Programa de Doctorado en Biociencias Moleculares

Departmento de Bioquímica

Facultad de Medicina

Universidad de Autónoma de Madrid

Using CRISPR/Cas9 based genetic screen to understand the role of *Flk1* in the specification of endothelial and blood lineages

Doctoral Thesis

Mayank Bansal
Licenciado en Biología
Madrid, 2019

Director: Dr. Rui Benedito

Cento Nacional de Investigaciones Cardiovasculares (CNIC)

SUMMARY

The molecular mechanisms responsible for the early development of endothelial and haematopoietic cell lineages are not fully understood. *Flk1 (Vegfr2)* is the main vascular endothelial growth factor (VEGF) receptor, and one of the most important genes for endothelial differentiation, proliferation and survival. However, little is known about the function of other genes that may cooperate, antagonize or partially substitute the function of *Flk1*. In this study, we investigated the role of *Flk1* in the differentiation of endothelial and blood progenitors by utilizing a wide range of functional genetic approaches. Our analysis showed that *in vitro* $Flk1^{KI/KI}$ (null) ES cells can differentiate to endothelial cells (ECs) but these proliferate less and do not sprout. *In vivo* E8.5 $Flk1^{KI/KI}$ embryos contain a large number of endothelial-hematopoietic progenitors, but these do not differentiate to ECs, presumably due to their distinct microenvironment.

Comparative transcriptional profiling between $Flk1^{KI/KI}$ and $Flk1^{KI/WT}$ ECs obtained *in vitro* revealed that the full loss of *Flk1* is compatible with endothelial development but these cells have lower proliferative and metabolic activity. Interestingly, a large subset of the $Flk1^{KI/KI}$ endothelial progenitors give rise to other alternative lineages through a process of endothelial to mesenchymal transition (EndMT). The *Flk1* loss of function was accompanied by the upregulation of Tgf-beta and the Tgf-beta-related bone morphogenetic proteins (Bmps) signaling pathways, extracellular matrix (ECM) deposit and remodeling. The study suggests the existence of a developmental barrier in $Flk1^{KI/KI}$ endothelial progenitor cells, which if given enough time is overcome by EMT and trans-differentiation into other cell types *in vitro*. To evaluate if similar biological processes also occured *in vivo,* we transcriptionally profiled *Flk1-null* ECs from E9.5 embryos. *In vivo* $Flk1^{KI/Flox}$ *Tie2-Cre+* ECs did not go into other lineages, and instead upregulated pathways related to apoptosis.

We also performed a genome-wide loss of function CRISPR/Cas9 screen in $Flk1^{KI/KI}$ endothelial cells to identify genetic alterations that can induce a growth advantage to endothelial progenitors with full loss of *Flk1*. The top hits from the screen revealed genes that regulate proliferation (such as the p53 pathway and tumor suppressor genes), differentiation and apoptosis indicating that the endothelial progenitor cells with *Flk1* loss of function may undergo growth arrest associated with cellular senescence or apoptosis and mutating these genes may give them a competitive advantage.

RESUMEN

Los mecanismos moleculares responsables del desarrollo temprano de los linajes celulares endoteliales y hematopoyéticos no se comprenden completamente. *Flk1* (*Vegfr2*) es el principal receptor del factor de crecimiento endotelial vascular (*VEGF*), y uno de los genes más importantes para la diferenciación, proliferación y supervivencia endoteliales. Sin embargo, se sabe poco sobre la función de otros genes que pueden cooperar, antagonizar o sustituir parcialmente la función de *Flk1*. En este estudio, investigamos el papel de *Flk1* en la diferenciación de los progenitores endoteliales y sanguíneos mediante la utilización de una amplia gama de enfoques genéticos funcionales. Nuestro análisis mostró que las células madre enbrionarias (ESCs) sin son cultivadas *in vitro* en ausencia del gen *Flk1* pueden diferenciarse a células endoteliales (EC), pero proliferan menos y no. Por otro lado, *in* vivo, los embriones *Flk1*$^{KI/KI}$ a estadío E8.5 contienen una gran cantidad de progenitores endoteliales-hematopoyéticos que no se diferencian a EC, presumiblemente debido a su microambiente distinto.

El perfil transcripcional comparado de EC *Flk1*$^{KI\ /\ KI}$ y Flk1$^{KI\ /\ WT}$ obtenidas *in vitro* reveló que la pérdida total de *Flk1* es compatible con el desarrollo endotelial, pero estas células tienen una menor actividad proliferativa y metabólica. Curiosamente, un gran subconjunto de progenitores endoteliales *Flk1*$^{KI\ /\ KI}$ dan lugar a otros linajes alternativos a través de un proceso de transición endotelial-mesenquimal (EndMT). La pérdida de función de *Flk1* estuvo acompañada por un incremento en*Tgf-beta* y las vías de señalización de proteínas morfogenéticas óseas (*Bmps*) relacionadas con *Tgf-beta*; depósito de matriz extracelular (ECM) y remodelación. El estudio sugiere la existencia de una barrera propia del desarrollo de estas células progenitoras , que con el tiempo suficiente son capaces de superar mediante EMT y diferenciación a a otros tipos celulares *in vitro*. Para evaluar si también se produjeron procesos biológicos similares *in vivo*, analizamos el transcriptoma de embriones embriones E9.5. con la delección del gen. Las célula sendoteliales con la deleción no entraron en otros linajes, y en su lugar aumentaron las vías relacionadas con la apoptosis

También realizamos un ensayo de pérdida de función de todo el genoma en células endoteliales Flk1$^{KI\ /\ KI}$ utilizando CRISPR / Cas9, con el fin de identificar alteraciones genéticas que pudieran inducir una ventaja de crecimiento para los progenitores endoteliales con pérdida total de Flk1. Los resultados más significativos revelaron genes que regulan la proliferación (como la vía p53 y los genes supresores de tumores), la diferenciación y la apoptosis. Esto, parece indicarque las células progenitoras endoteliales con pérdida de función de *Flk1* pueden sufrir un parón de crecimiento asociado con senescencia o apoptosis celular, permitiendo así la aparición de mutaciones que puedan suponer una ventaja competitiva.

TABLE OF CONTENTS

INDEX OF FIGURES AND TABLES

INTRODUCTION

RESULTS

ABBREVIATIONS

AGM: Aorta Gonad Mesonephros

BM: Bone Marrow

CFU: Colony Forming Unit

CFP: Cerulean Fluorescent Protein

CLP: Common Lymphoid progenitor

CMP: Common Myeloid Progenitor

DA: Dorsal Aorta

HSPCs: Hematopoietic stem progenitor cells

Dll4: Delta 4

EC: Endothelial cells

EHT: Endothelial to Hematopoietic Transition

EMT: Endothelial to Mesenchymal Transition

ES: Embryonic stem cells

Flk1: Vegfr2 receptor(Kdr: Kinase dependant receptor)

Flt-1: fms-related tyrosine kinase-1

Foxc2: forkhead box C1 and C2

HE: Hemogenic Endothelium

hESC: Human embryonic stem cells

HSC: Hematopoietic stem cell

IF: Immunofluorescence

LOF: Loss of Function

mESC: mouse embryonic stem cell

TGF-β1: Transforming growth factor β-1

VEGF: vascular endothelial growth factor

VEGFR1: VEGF receptor type I

WT: Wild type

Introduction

Background:

There is a close relationship between the cardiovascular and blood systems both in function as well as in their origins. Development of cell-based or pharmacological therapies is needed to overcome dysfunction in these organ systems. New hematopoietic, endothelial and cardiomyocytes can be generated from pluripotent or reprogrammed cells, but there are still many obstacles to overcome such as the low efficiency of differentiation or the inadequate function and phenotypic characteristics of the generated cells (Margariti et al., 2012; Morita et al., 2015; Levenberg et al., 2002; Adams et al., 2013; Wang et al., 2005). A deeper understanding of the molecular and cellular mechanisms involved in the generation of the cardiovascular and hematopoietic cells will certainly help to devise new therapeutic strategies for related diseases.

1. Development of cardiovascular and hematopoietic lineages

1.1. Mesoderm origin of the different cardiovascular cells

The embryonic primary germ layers, ectoderm, mesoderm and endoderm, arise during gastrulation and will give rise to all tissues. While the embryo is developing, specification of mesoderm commences with a primitive streak forming on the epiblast surface and generating a mesoderm layer connecting the endoderm and ectoderm (Gilbert, 2000a). Three components constitute the mesoderm: the initial being paraxial mesoderm, followed by the intermediate mesoderm and finally the lateral plate mesoderm. The paraxial mesoderm derived somites form the skeletal muscle, bone, cartilage and the skin dermis. Intermediate mesoderm, which links the paraxial and lateral plate mesoderm, gives rise to the kidneys, gonads and adrenal glands. The lateral plate mesoderm gives rise to the cardiovascular and blood systems (Gilbert, 2000b).

The cardiovascular cells derive from mesoderm that is marked by Brachyury/T box gene expression (Wilkinson et al., 1990). During gastrulation stage, Brachyury

expresses in the primitive streak, and is critical for the differentiation and movement of mesodermal cells from the primitive streak (Wilson et al., 1995). A subset of these cells while they differentiate also expresses *FLK1* (Fehling *et al.,* 2003). Flk1 was first found to express in the primitive streak's posterior region, then expressing in vascular ECs of the yolk sac, embryonic vasculature, and the endocardial tube (Yamaguchi et al., 1993).

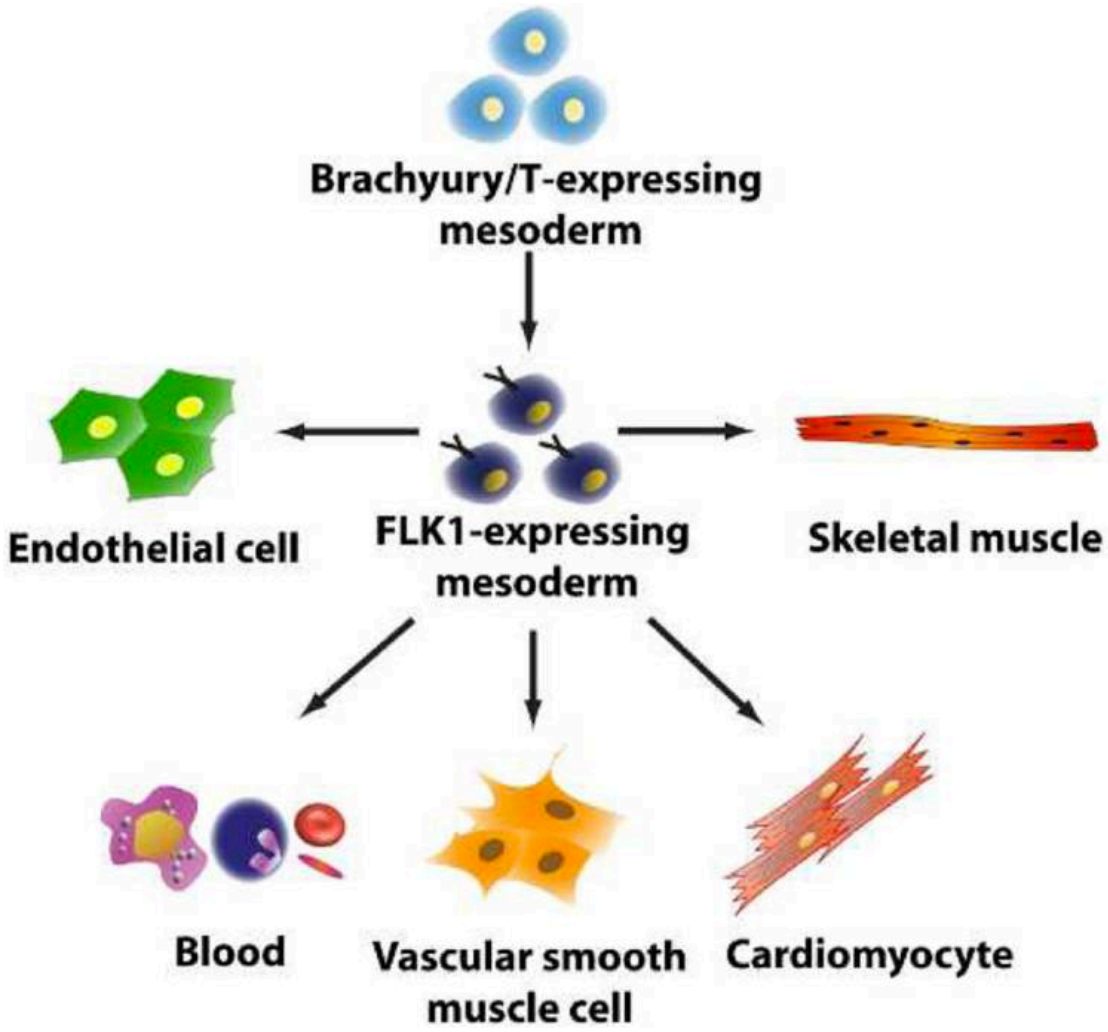


Figure 1: Developmental potential of FLK1$^+$ mesoderm: *Brachyury*/*T*-expressing cells present in the primitive streak of developing embryo form *FLK1*$^+$ mesoderm that further forms endothelial and blood cells. Also, mesodermal lineage cells that include vascular smooth muscle, cardiomyocyte and skeletal muscle cells are formed from it.

Distinct Flk1+ mesenchymal progenitors form endothelial, hematopoietic, vascular mural cells, skeletal muscle cells and cardiomyocytes (Motoike et al., 2003; Ema et al., 2006)(Fig.1). However, *Flk1* is only essential for the hemato-vascular development as deficiency of *Flk1* in mouse embryos leads to absence of hematopoietic and endothelial cells because the mesoderm is unable to relocate to its right spot or the inadequate differentiation and proliferation of these cells (Shalaby *et al.,* 1995). A mutated cloche gene, upstream of *Flk1*, also leads to flawed hemato-vascular specification (Stainier et al., 1995; Liao et al., 1998).

Flk1 mutant mice form smooth muscle cells, cardiomyocytes and skeletal muscle being dispesable for the differentiation of these cells (Motoike et al., 2003; Ema et al., 2006).

The fate of a cell and its differentiation is dependent on specific molecular programs that are activated only in some cells. Lineage specific transcription factors are key to determine the onset of cascades of gene expression that will determine the progressive acquisition of distinct phenotypic features (Fig. 2). For instance, hemato-vascular development is driven by the combined and sequential action of the transcription factors ETV2, TAL1/SCL and GATA2 (Sumanas et al., 2008; Kataoka et al., 2011a) and cardiac differentiation by HAND1, HAND2, GATA4 and GATA6 (Niu et al., 2008).

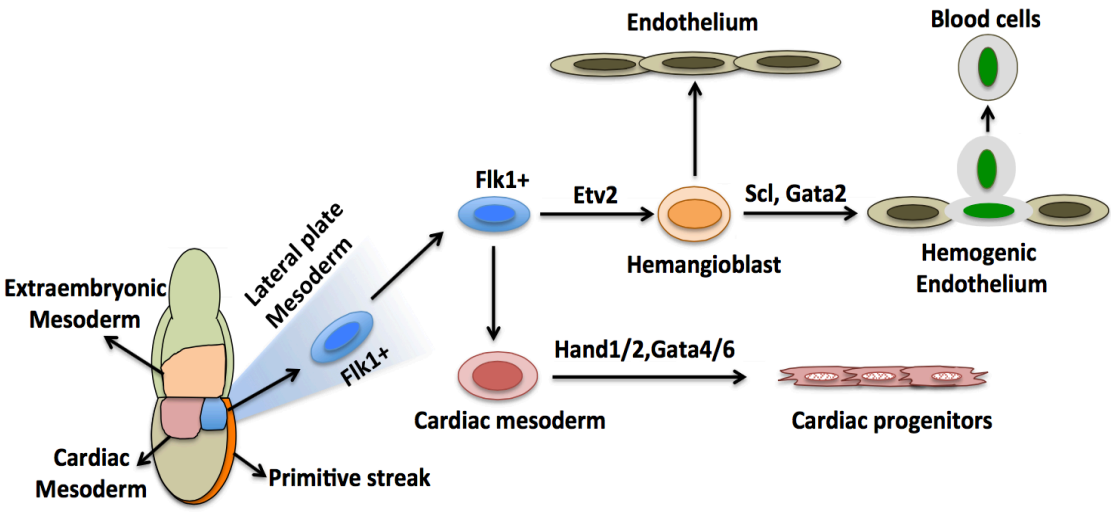


Figure 2: Mesoderm diversification: Mesoderm formation takes place during gastrulation. Hematopoietic, endothelial and cardiac progenitors arise in the lateral plate mesoderm of which *Flk1* is one of the markers. Different sets of transcription factors are responsible for the differentiation of the different lineages

1.2. Vascular development

The vascular system is among one of the first organs to be formed in the embryo and will provide the nutrients, oxygen and inter-organ communication routes that are essential for all organs to develop. Two independent processes give rise to the first vascular structures. The extra-embryonic tissues will form the yolk sac blood islands composed by endothelial and primitive blood cells that arise from distinct progenitor cells or from endothelial cells (Padron-Barthe et al., 2014; Choi, 2002). These primitive blood islands then coalesce and form a primitive vascular plexus containing blood. This plexus is later remodelled into a complex and hierarchical vasculature composed by arteries, capillaries and veins that will form the yolk sac vasculature connected with the embryo vessels.

In the embryo proper endothelial cells initially develop independently of blood cells, from lateral plate mesodermal precursors called angioblasts, which express *Flk1* (Flamme et al., 1997). These progressively acquire epithelial/endothelial characteristics, such as the expression of endothelial-specific adhesion and junctional molecules (*Pecam1* and *VE-cadherin*/*Cdh5)*. Once they acquire endothelial characteristics, they self-assemble a primitive vasculature including the cardiac endocardial tube, the dorsal aortas, brain vascular plexus, intersomitic vessels and vitelline vessels (Risau, 1997). This initial process of vascular formation directly from angioblasts is called **vasculogenesis.** As the embryo develops and needs more oxygen, the heart also starts to beat, which together with the growing tissue oxygen demands and hypoxia induces expression of VEGF and other factors that induce the remodelling and expansion of these primitive vessels, forming progressively a more complex network of arteries, veins and smaller size capillaries. This process of vascular expansion and remodelling of pre-existing vessels is called **angiogenesis** (Poole and Coffin, 1989; Flamme et al., 1997)(Fig. 3). VEGF signalling through *Flk1*

and the downstream activation of *Etv2* is essential for vasculogenesis (Lee et al., 2008; Ferdous et al., 2009), whereas VEGF signalling itself and several other signalling pathways (Dll-Notch, Tie-Ang, Wnt, Tfg-beta/Bmp-Alk, Fgf, Cxcr4/Cxcl12, etc...) regulate angiogenesis. *Etv2* expression is high during vasculogenesis and downregulated during angiogenesis. Most of the pathways important for angiogenesis are inactivated once angiogenesis is completed and the vessels become quiescent, whereas others have important functions for vascular homeostasis.

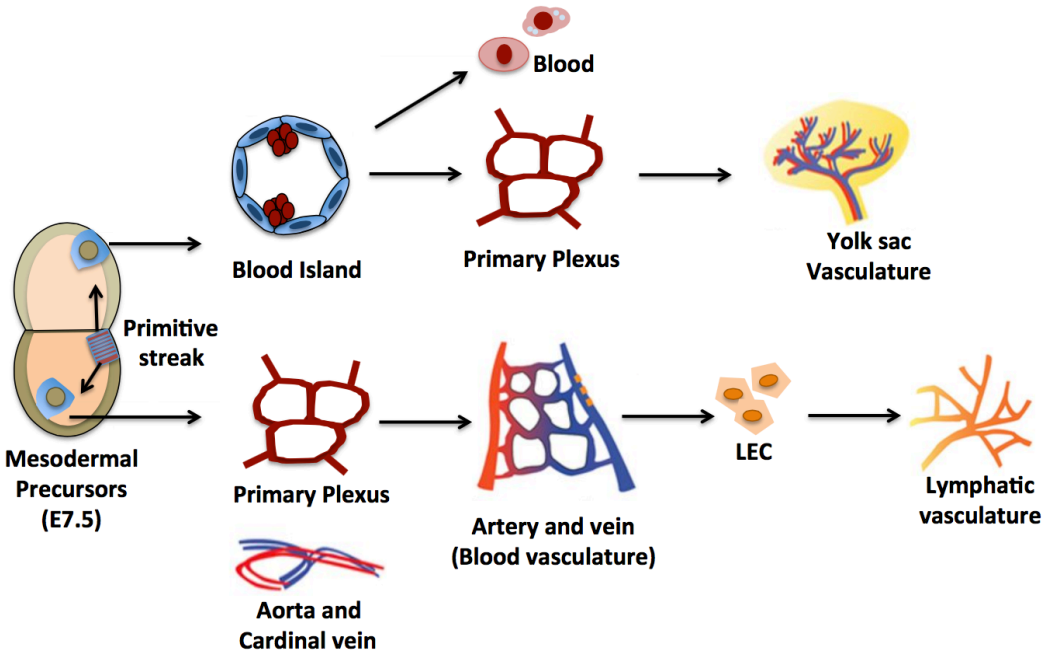


Figure 3: Stage wise development of vessels in extraembryonic and embryonic tissues: Blood islands containing endothelial (blue) and blood (red) cells are formed in extraembryonic tissues. In these islands, most endothelial and blood cells arise from independent progenitors, whereas some primitive blood cells are formed from hemogenic endothelial cells (ECs). In the embryo proper, angioblasts form the primary vascular plexus that later expands and remodels into major vessels (aorta, cardinal vein and capillaries). A portion of cardinal vein ECs at E9.5 forms lymphatic vessels after acquisition of lymphatic endothelial cell (LEC) fate.

1.3. Hematopoietic development

Three waves of differentiation occur during early development in order to form the transient embryonic blood and the definitive hematopoietic stem cells.

a) Primitive hematopoiesis- This first wave of differentiation gives rise to primitive erythroid cells, macrophages and megakaryocytes at day E7.5. It occurs in the blood Islands of the yolk sac. These cells provide the immediate necessary hematopoietic functions but die off before reaching adulthood.

b) Transient definitive wave- Starts at around day E8.25, soon after the first wave and leads to formation of definitive myeloid and red blood cells from multipotent myeloid and erythroid progenitors (Mikkola et al., 2003). These cells lack self-renewing and vigorous lymphoid potential.

c) Long-term definitive wave- The third and last wave of hematopoietic development starts at E9.5-E10.5 and produces multipotent hematopoietic stem cells (HSCs) that have self-renewing potential and metamorphose into erythroid, myeloid and lymphoid cells. A special hemogenic endothelium gives rise to hematopoietic progenitors and HSCs which is present in dorsal aorta's ventral wall and additional major arteries (Bertrand et al., 2010; Boisset et al., 2010). Eventhough the hemogenic endothelium is found in several anatomic sites as the aorta-gonad-mesonephros region (AGM) (Kyba et al., 2002; Galic et al., 2006), yolk sac (de Bruijn et al., 2000), and placenta (Rhodes et al., 2008), it is believed that only the DA endothelium localized in the AGM is able to give rise to definitive HSCs (Gritz et al., 2016). HSCs later expand and migrate to the fetal liver and after birth they enter the bone marrow and are mostly quiescent, unless they are challenged or hematopoiesis is needed (Kim et al., 2007) (Fig. 4).

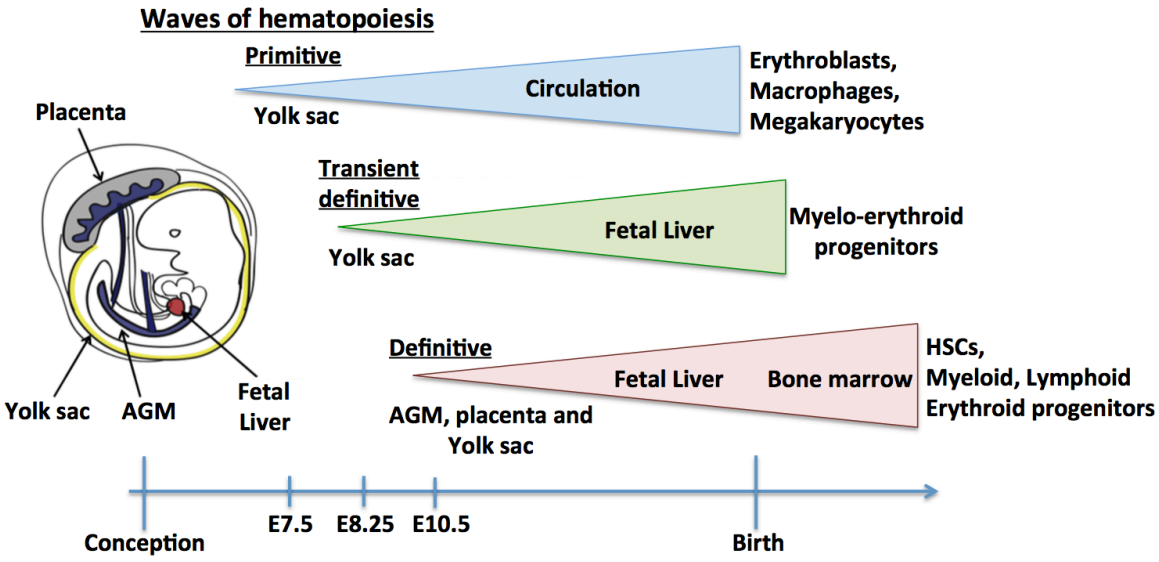


Figure 4: Three wave of Hematopoietic development: Primitive RBCs, macrophages and megakaryocytes are produced in E7.5 mouse yolk sac at the onset of the primitive wave and they enter the circulation. Myelo-erythroid progenitors cells are formed by the transient definitive wave in the yolk sac at E8.25 and these cells further differentiate after migration to the fetal liver. From E9.5-E10.5, multipotent self-renewing hematopoietic stem cells from AGM, yolk sac and placenta are produced and constitute the definitive wave of differentiation. The HSCs and other hematopoietic progenitors further expand in the fetal liver after migration there and after mobilize to the bone marrow.

