA1 and siRNA2 lines were able to form very few erythroid colonies, whereas the siRNA3 group was unable to form any erythroid colonies. siRNA1 on average represents 4% erythroid colonies and the siRNA2 1%. Cells with the scramble control vector were
able to form few erythroid colonies: scram1 15% and scram2 10%. The siRNA groups had a slight reduction in their ability to form erythroid colonies, however, no concrete conclusions can be made due to low number of erythroid colonies across all conditions. A low yield of erythroid colonies in all conditions became problematic because the differences between experimental conditions and control conditions were very small. For example, siRNA1 gave rise to 1.3 erythroid colonies and scram1 gave rise to 4, when the triplicate counts were averaged. It is difficult to make any concrete conclusion when colonies counts are so low overall because in theory each colony arose from a single progenitor and the difference between 1.3 and 4 is minuscule when assaying lineage commitment using colony counts. We cannot attribute this decrease due to knockdown of APOBEC3G due to the differentiation protocol being poor in yielding progenitors with erythroid potential. When the myeloid linages were analyzed, granulocyte and macrophage colonies, we saw a slight increase in the percentage of granulocyte/macrophage (myeloid) colonies in groups where APOBEC3G was knocked down (Fig. 6d). On average Scram1 and scram2 gave rise to 84.26% and 89.79% total myeloid colonies. The knockdown groups gave rise to slightly more myeloid colonies across the board on average: si1 84.26%, si2 99.10, and si3 100.0%. However, these slight increases were not statistically significant when compared to the scramble control (P>0.05). These data suggest that our differentiation protocol gives rise to progenitors with a biased for myeloid lineage commitment, yielding very little progenitors with erythroid potential. The use of these progenitors to test whether diminished APOBEC3G expression skews or influences the terminal lineage commitment of early hematopoietic progenitors is not sufficient due to the inability of the progenitors to give rise to robust
erythroid colonies. Since APOBEC3G expression is up-regulated in the MEP population, which gives rise to erythrocytes and megakaryocytes, along with previous unpublished data from our lab leads us to hypothesis that erythropoiesis may be impaired due to knockdown of APOBEC3G, therefore it was imperative that we establish differentiation parameters that would give rise to progenitors that have no bias for either myeloid or erythroid lineage commitment; if we were to see any alterations in lineage commitment in the absence of APOBEC3G, then these progenitors must have equal commitment potential.

**Developmental Stage-Specific Up-regulation of APOBEC3G Expression**

Although we were able to observe slight variations in the lineage commitment of hematopoietic progenitors derived from hESC when APOBEC3G was knocked down using previous protocols established in the lab, these data were inconclusive due to statistical insignificance and absence of robust formation of erythroid colonies, even in our controlled conditions. Subsequent experiments yielded similar results. This led us to conclude that the previous protocol illustrated by figure 4 was not optimal to assess erythroid development; progenitors emerging from this differentiation protocol inherently favored myeloid lineage commitment. Therefore, we established our own protocol that would give rise to progenitors that would be predetermined to commit equally to both erythroid and myeloid lineages. By combing the differentiation methods established by previous members of the lab (Galic et al. 2009) with the recently published literature (Kennedy et al. 2012), we were able to successfully differentiate hESCs into early hematopoietic progenitors that would allow us to assess their erythroid development. In this protocol we use the cell surface marker CD43 as a pan-hematopoietic marker to
detect emergence of early hematopoietic progenitors. The revised differentiation protocol is illustrated by figure 7.

Single lines from the siRNA and scramble control transduced hESC lines were arbitrarily chosen to carryout the subsequent experiments (siRNA2 and scram2). EBs were formed from hESCs and differentiated in hypoxic culture conditions for the first 8 days then transferred to atmospheric O₂ conditions starting on day 8. Time-point EBs from non-transduced hESCs were collected on days 3, 6, 8, 9, and 11 to measure APOBEC3G mRNA expression during early hematopoietic development. RNA was isolated from each time-point and cDNA was synthesized for subsequent measure of mRNA copy number; mRNA copy numbers were normalized to GAPDH standards (Fig. 8). There is an increasing trend in APOBEC3G expression as early hematopoietic development progresses. Interestingly enough, APOBEC3G expression levels slightly increase during the first 8 days when EB cultures are in hypoxic conditions. When cultures are moved to atmospheric O₂ conditions, APOBEC3G expression increases almost 100-fold by day 11 of differentiation when compared to day 3, which is in hypoxic conditions. As outlined by Kennedy et al. 2012, during the first 8 days of differentiation in hypoxic conditions, formation of mesoderm is induced and early hematopoietic commitment begins. When cultures are moved to atmospheric O₂ conditions, specification of early hemogenic endothelium to hematopoietic progenitors occurs along with the expansion and proliferation of these progenitors (Kennedy et al. 2012). It is interesting to observe a significant spike in APOBEC3G expression during the hematopoietic specification phase of the differentiation protocol. This developmental stage-specific up-regulation may be indicative of APOBEC3G’s importance during early
hematopoietic development or its role in predetermining progenitor terminal cell fate decisions.

**Hematopoietic Potential of CD43+ Populations**

On day 11, EBs were harvested and stained for flow cytometric analysis prior to sorting (Fig. 9). Since CD43+ cells represent early hematopoietic progenitors in these culture conditions, we first assessed the emergence of CD43+ cells within the whole EB harvest (Kennedy at al. 2012; Vodyanik et al. 2006). The analysis reveals equal percentages of CD43+ cells from all conditions, signifying that emergence of early hematopoietic progenitors was not impacted by the knockdown of APOBEC3G. The non-transduced H1 hESC control had 10.1% (fig. 9a) CD43+ cells whereas the scramble and siRNA had 5.52% (fig. 9b) and 6.28% (fig. 9c), respectively. There are no significant differences between the siRNA and scramble control, however, there is a slight decrease in the percentages of CD43+ cells between the non-transduced control and the scramble control. This can be attributed to vector selection of hESCs that have less hematopoietic potential. Transduced hESCs were further selected using mechanical and drug selection. These selection events may have selected for transduced hESCs that have less potential for hematopoietic differentiation due to the heterogeneity of the H1 hESC cell line. Our scramble control vector allows us to control for these phenomenon and we can therefore compare the knockdown of APOBEC3G to the scramble control. Retention of LNGFR expression of the scramble control (Fig. 10a) and siRNA (Fig. 10b) shown by flow cytometry indicates that vectors are not being down regulated during differentiation and expression of both the shRNA against A3G and the scramble sequence are being expressed. Representative flow plots for LNGFR expression gated on CD43+ populations
show that 97.4% of the scramble control and 98.8% of the siRNA populations retain constitutive expression of our vectors. From these data we can conclude that knockdown of APOBEC3G does not impair the ability of hESCs to differentiate into early hematopoietic progenitors.