1.4. Hemato-vascular progenitors: The tussle between Hemangioblast and Hemogenic endothelium.

Studies going back a century, spawned the concepts of hemangioblast and hemogenic endothelium. In the year 1920, observed by Sabin, ECs in a chick embryo showed budding of 'red blood corpuscles' and Murray coined this cell population as 'hemangioblast'(Murray et al., 1932; Sabin et al., 1929). Now, hemogenic endothelium, which produces blood, is thought to link better with this description.

Hemangioblast: Hematopoietic cells and ECs are mesodermal derivatives and were first thought to arise from a common bipotent progenitor named as hemangioblast, due to the close link between vessel formation and hematopoiesis in the embryo, shared gene expression profiles and the ability to differentiate progenitor cells to both endothelial and hematopoietic lineages *in vitro* (Murray et al., 1932; Sabin et al., 1929; Choi et al., 2002). *In vitro* blast colony-forming cells (which are a transient cell population) express markers of hematopoietic and endothelial lineages and are formed when ES cells are stimulated by vascular endothelial growth factor (VEGF) that activates the receptor *Vegfr2/Flk1/Kdr* (Choi et al., 1998; Faloon et al., 2000; Chung et al., 2002). These cells undergo clonal expansion and differentiation into distinct populations of hematopoietic and EC upon re-plating. Blast colonies formed by Brachyury/green fluorescent protein $(GFP)^{+}FLK1^{+}$ cells that appeared at E7.5 embryos were reported in mice to be similar to those formed by ES cells and differentiated into hematopoietic and ECs in the yolk sac (Huber et al., 2004). However, more advanced cell lineage/tracing experiments by transplantation or genetic pulse and chase clonal analysis showed that most hematopoietic and endothelial cells in the yolk sac arise from independent progenitors localized in the lateral plate mesoderm (Padron-Barthe et al., 2014; Kinder et al., 1999; Uneo et al., 2006). According to the latest single cell gene expression analysis, the hematopoietic program initiates at the stage where cells being immersed in an endothelial layer lack endothelial potential (Swiers et al., 2013). This suggests that segregation of hematopoietic and endothelial lineages occur earlier than was initially thought to be, and that common progenitors or hemangioblasts are either non-existent or very infrequent.

Hemogenic Endothelium: It is a population of ECs with the ability to directly generate hematopoietic cells.

Lineage tracing studies using *VE-cadherin-CreERT2* and *Rosa26-LSL-LacZ* alleles showed for the first time that VE-cadherin+ lineage cells present in the aorta-gonad-mesonephros generated hematopoietic stem cells (Zovein et al., 2008). This observation was reconfirmed later *in vitro*, in mice and zebrafish by direct imaging of endothelial cells transdifferentiating to blood cells (Eilken et al., 2009; Bertrand et

al., 2010; Boisset et al., 2010; Kissa et al., 2010. These studies have shown that hemogenic ECs are flat in shape and form adherent or tight endothelial junctions. Later, they go through an endothelial-to-hematopoietic transition (EHT) and acquire the round shape characteristic of blood cells.

Endothelial cells with hemogenic capacity were found in the yolk sac, aorta-gonad-mesonephros (AGM), allantoi and the endocardium region of a developing embryo (Nakano et al., 2013; Boisset et al., 2010; Zovein et al., 2008; Jaffredo et al., 1998). These cells were seen to undergo the same process observed during differentiation of ES cells in vitro (Lancrin et al., 2009; Ahmed et al., 2016; Eilken et al., 2009; Nishikawa et al., 1998). Recent studies have also confirmed that hemogenic endothelium ocurring in the Dorsal aorta AGM can give rise to definitive multilineage hematopoietic stem cells that later are able to produce all the different kinds of blood cells (Hirschi et al., 2012; Zape et al., 2011).

Hemogenic endothelium with its specific endothelial morphology was thought to be able to give rise to both hematopoietic and endothelial cells as it expresses markers for both lineages; but advanced lineage tracing studies conducted recently report that in these cells the endothelial potential is completely lost and they only form blood (Swiers et al., 2013; Ditadi et al., 2015).

1.5 The role of VEGFA-VEGFR2/Flk1 and Etv2 in endothelial-hematopoietic differentiation and development

Many aspects of endothelial cell biology are controlled by VEGF signaling which includes functions such as endothelial differentiation, survival, proliferation, migration and vascular permeability (Fig. 5). *Flk1/VEGFR2/Kdr* is the major receptor involved in these functions. Binding of the ligand VEGF-A to the receptor *VEGFR2(Flk1*) triggers the activation of downstream molecules such as *PI3K/Akt/protein kinase B, MAPK-ERK, p38-MAPK* and *phospholipase C* and these after execute the diverse endothelial functions of VEGF. Mice deficient in *Vegfa* or *Vegfr2/Flk1* die in utero at 8.5-9.5d postcoitum due to absence of endothelial and hematopoietic cells (Shalaby et al., 1995; Carmeliet et al., 1996; Shalaby et al., 1995).

VEGF and *Flk1* are essential for endothelial progenitor cells proliferation, migration and differentiation (Shalaby et al., 1997 Carmeliet et al., 1996). Besides *Vegfr2/Flk1*, *Vegfr3* and *Vegfr1* are also expressed but they have distinct functions. *Vegfr1* has a very weak kinase activity and its soluble form is released to the extracellular space, acting mostly as a *VEGF* decoy molecule and as a negative regulator of VEGF signalling. *Vegfr3* does not bind *VEGFA*, it binds mostly *VEGFC*, and has no role in early vasculogenesis.

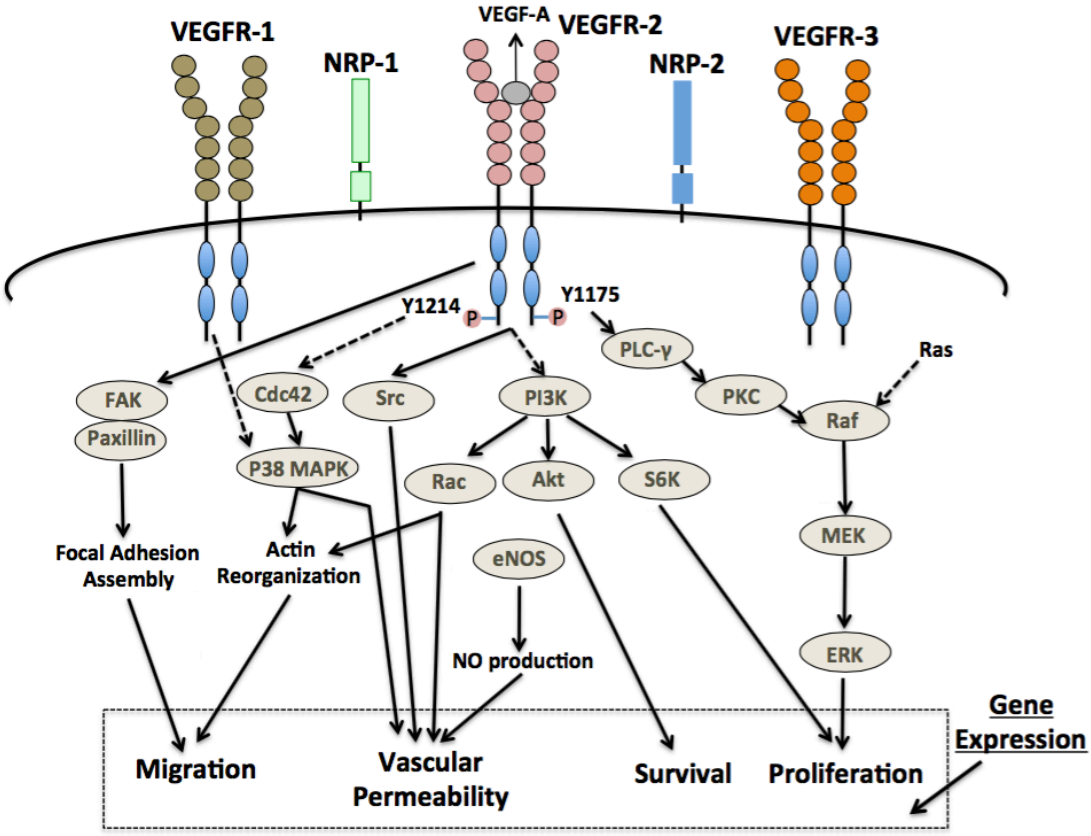


Figure 5: Signal transduction downstream of *Flk1/VEGFR2* receptor: *Flk1/VEGFR2* mediates vasculogenesis and angiogenesis through the activation of multiple signalling cascades that are essential for EC proliferation, migration, survival and vascular permeability**.**

An increase in VEGF signaling in mesodermal progenitor cells is thought to activate the expression of several ETS transcription factors (Kataoka et al., 2011; Rasmussen et al., 2012) of which *ERG, Fli1* and *Etv2* are quite specific to the endothelial lineages. The ETS family is present throughout the body and is involved in a wide variety of functions including the regulation of cellular differentiation, cell cycle control, cell migration, cell proliferation, apoptosis (programmed cell death) and angiogenesis.

From these, the only single ETS factor that is essential for all endothelial and hematopoietic development is Etv2. *Etv2* KO embryos or ES cells in vitro cannot form any endothelial or hematopoietic cells (Lee et al., 2008) whereas endothelial and hematopoietic progenitors can still arise in Flk1 knockout embryos or ES cells (Shalaby et al., 1995; Shalaby et al., 1997; Schuh et al., 1999). Forced *Etv2* expression in wildtype or *Flk1* KO cells, can also induce hematoendothelial differentiation (Lee et al., 2008; Sumanas et al., 2008; Liu et al., 2015). Therefore, Etv2 is considered to be the master regulator of endothelial/blood cell lineages development, downstream of VEGF/Flk1 activation.

Production *of Flk1+/Pdgfra+* double positive unpatterned progenitors takes place in Etv2 mutant ES/EB system and embryos, but their progression to *Flk1+/Pdgfra-* vascular mesoderm does not come about. When *Etv2, Scl* or *Fli1* is over-expressed, this block is removed (Liu et al., 2012; Katoka et al., 2011; Rasmussen et al., 2013). Over-expression of *Etv2* rescues haemato-endothelial deficiencies of $Flk1^{-/-}$ ES cells (Rasmussen et al., 2013). These results prove that *Flk1* acts upstream of Etv2 at early mesoderm differentiation stages and identify the nodal point for Etv2 regulation as the *Flk1+/Pdgfra+*→*Flk1+/Pdgfra-* transition (Katoka et al., 2011; Rasmussen et al., 2013). Confirming this, *Flk1* mutant embryos had significantly lowered *Etv2* expression but Etv2 mutants had little affected *Flk1* expression (Rasmussen et al., 2012; Ishitobi et al., 2011). In differentiated endothelial cells, *Etv2* reciprocally feedbacks and retro-regulates *Flk1* expression (Lee et al., 2008; De Val et al., 2009). During development, recurring regulatory loops are common for induction of haemato-endothelial genes (Wareing et al., 2012). Haemato-endothelial progenitors can be independently induced by *Etv2, Gata2*, and *Scl* (Lugus et al., 2007; Liu et al., 2013; Ismailoglu et al., 2008). But overexpression of the 3 factors together worked

better for progression of *Flk1+Pdgfra-* endothelial lineage (Liu et al., 2013). On the other hand, over-expression of *Fli1, Erg* or *Ets1* (other Ets factors) with *Gata2* and *Scl* could not effectively induce the above lineage suggesting that is the interaction between *Etv2, Gata2* and *Scl* that drives differentiation (Koyano et al., 2012; liu et al., 2015).

Despite identification of several transcriptional activators as downstream targets of Etv2, it is not known what transcriptional mechanisms cause down-regulation of *Pdgfra* in the *Flk1+/Pdgfra+→Flk1+/Pdgfra-* transition. *Etv2* might be initiating the hematopoietic/endothelial lineage by activation of their respective factors and establishment of Ets hierarchy (Liu et al., 2015).

Etv2 expression is partially independent of *Flk1* or *Vegf,* but *Vegf* is able to induce expression of Etv2 to the levels required for endothelial differentiation. *Etv2* and *Flk1* regulate each other in a positive feedback mechanism to generate $Etv2^{high}Flk1^{high}$ state cells, the cells that can later efficiently differentiate to the endothelial-hematopoietic lineages (Rasmussen et al., 2012; Rasmussen et al., 2013). In the Etv2-deficient embryos, only $Flk1^{high}$ cells were absent (Liu et al., 2015). It was also proposed that *Etv2* expression initiation can take place independently of VEGF-Flk1 signaling, but this pathway is later required to achieve the high threshold of *Etv2* expression. Thus VEGF signaling enforces the hematoendothelial fate via induction of *Etv2* above the required threshold of expression.

The embryo hematopoietic cells can be formed directly from hemogenic endothelium or from Hematoblasts. Expression of the key regulators of hematopoiesis, i.e. *Tal1 (Scl), Runx1, GATA1,* and *GATA2* is induced by *Flk1* and *Etv2* in hematoblasts (Shi et al., 2014; Morita et al., 2015; Xu et al., 2017). *Gata2* mutants developmental defects are restricted to the blood lineages (Tsai et al., 1994) whereas *Flk1* and *Etv2* mutants have both hematopoietic and vascular defects (Kataoka et al., 2011). *Tal1* regulates *Runx1* directly upstream and is crucial for primitive and definitive hematopoiesis (Landry et al., 2008; Shalaby et al., 1997). *Tal1* is essential for erythropoiesis in the yolk sac, whereas both *Runx1* and *Tal1* are required for the endothelial to hematopoietic transition (Shivdasani et al., 1995; Lancrin et al 2009; Zhen et al., 2013).

1.6 CRISPR-Cas9 pooled genetic screens

Precise genetic perturbations can be achieved with the recently found CRISPR-Cas9 system. This precision comes from a component of microbial CRISPR immune system i.e. RNA-guided endonuclease Cas9 from *S. pyogenes* (Mali P. et al., 2013; Cong L. et al., 2013). In its application for genome editing, the endogenous prokaryotic RNA components are replaced with synthetic single guide RNA (sgRNA) which is complementary to a given desired target site having a 20-bp sequence followed by a PAM sequence (Jinek M. etal., 2012). A precise double strand break (DSB) is created by the Cas9-sgRNA ribonucleoprotein complex after its recognition of the target sequence in the genome. These DSBs created at target loci undergo repair using the rare homology-directed repair (HDR) mechanism or the more frequent nonhomologous end-joining (NHEJ) pathway. The first is a precise repair system that uses homologous DNA template, whereas the second is far more frequent and introduces indels at the DSB. Therefore, in cells with Cas9 activity and expressing a given synthetic guide RNA, loss-of-function (LOF) mutations can occur if the target sequence is in a protein-coding region. This happens because the introduced INDEL, frequently causes a premature frameshift generating a truncation or alteration of the protein coding sequence, leading to non-sense mediated decay of the RNA or the generation of a non-functional protein.

The CRISPR pooled genetic screen approach consists in delivering to every cell, in a pool of cells, a distinct genetic perturbation, such as a guide RNA recognizing a specific target sequence, that will disrupt the function of a specific gene. When thousands of different cells contain different genetic perturbations, they will proliferate or differentiate differently, according to the perturbation they received.
By analyzing how the pooled genetic modifications evolve, is possible to identify which perturbations/genes influence a particular cellular phenotype. These forward genetic screens use a 'phenotype-to-genotype' approach allowing us to identify specific genetic perturbations that induce a desired phenotype (e.g. induce drug resistance, or expression of a specific cell surface receptor, or a specific cell lineage differentiation or increased proliferation). In this pooled approach, there is usually a

strong selection pressure that leads to an amplification or depletion of perturbations relevant to the analysed phenotype. These pooled genetic screens are high-throughput and scalable, allowing the screening of thousands to millions of perturbations, which is compatible with whole genome forward genetic screens (Shalem et al., 2014; Wang et al., 2014; Koike-Yusa et al., 2014; Zhou et al., 2014).

Most frequently, CRISPR pooled genetic screens start by transducing millions of cells at low multiplicity of infection (MOI) with a lentiviral-pooled library consisting of thousands of unique sgRNAs. Each of these virus expressing a given sgRNA will integrate in the cell genome and mediate a unique gene mutation (perturbation). The second step is to apply a selective pressure to these cells (of different natures) which identifies sgRNAs targeting genes that elicit a specific biological response. This pressure ensures that a variable fraction of sgRNA-containing cells to be either amplified or reduced in the cell population. Genomic DNA (gDNA) is isolated from the starting transduced cells (reference cells) and final transduced cells (experimental cells) for PCR amplification of the barcoded and infected cells and identification of guideRNAs/gene mutations that favour or inhibit a given biological process. PCR amplification of sgRNA sequences from this isolated gDNA is carried out using Illumina-adapted Forward and Reverse Indexed primers for minimal amplification bias. The amplified indexed PCR products are loaded onto Illumina flow cells and sequenced employing Decode sequencing primers. Reference and experimental cells reveal differences in the relative sgRNA abundance between them (Fig. 6). Therefore, one screen can provide a plethora of information on genotype-phenotype interactions which can be followed up after.

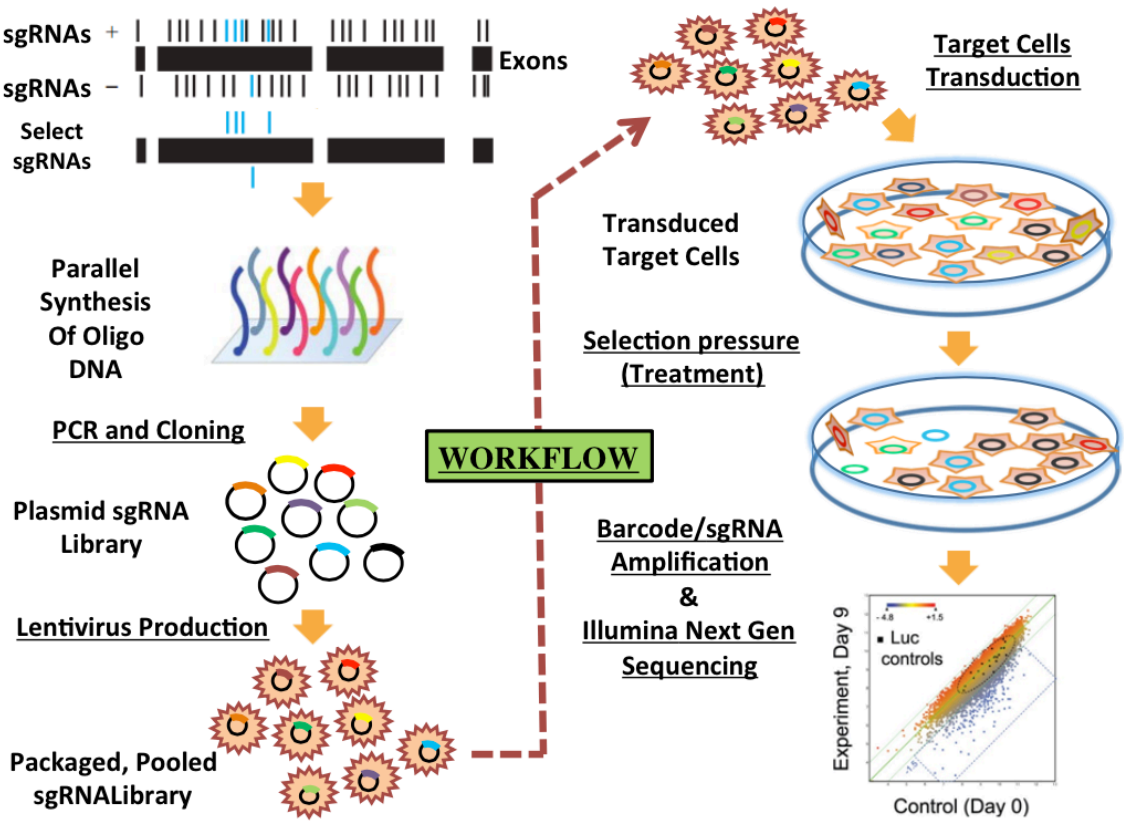


Figure 6: Workflow of a typical pooled approach for Genetic screen utilizing CRISPR/Cas9 genome editing tool.

1.7 Embryonic stem cells and *in vitro* endothelial differentiation

The inner cell mass (ICM) of a pre-implantation embryo can give rise to embryonic stem cells (ESCs) *in vitro* (Thomson et al., 1998). Several pharmacological compounds (2i) or the Leukemia inhibitory factor (LIF) can be used to maintain them in an undifferentiated state *in vitro*. Alternatively, they can be cultured on mitotically inactivated mouse embryonic fibroblasts (MEFs) layer (Reubinoff et al., 2000). These ICM derived cells possess pluripotency, i.e. the ability to generate all germ layer cells and differentiation into all cell types.

The study of ESC differentiation can provide insights into early embryonic development mechanisms and the underlying molecular regulations. When pluripotency maintenance factors are removed from the culture medium, ESCs form

multi-cellular aggregates called as embryoid bodies (EBs), with formation of ectoderm, mesoderm, endoderm and their derivatives (Doss et al., 2012; Stevens, 1960) (Figure: 7). EBs formed from ES cells can thus aid in the study of the mechanisms responsible for mesoderm diversification and cell lineages development.

A more efficient way to induce endothelial/hematopoietic differentiation is to co-culture ES cells with OP9 cells (Figure: 9). The OP9 stromal cell line comes from calvaria of osteopetrosis mutant mice in which M-CSF is non-functional. OP9 cells secrete factors that enhance the development of different endothelial and hematopoietic progenitors

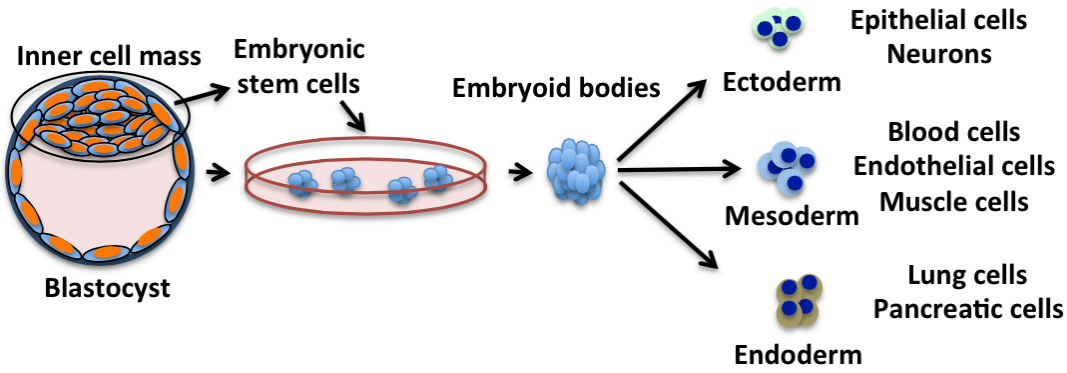


Figure 7: Embryoid body differentiation from Blastocyst derived ES cells: Inner cell mass of blastocyst form ES cells that can be cultured to form embryoid bodies. They later differentiate into endoderm, mesoderm and ectoderm, the three germ layers that further give rise to respective lineages.

Given the relative ease of generating genetically modified ES cells *in vitro*, and the potential to differentiate these cells *in vitro* towards the endothelial and hematopoietic cell lineages when co-cultured with OP9 cells, it is a very convenient system to study the impact of genes in early stages of endothelial and hematopoietic differentiation. It can also be combined with CRISPR-Cas9 pooled lentiviral libraries, to perform large-scale genetic screens.

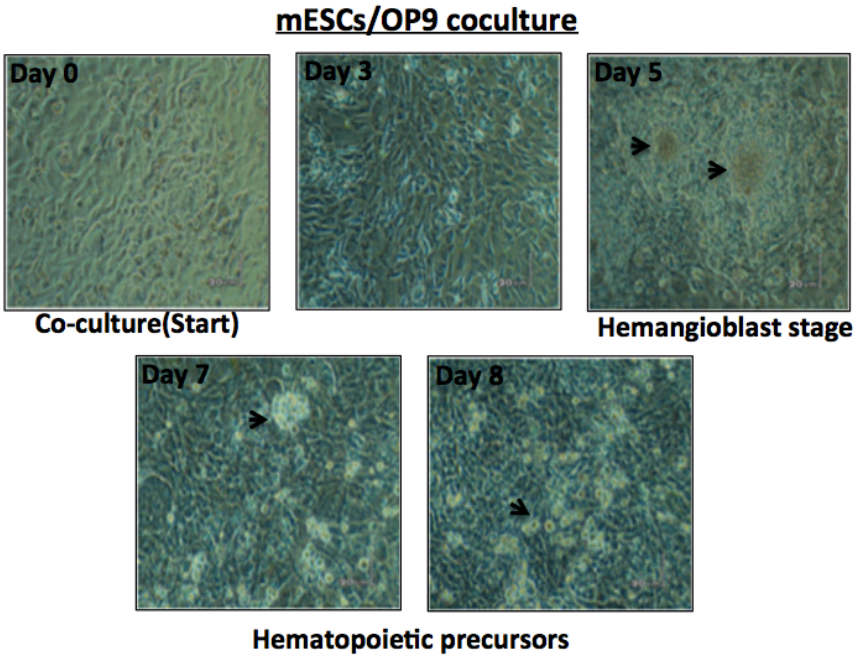


Figure 8: mESC/OP9 coculture differentiation kinetics: **Day0:** ES cells are adsorbed on a confluent OP9 cell layer; **Day3:** Visibility of ES-derived progeny which look like a cell cluster in a whorl pattern; **Day5:** By this day, hemangioblasts, the common mesodermal precursors for endothelial and hematopoietic lineages are seen as piled up cell whorls. Beyond 5 days of co-culture, ES-derived progeny appear as hematopoietic cell clusters; **Day 6/7**: 2-4 celled clusters appear which grow larger by Day 8/9 of co-culture. The stromal cell layer also holds other ES-derived progeny bound to it.