To further assess the efficiency of our modified protocol, we compared CD34 and CD45 expression of whole EB harvest to the previous protocol (Fig. 11). Double positive populations are highlighted in red. Percentages of double positive populations from the previous protocol are in blue and the percentages in red represent the modified protocol double positive populations. We observed a dramatic increase in the percentage of double positive cells emerging from our modified differentiation protocol when compared to the previous protocol: non-transduced 11.3%, scramble 6.16%, and siRNA 7.56%. On average, we observed a 6% increase in CD34 and CD45 double positive cells. Moreover, the appearance of the flow cytometry data displays more distinct populations when compared to the flow plots produced by the previous protocol. The CD45 expression in these populations is higher, which indicates that these progenitors are more developed and have advanced from the hemogenic endothelial stage of development. In the previous protocol, most cells were CD34 positive and the appearance of the flow plot population distributions indicates that these single positive cells were still undergoing progenitor commitment; cells that are only express CD34 represent a hemogenic endothelium population (Vodyanik et al. 2006). Co-expression of both hematopoietic markers are indicative of hematopoietic potential, thus the increase we see in our modified protocol are promising in that this protocol is more efficient in inducing hESCs to differentiate into early hematopoietic progenitors.
EB harvests were then sorted for CD43 using magnetic activated cell sorting (MACS). Portions of the CD43 positive fractions were retained and stained for flow cytometry analysis, CD43 gates were set at $10^3$. Cells were then examined for their hematopoietic potential by gating on CD43+ populations, gates were set at $10^3$ for CD45 and $10^2$ for CD34; the representative flow plots for CD34 and CD45 expression are shown in panel A of figure 12 and are highlighted in red. This reveals that CD43+ cells express both hematopoietic markers, indicating high hematopoietic potential in this population. There are no significant changes in the expression of these markers across all conditions. Non-transduced cells are 45.8% double positive, scramble is 40.9%, and the siRNA is 40.8%. Furthermore, since we are interested in studying the impact of APOBEC3G knock down in terminal erythroid commitment, we examined the expression of CD235a (Fig. 12b), a common erythrocyte marker (Sato et al 2000). Cells were gated on CD34 and CD45 double positive populations from the CD43+ sorted fractions. All cells that are double positive for CD45 and CD34 express CD235a: non-transduced 100%, scramble 99.9% and siRNA 99.9%. These data indicate that CD43+ cell populations have high hematopoietic potential along with erythroid and myeloid potential. Using CD43 as a pan-hematopoietic marker for progenitors that arise from hESCs, along with their co-expression of hematopoietic specific and erythroid specific markers, we aim to further test their terminal lineage commitment by subsequent assays based on the sorted and enriched CD43 population. (Kennedy at al. 2012; Vodyanik et al. 2006). Together, these data demonstrate that knockdown of APOBEC3G does not impair the emergence of these progenitors; moreover, this further demonstrates that our modified protocol is sufficient enough to test our hypotheses.
Colony Forming Assay of CD43+ Sorted EB Harvest

Sorted CD43 positive cells were plated into complete methylcellulose and scored 14 days later. This allowed us to test the lineage commitment of these cells. Sorted progenitors were plated at a density of 10,000 cells per 100mm grid dish in triplicate. First, total colonies were counted and then the colonies were phenotyped as erythroid, granulocyte, macrophage, or mixed granulocyte/macrophage colonies. On average the non-transduced culture gave rise to 52±4.6, scramble 33.67±7.2, and siRNA 30±5.2 (Fig. 13a). There were no significant difference in total colonies formed between the siRNA and scramble conditions; however, a slight reduction of total colonies can be seen between the non-transduced and the scramble control. Again, this can be attributed to transduction selection of hESCs with reduced hematopoietic potential. When colonies are examined by phenotype, there are apparent differences in the types of colonies that are being formed (Fig. 13b). These data are represented as a percentage of total colonies that are either erythroid and myeloid. The percentage of erythroid colonies in the siRNA condition is significantly (p=0.02) reduced by almost 50% when compared to the scramble control (Fig. 13c). There are no significant differences between the non-transduced and scramble conditions. When the myeloid colonies are examined, the inverse can be seen. Knockdown of APOBEC3G (siRNA) significantly increases (p=0.02) the percentage of myeloid colonies when compared to the scramble control (Fig. 13d). The non-transduced and scramble control give rise to similar percentages of myeloid colonies. Since total colonies do not change significantly between conditions, but the phenotypes of the colonies decrease or increase, particularly in the erythroid and myeloid lineages, we can therefore conclude that knockdown of APOBEC3G impairs the
formation of erythroid colonies and favors the formation of myeloid colonies. This indicates that APOBEC3G is an important factor in determining lineage commitment of early hematopoietic progenitors.
IV

Discussion

In this study, we sought to harness the power of hESCs in studying the role of APOBEC3G in early hematopoietic development. Although much is known about human hematopoiesis, there are many holes in the knowledge base of this essential process. Through the advancement of genomics and proteomics research and technologies, we now know that many blood disorders are complex and are influenced by an intricate system of factors that are important in regulating normal development. Previous work done in our laboratory has directed our interest in understanding the role of APOBEC3G in early hematopoietic development.

We previously demonstrated that APOBEC3G is important in the lineage commitment of the common myeloid progenitor during hematopoiesis. Briefly, when APOBEC3G is knocked down in the common myeloid progenitor, erythropoiesis is impaired and myeloid lineage commitment is favored. This led us to ask if the same phenotype could be observed in the commitment of early hematopoietic progenitors when APOBEC3G is knocked down, and if other stages earlier in hematopoiesis might be influenced.

This study aimed to answer three distinct questions: 1) is APOBEC3G important in maintaining the pluripotency of hESCs, 2) is APOBEC3G important in hESC differentiation into early hematopoietic progenitors, and 3) is APOBEC3G important in regulating the terminal differentiation of early hematopoietic progenitors? To answer these questions, we employed a loss-of-function strategy starting at the hESC stage and followed their hematopoietic development to the terminal stages. We hypothesized that
knockdown of APOBEC3G will impact the pluripotency of hESC and alter the expression of essential pluripotency markers, in-turn impairing the emergence of progenitors when hematopoietic differentiation is induced. We further hypothesized that early knockdown of APOBEC3G would predetermine the terminal lineage commitment of the emerging hematopoietic progenitors and would skew colony phenotype, but would not change the quantitative colony forming potential of these cells.

To address our first aim, we first examined APOBEC3G expression in H1 hESCs and through the process of hematopoietic development in EB cultures. Our data show that APOBEC3G, although at a low level, is in fact expressed in H1 hESCs and that its expression increases significantly as development towards hematopoietic progenitors in EB cultures progresses. This is interesting in that APOBEC3G is currently primarily characterized as a host-restriction factor and that hESCs are not a primary target for HIV or other retroviruses. This leads us to believe that APOBEC3G may have a unique role in hESCs. Although one other group has shown that knockdown of APOBEC3G does not impair pluripotency of H1 and H9 hESCs, the characterization was brief and incomplete (Wising et al. 2011). It has also been shown that APOBEC3G does not impact the mobility of retrotransposable elements in H1 and H9 hESC, which is increased in hESCs (Wissing et al. 2011). Therefore, APOBEC3G may have an alterative role in these cell types.

Our data show that knockdown of APOBEC3G does not change the pluripotent state of hESCs shown by retention of imperative pluripotency factors and markers. Moreover, confirmation of hESC pluripotency ensured that if any downstream alterations were to be observed in progenitor and terminal phenotypes due to knockdown of APOBEC3G, they
were not due to the loss of the ability of hESCs to give rise to cells of all three germ layers and that the observed APOBEC3G phenotype is restricted to the hematopoietic compartment. Recently, there have been published reports that claim that APOBEC3G interacts with proteins that regulate expression of pluripotency gene transcripts via protection from miRNA-mediated degradation (Ali et al. 2013; Zhu et al. 2011). DND1 binds to these transcripts and blocks binding of miRNA. Notably, the transcripts that fall under the protection of DND1 include Oct-4 and Sox2, which are important regulators of the transcriptional pluripotency network of hESCs (Zhu et al. 2011). APOBEC3G binds to DND1 to reverse this blocking effect. The tight regulation of these transcripts is important in maintaining pluripotency (Bohler 2009; Thompson et al. 1998). The slightest dysregulation in the expression of these genes alters pluripotency, however, our data show that diminished quantities of APOBEC3G mRNA do not change the expression of these genes on a protein level, nor do they hinder the ability of these cells to form all three germ layers. It has been shown the miRNA post-transcriptional control of Oct-4 and Sox2 is important in maintaining pluripotency and regulating hESC differentiation (Wang et al. 2013). If APOBEC3G were in fact involved with regulating the binding of these miRNAs, then altering its expression would be detrimental to the tight regulation of these pluripotency genes and lead to the loss of pluripotency markers, however, our results suggests that this is not the case and that APOBEC3G is not be involved in maintaining pluripotency. This does not rule out a functional importance of APOBEC3G in hESCs; there may be an alternative role for APOBEC3G expression in these cells, not explored here.
Our second aim was to assess the ability of hESCs with APOBEC3G knocked down to differentiate into early hematopoietic progenitors. Previously, our lab had established a differentiation protocol that yielded hematopoietic precursors that ultimately gave rise to lymphocytes when introduced to humanized mice (Galic et al. 2009) and a feeder-free system for generating similar precursors, that when stimulated with a cytokine cocktail, would yield phenotypically accurate and functional macrophages (Subramanian et al. 2009). Using the reasoning that lymphoid cells arise from the same hematopoietic progenitors that give rise to other lineages of the blood, we sought to employ the same differentiation protocol, however, we were primarily interested in the erythroid and myeloid lineages. To enhance erythroid potential, we modified the protocol by adding the cytokine VEGF.