OBJECTIVES

The differentiation of blood and endothelial progenitors is tightly dependent on the *FLK1- VEGF* signalling axis. However little is known about the function of other genes that may cooperate, antagonize or partially substitute the function of *Flk1* during this process. The main objective of this PhD thesis was to obtain a more comprehensive and deeper overview of the genetic networks linked with the *Flk1* function in this early and important developmental process.

Main objectives of this thesis were as follows:

1) Generation of cell lines and global transcriptional profiling of *Flk1* heterozygous (*Flk1*$^{KI/WT}$) and homozygous (*Flk1*$^{KI/KI}$) reporter knock-in differentiated cells, to see which genes and pathways are regulated by *Flk1* during the endothelial-hematopoietic differentiation process *in vitro*.

2) Generation of mice and global transcriptional profiling of *Flk1* heterozygous (*Flk1*$^{KI/WT}$) and homozygous (*Flk1*$^{KI/KI}$) reporter knock-in endothelial cells from E9.5 mouse embryos.

3) Large scale CRISPR/Cas9 based mutagenesis screen on mESCs and differentiated ECs with or without loss of *Flk1* function **(***Flk1*$^{KI/KI}$**)** to identify new genes or pathways that can control the normal endothelial differentiation process and rescue the *Flk1* loss-of-function phenotype.

Material And Methods:

ES cell culture and genotyping

The basic elements of the DNA constructs for CRISPR targeting (Fig. 1 and Fig. 2) were obtained from varied sources and assembled by standard DNA. Mouse ES cells on the G4 background50 were cultured in standard ES cell culture medium (DMEM containing Glutamax (31966–047, Gibco), 15% FBS (tested for germline transmission), 1x NEAA (Hyclone, SH3023801), 0,1% ß- mercaptoethanol (Sigma, M7522), 1x Pen/Strep (Lonza, DE17-602E) and LIF) in dishes covered with a feeder layer of mouse embryonic fibroblasts (MEFs). For Cas9 gene targeting with the large plasmids for Flk1 targeting and Rosa26 Cas9 targeting, 25 ug of linearized DNA along with Cas9 vector plasmid PX330 (Addgene) was used to electroporate 5 million ES cells. Selection in 200 ug/ml G418 (Geneticin) was performed for 6 days, after which individual colonies were picked for storage, PCR, and Southern blot screening. Selected positive clones were expanded and used for microinjection in host blastocysts of the C57Bl/6 J strain. Chimeras with a high percentage of agouti coat color were then crossed with mice to obtain germline transmission of the targeted insertion. PCR with Flk1 and ROSA26 5'homology arm flanking primers allowed us to identify ES cell clones with precise homologous recombination and insertion of the constructs. Selected ES cell clones were expanded for further analysis and mouse generation.

Embryoid Bodies and OP9/mESCs coculture for endothelial differentiation

To generate embryoid bodies (EBs) from ES cells we used the standard hanging drop method. Briefly ES cells were first grown for 2 days on gelatinized plates and after, tripsinized and ressuspended at a density of 60.000 cells per ml in embryoid bodies media (DMEM Glutamax (Gibco, 31966-047), 15% FBS, HEPES (Biowhittaker EE17-737) and Monotyoglycerol (Sigma, M6145). For each EB, 20ul drops of this solution were pipetted onto the lid of a petri dish. This lid was inverted, to form the hanging drops, and the dish further filled with PBS to prevent evaporation. Four days after differentiation the embryoid bodies were plated on an OP9 cells mono- layer and differentiated in basal media (MEM alpha (Gibco, 11900-016), supplemented with

20% FBS and 7.5% Sodium bicarbonate (Gibco, 25080-060)) containing 30ng/ml of VEGF to further induce endothelial differentiation and proliferation for 5 days.

ES cells upon culture on OP9 stromal cells (aMEM media with OP9 tested 10%FBS), form Flk-1+ endothelial and hematopoietic progenitors, and in the end mature, terminally differentiated lineages. The protocol for this system includes plating of ES cells onto a merging OP9 cell layer and regular re-plating to get rid of old OP9 cells to prevent contamination. If ES cells properly differentiate on the OP9 cell layer, Flk1+ hemangioblasts are formed as piles of cell whorls on day 5 of co-culture. Later in co-culture, progeny of ES cells are visible in the form of clustered hematopoietic cells. At Day 6/7 of co-culture, there are 2-4 clusters which grow larger (both adherent and in suspension) by Day 8/9. The adherent stromal cell layer contains tightly bound other progeny of ES cells. But these OP9 cells are sensitive to changes in surrounding conditions of medium source and serum lot, which can change their ability to support hematopoiesis.

Immunostainings

For immunostaining of ES cells or ECs derived from embryoid bodies and Op9 co-culture, cells were fixed for 10 minutes in PBS containing PFA 4% and Sucrose 4%. After a brief rinse in PBS, cells were permeabilized in 0.1% Triton for 10 minutes and then immersed in a blocking solution (10% Fetal bovine serum in PBS). Primary antibodies were diluted in blocking solution and incubated for 2 hours at room temperature or overnight, followed by three washes in PBS of 10 minutes each and incubation for 1 to 2 hours with conjugated secondary antibodies at room temperature. After three washes in PBS, cells were mounted with Fluoromount-G (SouthernBiotech).

For whole-mount Embryo Staining: Once isolated the embryos were collected in PBS and fixed in 2% PFA for 20 mins at 4C. Embryos were then washed three times for 15 mins each round in PBS. They were then subjected to a serial dehydration with increasing methanol/PBS dilution (50%,75%,100%; 30 mins each) and stored in 100% methanol until ready for processing. For processing, they are rehydrated back to PBS using the increasing PBS/methanol dilutions. Embryos were then blocked and permeabilized overnight at 4C in blocking solution(1% skim milk, 0.4% Triton and

0.5% BSA in PBS). Primary antibodies were diluted in working solution(1% skim milk, 0.4% Triton in PBS) and tissues were incubated overnight at 4C. After three washes in working solution, each 30 mins on ice, Alexa conjugated secondary antibodies were incubated in working solution overnight at 4C. After three washes in PBS, embryos were subjected to serial dehydration in glycerol(20%, 40%, and 60%. Embryos were then mounted in glycerol solution on a Fast well chamber. A coverslip was placed on the Fast well and the sample was stored in 4C until ready for imaging.

Primary antibodies were used against the following proteins: Erg (Abcam, ab110639 1:500); GFP/YFP/Cerulean (Acris, R1091P, 1:400); Ki67 (Thermo Fisher, RM-9106-S0, 1:400); Tomato (Clonetech, 632496, 1:400); Icam2 (Pharmingen, 553325, 1:200); Flk1 (Pharmingen, 550549, 1:200). To combine multiple antibodies in the same immunodetection, in some instances the following proteins were detected with conjugated primary antibodies: GFP/YFP/ Cerulean (AF-488, Invitrogen A213111, 1:100); Erg (AF-647, Abcam ab196149, 1:100); HA (AF-647, Cell Signalling 3444S, 1:100 or AF-594, Thermo Fisher A- 21288 1:200); Tomato (CF-594, Biotium 20422, 1:400). To detect two primary antibodies from the same host, we combined detection with conjugated Fab fragment secondary antibodies and directly conjugated primary antibodies. The following secondary antibodies were used: donkey anti-rabbit (Thermo Fisher, AF-488, A-21206 or AF-680, A-10043); donkey anti-goat (Thermo Fisher, AF-488, A-11055 or AF-633, A21082 or AF-647, A-21447 or AF-680, A-21447); donkey anti-rabbit Fab fragment (Jackson Immunoresearch, Cy3, 711-167-003 or AF-647 711-607-003); and streptavidin 405 (Thermo Fisher, S-32351, 1:200).

Microscopy

Depending on the complexity of the immunostainings and the combination of FPs to detect, we used different laser-scanning confocal microscopes. For up to 4 channels acquisition of large fields we used the ZEISS LSM700 inverted microscope with laser lines 405, 488, 546 and 633nm. For multi-color detection of up to 7 different signals we used the inverted Leica SP5 confocal (405, 488, 514, 546, 594, 633nm) or the

Leica SP8 confocal with a 405nm laser and a white laser that allows excitation at any wave- length from 470nm to 670nm.

Flow cytometry and qRT-PCR analysis

ES cells were differentiated to embryoid bodies and endothelial cells, as shown above, and after trypsinization, cells were analysed in a FACSAria containing lasers to detect Cerulean and Alexa secondary antibodies

To detect, separate, and profile the different endothelial cell populations in Flk1KI/WT-Tie2Cre and Flk1Flox/Flox-iSuRe-Cre mice, we interbred these mice animals to activate recombination and fluorescent protein expression in the endothelium. Embryos were collected in PBS, minced, and digested with 2.5 mg/ml type I collagenase (Thermo Fisher, 17100017), 2.5 mg/ml dispase II (Thermo Fisher, 17105041), and 50 ng/ml DNAseI (Roche) at 37 °C for 20 min to create a homogeneous cell suspension. Cell suspensions were passed through a 70 µm filter to remove any undigested tissue. To remove erythroid cells, cell suspensions and blood samples were incubated for 10 min on ice in blood lysis buffer (0.15 M NH4Cl, 0.01 M KHCO3 and 0.01 M EDTA in ddH20). Before analysis, cell suspensions were incubated at 4 °C for 30 min with APC rat anti-mouse CD31 (BD Pharmigen) diluted 1:200. The Flow cytometry analyses were performed either with a FACS Aria Cell Sorter (BD Biosciences) or a Synergy4L Sorter. Viable (DAPI-) ECs (APC-CD31+) were sorted and analysed according to their endogenous fluorescence (GFP, Cerulean and MbTomato). The resulting cell suspension was incubated with conjugated APC-anti-CD31 antibody (BD Pharmingen, 551262), followed by DAPI staining to exclude dead cells. Viable (DAPI−) ECs (APC-CD31+) was FACS sorted and processed for total RNA extraction with the RNAeasy Micro kit (Qiagen).

RNA extraction for the sorted endothelial cells was done using the RNAeasy Micro Kit. For quantitative real time PCR (qRT-PCR), RNA extracted from the ECs obtained as above was retrotranscribed with the High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Thermo fisher, 4368814). cDNA was preamplified with Taqman PreAmp Master Mix containing a Taqman Assay-based pre-amplification pool containing a mix of the following Taqman assays (Applied Biosystems): Actb, Gapdh,

Pecam1, Cdh5, Esm1 and Flk1. Preamplified cDNA was used in standard qRT-PCR with gene- specific Taqman Assays (Thermo fisher) in a AB7900 thermocycler (Applied Biosystems).

Western blot analysis

For the analysis of Cas9 protein expression, Rosa 26 Cas9 targeted ES cells were lysed were lysed with RIPA buffer (Sigma R0278) containing protease inhibitors (Sigma P-8340) and phosphatase inhibitors (Calbiochem 524629) and orthovanadate-Na 1 mM. Tissue/ cell debris was removed by centrifugation, and the supernatant was diluted in 3xSDS loading buffer and analysed by SDS–PAGE and immunoblotting. Membranes were blocked with BSA and incubated with primary antibodies diluted 1/1000 against Flag(Merck, 69050–3) and β-Actin (Santa Cruz Biotechnologies, sc-47778).

Bioinformatics Filtering and Enrichment analysis

Global analysis of the genome wide data was performed on usegalaxy.org and on the CNIC High Performance Computing cluster. Acquired RNAseq reads were mapped to the mouse genome (GRCm38/mm10) using TopHat 2.0.9. Transcripts were assembled using Cufflinks 2.0.0 based on the reference genome with quartile normalisation and effective length correction. A combined gtf file was produced using Cuffmerge 2.1.1 and used for determination of the differential gene expression using Cuffdiff 2.0.1. Principle component analysis and hierarchical clustering were performed on all RNAseq samples to assess inter-sample variation. Gene differential expression and FPKM gene expression files were used to select the gene IDs that were significantly differentially expressed, with $P < 0.05$, longer than 200 bp and FPKM >10 in at least one of the samples. Heat maps, Hierarchical Clustering were computed by MultiExperiment Viewer v4.9.0 based on Pearson correlation with complete linkage. Self Organizing Tree Analyses determined the number of clusters. Gene ontology enrichment was performed using DAVID 6.7, GSEA, GREAT and Enrichr API. GO terms for biological processes with *P*-value <0.05 were considered. Bar charts representing the *z*-score values of the terms in each category (biological

process, cellular component, molecular function, and KEGG pathways) were generated with GraphPad Prism 7.

Results:

1. Transcriptional profiling of *Flk1* heterozygous (*Flk1*$^{KI/WT}$) and homozygous (*Flk1*$^{KI/KI}$) reporter knock-in differentiated cells in vitro.

1.1. Generation of *Flk1* Knock-in mESC lines by CRISPR/Cas9

In order to label, target and isolate endothelial progenitor cells with expression of Flk1 we inserted by CRISPR/Cas9 assisted HDR the cassette H2B-Cerulean-V5-2A-CreERT2 downstream of the mouse Flk1 promoter (Fig. 9). We generated ES cell clones having half (Flk1 KI/Wt) or no Flk1 expression (Flk1 KI/KI). The cassette H2B-Cerulean-V5-2A-iCreERT2-Sv40pA-FRT-PGK-Neo-pA-FRT was inserted in the Flk1 starting codon (ATG) by CRISPR/Cas9-induced double-strand break and homologous recombination-dependent repair (Fig. 9). The donor vector contained, two homology arms (HA) with 394bp and 1.17kb, respectively, and was electroporated into ES cells. After selection in G418 (neomycin resistance given by the Neo selection cassette), the resistant clones were screened for correct vector integration, by PCR. 7 out of 24 picked clones had only one allele targeted (heterozygous) and 5 had both alleles targeted (homozygous). To avoid any interference of the PGK-Neomycin-pA (Neo) cassette on the endogenous Flk1 locus gene expression, the FRT-Neo-FRT cassette was flipped out by transient lipofection of these cells with plasmids expressing FlpO.

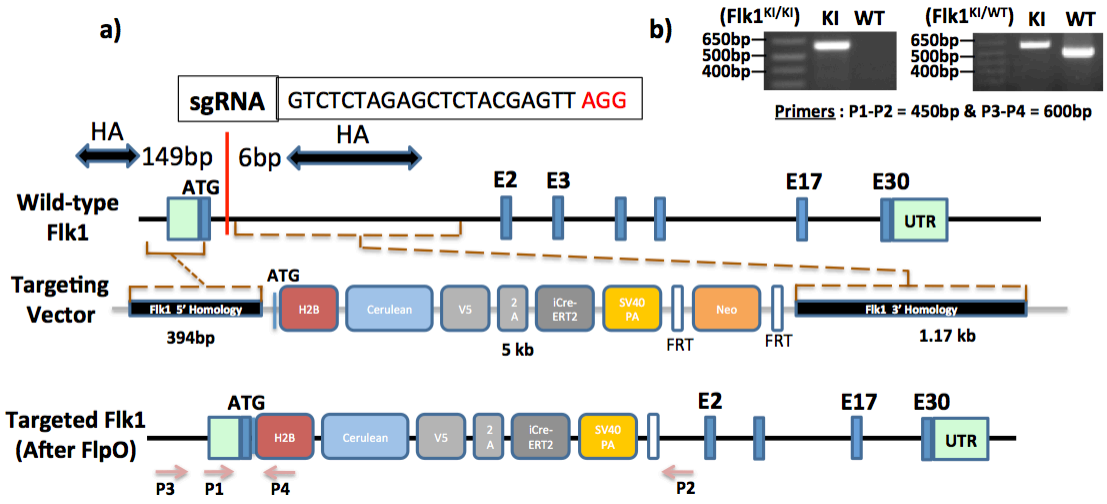


Figure 9: Generation of mouse ES cell lines with one or two knock-ins in the

endogenous mouse Flk1 gene by CRISPR/cas9: (a) Schematic illustration of the wild type Flk1 locus with its exons (E), targeting vector and the finally modified (knocked-in) Flk1 gene. Representative sequences of Wild type Flk1, targeted vector and targeted Flk1 are shown. Dotted line indicates recombination sites. Orange arrows show the position of primers for PCR. **(b)** PCR analysis of genomic DNA from the selected ES cell clones. Upper band (600bp) with primers P3 and P4 indicate the knock-in allele. Shorter band (450bp) with primers P1 and P2 indicate the wild type, non-targeted allele.

1.2. Characterization of *Flk1*$^{KI/WT}$ and *Flk1*$^{KI/KI}$ mESCs

As mentioned in the introduction, when ES cells are cultured on top of a monolayer of OP9 stromal cells, they differentiate into Flk1+ endothelial and hematopoietic progenitors and its lineages. To test the effect of Flk1 targeting on mESC differentiation to endothelial cells, we co-cultured Flk1$^{WT/WT}$ (wild type), Flk1$^{KI/WT}$ (heterozygous) or Flk1$^{KI/KI}$ (homozygous null) mESCs with OP9 stromal cells. Flk1 targeting did not affect mESCs viability or proliferation rate as expected (data not shown). However, it significantly affected endothelial differentiation of Flk1$^{KI/WT}$ and particularly Flk1$^{KI/KI}$ cells, in this 6 days co-culture assay. Immunofluorescence analysis of Cerulean expression revealed its correct labeling of the nuclei of endothelial progenitor and differentiated cells in Flk1$^{KI/WT}$ and Flk-1$^{Ki/Ki}$ clones. Flk1 immunostaining revealed its stronger expression in fully differentiated endothelial cells of Flk1$^{KI/WT}$ and Flk1$^{WT/WT}$ clones, but not in the Flk1$^{KI/KI}$ clone. Pecam1 in this 6 days co-culture labels both endothelial progenitors and differentiated endothelial cells, and was decreased in Flk1$^{KI/WT}$ and particularly the Flk1$^{KI/KI}$ clone (Fig. 10).

These results are in line with previous observations demonstrating that Flk1–null embryonic stem cells generates both ECs and HPCs *in vitro*, but at much lower efficiencies (Hidaka, M. *et al.*, 1999 and Schuh *et al.*, 1999). This population of Flk1-Cerulean+/Pecam+ cells most likely represents endothelial/hematopoietic precursors, arrested in their development because of suboptimal VEGF-Flk1 signaling at critical stages of vasculogenesis. In contrast to the Flk1$^{KI/WT}$ cells, the Flk1 null

(KI/KI) PECAM+ cells do not show any sprouting morphology. Note that in vitro, PECAM1 is expressed by both endothelial and non-endothelial progenitor cells.

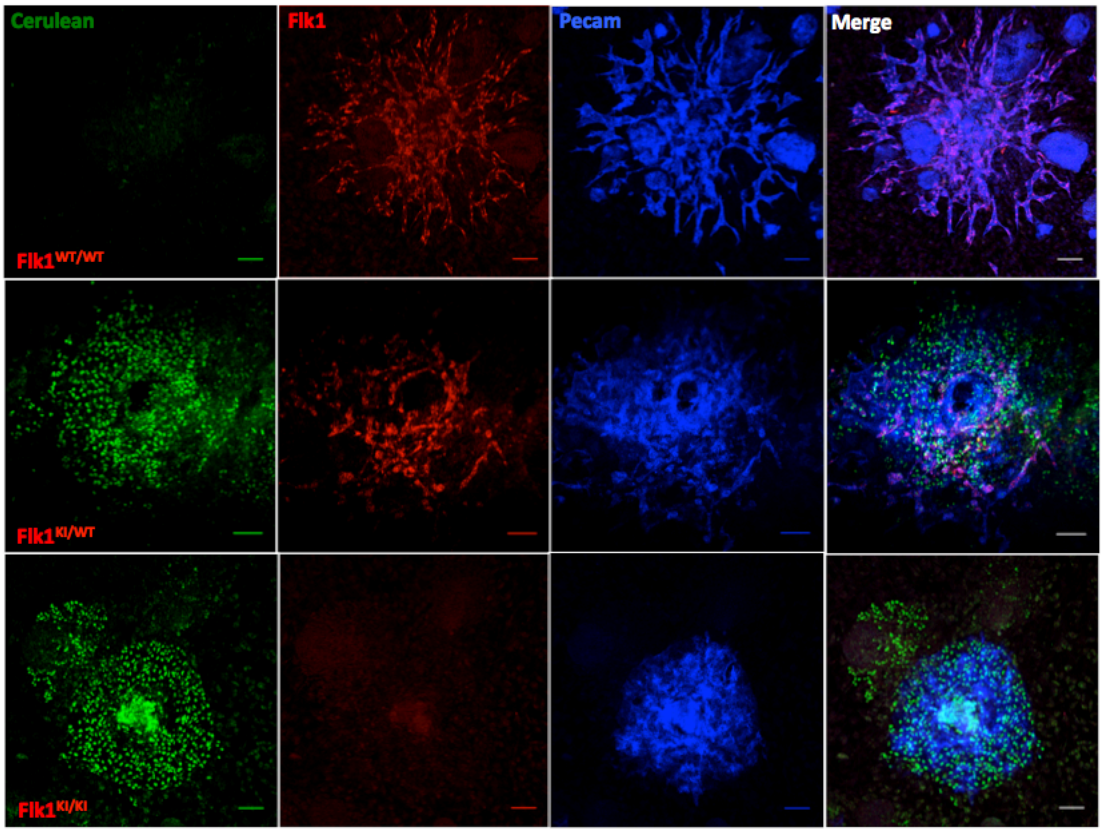


Figure 10: Immunostaing of mESC lines Flk1$^{WT/WT}$, Flk1$^{KI/WT}$and Flk1$^{KI/KI}$ differentiated for 6 days on OP9 stromal cell line (mESC/OP9 coculture) for Pecam1, Flk1 and the cerulean fluorescent protein. (Scale bar = 100um)

1.3. Isolation and transcriptional profiling of endothelial progenitor cells with distinct *Flk1* levels

Despite the importance of the Flk1 receptor for endothelial specification and in angiogenesis, little is known about the molecular mechanisms elicited by Flk1 in progenitor and differentiated endothelial cells, and more importantly, what alternative cell differentiation paths and molecular mechanisms are elicited when the Flk1 function is absent from these cells. The understanding of these mechanisms could offer us clues not only about Flk1 biology, but also about the biology of

endothelial progenitor cells that loose its function, which can be relevant to understand alternative pathways of endothelial expansion or mechanisms of resistance to anti-VEGF, which is a highly prevalent process in cancer (Itatani et al., 2018).

To gain deeper insight on these mechanisms, we first decided to transcriptionally profile differentiated Flk1$^{WT/WT}$, Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ endothelial cells. After 8 days of differentiation of mESCs on OP9 stromal cells, we sorted the differentiated endothelial population with the endothelial marker Icam2, which labels fully committed ECs. We noted that the levels of ICAM2 expression decreased with a decrease in the levels of Flk1 (See anti-Icam2-APC axis levels on the FACS plots (Fig. 11), however all sorted and analyzed cells were clearly positive for Icam2. Given the genetics of our clones, we could identify and separate 2 populations of cells in differentiated Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ clones. A population having strong ICAM2 and strong Flk1-Cerulean expression (green population on the FACS plots Fig. 12, and a population having ICAM2 expression only and very weak or no-expression of Flk1 (blue population on the FACS plots Fig. 12). Thus we were able to capture five different populations: Flk1$^{KI/KI}$(C+I+); Flk1$^{KI/KI}$(C-I+); Flk1$^{KI/WT}$(C+I+); Flk1$^{KI/WT}$(C-I+) and Flk1$^{WT/WT}$(I+) (Figure: 11).