We used the expression of two well-defined hematopoietic cell surface markers CD34 and CD45 to indicate development of early hematopoietic progenitors. Using the protocol illustrated in figure 4, we were able to differentiate hESCs into early hematopoietic progenitors. We observed no change in the percentage of cells that expressed both hematopoietic markers, indicating that knockdown of APOBEC3G did not alter early hematopoietic development. However, only a small portion of cells expressed CD45 overall. Moreover, methylcellulose-based colony forming assays yielded inconclusive results. We observed a statistically insignificant decrease in the percentage of erythroid colonies and increase in myeloid colonies, largely due to the fact that the developmental protocol does not support robust erythroid progenitor development. As our third aim was to address terminal differentiation based on colony forming assays, it was imperative to be able to examine erythroid colony formation.
To be able to better test our hypothesis that erythropoiesis may be impaired from hESCs with impaired APOBEC3G expression; we sought to improve our differentiation protocol. A recent report by Kennedy et al. established a serum-free protocol that allows for phenotypic distinction between primitive and definitive programs of early hematopoietic development, which would provide the tools needed to test erythropoietic development from hESC more efficiently by selecting a population of cells with greater erythroid potential. Combining the hypoxic conditions, time frame, and a simplified cytokine cocktail of their protocol with our previous protocol, we were able to observe the emergence of hematopoietic progenitors of similar phenotypes. The most notable modification to our differentiation protocol was the use of hypoxic conditions during the first 8 days of differentiation (figure 7). During the first 8 days, nodal signaling is induced by the addition of activin A along with hypoxic conditions; this signaling pathway is important for mesodermal development (Kennedy et al. 2012). To derive early hematopoietic progenitors from hESCs, we must induce development stage-specific differentiation. Hemogenic endothelium is of mesodermal origin, which gives rise to early hematopoietic progenitors (Kennedy et al. 2012). Using this reasoning, EBs were cultured in hypoxic conditions starting on the first day of formation. It has been shown that early development of hematopoietic progenitors relies heavily on hypoxia-inducible factors (HIFs), which are transcription factors important in regulating erythropoietin (EPO) production and iron metabolism (Yoon et al. 2011). Once differentiation parameters were established, we moved forward in using this modified protocol to test our hypothesis.
CD43 expression indicates early hematopoietic progenitors that emerge from the EB cultures, which also express hematopoietic markers, CD34, CD45, and CD235a. Further, colony-forming assays confirmed that these progenitors had greater hematopoietic potential and erythroid potential by their ability to give rise to greater number of erythroid colonies. Similar to the results of the previous protocol, we observed no changes in the ability of hESCs with APOBEC3G knocked down to be able to differentiate and develop into early hematopoietic progenitors. Recent reports have shown that although cells express a similar phenotype, their global gene expression patterns differ; along with their self-renew capacity (Qiu et al. 2014). Although two cells might express the same cell surface markers and resemble each other phenotypically, their hematopoietic potential may differ due to changes in global gene expression and in-turn alter the terminal cell type it can differentiate into (Qiu et al. 2014). To be able to make concrete conclusions at this point, we would have to further evaluate gene expression patterns of these cell types to see if there are any variations due to diminished levels of APOBEC3G. Global gene expression may have been altered due to the absence of APOBEC3G early on during development, which would predetermine these progenitors to favor developing into certain lineages of the blood. APOBEC3G does not seem to be an important regulator of early development, as we observe no changes in expression of hematopoietic markers, however, that does not denounce the hypothesis that these progenitors do not have the same potential for every lineage of the blood when APOBEC3G expression is altered. Further analysis of their terminal differentiation revealed that the ability of these progenitors with APOBEC3G knocked down to give rise to erythroid colonies is impaired and myeloid colonies are favored. This may be due to
changes in global gene expression up-stream in the developmental process, which predetermines the terminal cell type these progenitors are able to give rise to, changing their potential and down-stream activity. Due to the greater number of erythroid colonies provided by a less lymphoid-biased protocol, we were able to make statistically significant observations that demonstrate knockdown of APOBEC3G alters early hematopoietic progenitor potential by skewing lineage commitment, which is evident by impaired ability to give rise to erythroid colonies and an increase of myeloid colonies with no change in quantitative colony-forming potential.

The next logical step in this study is to uncover the underlying mechanisms of action, which would require a genomics approach utilizing microarray gene expression assay and RNA sequencing of the transcriptome. Although we do not have the data of these approaches in hESC derived progenitors, we can postulate the processes in which APOBEC3G may be involved. We believe that APOBEC3G does not alter gene expression by directly acting as a transcriptional regulator. This is due to APOBEC3G being exclusive to the cellular cytoplasm and the fact that it is kept away from the cellular genome, even during mitosis (Mangeat et al. 2003). We do, however, believe the APOBEC3G maybe involved in post-transcriptional regulation of gene expression. This can be attributed to APOBEC3G directly binding mRNAs or by regulating miRNAs that target transcripts involved in erythropoiesis. APOBEC3G exists in two forms in the cell, high-molecular mass (HMM) and low-molecular mass (LMM) (Wissing et al. 2007). The HMM consists of RNPs that bind to APOBEC3G due to both RNA affinity and RNA binding domains (Gallois-Montburn et al. 2007). This could indicate that APOBEC3G may be responsible in sequestering mRNA after transcription to hinder translation, or it
may be protecting transcripts from miRNA-mediated degradation. Furthermore, APOBEC3G localizes to processing bodies, which are sub-cellular foci that are active centers of miRNAs and mRNA editing (Gallois-Montburn et al. 2007). Moreover, APOBEC3G has been shown to directly influence miRNA activity by inhibiting the assembly of RNA-induced silencing complexes (RISC) (Liu et al. 2012). Therefore, we cannot dismiss that APOBEC3G may be involved in miRNA regulation and that it may in fact be regulating miRNAs that are imperative for both determining hematopoietic potential during early development and further down-stream development processes, such as erythropoiesis. Throughout the process of developmental hematopoiesis, miRNAs play an important role in determining terminal cellular differentiation. It has been shown that miR-233 is expressed in a dose-dependent manner in the MPP population until the commitment of the CMP during granulopoiesis by inhibiting commitment of the CMP to the MEP and promoting GMP commitment (O’Connell et al. 2011). Furthermore, studies have shown that defects in miR-451 expression in mice contribute to the loss of erythrocytes circulating in the blood due to defective erythropoiesis and that miRNA-155 is down regulated during erythropoiesis (O’connel et al. 2011; Undi et al. 2013). These miRNAs may be a part of the APOBEC3G interactome, hence providing an explanation of the phenotype seen when APOBEC3G expression is diminished. Further study is warranted to uncover the underlying mechanisms of APOBEC3G in cellular processes during various time points in hematopoietic development.

Interestingly enough, similar phenotypes have been seen in other studies that do not directly involve APOBEC3G, but do have logical connections. During HIV infection, human hematopoiesis is impaired, particularly the erythroid compartment (Nixon et al.
2013). These studies were carried out using wild-type HIV, which had all components intact, most notably *vif*. For HIV to overcome APOBEC3G restriction, *vif* must degrade APOBEC3G via polyubiquitination and proteasomal degradation (Stopak et al. 2003; Sheehy et al. 2003). This step is imperative in order for HIV to successfully infect a target cell. This degradation of APOBEC3G may be the cause of the phenotype seen in these studies. It would be interesting to see if erythropoiesis is restored when hematopoietic progenitors are infected with HIV Δ*vif*. This would further create strong evidence that APOBEC3G is a key player during erythropoiesis and that it is involved in hematopoietic development.