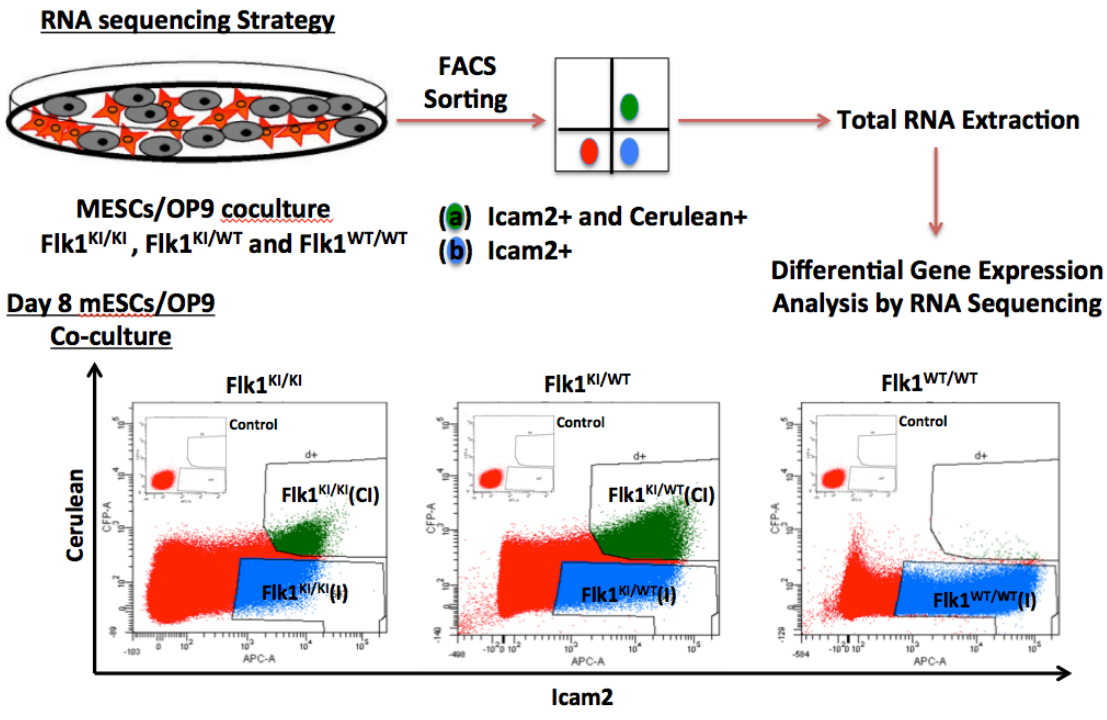


Figure 11: Strategy for cell differentiation and isolation by FACS: Flk1 ES cell lines Flk1$^{WT/WT}$, Flk1$^{KI/WT}$and Flk1$^{KI/KI}$ were differentiated on OP9 for 8 days. Scatter plots with the gating strategy used for cell sorting, according to its signals for Cerulean and Icam2 immunostaning. The Icam2+ populations were gated from live cell population after removing doublets.

1.3.1. Validation of RNAseq raw data and selection of the control experimental groups for comparative analysis

First we analyzed if has expected the levels of Flk1 mRNA were lower in Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ cells when compared with Flk1$^{WT/WT}$ cells. Flk1$^{KI/KI}$(C+/I+) expressed the higher amount of reads (av Reads:2457) followed by Flk1$^{KI/WT}$(C+/I+) with average Reads = 1462 and Flk1$^{WT/WT}$ with 1250 average number of reads. Flk1$^{KI/KI}$(C-/I+) and Flk1$^{KI/WT}$(C-/I+) expressed the minimum amount of reads in the range of 700 average reads. Since the Flk1$^{WT/WT}$ cells did not have the same Flk1 knock-in reporter, and the ICAM2+ population could be a mixed population of the two different (C-I+ and C+I+) populations detected in the Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ cells, we decided to first compare and validate that the Flk1$^{WT/WT}$ and the mixed Flk1$^{KI/WT}$ populations had a similar gene expression profile. When the Flk1$^{KI/WT}$ (C+I+) population or the Flk1$^{KI/WT}$ (C-I+) was compared with Flk1$^{WT/WT}$ (C-I+) population we saw 650 and 566 differentially expressed genes respectively (Fig. 13). However, when we combined the RNAseq profile of the Flk1$^{KI/WT}$ (C-I+) and Flk1$^{KI/WT}$ (C+I+) populations with the Flk1$^{WT/WT}$ wild type populations we saw 0 DEG (Fig. 12). This shows that the differentiated Flk1$^{KI/WT}$ cells are molecularly similar to Flk1$^{WT/WT}$ cells (Fig. 13), even though they have half the Flk1 dose and proliferate/differentiate at a lower rate (Fig. 10).

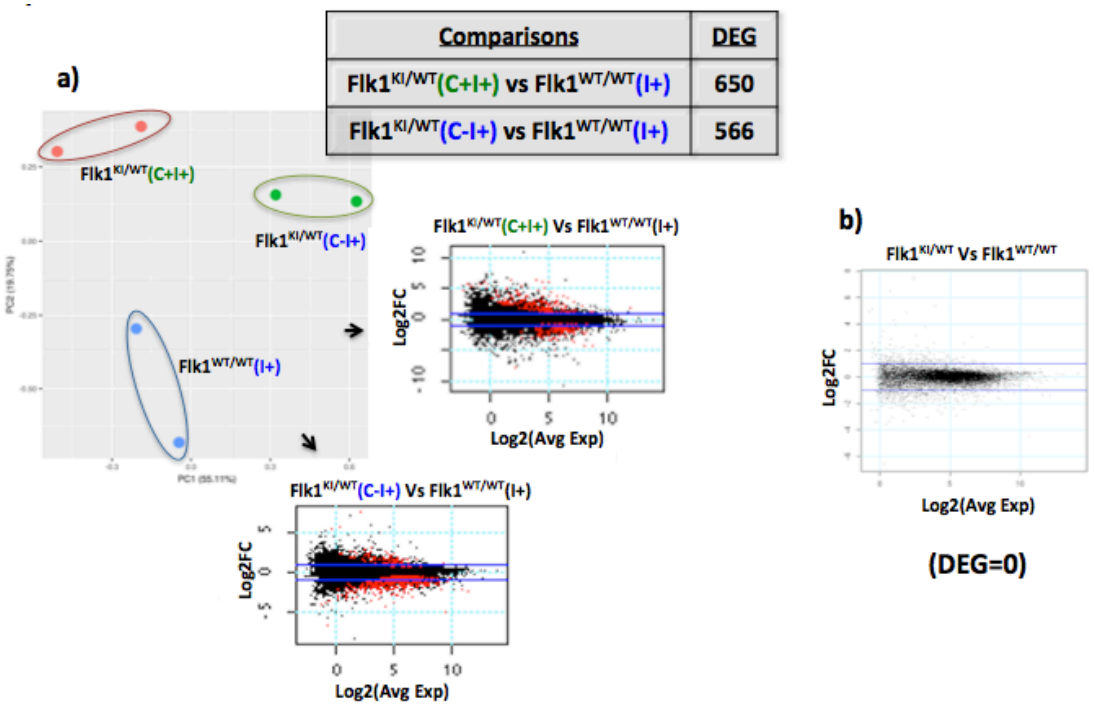


Figure 12: Comparison of gene expression profiles of Flk1$^{WT/WT}$ and Flk1$^{KI/WT}$ mixed population (a) PCA and smear plots showing differences between Flk1$^{KI/WT}$(C+/I+) and Flk1$^{KI/WT}$(C-I+) and Flk1$^{WT/WT}$(I+)populations. In the smear plot Log fold change values are plotted against average log expression values (standardized read counts) and differentially expressed trasncripts are represented by red dots The table shows the differentially expressed genes between the different contrasts; **(b)** Smear plot showing differences between the combined Flk1$^{KI/WT}$ two populations and Flk1$^{WT/WT}$population. (No red dots denote zero differentially expressed genes (DEG)).

1.3.2. Comparison of molecular profiles in *Flk1$^{KI/WT}$(C+I+)* and *Flk1$^{KI/WT}$(C-I+)* populations with published endothelial-hematopoietic development datasets.

The two cell populations analysed (C-I+ and C+I+) express ICAM2 at similar levels but are distinct in the expression of the Flk1 knock-in cerulean reporter. The population I+C- does not express the Flk1-Cerulean reporter, or express it at very low levels, whereas the population C+I+ expresses the Flk1-Cerulean at high levels. To identify if the cells of these two populations were at a different status of differentiation, we

compared our transcriptomic data with the published transcriptome profiling of multiple stages in ES cell-endothelial differentiation (Goode et al., 2016). Briefly, we performed gene set enrichment analysis (GSEA) using this previously published data. We selected all genes, or genes coding for transcription factors, that were upregulated (absolute LogFC>1) at distinct differentiation stages. **ESC=> Mesoderm**(MES); **Mesoderm=> Hemangioblast**(HB); **Hemangioblast=> Hemogenic endothelium**(HE); **Hemogenic endothelium=> Hematopoietic progenitor**(HP); **Hematopoietic progenitor=> Megakaryocytes**(MAC). These genes defined unique gene sets/clusters (C1 All, C1TF, C2 All, C2TF, C3All, C3TF, C4All, C4TF, and C5All, C5TF respectively), for each of the differentiation stages. We than analysed which of these gene sets were enriched in the $Flk1^{KI/WT}$populations $Flk1^{KI/WT}$(C+I+) and $Flk1^{KI/WT}$(C-I+). Interestingly, the gene set C3All and C3TF comprising genes and transcription factors upregulated in hemogenic endothelium or endothelial cells (HB=>HE) was enriched in $Flk1^{KI/WT}$(C+I+) and the gene set C4All and C4TF comprising genes and transcription factors upregulated in hematopoietic progenitor cells (HE=>HP) was enriched in $Flk1^{KI/WT}$(C-I+). Figure 13 shows the enriched gene set enrichment plot for the corresponding populations. Thus our GSEA analysis shows that the $Flk1^{KI/WT}$(C+I+) population has a higher fraction of endothelial cells or its population is expressing higher levels of genes enriched in endothelial cells, whereas the $Flk1^{KI/WT}$(C-I+) population has a higher fraction of ICAM2+ sorted cells with hematopoietic progenitor characteristics.

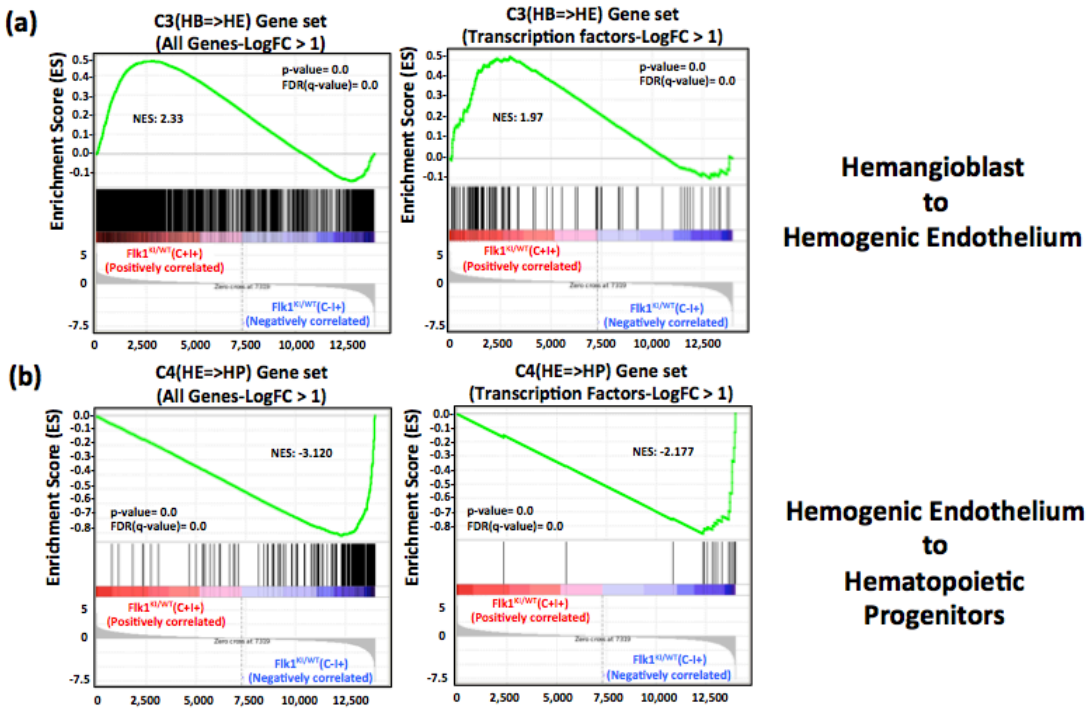


Figure 13: Comparison of DEG in Flk1$^{KI/WT}$(C+I+) and Flk1$^{KI/WT}$(C-I+) cells: **(a)** Enrichment plots of all genes (left) or only transcription factors (right) of the C3 cluster (HB=>HE) in the two populations; **(b)** Enrichment plots of all genes (left) or only transcription factors (right) of the C4 cluster (HE=>HP) in the two populations. NES indicates normalized enrichment score.

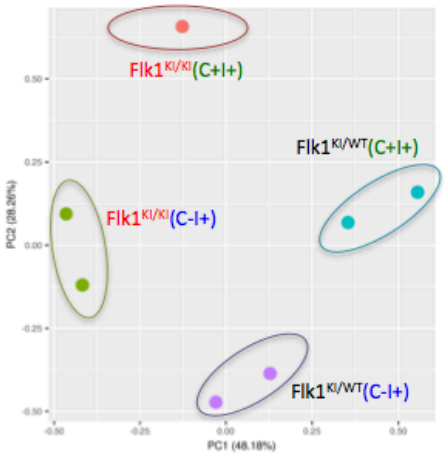
1.3.3. Global comparison of the transcriptional differences among the four populations reveals the biological roles of *Flk1* in distinct cell populations

For a more comprehensive and non-biased overview of the differentially expressed genes in each cell population, we compared the four cell populations (Flk1$^{KI/WT}$(C+I+), Flk1$^{KI/WT}$(C-I+), Flk1$^{KI/KI}$(C+I+) and Flk1$^{KI/KI}$(C-I+)) with each other.

In the principle component analysis (PCA), the gene expression of all four populations showed significant distance to each other at PC1 and PC2 (Fig. 14(a)), which represents 48.18% and 28.26% variance respectively. For further analysis, differentially expressed genes (DEGs) were evaluated by applying a stringent statistic

threshold of a greater than or equal to two-fold change (FC), a false discovery rate (FDR) of <0.05 and a p-value of <0.05 (student t-test) between different contrasts. According to these stringent criteria, the following numbers of DEG were identified (Fig. 14(b)).

a)



b)

Comparisons	DEG
$Flk1^{KI/KI}$(C+I+) vs $Flk1^{KI/KI}$(C-I+)	32
$Flk1^{KI/WT}$(C+I+) vs $Flk1^{KI/WT}$(C-I+)	826
$Flk1^{KI/KI}$(C+I+) vs $Flk1^{KI/WT}$(C+I+)	259
$Flk1^{KI/KI}$(C-I+) vs $Flk1^{KI/WT}$(C-I+)	1495

Figure 14: Transcriptional differences among the four cell populations. (a) The plot shows principle component analysis (PCA) showing separation between the $FLK1^{KI/KI}$ and $Flk1^{KI/WT}$ populations. **(b)** The table shows number of differentially expressed genes between different populations and contrasts.

To get a better picture of variation of gene expression among the different groups, we performed unsupervised (K-means) clustering on the collection of genes (2062) identified as DEG in at least one of the above contrasts. This resulted in the identification of 20 different clusters of genes that had a specific and well-recognized pattern of gene expression among the four groups. Clusters with a more similar regulation pattern are clustered into three main metaclusters (MC1, MC2, MC3). We than performed Gene Set Enrichment Analysis using the Enrichr API to explore the function of the genes contained in the three metaclusters (Fig. 15 and Fig. 16)

MC1 cluster: This cluster contains genes that are mostly upregulated in the $Flk1^{KI/WT}$(C+I+) population versus all the others, particularly the $Flk1^{KI/KI}$(C-I+)

population. The GO terms of biological processes (BP) that were enriched include: cell migration, protein phosphorylation, vasculogenesis, sprouting angiogenesis and VEGFR signalling. All these processes are well known to be regulated by Flk1. The KEGG pathways analysis also revealed an increase in genes related with MAPK signalling, PI3K-AKT signalling and Focal Adhesion, which correlates well with the known function of Flk1 as one of the most important endothelial tyrosine kinase receptor molecules, that when activated elicits a cascade of phosphorylation that results ié higher MAPK and PI3K activity.

MC2 cluster: This cluster contains genes that are mostly upregulated in the $Flk1^{KI/WT}$(C-I+) population versus all the others, particularly the $Flk1^{KI/KI}$(C+I+) population. GO terms of biological processes (BP) and pathway information from KEGG database shows the enrichment of categories that are related with hematopoiesis and blood lineage differentiation. This suggests that $Flk1^{KI/WT}$(C+I+) or cells with full loss of the Flk1 function have decreased blood lineage commitment.

MC3 cluster: This cluster contains genes that are mostly upregulated in $Flk1^{KI/KI}$(C-I+) cells and also to a certain degree in $Flk1^{KI/KI}$(C+I+) populations. The GO terms of biological processes (BP) that were enriched include: Extracellular matrix remodelling, Epithelial to mesenchymal transition (EMT), heart, skeletal and mesenchyme development. All these biological processes are related with the deregulation of the KEGG pathways: extracellular matrix-receptor interactions, Tgf-beta/Bmp signalling and cardiomyopathy pathways, likely suggesting that loss of Flk1 leads to a loss of endothelial-hematopoietic fate and an increase in mesenchymal, smooth-muscle and cardiomyocyte lineages.

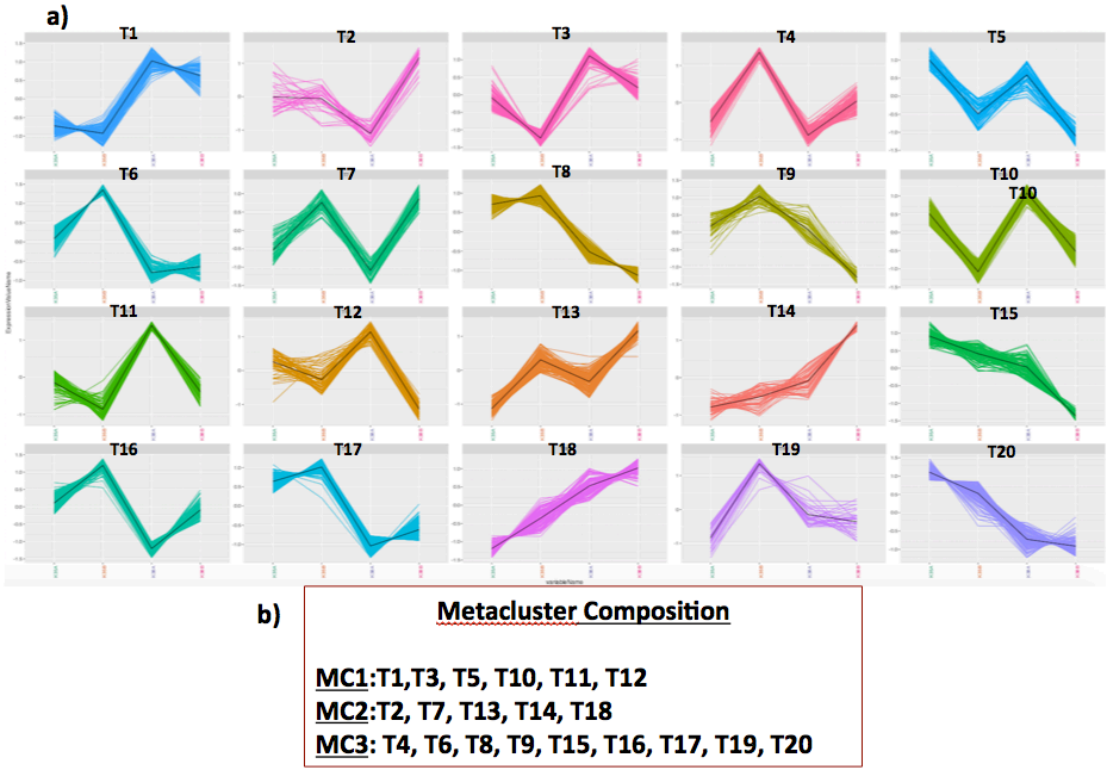


Figure 15: The expression dynamics of differentially expressed genes in Flk1$^{KI/KI}$(C+I+), Flk1$^{KI/KI}$(C-I+), Flk1$^{KI/WT}$(C+I+), Flk1$^{KI/WT}$(C-I+). Hierarchical clustering (K- means clustering) of the normalized expression values of the genes in the four populations generated **(a)** 20 clusters; **(b)** 3 Metaclusters (MC1, MC2 and MC3) composition.

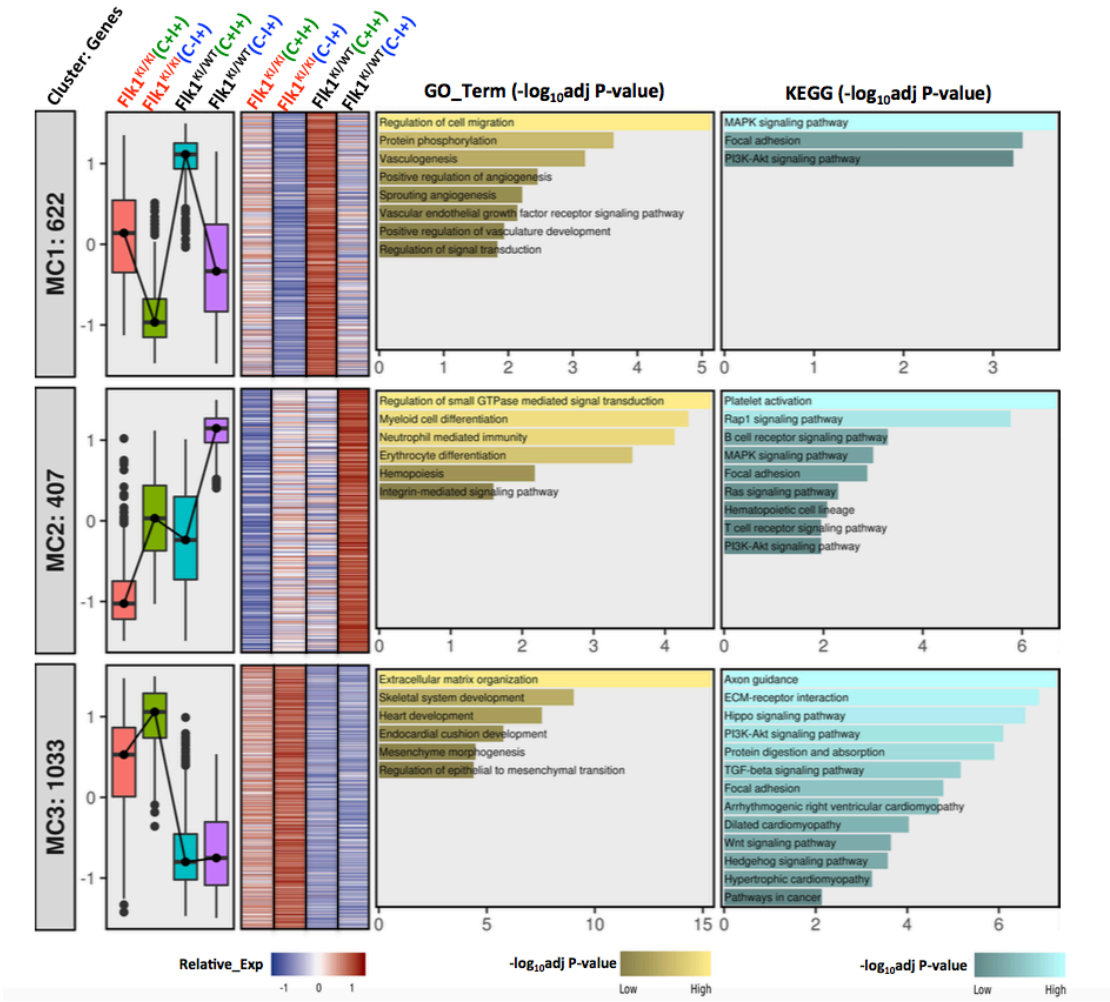


Figure 16: **Differential expression of Metaclusters and enriched GO Term and KEGG**: The boxplot shows the variation in expression profiles of the 3 identified metaclusters in the four cell populations. Heat maps of the differentially expressed genes within each metacluster. Enrichment for genes related with a given biological process (GO_Term) and the list of top canonical pathways (KEGG) were calculated using Benjamini & Hochberg (BH) adjusted p-value and set to less than .05.

1.3.3.1. Comparison of molecular profiles in *Flk1$^{KI/WT}$(C+I+)* and *Flk1$^{KI/WT}$(C-I+)* populations reveals distinct endothelial-hematopoietic profiles

After performing the global transcriptomic analysis, we next compared directly the different Flk1 heterozygous cell populations. The two cell populations analyzed

(C+I+) and (C-I+) express ICAM2 at similar levels but are distinct in the expression of the Flk1 knock-in cerulean reporter. The population (C-I+) does not express the Flk1-Cerulean reporter, or expresses it at very low levels, whereas the population (C+I+) expresses the Flk1-Cerulean at high levels. 826 genes were differentially expressed among these two cell populations.

$Flk1^{KI/WT}$(C+I+) and $Flk1^{KI/WT}$(C-I+) cells were significantly different according to the heatmap and cluster analysis(Fig. 17).

Gene set enrichment analysis revealed that genes significantly more expressed in $Flk1^{KI/WT}$(C+I+) than $Flk1^{KI/WT}$(C-I-) cells were predominantly related to endothelial development and angiogenesis and vasculogenesis (Fig. 17(a)). Further Gene enrichment analysis of 660 differentially expressed genes using Enrichr revealed biological Go term such as regulation of cell migration, angiogenesis , sprouting angiogenesis and vasculogenesis. These results suggested that $Flk1^{KI/WT}$(C+I+) population comprise of endothelial progenitors and differentiated endothelial cells.

In contrast, $Flk1^{KI/WT}$(C-I+) population upregulated gene sets of hematopoietic cell lineage, heme metabolism, Myc and E2F targets. Enrichment of Myc and E2F target hallmark genesets in $Flk1^{KI/WT}$(C-I+) suggest their more proliferative nature, which is in line with the fact that differentiated ECs (in $Flk1^{KI/WT}$(C+I+)) are less proliferative than hemogenic endothelial cells or hematopoietic progenitor cells, more enriched in $Flk1^{KI/WT}$(C-I+)(Fig. 17(b)). The GO analysis revealed hematopoiesis including erythrocyte and myeloid differentiation and upregulation of integrin mediated signaling pathway and cell migration. Genes of the integrin-mediated signaling pathway (Itgam, Itga3, Itgb2, Itgb3, Itgb7, Col1a1, Col3a1, Col5a1, Rac2, etc) were enriched in the $Flk1^{KI/WT}$(C-I+) group, reflecting that cells which had gone through the Endothelial to hematopoietic transition process were more motile (Fig. 17(d)). These receptors play a crucial role in mediating platelet aggregation for blood coagulation system and immune responses.