The majority of hematologic disorders are due to developmental abnormalities. A complete understanding of the various factors involved during early hematopoietic development is crucial in uncovering how interruption of these normal molecular processes may contribute to hematologic disorders. Early hematopoietic progenitors are responsible for the formation of primitive erythrocytes found in yolk sac blood islands that are important for circulating oxygen in the embryo and for the formation of embryonic vasculature (Stugeon et al. 2012; Shalaby et al. 1995). Although a transient population, the absence of primitive erythrocytes during early development is fatal and interruption of key regulatory elements in these early progenitors leads to the impairment of vasculature development (Shalaby et al. 1995). Therefore, this study identifies a novel factor, which when interrupted, that may contribute to these phenotypes seen and may be involved in the progression of various hematologic disorders. Further examination of APOBEC3G dysregulation in patient samples from various hematologic disorders may shed further light onto the extent of its involvement in various cellular processes.
Conclusion

Overall, these data provide sufficient evidence to either support or oppose our hypotheses. We had reason to believe that APOBEC3G would be of importance in maintaining the pluripotency of hESCs, however, our data rejects our original hypothesis. With the knockdown of APOBEC3G, hESCs are retaining expression of essential pluripotency transcription factors and cell surface markers with the ability to form all three embryonic germ layers, demonstrating that diminished levels of APOBEC3G expression do not result in a loss of pluripotency. Furthermore, we had postulated that in the absence of APOBEC3G hESCs would be able to differentiate into early hematopoietic progenitors, but that their lineage commitment to terminal cell types would be skewed; our data supports that APOBEC3G is not of importance in the development of early hematopoietic progenitors. When progenitor ability to differentiate into terminal cell types is evaluated by colony forming assay there is an apparent skewing of lineage commitment, primarily during erythropoiesis. This is in agreement with the fluctuating levels of APOBEC3G expression observed through the dMAP database (http://www.broadinstitute.org/dmap/home; Novershtern et al. 2011). We can therefore conclude that absence of APOBEC3G predetermines early hematopoietic progenitor potential for terminal lineage commitment by favoring the myeloid lineages and in-turn impairs erythropoiesis.
References


Appendix: Figures

Figure 1: shRNA (A) and scramble control (B) cassette are constructed using a FG12 based vector. A strong CMV promoter for vector integration drives the vector. An H1 promoter drives the shRNA and scramble cassettes. Reporter gene dLNGFR is used to identity cells that have been transduced. Puromycin resistance allows transduced cells to be selected using drug selection.
Figure 2: Successful transduction of hESC colonies to achieve A3G knockdown by RNAi. Population shifts from non-transduced colonies (red) can be seen in both the scramble control (A) and cells expressing the siRNA knockdown (B) of A3G (blue). Expression of our reporter gene dLNGFR indicates stable transduction of H1 hESC colonies with no overt toxicity. (C) qRT-PCR analysis of APOBEC3G expression showing significant decrease in APOBEC3G transcript in hESCs transduced with A3G knockdown vector and measurement of endogenous APOBEC3G expression levels in H1 hESC cell line.
**Figure 3:** (A) Western blot analysis of transcription factors that are essential for hESCs to remain in a pluripotent state. There are no apparent differences in Oct-3/4 and Sox2 protein between transduced cells. A3G knockdown does not affect pluripotency transcription factors. Transduced colonies were also negative for SSEA 1, which indicates onset of differentiation (B). (C) Flow analysis reveals high expression of embryonic stem cell surface markers SSEA 4 and TRA-1-60. This indicates that transduction of H1 hESCs does not effect the stem cell state of the colonies. Teratoma assays of hESC (D) demonstrate ability of transduced hESCs to give rise to all three embryonic germ layers, further confirming pluripotency retention.
**Figure 4:** Schematic illustration of EB differentiation protocol from day 0-14. EBs formed from hESCs were cultured in atmospheric O$_2$ with subsequent exposure to cytokine cocktails pushing for hematopoietic differentiation.