We also analyzed in detail the expression pattern of key factors endothelial, hemogenic endothelial and hematopoietic system development; the results are presented in heatmap(Fig. 17(d)). The upregulated genes in the $Flk1^{KI/WT}$(C+I+) population include EC surface markers such as Cdh5, Pecam1, Cldn5, Emcn;

receptors such as AplnR, Notch1, Notch4, Nrp1, Flt1, Flt4 and transcription factors such as Erg, Fli1, Ets1, Sox17, Sox18, Foxc1, etc (Fig. 17(d)). All of them have been shown to play an important role in endothelium development and angiogenesis. PI3K and MAPK pathways are required for EC differentiation from Pluripotent stem cells is activated by VEGF(Harding, A. etal., 2017). Several ETS family transcription factors (ERG and FLI1) essential for EC formation were found before to be downstream of MAPK and PI3K signaling (Harding. A. etal., 2017; Hart, A. etal., 2000; Xu, J. etal., 2008). PI3K and MAPK pathway along with Focal adhesion are responsible for sprouting angiogenesis and migration. In contrast critical genes for hematopoietic development such as Tal1, Runx1, Spi1, Gata1, Myb, Runx3, Gfi1, Gfi1b, Ikzf1, klf1, Tyrobp, etc were enriched in the $Flk1^{KI/WT}$(C-I+) population. Hematopoietic surface marker genes such as Spn(CD43), Itga2b(CD41), Ptprc(CD45), Kitl and so forth were highly expressed in $Flk1^{KI/WT}$(C-I+) cells.

Thus the comparative transcriptomic analysis of these two populations revealed that the expression of the Flk1-cerulean reporter enables the separation of cells with a marked endothelial phenotype ($Flk1^{KI/WT}$(C+I+)), from cells with a endothelial-to-hematopoietic transition, or hematopoietic progenitor phenotype ($Flk1^{KI/WT}$(C-I+)).

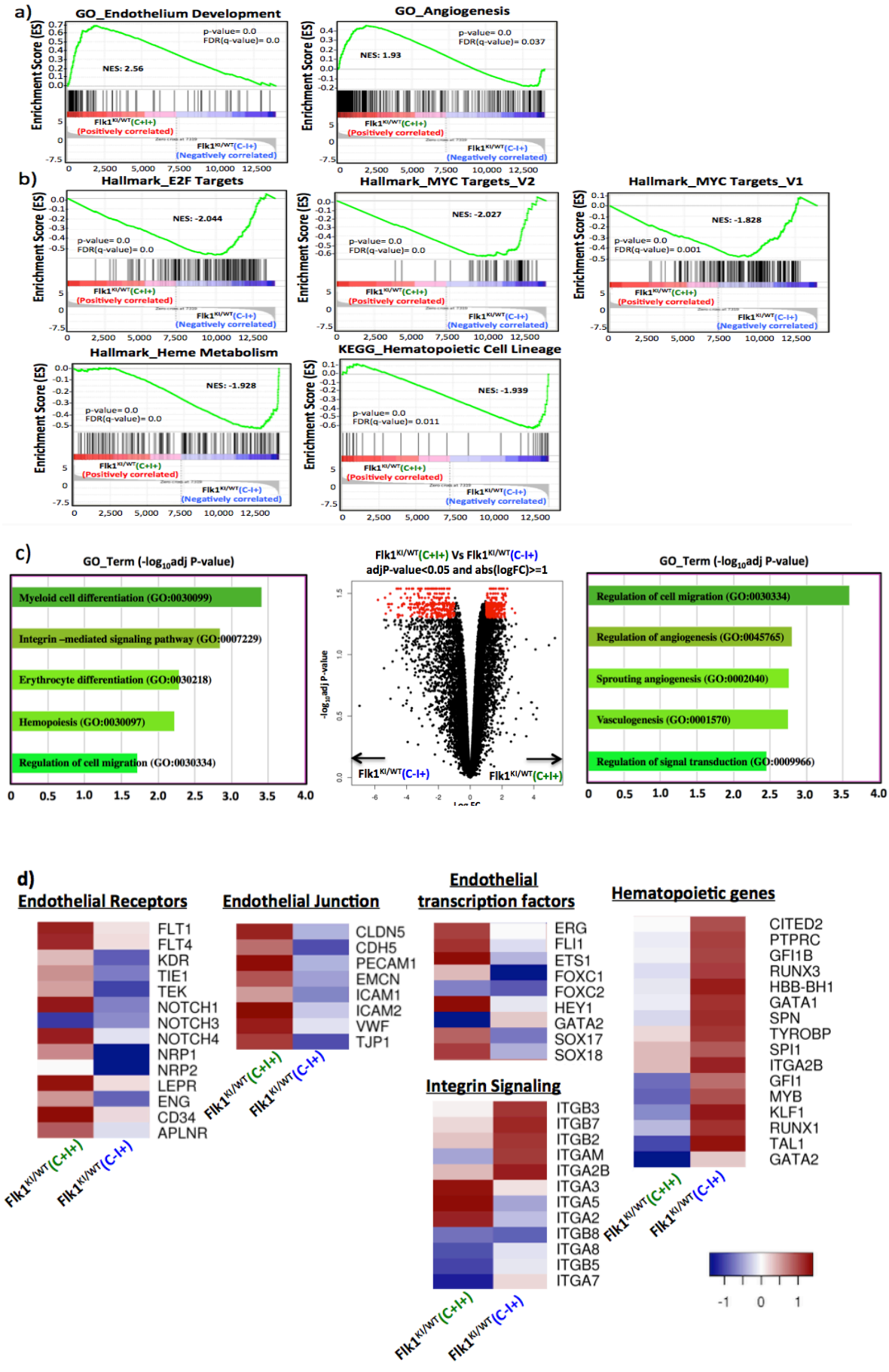

a)
GO_Endothelium Development
p-value= 0.0
FDR(q-value)= 0.0
NES: 2.56
GO_Angiogenesis
p-value= 0.0
FDR(q-value)= 0.037
NES: 1.93
b)
Hallmark_E2F Targets
NES: -2.044
p-value= 0.0
FDR(q-value)= 0.0
Hallmark_MYC Targets_V2
NES: -2.027
p-value= 0.0
FDR(q-value)= 0.0
Hallmark_MYC Targets_V1
NES: -1.828
p-value= 0.0
FDR(q-value)= 0.001
Hallmark_Heme Metabolism
NES: -1.928
p-value= 0.0
FDR(q-value)= 0.0
KEGG_Hematopoietic Cell Lineage
NES: -1.939
p-value= 0.0
FDR(q-value)= 0.011
Enrichment Score (ES)
Flk1KI/WT(C+I+)
(Positively correlated)
Flk1KI/WT(C-I+)
(Negatively correlated)
c)
GO_Term (-log10adj P-value)
Myeloid cell differentiation (GO:0030099)
Integrin –mediated signaling pathway (GO:0007229)
Erythrocyte differentiation (GO:0030218)
Hemopoiesis (GO:0030097)
Regulation of cell migration (GO:0030334)
Flk1KI/WT(C+I+) Vs Flk1KI/WT(C-I+)
adjP-value<0.05 and abs(logFC)>=1
-log10adj P-value
Flk1KI/WT(C-I+)
Flk1KI/WT(C+I+)
GO_Term (-log10adj P-value)
Regulation of cell migration (GO:0030334)
Regulation of angiogenesis (GO:0045765)
Sprouting angiogenesis (GO:0002040)
Vasculogenesis (GO:0001570)
Regulation of signal transduction (GO:0009966)
d)
Endothelial Receptors
FLT1
FLT4
KDR
TIE1
TEK
NOTCH1
NOTCH3
NOTCH4
NRP1
NRP2
LEPR
ENG
CD34
APLNR
Endothelial Junction
CLDN5
CDH5
PECAM1
EMCN
ICAM1
ICAM2
VWF
TJP1
Endothelial transcription factors
ERG
FLI1
ETS1
FOXC1
FOXC2
HEY1
GATA2
SOX17
SOX18
Integrin Signaling
ITGB3
ITGB7
ITGB2
ITGAM
ITGA2B
ITGA3
ITGA5
ITGA2
ITGB8
ITGA8
ITGB5
ITGA7
Hematopoietic genes
CITED2
PTPRC
GFI1B
RUNX3
HBB-BH1
GATA1
SPN
TYROBP
SPI1
ITGA2B
GFI1
MYB
KLF1
RUNX1
TAL1
GATA2
Flk1KI/WT(C+I+)
Flk1KI/WT(C-I+)
-1
0
1

Figure 17: Comparative transcrptome analysis of Flk1$^{KI/WT}$(C+I+) and Flk1$^{KI/WT}$ (C-I+) **(a)** Gene sets upregulated in Flk1$^{KI/WT}$(C+I+) **(b)** Gene sets upregulated in Flk1$^{KI/WT}$(C+I+). Enrichment plots generated by GSEA using GO_BP, Hallmark and KEGG (NES: Normalized enrichment score; FDR: False discovery rate; **(c)** Volcano plot showing 660 DEG between **Flk1$^{KI/WT}$(C+I+)** and **Flk1$^{KI/WT}$(C-I+)** with adjP-value<0.05 and abs(logFC)>=1 with the enrichment analysis using Enrichr (False discovery rate(q value)<.05) **(d)** Heatmaps for selected genes related with different cell fates or biological processes. Color Scale bar represents relative counts that have been normalized by library size (TMM method), averaged by condition, log transformed, and scaled by gene.

1.3.3.2. Loss of Flk1 inhibits endothelial and hemogenic endothelium specification and induces endothelial to mesenchymal transition

Having established the distinction between the Flk1$^{KI/WT}$(C+I+) and Flk1$^{KI/WT}$(C-I+) populations, we than wanted to compare the transcriptome of Flk1$^{KI/KI}$(C+I+) and Flk1$^{KI/KI}$(C-I+) populations. Interestingly this contrast resulted in only 32 differentially expressed genes, which may be explained by the full loss of Flk1 function in these cells. In this case, cells expressing the Flk1-reporter Flk1$^{KI/KI}$(C+I+), will not have expression of Flk1, and therefore they are more similar to Flk1$^{KI/KI}$(C-I+) cells. We have previously observed in Flk1$^{KI/KI}$(C+I+) an upregulation of Metacluster MC1 compared to Flk1$^{KI/KI}$(C-I+) (Fig. 16). By analyzing this data in more detail, we found that Flk1$^{KI/KI}$(C+I+) expresses higher levels of most of the endothelial receptors, endothelial secreted molecules, endothelial transcription factors and endothelial junction markers (Emcn, Cldn5, Icam2, Pecam1, Cdh5, Sox17, Sox18, AplnR, Notch1, Notch4, Dll4, etc) when compared with Flk1$^{KI/KI}$(C-I+) cells (Fig. 18(c)), however this endothelial signature is less pronounced than in Flk1$^{KI/WT}$(C+I+) cells, suggesting that indeed Flk1 is essential for the full acquisition of endothelial characteristics.

We also saw that Flk1 is essential for endothelial to hematopoietic transition, since Flk1$^{KI/KI}$(C+I+) and Flk1$^{KI/KI}$(C-I+) cells expressed significantly less hematopoietic genes than the Flk1$^{KI/WT}$(C+I+) and Flk1$^{KI/WT}$(C-I+) cells.

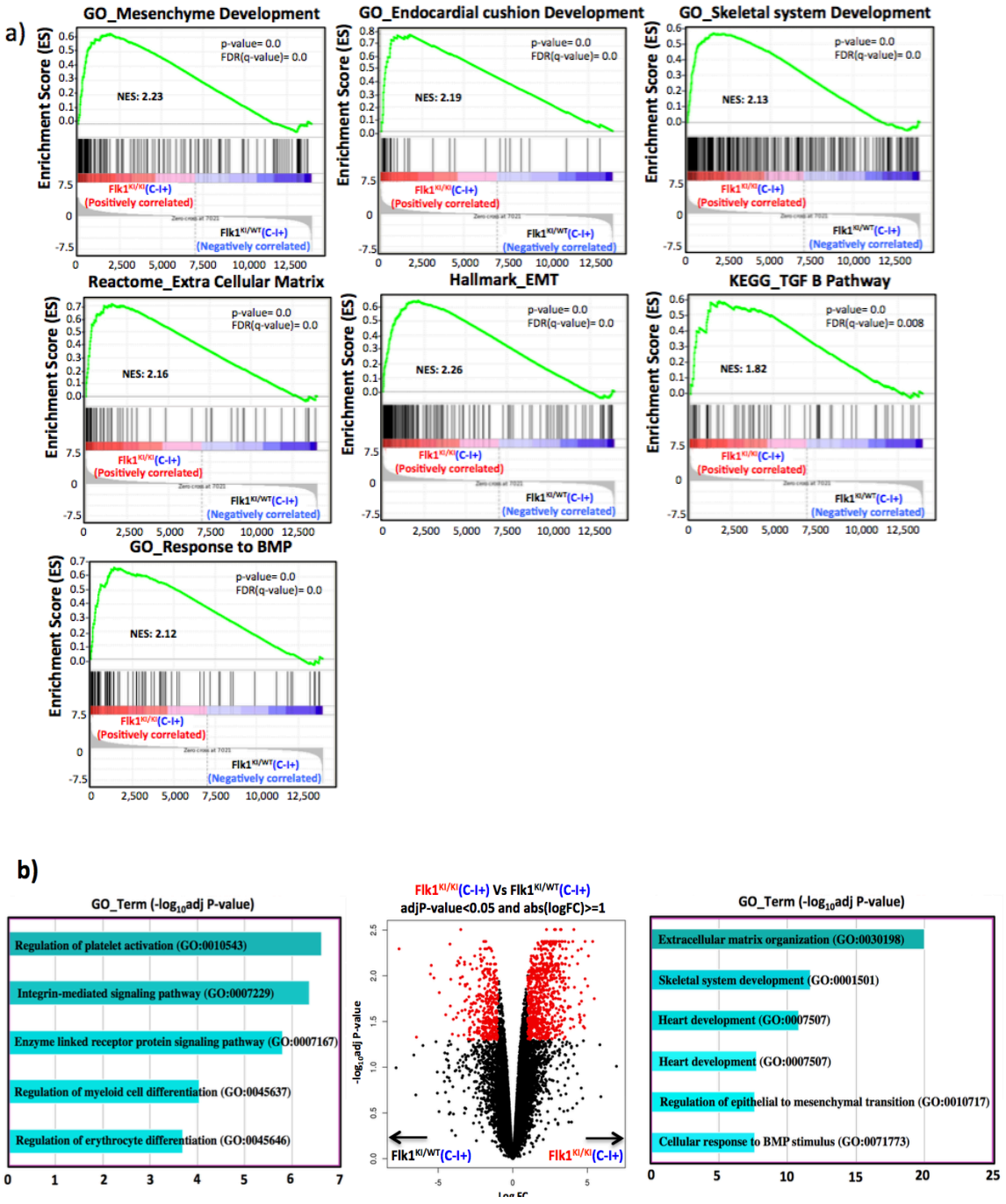
Next, we compared the transcriptome of Flk1$^{KI/KI}$(C-I+) cells with the transcriptome of Flk1$^{KI/WT}$(C-I+) cells, which resulted in 1495 DEG according a false discovery rate (FDR) of <0.05 and a p-value of <0.05 (student t-test). The gene set enrichment analysis revealed the upregulation of gene set categories in biological process such as Heart development, endocardial cushion development, mesenchyme development and skeletal system development. Hallmark, KEGG and Reactome shows enrichment of EMT, TGF-B signaling pathway and categories related to extra-cellular matrix respectively (Fig. 18(a)).

Interestingly, the loss of Flk1 leads to increased expression of genes related to heart development (Hand1, Hand2, Gata4, Gata5, Gata6, Pitx2, Isl1, Tbx2, Tbx20)(Fig. 18(d)). Flk1$^{KI/KI}$(C-I+) endothelial cells also upregulated the expression of genes related with the mural cell fate. This was evident from the upregulation of canonical pericyte markers (Pdgfrb, Cspg4, and Des); smooth muscle cell markers (Acta2 and Tagln) and Fibroblast markers (Pdgfra, Lum and Dcn). The upregulated transcripts also involved Fibroblast specific transcripts such as Fibrillar collagens(Col1a1, Col1a2, Col3a1, Col5a1, Col5a2 and Col5a3) and non fibrillar collagens (Col6a1, Col6a2, Col6a3, Col8a1, Col8a2, Col11a1, Col13a1, Col14a1, Col16a1) collagen modifying enzymes (Loxl1, Mmp2, Mmp9,Timp1, Timp2) proteins involved in collagen fibril spacing, such as the small leucin-rich proteoglycans lumican (Lum) and decorin (Dcn). There was also a significant over representation of ECM- remodeling and its components (ECM production, collagen fibrillogenesis, collagen crosslinking, ECM binding and remodeling enzymes)(Fig. 18(e,f)).

This data overall suggested that the loss of Flk1 induces endothelial to mesenchymal transition, particularly in the Flk1$^{KI/KI}$(C-I+) population. One of the key pathways involved in EMT, is the Tgf-beta and the Tgf-beta-related bone morphogenetic proteins (Bmps) signalling pathways. Indeed we also found a remarkable deregulation of genes that belong to the Tgf-Beta pathway (Dcn, Tgfb2, Tgfb3, Bmp4, Bmp5, Bmp7, Id1, Id2, Id3, etc) and EMT markers (Snai1, Snai2, Twist1, Hand1, Hand2, Tgfb2, Tgfb3, Ids, BMPs, Cdh2, etc) (Fig. 18(f)). In addition, Integrins, such as Itgb5, Itgb8, Itga7 and Itga8 mediating a wide spectrum of cell-cell and cell-matrix

interactions are also upregulated, as well as extracellular matrix secreted and remodelling proteins such as fibrins, collagens and MMPs (Fig. 18 (c,e)).

This data suggest that in the absence of Flk1, endothelial progenitor cells acquire an alternative fate that is more characteristic of mesenchymal cells.



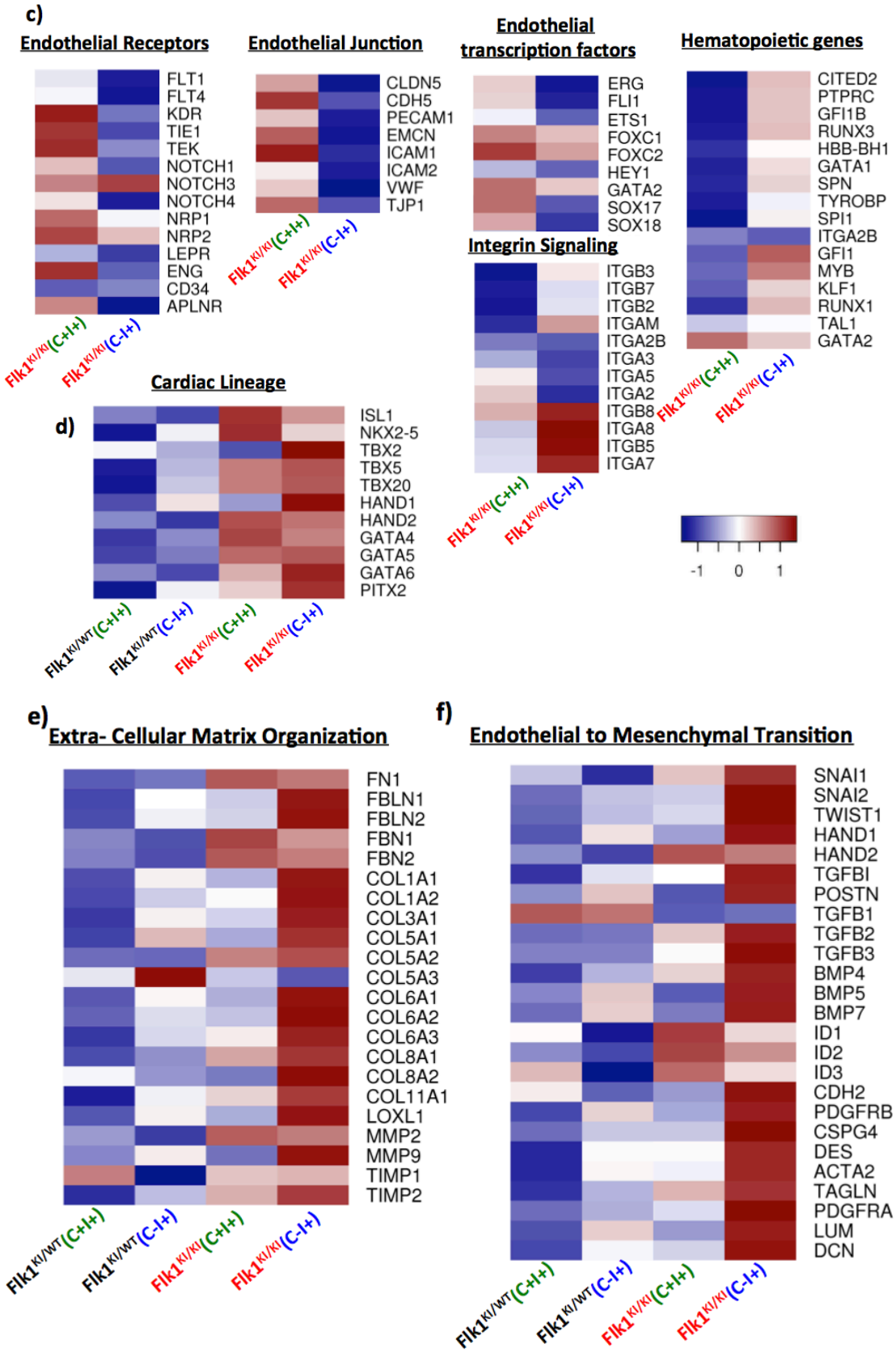


Figure 18: Comparative transcrptome analysis of Flk1$^{KI/WT}$(C-I+) and Flk1$^{KI/Ki}$(C-I+)

(a) Gene sets upregulated in Flk1$^{KI/Ki}$(C-I+). Enrichment plots generated by GSEA using GO_BP, Hallmark and KEGG (NES: Normalized enrichment score; FDR: False

discovery rate; **(b)** Volcano plot showing 1132 DEG between Flk1$^{KI/WT}$(C+I+) and Flk1$^{KI/Ki}$(C-I+) with adjP-value<0.05 and abs(logFC)>=1 with the enrichment analysis using Enrichr (False discovery rate(q value)<.05); **(c,d,e,f)** Heatmaps for selected genes related with different cell fates or biological processes. Color Scale bar represents relative counts that have been normalized by library size (TMM method), averaged by condition, log transformed, and scaled by gene. **(c)** Heatmaps for Endothelial lineage, hematopoietic lineage and integrin signaling for Flk1$^{KI/Ki}$(C+I+) and Flk1$^{KI/Ki}$(C-I+). **(d)** Cardiac lineage **(e)** Extracellular matrix **(f)** Endothelial to mesenchymal transition for Flk1$^{KI/WT}$(C+I+), Flk1$^{KI/WT}$(C-I+), Flk1$^{KI/Ki}$(C+I+) and Flk1$^{KI/Ki}$(C-I+).

2: Transcriptional profiling of *Flk1* heterozygous (*Flk1*$^{KI/WT}$) and homozygous (*Flk1*$^{KI/KI}$) reporter knock-in endothelial cells in vivo

2.1: Generation and characterization of *Flk1* knock-in mice

To study the *in vivo* function of Flk1, we generated mice with the same ES cells Flk1$^{KI/WT}$ ((Gt(Flk1)-H2B- Cerulean- 2A-iCreRT2) we studied before (Fig. 19). These ES cells have the G4 genetic background and therefore a high germline contribution competence. These ES cells were expanded by me *in vitro* and after injected in C57Bl6 blastocists by the CNIC transgenesis unit. We obtained several chimeras with a high contribution from the Flk1$^{KI/WT}$ ES cells (Agouti coat color). After intercrossing these chimeras with wildtype C57Bl6 mice, we observed that Flk1$^{KI/WT}$ mice, like Flk1+/- heterozygous mice (Shalaby, F. et al., 1997), are viable and fertile. These Flk1$^{KI/WT}$ heterozygous mice were interbred to generate Flk1$^{KI/KI}$ null embryos. As expected, all Flk1 null embryos were clearly small and abnormal when compared with their Flk1 heterozygous and wild type littermates, around E9.5 (Data not shown).

To visualize the endothelial cells in the Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ embryos expressing Flk1-Cerulean, we performed co-immunostaining for Cerulean and the endothelial

markers ERG and Icam2 (Fig. 20). Icam2 and ERG positive blood vessels were clearly visible in Flk1$^{KI/WT}$ but not in Flk1 $^{KI/KI}$ embryos (Fig. 20). In Flk1$^{KI/WT}$ embryos, the Flk1-Cerulean expression faithfully recapitulated the temporal and spatial expression patterns of endogenous Flk1 published before (Shalaby, F et al., 1997) In Flk1$^{KI/KI}$ embryos, we did not find any Icam2 positive staining, eventhough we could observe some cells expressing ERG. The strong expression of Cerulean in ERG-Icam2- cells indicates that there are many Flk1 expressing progenitors in these mutant mice, but in contrast to the results obtained *in vitro*, the large majority is not able to differentiate to Icam2+ERG+ endothelial cells. We ignore the identity of the few ERG+ cells in Flk1 $^{KI/KI}$ embryos.