**Figure 5:** H1 hESC lines were derived from single colonies of transduced hESC cultures. Cultures were differentiated using embryoid body (EB) method in differentiation media with cytokines pushing for hematopoietic progenitor differentiation. Hematopoietic potential analyses were done using CD34 and CD45 markers. There are no apparent changes in the percentage of CD34$^+$45$^+$ cells (highlighted in red) in EB harvests when A3G is knocked down. This indicates that A3G does not play a role in the ability of hESC to differentiate into hematopoietic progenitors.
Figure 6: hESC derived hematopoietic progenitors from EB harvests were seeded into methycellulose. The ability of the progenitors to form colonies is not affected. There are no significant differences in number of total colonies (A) when A3G knock down is compared to the scramble control. Phenotypes of colonies show slight differences in the commitment of progenitors. Erythroid colony formation is slightly decreased with A3G knock down and myeloid colonies are slightly increased, however, statistical significance is not consistent across all experimental groups and controls.
Figure 7: Schematic illustration of EB differentiation protocol from day 0-11. EBs formed from hESCs were cultured in hypoxic conditions with subsequent exposure to cytokine cocktails pushing for hematopoietic differentiation. First mesoderm was induced (days 0-4) and then specification of hematopoietic progenitors was induced (day 5-11). Cell were harvested and sorted on day 11 for CD 43. For the first 8 days of differentiation, cultures were maintained in hypoxic conditions and then transferred to an incubator with atmospheric oxygen.
Figure 8: Non-transduced EBs were harvest at 5 different times points to examine APOBEC3G expression as hematopoietic development progresses. On days 3-8, EBs were cultured in hypoxic conditions then moved to atmospheric O₂ conditions. As development progresses, there is a significant up-regulation of APOBEC3G expression, most notably during the phase of differentiation that induces hematopoietic specification, which begins on day 9.

Figure 9: EBs were harvested on day 11 of differentiation and their hematopoietic potential was assessed through the expression of CD43. Whole EB harvests was stained for CD 43 and compared to non-transduce H1 hESCs, revealing no apparent changes in the emergence of CD 43 positive hematopoietic progenitors from the experimental groups (scramble control and siRNA).
Figure 10: Expression of dLNGFR (blue) gated on CD43 positive populations compared to non-transduced H1 hESC negative control (red). LNGFR expression is retained after hematopoietic differentiation of hESCs indicating no loss of vector expression in the scramble control (A) and siRNA (B) group.

Figure 11: Hematopoietic potential of whole EB harvest was assessed by CD34 and CD45 expression. Double positive populations are highlighted in red. The numbers in blue represent percentage of double positive cells from the previous protocol illustrated by figure 4 and the numbers in red represent percentage of double positive cells from the modified protocol. A significant increase in the double positive populations can be seen in the modified protocol, indicating that these cells have a higher hematopoietic potential.
Figure 12: Hematopoietic potential of CD43+ was assessed by CD34 and CD45 expression (A). Double positive populations are highlighted in red. High expression of these markers can be seen in the CD43+ population. To further assess erythroid potential of these cells, cells were gated of CD34 and CD45 double positive populations and expression of an erythroid specific marker, CD 235a, was measured shown in panel B. All cells that are CD43+CD34+CD45+ express the erythroid specific marker.
Figure 13: CD43 sorted cells were seeded into methycellulose. Total colonies were counted (A) and phenotyped (B). There are no apparent changes in the ability of hematopoietic progenitors to give rise to colonies, however, there are apparent differences in the types of colonies that are being formed. Percentage of erythroid colonies in the knockdown of A3G (siRNA) is significantly decreased when compared to the scramble control (C) and there is a significant increase in the percentage of myeloid colonies (D) in the absence of A3G.