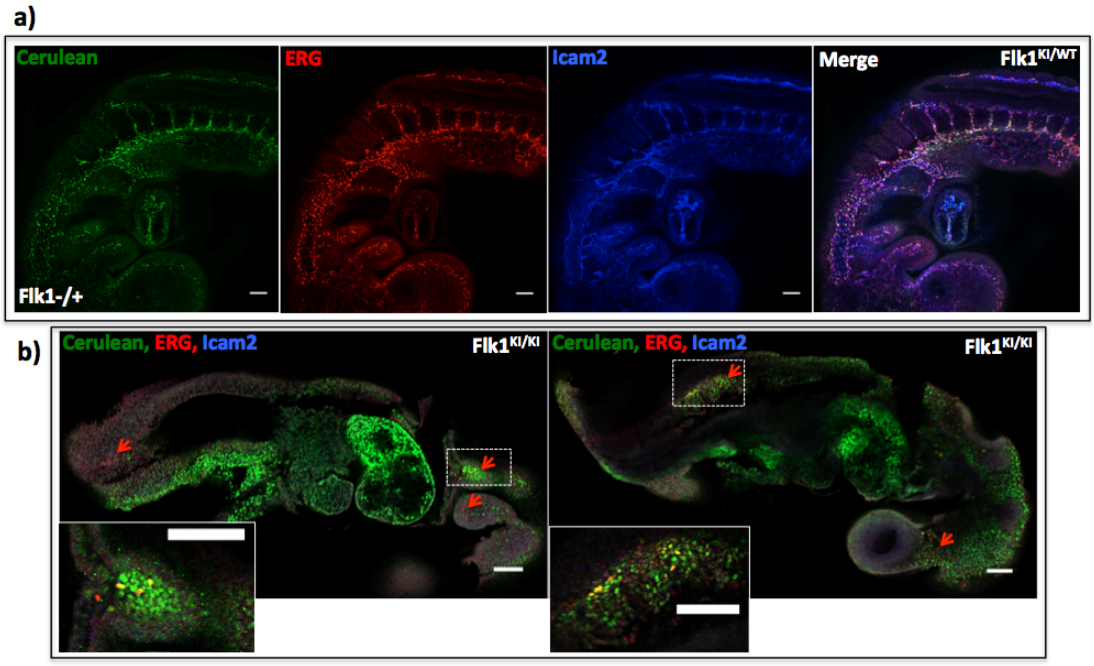


Figure 19: Wholemount Immunostaining for endothelial markers Icam2 and ERG in Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ E9.5 embryos.

2.2. Comparative transcriptional profiling of *Flk1$^{KI/WT}$* and *Flk1$^{KI/KI}$* cells

Since Flk1$^{KI/KI}$ embryos did not have any detectable ICAM2+ ECs, we thought that the use of conditional Flk1 flox Tie2-Cre mice would allow us to obtain a few more ICAM2+ ECs since Tie2-Cre is expressed later than Flk1. We thought with this approach we would be able to delete Flk1 only after endothelial progenitors undergo some differentiation to the endothelial lineage, and be able to isolate more Icam2+ cells.

To compare the transcriptome of ICAM2+ endothelial cells from Flk1$^{KI/WT}$ and Flk1$^{KI/KI}$ embryos, we had to generate mice with a complex mix of alleles**.** This is because we could only intercross Flk1$^{flox/flox}$ or Flk1$^{KI/WT}$ mice (Flk1$^{KI/KI}$ mice die in utero) and expected a very small amount of Flk1$^{KI/KI}$ cells, and therefore we had to FACS sort the entire pooled litter embryos in one go, and without separating and genotyping the embryos individually. We also had to make sure that we could separate Flk1KO (KI/flox Tie2-Cre+) and Flk1Het (KI/flox Tie2-Cre-) cells by FACS. In order to do that, we decided to combine male and female mice with the genotypes indicated below (Fig. 20(a)). The Tie2-Cre allele induces recombination in all endothelial and hematopoietic lineages starting at E8.0-E8.25. The iSuRe-Cre allele assures genetic deletion in Cre-reporter expressing cells and their lineages, due to the co-expression of MbTomato and Cre. The Flk1-Cerulean knock-in allele allow us to distinguish Flk1 expressing cells, and in this case also allows the complete separation of the Flk1$^{KI\text{-}Cerulean/flox}$ from the Flk1$^{WT/flox}$ populations.

This strategy allowed us to perform RNA seq analysis of the endothelial cells extracted from pooled E9.5 embryos without the need for genotyping.

For the analysis we isolated Icam2+/Cerulean+/Tomato+ (I+C+T+) cells and Icam2+/Cerulean+/Tomato- (I+C+T-) cells from the pooled embryos. Given the breeding strategy used (Fig. 20(a)), the I+/C+/T+ cells can only be derived from Flk1$^{KI/Flox}$ Tie2-Cre+ R26-Tomato+ embryos and have full loss of Flk1 function, whereas the I+/C+/T- cells can only be derived from Flk1$^{KI/Flox}$Tie2-Cre- embryos and have one intact copy of Flk1 (Flk1 heterozygous). As expected, using this breeding

strategy we got far more Flk1 heterozygous cells (I+/C+/T-) than Flk1 null (I+/C+/T+) cells (Fig. 20(b)).

For simplicity, from here on we will refer to (I+/C+/T+) cells as Flk1-null and (I+/C+/T-) cells as Flk1 het cells only.

Next, we did a pre-analysis and validation of the RNA samples obtained by estimating the relative abundance of the transcripts for 4 genes (Flk1, Esm1, Cdh5 and Pecam1). The endothelial markers Pecam1 and Cdh5 were equally expressed in Flk1 null (replicates R3, R5 and R9) when compared to Flk1 heterozygous control replicates (R4, R6 and R10). However, there was no qPCR amplification for Flk1 and its downstream target gene Esm1 (Figure: 20(c)). These results confirm that the sorted cell populations are pure and have the expected relative difference in the expression of Flk1 and Esm1.

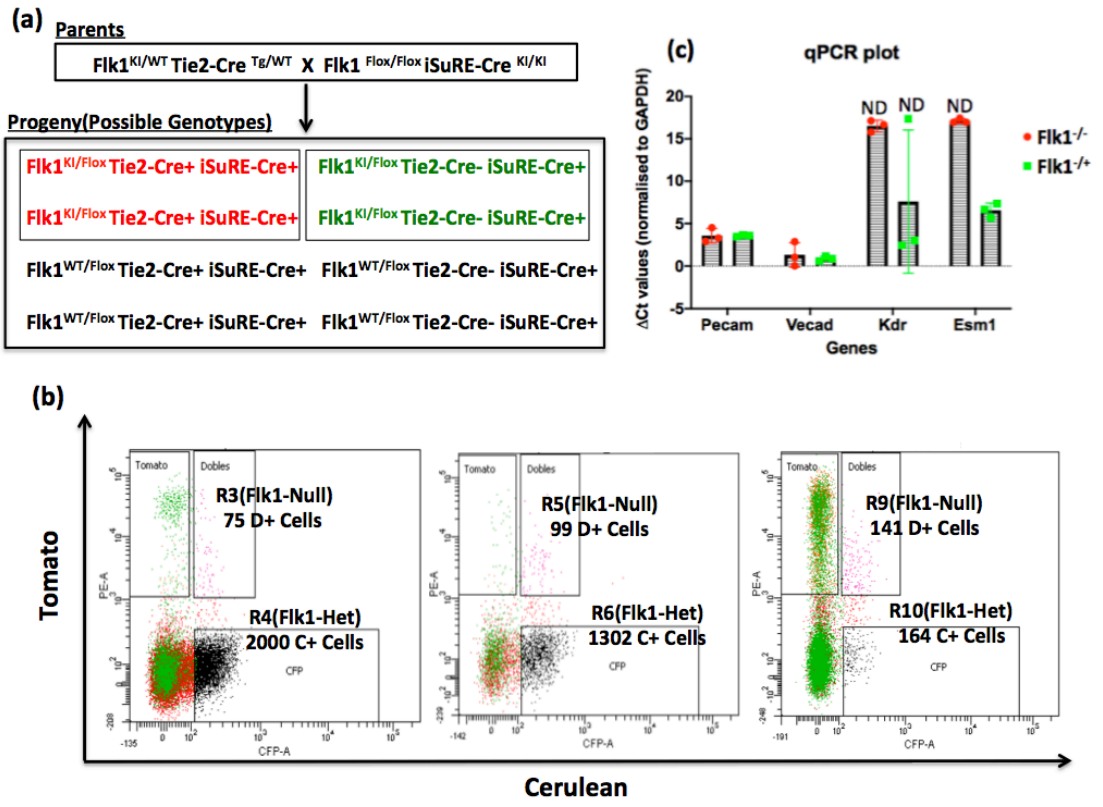


Figure 20: Mouse breeding strategy, isolation and validation of Flk1 heterozygous and knockout Icam2+ endothelial cells: (**a**) Breeding strategy and alleles of

compound male and female mice used to obtain the endothelial cells for RNA Sequencing (Progenies in red are the source for endothelial cells in Flk1-null and progenies in green are the source for Flk1 heterozygous); **(b)** Fluorescence scatter plots showing the gatings used to isolate the Flk1 null and heterozygous cells. The gatings shown were done on a live endothelial cell population (Icam2+ Dapi-). R3, R5, R9 and R4, R6 and R10 are biological replicates for Flk1-null and Flk1 heterozyguos respectively; **(c)** q-PCR analysis of canonical endothelial markers (Pecam1, Cdh5), Flk1 and its downstream target gene Esm1.

2.2.1 Bioinformatics analysis of the *in vivo* transcriptome of *Flk1-null* and *Flk1* heterozygous cells

After performing several types of bioinformatic analysis, we did not obtain any significantly differentially expressed genes between Flk1 null and Flk1-Het endothelial cell transcriptomes from E9.5 embryos, due to the relatively low cell numbers numbers of the Flk1 null samples and the varying amounts of mRNA across samples. This led to variations in transcriptome coverage as well as non-uniform coverage along the length of the transcript among replicates, which impacts the stringent statistical analysis (Fig. 21).

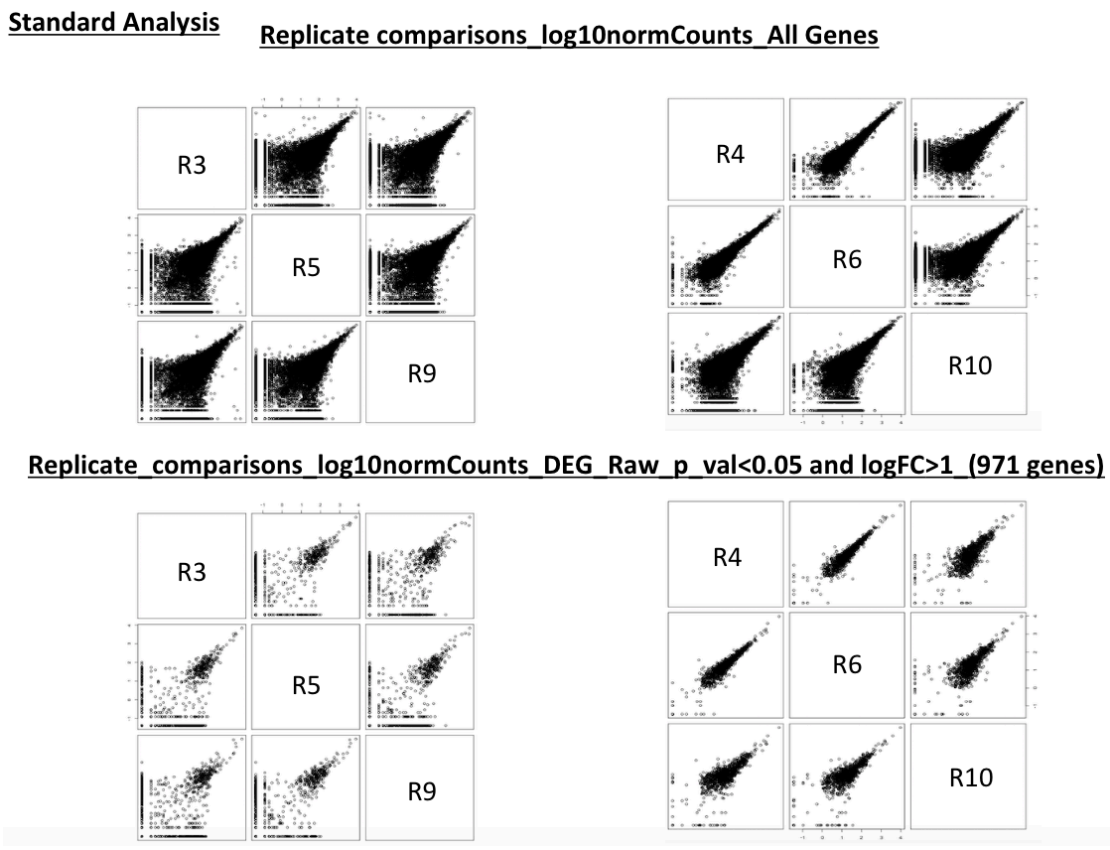


Figure 21: Scatter plots are drawn comparing biological replicates within Flk1 null and Flk1 heterozygous replicates: The Flk1 null replicates (R3, R5 and R9) show much more scattered counts, whereas Flk1 heterozygous replicates (R4, R6 and R10) show a higher correlation. It is also possible to see that the correlation between replicate samples significantly increases for the genes that are highly expressed and with a p value lower than 0,05.

Interestingly, canonical endothelial and hematopoietic genes did not change much between Flk1 null and Flk1 heterozygous cells. The endothelial specific genes including adhesion molecules such as Cdh5, Pecam1, Emcn, Icam2, transcription factors such as Ets1, Hey1 and Sox17 and receptors such as Flt1, Tie1, Tie2, Nrp1 and Eng did not show any significant variation between Flk1 null and heterozygous cells. This suggests that the sorted Icam2+ Flk1-cerulean+ expressing cells are true endothelial escapees. These cells could differentiate in the absence of the Flk1 function in the Tie2-Cre lineage. Also, CD38, CD34 and Fli1, genes implicated in hematopoiesis didn´t show any significant change (Fig. 22).

Since we know that the stringent whole transcriptome statistical analysis of low input samples has several caveats, and did not identify the gene Esm1 as a DEG, a gene clearly differentially expressed among the analysed samples (Avg reads Flk1 het samples = .071418 versus Avg reads Flk1 null Samples = 87.09, a logFC of 10.1), we decided to instead focus on the list of upregulated and downregulated genes with raw pvalue >.05 and 1<log FC<1. Examining their gene ontology (GO) enrichment analysis gave us a hint at the underlying biological differences in the two cell populations. The top functional categories upregulated in Flk1 null cells were apoptosis, immune responses and proteolysis involved in cellular protein catabolic process. The downregulated genes were cell cycle, focal adhesion and metabolic pathways.

Enrichment analysis using Enrichr API of the upregulated genes in Flk1 null cells revealed categories related to endoplasmic reticulum organization, protein processing, cell cycle control and cholesterol metabolism and biosynthesis (Fig. 23(a)). The interconnection between ER stress and Cholesterol metabolism leading to programmed cell death has already been shown (Rohrl C., 2018).

Among the upregulated genes in Flk1 null cells were Cathepsins such as CTSZ, D, H, B that are implicated in cellular protein catabolism and autophagy. Autophagy is a degradation pathway that plays critical roles in cellular stress response and cell death. Cathepsins have an integral role in autophagy by degrading autophagic material. Genes such as BCL2a1D, BCL2A1B, BNIP3 and BNIP3L, which are regarded as proapoptotic genes, were also highly upregulated. Interestingly BNIP3 has been shown to have dual roles on apoptosis and autophagy.

Apart from this, MGST1 and MT1, proteins that reside in outer membrane of mitochondria and Endoplasmic reticulum and are upregulated under oxidative stress (to protect against oxidative stress) were also upregulated in Flk1 null cells.
Further, Gene set enrichment analysis also showed that the p53 pathway was upregulated, suggesting DNA damage or cellular stress. The pathways related to the promotion of cellular proliferation (PI3K/AKT/MTOR signaling, Myc and E2f Target

genes) and protein biosynthesis/secretion and were downregulated in Flk1 null cells (Fig. 23(b,c)

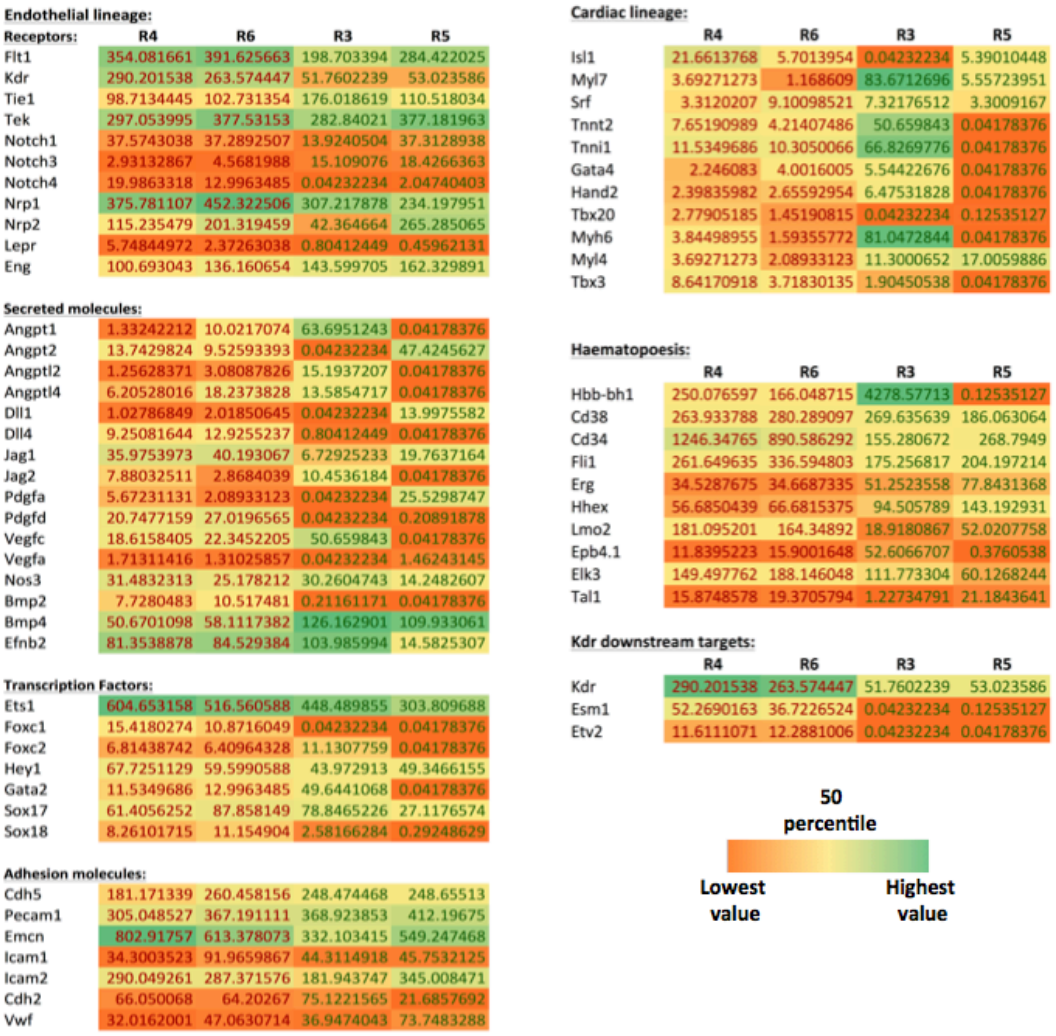


Figure 22: Heatmap of selected genes for the different categories: The values used for generating heatmap were normalized counts per million reads. R3 and R5 are Flk1-null replicates; R4 and R6 are Flk1 heterozygous replicates (We decided to discard the Flk1-Null replicate R9 sample and Flk1 heterozygous replicate R10 sample for lack of consisitency in their expression profiles)

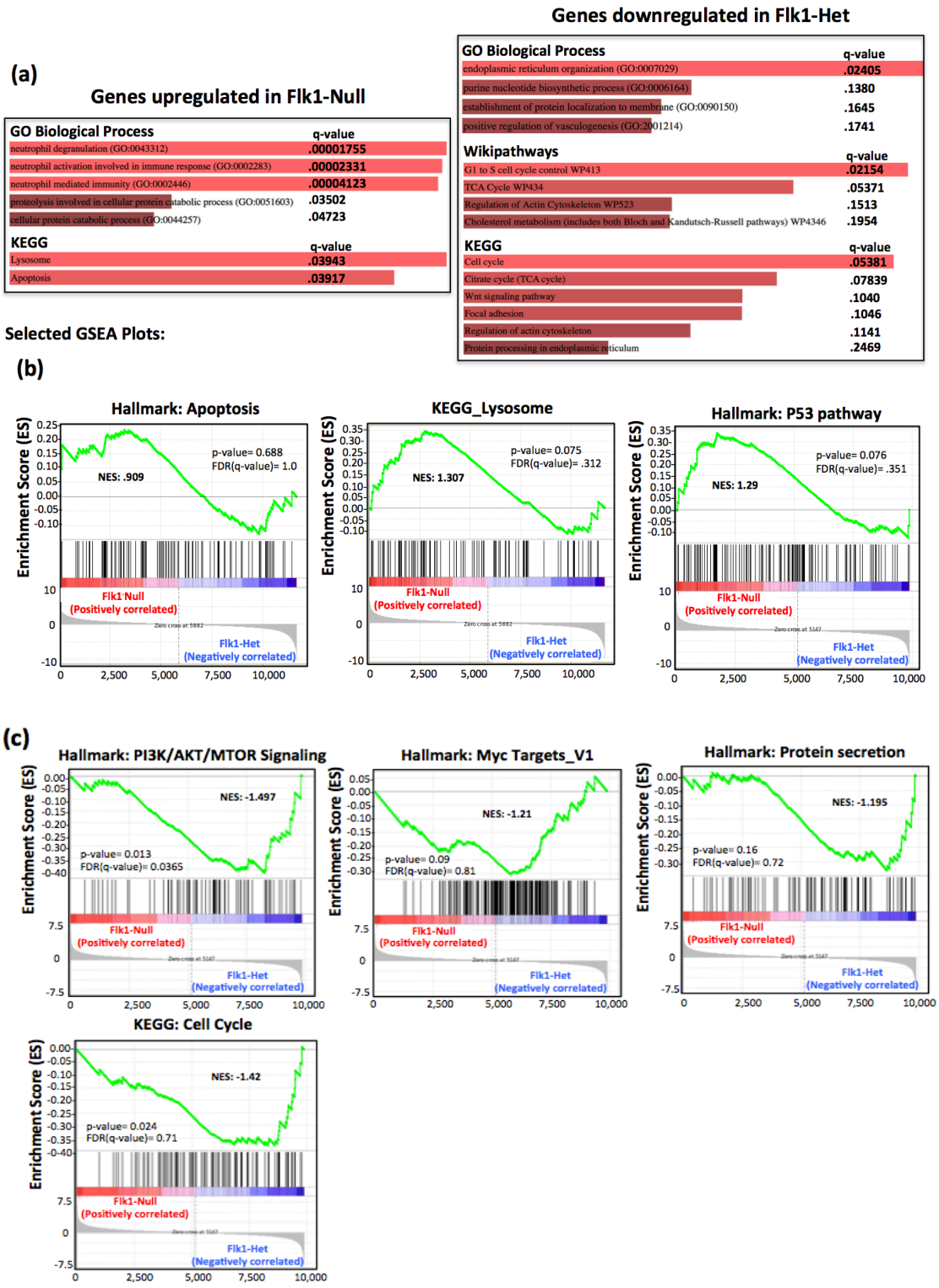


Figure 23: Comparative analysis of the *in vivo* transcriptome of Flk1 null and Flk1 heterozygous cells: (a) Gene ontology (GO) and KEGG analysis of processes upregulated and downregulated at raw pvalue <.05 and 1<logFC>1;q value(FDR) is set at a threshold of 0.05 for upregulated genes and 0.25 for downregulated genes

in FLK1 null; **(b)** Selected GSEA plots showing upregulation of apoptosis, lysosome and p53 pathway reated genes in Flk1 null endothelial cells; **(c)** Selected GSEA plots showing downregulation of PI3K/AKT/MTOR signaling, Myc targets, protein secretion and cell cycle in Flk1 null endothelial cells.

3. A genome wide loss of function genetic screen to identify mutations that favor endothelial expansion in the presence or absence of the *Flk1* function:

Flk1 is one of the few essential genes for endothelial and hematopoietic differentiation. To date, no other mechanisms or genetic mutations were shown to overcome the loss of Flk1 function during endothelial differentiation. The host group has shown several years ago that the effect of pharmacological targeting of VEGF/VEGFR2 signalling can be rescued by targeting the Notch signalling pathway, during retina angiogenesis (Benedito et al., 2012). However, in those studies only short-term pharmacological or genetic loss-of-function experiments were used. It is still not known if Notch or any other signalling pathway blockade could overcome the Flk1 loss-of-function during endothelial differentiation and proliferation *in vitro*. Understanding the molecular mechanisms by which other genes can functionally interact with Flk1, or replace its function, is of great relevance, particularly given the resistance to anti-VEGF in cancer therapies (Itatani et al., 2018).

To identify which genes may control the differentiation and proliferation of ECs, particularly in the absence of Flk1, we utilized a whole genome CRISPR/Cas9 based genetic screen, which was recently published **(**Doench et al., 2016**).** The screening is conducted by introducing a pooled lentiviral library encoding 78,637 specific sgRNA sequences (4 sgRNAs per gene) targeting 19,675 genes in the mouse genome. This lentiviral library from David Root group is available through AddGene (cat# 736333-LV) under the name Mouse CRISPR Knockout Pooled Library (Brie). This format requires Cas9 expression in the desired cell lines.

3.1. Generation of *Flk1*$^{KI/KI}$ ES cells expressing Cas9

We first generated a Rosa26 targeting vector containing the sequences for ubiquitous and strong CAG promoter followed by Cas9 nuclease protein having a nuclear localization signal, WPRE, Woodchuck hepatitis virus posttranscriptional regulatory element to enhance expression, and a transcriptional stop signal. The vector construct also contains PGK driven neomycin cassette for selection. This vector, and a px330 vector containing a guide sequence (sgRNA) for targeting the Rosa26 locus was used to nucleofect the previously generated ES cells Flk1$^{WT/WT}$ and Flk1$^{Ki/Ki}$ cell lines. After neomycin selection, several ES cell clones were picked and expanded. Some had a single copy integration of the donor vector in the Rosa26

locus (heterozygous), whereas others had a homozygous knock-in of the donor vector (Fig. 24). We also characterized these cell lines for the expression of the Flag-tagged Cas9 protein by immunostaining and western blot (Fig. 24).

For the lentiviral genomic screen we used the Flk1$^{Ki/Ki}$ R26$^{Cas9/Cas9}$ homozygous cell line (KC7), given its ubiquitous and higher expression of Cas9 (Fig. 24).

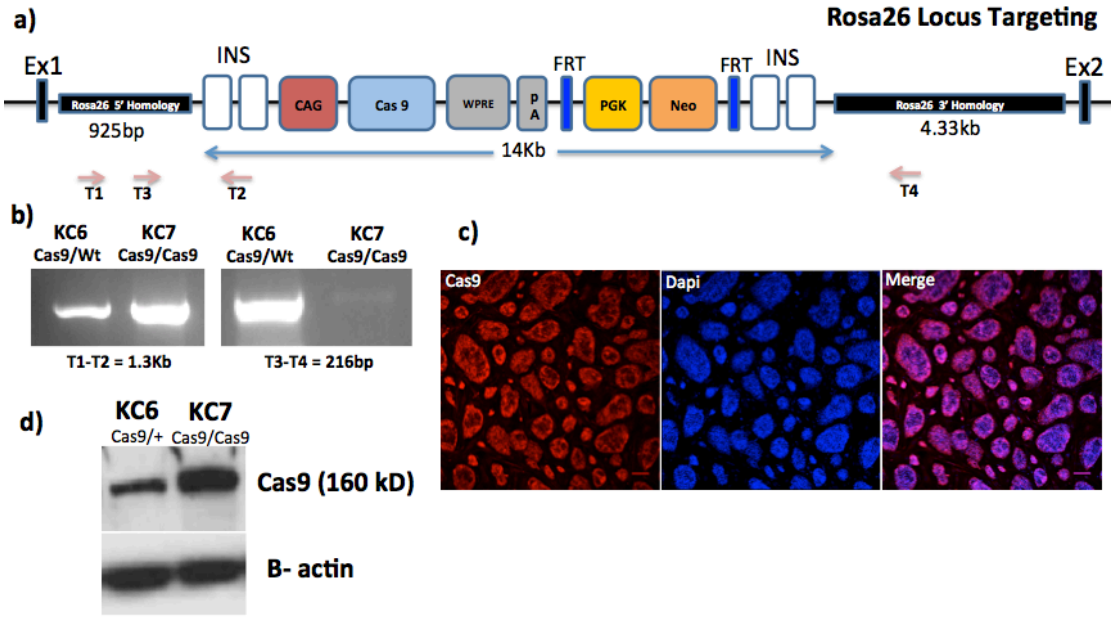


Figure 24: Generation of Flk1$^{KI/KI}$ R26$^{cas9/cas9}$Cell line by CRISPR/cas9: (a) Rosa 26 locus targeting of Cas9; **(b)** PCR analysis of genomic DNA from the selected ES cell clones. Left gel picture indicate the knock-in allele (Primers T1 and T2: 1.3kb) and Right gel picture with the wild type Rosa26 primers (primer T3 and T4: 216bp) indicate wildtype, non-targeted allele; **(c)** Immunostaining for Flag-tagged Cas9 protein for the clone KC7 (Rosa26:Cas9/Cas9); **(d)** Western blot for Flag-tagged Cas9 protein for clones KC6(Rosa26:Cas9/+) and KC7(Rosa26:Cas9/Cas9).

3.2. Determination of multiplicity of infection to achieve one lentivirus infection per cell

In order for the genomic screen assay to work, the mouse embryonic stem cell line Flk1$^{KI/KI}$:Rosa26$^{cas9/cas9}$ must be infected with the lentiviral library at a pre-determined MOI (multiplicity of infection) such that on average each mESC is infected with only one lentivirus. This guarantees a single sgRNA expression (and mutation) per infected cell. Therefore, we first had to calculate the functional Lentiviral titer. As can be seen in the titer chart, when the percentage of infected cells is equal to or less than 20%, the number of integrations is approximately equal to the number of transduced cells. At higher transduction levels the number of integration events per cell becomes non-linear owing to multiple integrations. Using this chart and the other optimized parameters for infection (Fig. 25 (b)) we were able to calculate MOI

(multiplicity of infection=integrations/cell) with good accuracy in the range of 0.1-0.3 MOI (Fig. 25 (c, d)).

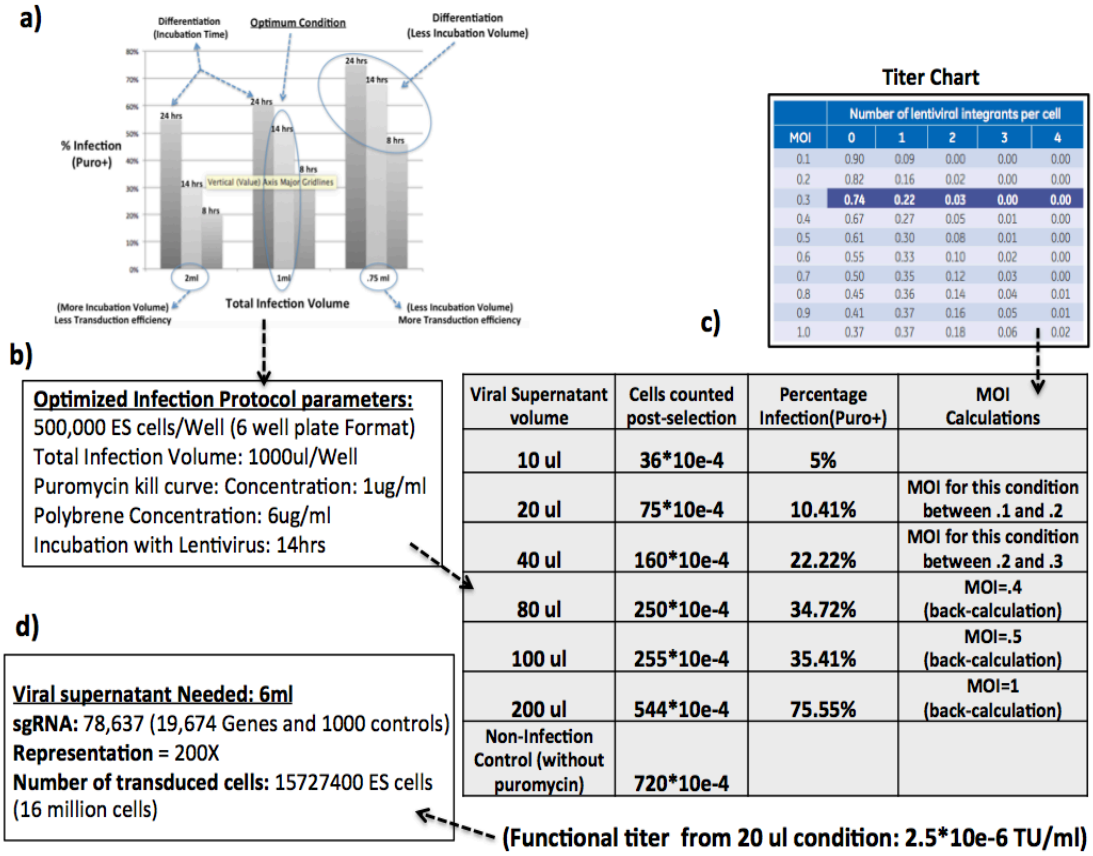


Figure 25: Optimization of pre and post assay parameters: (a) Length of Virus Incubation on cells and Influence of total media volume on Infection efficiency. **(b)** Optimized Infection parameters **(c, d)** Calculation of functional titer for Lentivirus library

3.3. Optimization of mESCs/OP9 coculture system to maintain a 200x library representation

For the efficient and reproducible differentiation of Flk1 loss of function ($Flk1^{KI/KI}$) mESCs into endothelial cells we had to optimize different parameters given the observation that $Flk1^{KI/KI}$ mESCs differentiate poorly to endothelial cells. After several trials, we found that a plating density of 10,000cells/cm^2 and 15 cm plate dishes gave us optimal differentiation (data not shown). After calculating several factors, such as library representation and cost of sequencing, we decided to maintain an sgRNA lentiviral library representation of 200X throughout the screening. This library contains 78,637 lentiGuide sgRNA constructs. Thus to achieve a 200x representation we needed eleven 15cm dishes plates to maintain our representation of 16 million post-infection cells at all steps of the protocol. For sorting the huge amount of differentiated endothelial cells needed, from a heterogenous pool of $Flk1^{KI/KI}$ cells, in a reliably, selective and fast way, we tried different antibodies (anti-Pecam, anti-Ve-

Cad, anti-Icam2) and fast magnetic-beads based protocols. After trying different combinations depending on their compatibility, binding efficiency and selectivity, we decided to use rat anti-Icam2 as an endothelial marker and anti-rat IgG magnetic beads for a first and quick round of cellular enrichment, followed by anti-Pecam-APC labeling for high quality/purity sorting of the endothelial population (Fig. 26).

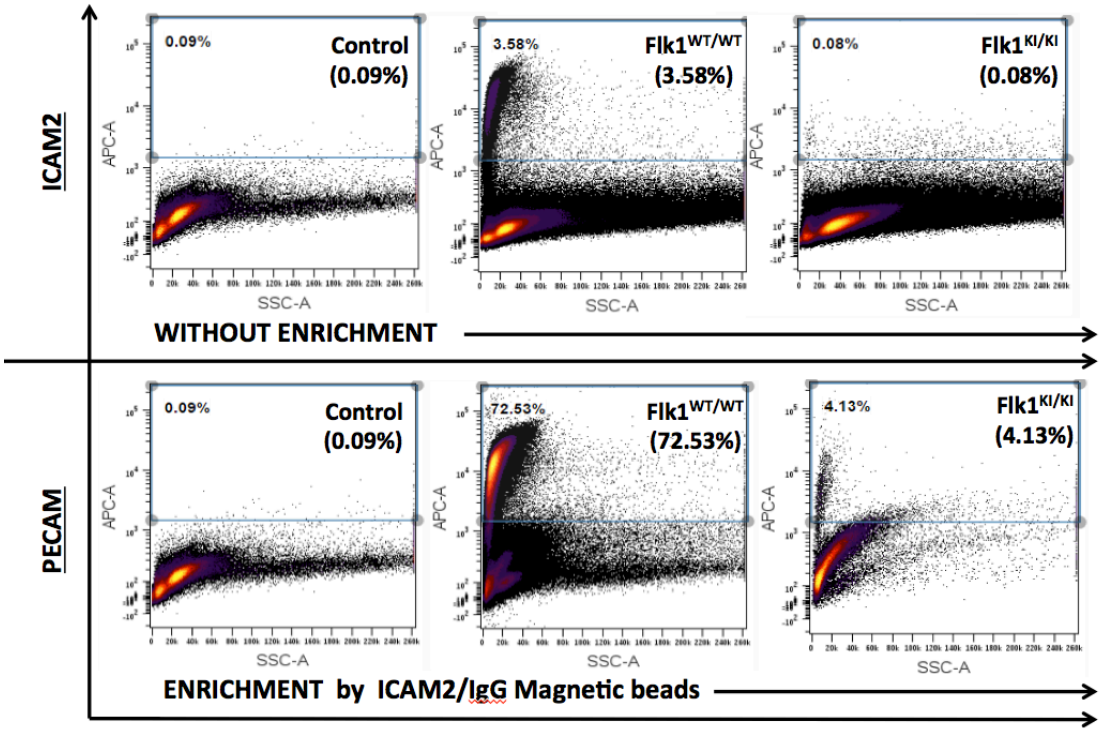


Figure 26: Optimisation of Enrichment protocol for Post-assay Selection Read out: Representative dot plot of the cytometry analysis. The cells were plotted in side scatter(SSC) versus Icam2+ and Pecam+ with their respective percentages. The above plots for Flk1$^{WT/WT}$ and Flk1$^{KI/KI}$ are without enrichment. The lower plots for Flk1$^{WT/WT}$ and Flk1$^{KI/KI}$ are after enrichment with Icam2/IgG Magnetic beads.

3.4. Infection of mESCs with the CRISPR Knockout Pooled lentiviral library and experimental samples

50 million mESCs were infected with LentiGuide sgRNA puro library at 0.4 MOI (200 mESCs per sgRNA construct unit). One day after, the puromycin selection was carried out for 5 Days to select the mESCs with lentiviral integrants. After selection, the pool of infected cells was trypsinized and divided into three samples or pools of cells. One pool **(Reference Infection Control Sample)** was subjected to immediate lysis and DNA extraction for library preparation. The second pool was plated on MEFs and maintained in mESC media for another 10 Days until the collection of the cells and DNA isolation **(Reference mESCs Sample)**. The third pool was plated on eleven 15 cm plates containing OP9 cells to induce endothelial differentiation. After 10 Days, these cells were trypsinzed and endothelial cells enriched using anti-Icam2 and magnetic beads. These cells were further sorted for PECAM-APC+ to obtain a highly pure mix of endothelial cells. Finally, and due to the requirement of Flk1 for endothelial

differentiation, we could only obtain 1 million Flk1$^{KI/KI}$ endothelial cells for DNA extraction and analysis (**Endothelial sample),** which is less than the 16 million we initially planned, but it still maintained library representation. (Fig. 27, 28)

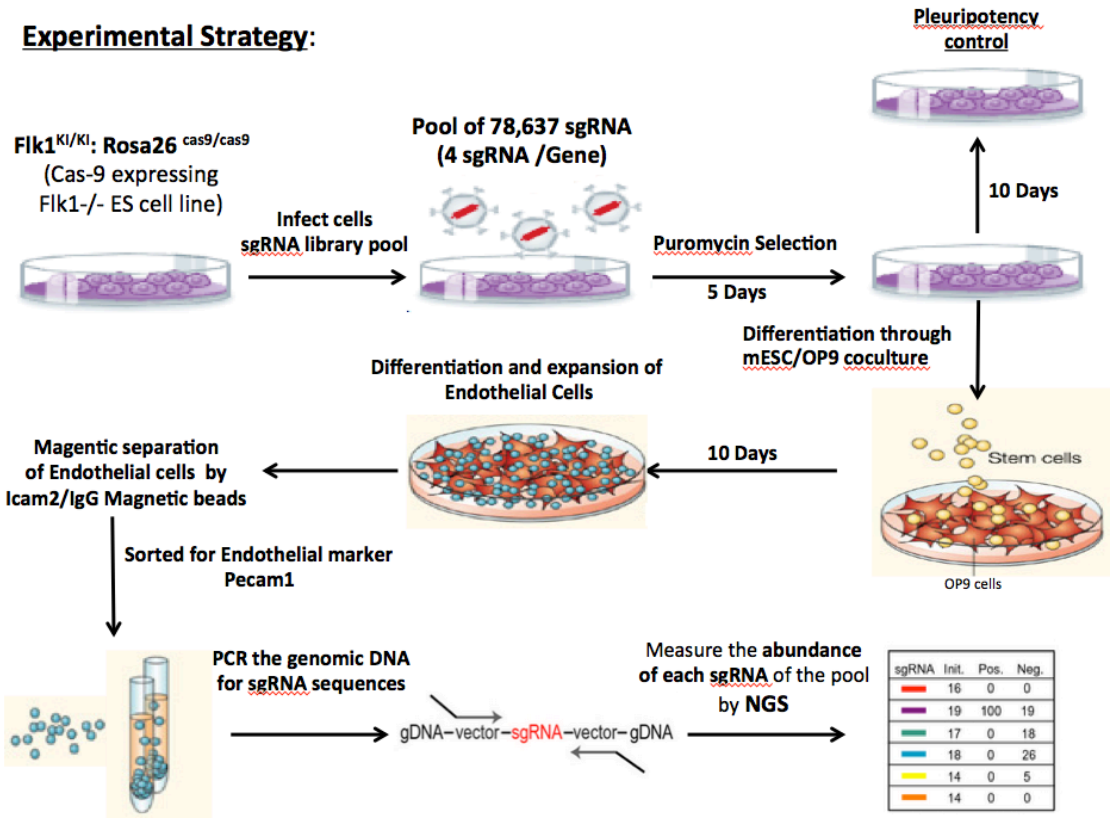


Figure 27: The experimental workflow of the Flk1$^{KI/KI}$ genomic screen explaining the timeline of experiment

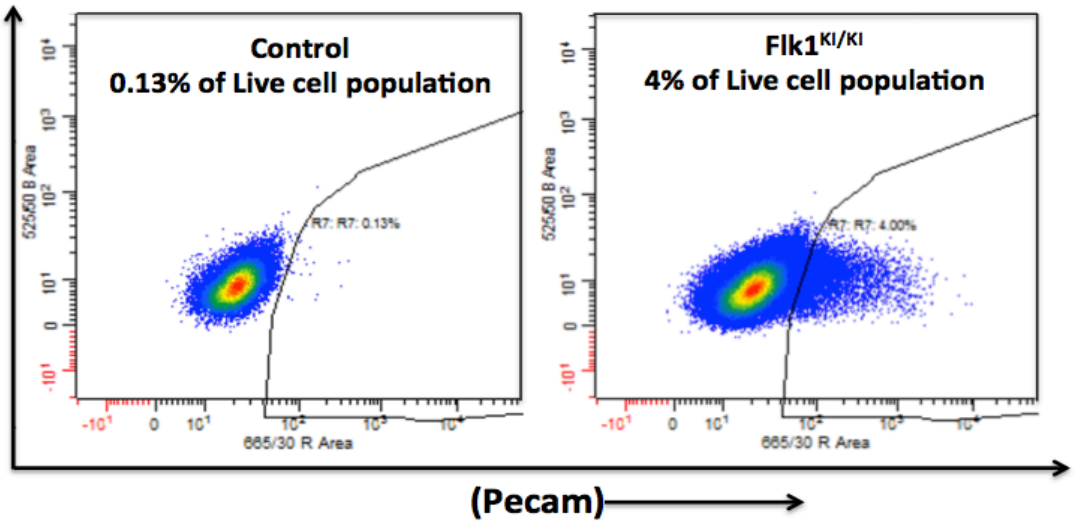


Figure 28: Representative scatter plot of the Flk1 $^{KI/KI}$ sorting for Pecam positive population after magnetic enrichment with Icam2: Representative dot plot with the

original gatings. The Pecam population is gated from live population (The screen is performed for two biological replicates)

Reference Infection Control Sample: This corresponds to DNA extracted from $Flk1^{KI/KI}$ mESCs 5 days after infection and selection in Puromycin. This sample allowed us to determine the baseline sgRNA distribution and library representation in the infected cells.

Reference mESCs Sample: This corresponds to DNA extracted from $Flk1^{KI/KI}$ mESCs maintained in mESC media for 10 days more after selection with Puromycin, to determine which sgRNAs/mutations changed the relative proliferation/ differentiation rate of $Flk1^{KI/KI}$ mESCs. Note that Flk1 is not expressed in mESCs, only in endothelial cells. Therefore the enrichment or underrepresentation of sgRNAs at this step is exclusively related with the targeting of genes that are essential for mESC (or any other cell) biology.

Endothelial sample: This corresponds to DNA extracted from $Flk1^{KI/KI}$ Icam2+/Pecam+ endothelial cells after 10 days of ES to endothelial-hematopoietic differentiation using the mESC/OP9 co-culture method. The sgRNAs/Cas9 induced mutations that favor the differentiation of $Flk1^{KI/KI}$ mESCs into endothelial cells or the proliferation of $Flk1^{KI/KI}$ endothelial cells will get enriched and the sgRNAs/Cas9 induced mutations that further block the differentiation or proliferation of $Flk1^{KI/KI}$ cells will be under-represented, in relation to their relative distribution in the mESCs and infection control samples.

3.5. Lentiviral sgRNA/CRISPR DNA library preparation

After the genomic DNA extraction from the different samples mentioned above, a two-step PCR was performed to prepare NGS libraries for sequencing. The first PCR with F1_PCR1 and R1_PCR1 primers were used to amplify all the different lentiviral sgRNA DNAs integrated in the genome of the collected cells. This PCR conditions depends on the sample cell number (library size). The second PCR with F0_PCR2_GeCKO and R0_PCR2_GeCKO was done for attaching illumina adapters and barcodes for library multiplexing. The strategy for preparing the NGS DNA library was borrowed from the Broad Institute´s Genetic perturbation platform and is described in Shalem et al..(Shalem et al., 2014) and in Figure: 29(a). The PCRs containing the libraries were gel purified and the obtained DNA checked for quality control and concentration with a High Sensitivity DNA chip on a Bioanalyzer (Agilent Technologies). To our surprise, we observed a DNA size distribution between 400 and 1000 base pairs. However, the library size should be only 370 bp. The only sample that fair in this analysis was endothelial sample from the second biological replicate (Fig. 29(c)). These PCR artefacts are partly double-stranded and partly single-stranded DNA, often called PCR bubbles. These libraries are however suitable

for sequencing as the abnormal DNA structures are merely mis-annealed PCR products. Therefore we went ahead with the DNA sequencing.

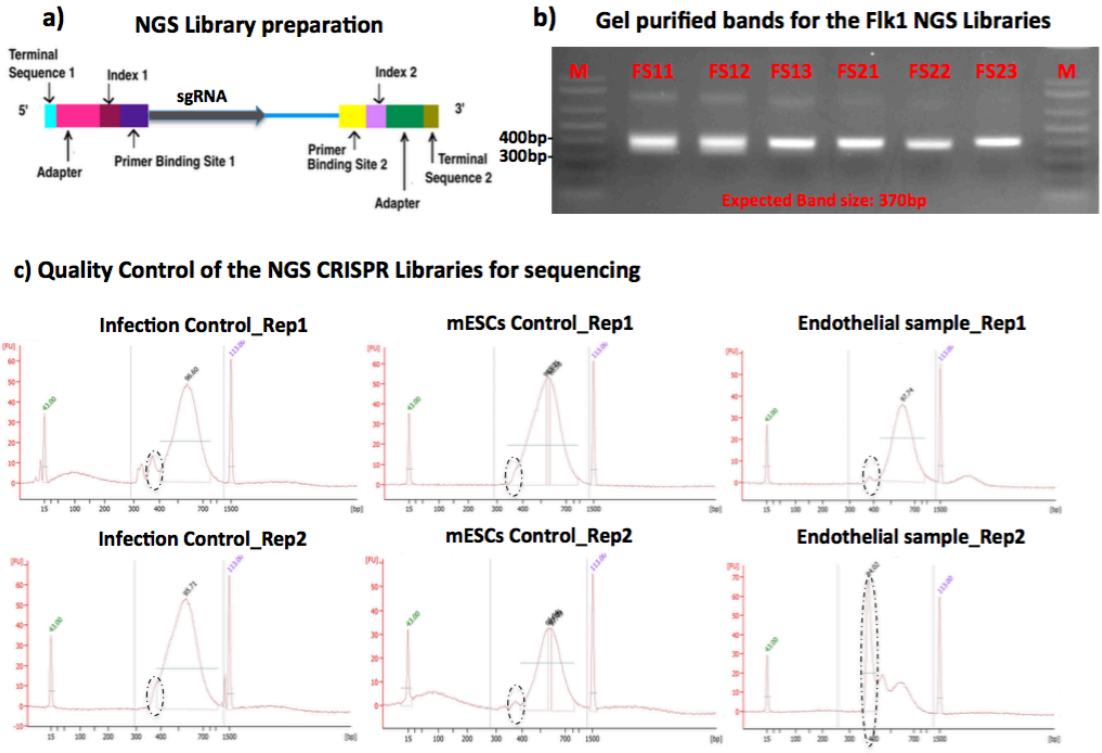


Figure 29: CRISPR DNA Libraries preparation for sequencing (a) Next Generation Sequencing (NGS) Library prep scheme **(b)** Quality control Run after Gel extraction of Flk1 NGS Libraries **(c)** Quality control of CRISPR Libraries by bioanalyzer.
***(**Biological replicate: 1**= **FS11**(Infection control), **FS12**(Pluripotency control) and **FS13**(Differentiation arm); **Biological replicate: 2 = FS21**(Infection control), **FS22**(Pluripotency control) and **FS23**(Differentiation arm))

3.6. Deep Sequencing the Lentiviral sgRNA/CRISPR DNA library

After the DNA library preparation we used the HiSeq 2500 (Illumina) platform/chip for the deep sequencing. Due to an unexpected technical problem, we got a relatively low number of reads (50% of the capacity of flow cell)**.** This might be due to under loading of the sample due to an initial inaccurate determination of its concentration. A very important step in the Next generation sequencing (NGS) workflow is the library quantification. Briefly, as you add your library on a flow cell, the single stranded, adapter-ligated fragments hybridize to the immobilized primers on the flow cell lane. Each hybridized DNA strand undergoes multiple rounds of amplification to produce upto 1000 copies of the same molecule and in the same location. These are called clusters. Too little DNA will result in an under-clustered Flow cell lane and fail to make an optimal use of the space thereby giving low reads.

Too much DNA on the other hand will result in an over-clustered Flow cell giving densely packed clusters and difficult to interpret the sequencing data due to poor resolution. These clusters are rejected during the quality control (reads passing filter) giving you less reads. We were not able to quantify our library properly due to over-amplification, as discussed in the above section, resulting in an under-clustered Flow cell and low number of reads.

Sequencing reads from two replicate screenings were further reduced after demultiplexing. To run all the samples together, we assigned a unique barcode to all the samples so that we can distinguish the reads coming from different samples. However due to mismatches in the barcode sequences, we were not able to assign the right read to the right sample and therefore ended up loosing a lot of reads in the process. This low mappability could be due to sequencing error or the sample quality.

In the end, we got very low number of reads for the first replicates (Infection Control_Rep1, mESCs Control_Rep1 and Endothelial sample_Rep1). However, we were able to recover a good amount of reads for the other replicate experiment containing the samples Infection Control_Rep2, mESCs Control_Rep2 and Endothelial sample_Rep2. We reasoned that although we lost one of the replicate experiments we still had multiple sgRNA reads for every gene in one single experiment, which can serve as an internal control. So, we went forward with our analysis using one replicate experiment. After preprocessing and aligning using MAGeCK v. 0.5.2 (Wei Li and Liu. 2014), the abundance of each sgRNA was assessed and normalized among samples. We were able to correctly map around 75% sgRNA sequences across this replicate sample points. Their QC metrics faired on every front. Despite having a lower than expected number of reads, we were able to detect almost 100% of the target genes and around 99% of sgRNA sequences. The samples showed an even/gaussian distribution of sgRNA, as evident from their low gini index (a common measure of income inequality in economics, can measure the evenness of sgRNA read counts)(Figure: 30)

a)

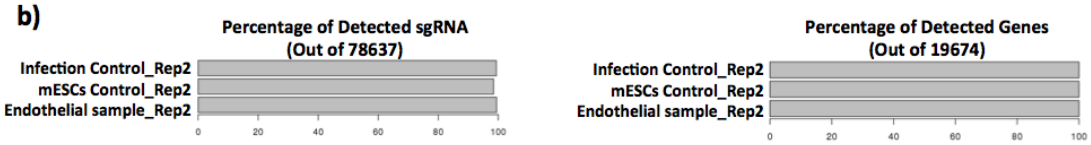
Sample Name	Reads	Mapped	Percentage	Expected
Infection Control_Rep2	1956212	1509745	77	65
mESCs Control_Rep2	3300029	2492364	76	65
Endothelial sample_Rep2	3830126	2855260	75	65

Sample Name	Total sgRNAs	Zero Counts	Percentage	Expected (upto)
Infection Control_Rep2	78637	508	0.64	1
mESCs Control_Rep2	78637	1234	1.55	1
Endothelial sample_Rep2	78637	442	0.56	1

Sample Name	Gini Index	Expected (upto)
Infection Control_Rep2	0.12	0.10
mESCs Control_Rep2	0.15	0.20
Endothelial sample_Rep2	0.10	0.20

The total number of reads should be 5-10 times more.....
200 * 78637 = 15727400

b)



c)

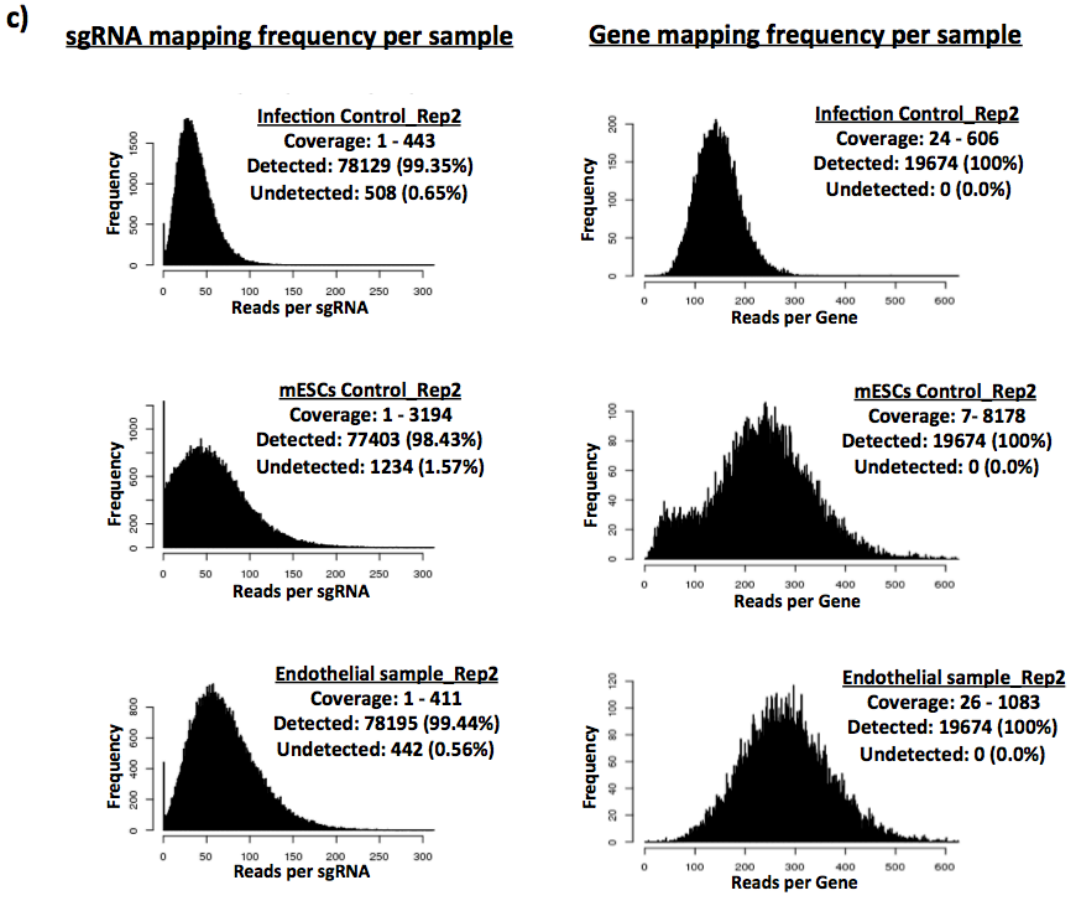


Figure 30: Table and plots showing sgRNA and Gene mapping frequency and their whole genome coverage (a) Measurement showing mapped reads, zero counts and Gini index across the samples in Replicate 2; **(b, c)** Number of detected and undetected sgRNAs and genes and their genome-wide coverage.

3.7. Bioinformatic analysis of differences in sgRNA/mutations distributions across samples

Identifying sgRNA/mutations that get enriched or underrepresented and exploring their biological functions is the next step in the data analysis. And, for this we used MAGeCK. MAGeCK calculates the Beta score, a metric that evaluates sgRNA/gene mutation selection performance. A positive or negative value for the Beta score indicates a positive or negative selection respectively (Figure: 31)

To perform the first global evaluation of the library mutational screening performance, we analyzed the distribution of sgRNA/mutations for a curated list of cell core essential genes (2744 sgRNA targeting 686 genes)(Wang et al., 2014) and 1000 negative control sgRNAs that do not target any gene function (included also in the Lentiviral Library).

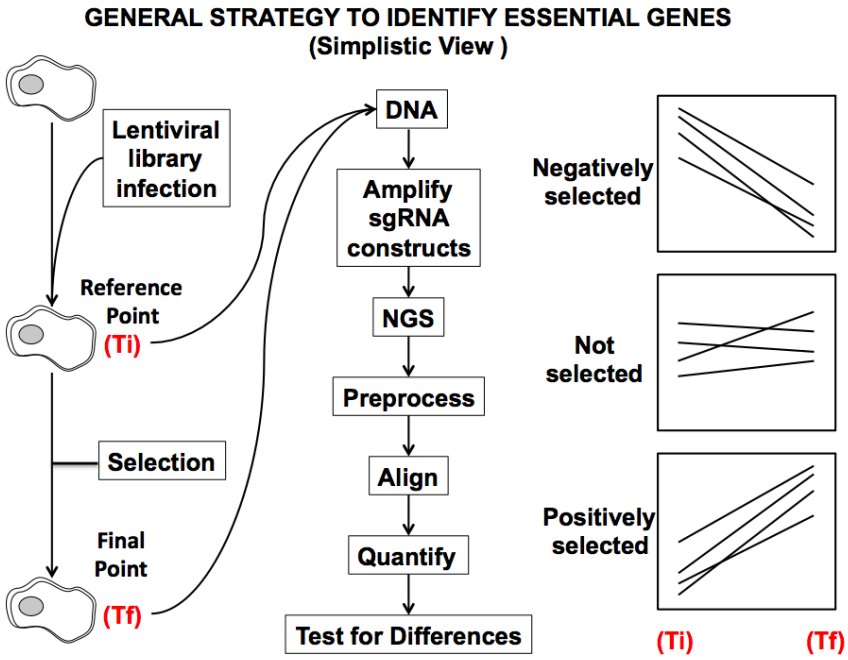
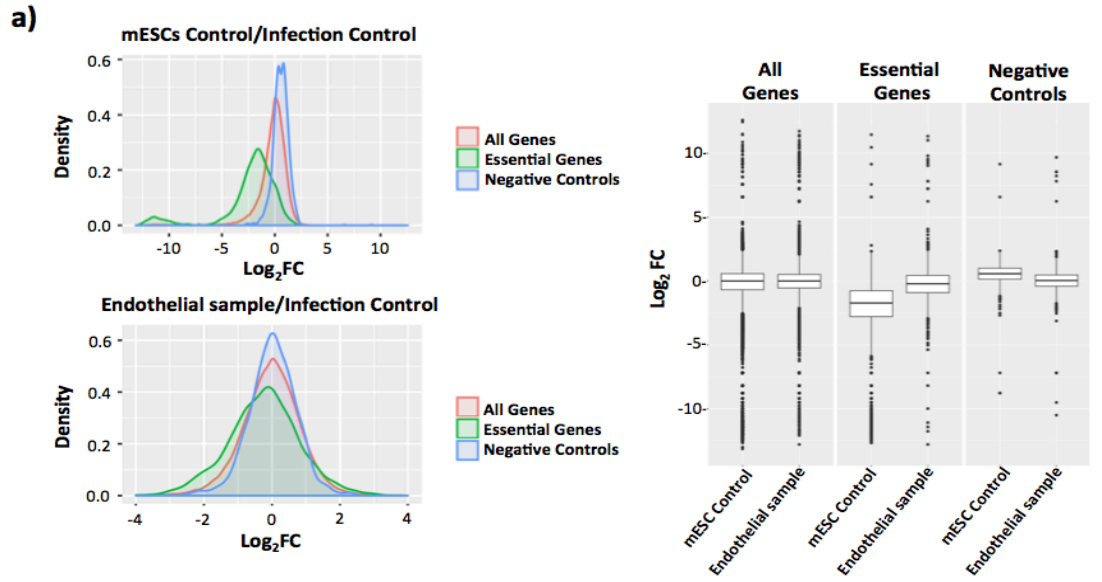
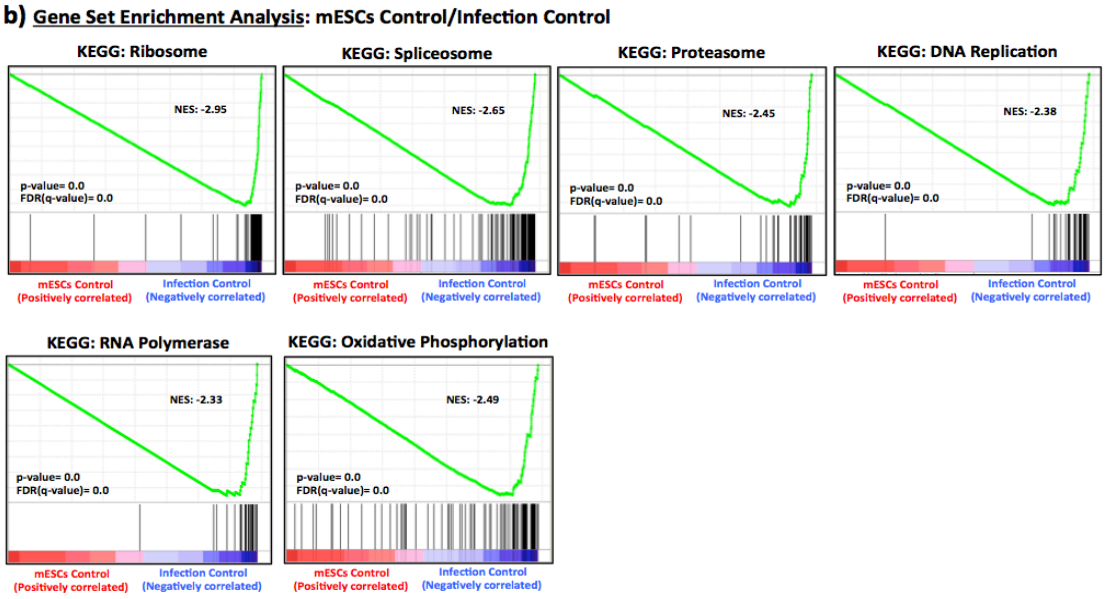


Figure 31: General lentiviral library screening strategy and MAGeCK workflow for identifying sgRNA/mutations that are positively or negatively selected.

These core essential genes are related to fundamental biological processes including DNA replication, gene transcription, and protein degradation. These genes are normally required for the survival of cells and when mutated will lead to cellular/sgRNA under-representation from the pool of cells/library (Wang et al., 2014; Hart et al., 2014, 2015). In the case of our samples, we performed 10 days of selection. As it can be seen in Fig. 32 the representation of the reads for the sgRNAs targeting the 686 essential genes significantly decrease when comparing the mESCs Control sample versus the Infection Control sample, whereas the 1000 negative control sgRNAs representation does not change. Further Gene sets related to

fundamental biological processes—including DNA replication, gene transcription, and protein degradation—showed strong depletion, which is consistent with their essentiality. However, in the case of endothelial sample, we did not observe such a significant under-representation of sgRNAs targeting essential genes, compared to mESCs control sample, although there was a noticeable difference in the drop-out efficiency of these gene. This could be due to the mutations, in these fundamental biological processes, providing growth advantage to $Flk1^{KI/KI}$ endothelial population. The proteasome, mitochondria and ribosomes have been implicated in the control of apoptosis by modulating the levels of both pro- and anti-apoptotic molecules (Gupta et al., 2018; Friedman et al., 2004; Wang et al., 2001; Zhou X., 2015; Golomb L, et al., 2014; Eberhard et al., 2013)





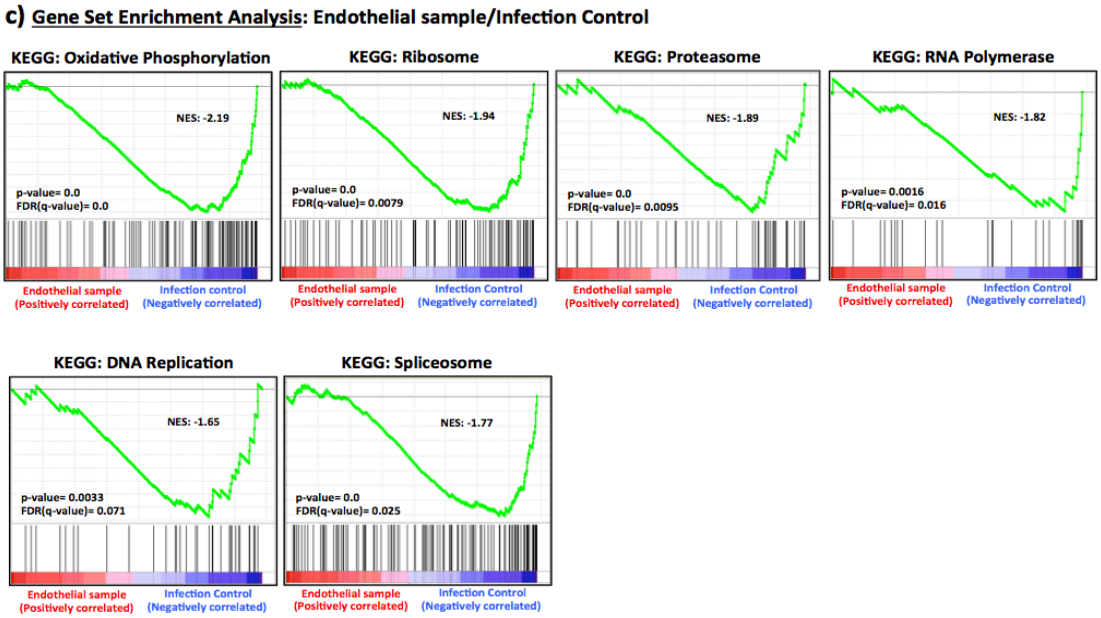


Figure 32: Global evaluation of Genetic screen comparing depleted sgRNAs target genes involved in fundamental biological process: (a) Fold change distribution of sgRNA for all genes, essential genes and negative control**.** The fold change distribution of sgRNAs targeting essential genes is shifted compared to the fold change distribution of the negative control sgRNAs; **(b, c)** Gene set enrichment analysis was performed on genes ranked by their combined depletion scores from screen in $Flk1^{KI/KI}$ comparing mESCs Control and Endothelial sample w.r.t Infection Control. Y-axis shows the enrichment score and Vertical lines underneath the x-axis denote members of the gene set analyzed. FDR(q value)<0.05, NES=normalized enrichment score.

3.7.1. Biological functional analysis of the Enriched and Underrepresented sgRNAs/Genes in the mESCs Control sample versus Infection control

In this comparison we got mutations in 16 genes that were enriched and mutations in 5365 genes that were underrepresented after applying a stringent statistic threshold of an FDR of <0.05. These are just the most deregulated genes, for which statistical significance was obtained. These genes respectively suppress (16 genes) or enhance (5365) mESCs proliferation, adhesion to the well, survival or overall fitness during the 10 days in culture. It is interesting to see that as expected, most of the mutations that gave hits lead to a loss of gene and consequently cell function/viability. Whereas few mutations induced an increase in cell proliferation or viability. Many of the 5365 genes that when mutated reduce cell survival belonged to the 686 genes of the list of core essential genes (Wang et al., 2014; Hart et al., 2014, 2015) (Fig. 32 (a,b)). Biological pathway or process enrichment analysis for the top 500 ranked genes showed that these genes belonged to categories related with house keeping functions and cellular homeostasis, such as ribosome function and biology, protein degradation, DNA replication, spliceosome and oxidative phosphorylation (Fig. 33(a)). In contrast, the enriched gene mutations that enhanced

cellular survival and proliferation belonged to the p53 pathway and other known tumor suppressors. In fact the second strongest hit in our screen was Trp53/p53. A recent report from Merkle et al., 2017 showed that hPSCs carrying dominant negative TP53 mutations expand within a population of hPSCs. The top 20 ranked genes also include Myh9, Myl6, Rock1, Rock2, Gna13, Dapk3, Flii and Pawr. Most of these genes have a role in the actin and myosin network or cytoskeleton (Fig. 33(b,c)). It has been demonstrated that ROCK/myosin /actin pathway triggers cell death, similar to anoikis (Chen et al., 2010; Ohgushi et al., 2010). Also, ROCK and myosin inhibitors are routinely used when passaging hPSC, as ROCK and Rho gets activated on dissociation of hPSCs that leads to myosin phosphorylation causing membrane blebbing and ultimately cell death (Watanabe et al., 2007; Chen et al., 2010). Another top ranked hit, was Gna13, which is known to activate RhoA/ROCK signaling pathway. We passed our mESCs 4-5 times during the whole screening and we believe that this is what might have led to the enrichment of mutations in these genes. For most of the top ranked hits multiple sgRNA targets were recovered with a significant Log FC giving strength to our data. There was also an enrichment of mutations in proapototic regulator genes like Pawr. Its role as an actin-cytoskeleton component leading to membrane blebbing and cell death has been demonstrated (Burikhanov et al., 2009).

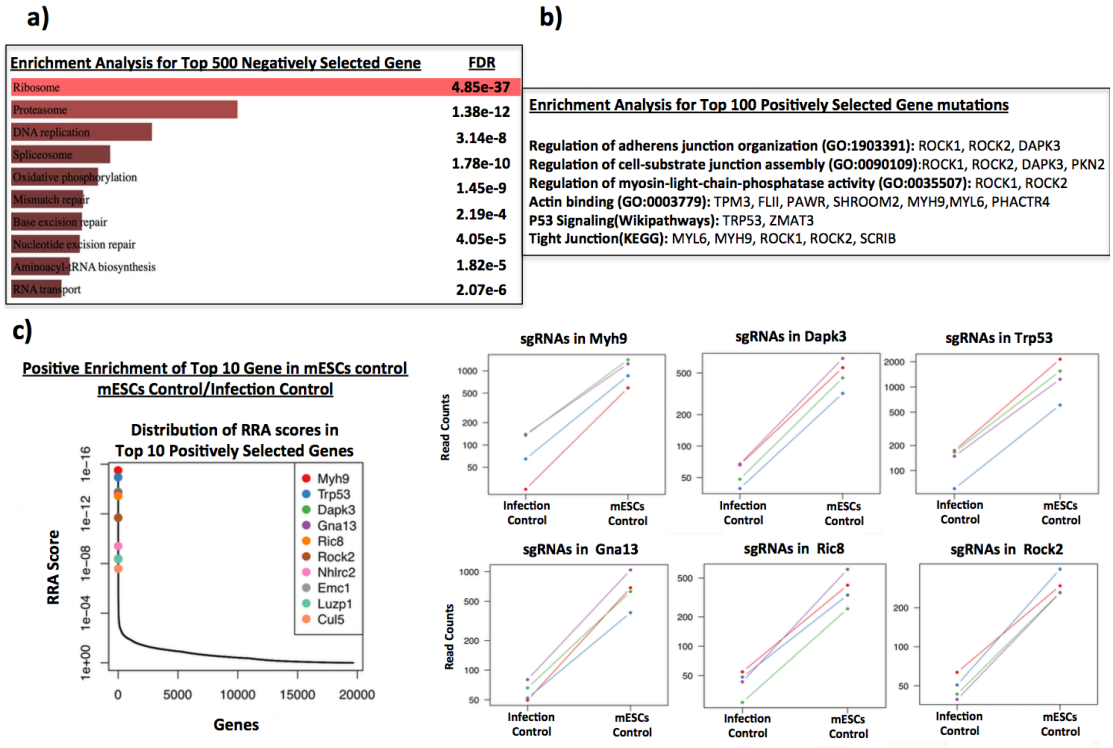


Figure 33: Functional enrichment analysis of the Top Positively/Negatively selected gene mutations in mESCs Control: (a) Functional enrichment analysis of the top 500 negatively selected gene mutations using Enrichr API. **(b)** Functional enrichment analysis of the top 100 positively selected gene mutations using Enrichr API **(c)** Plot showing top 10 positively selected gene mutations and a few selected one´s with their sgRNA behavior.

3.7.2. Biological functional analysis of the Enriched and Underrepresented sgRNAs/Genes in the Endothelial sample versus Infection control

As mentioned above, the number of cells obtained for the Endothelial sample was lower than ideal. We obtained only 1 million, instead of 16 million cells. This affected the amount and depth of the sequencing analysis, which impacts on the bioinformatic statistical analysis. Nonetheless, we analyzed the results carefully to see if in the surviving $Flk1^{KI/KI}$ Icam2+/Pecam+ Endothelial sample there was any particular mutation giving any advantage or disadvantage to the $Flk1^{KI/KI}$ population during endothelial differentiation. To this end, we got only 8 positively enriched genes and 15 negatively enriched genes that passed the threshold of an FDR of <0.25. Given the lower input and quality of sequencing of the FS23 sample, we believe these are only the top hits that passed the statistical test. The top 10 positively selected gene mutations such as Eif3f, Dcaf6, Ubtd1, Itch and Ssna1 are regulators of cell proliferation, differentiation and apoptosis. Eif3f is a part of complex that initiates translation of mRNAs involved in cell proliferation, differentiation and apoptosis to activate or repress these processes translationally. Ubtd1 is a p53 downstream target gene involved in the induction of cellular senescence through a positive feedback loop with p53. Dcaf6 may act as tumor promoters and tumor suppressors. Endothelial cells without the Flk1 function in endothelial sample may undergo cellular senescence or apoptosis and mutations in these genes may give them a competitive advantage. Itch is involved in the repression of apoptosis and reactive oxygen species levels through the ubiquitination and proteasomal degradation of Txnip (Fig. 34(a)). Other genes in the top 50, such as Hira, mediates the irreversible cell cycle changes that occur in senescent cells and efficient senescence-associated cell cycle exit, Ptprq catalyzes catalyzes the dephosphorylation of phosphotyrosine and phosphatidylinositol and plays roles in cellular proliferation and differentiation. Dact1, Agpat1, Itm2a and Vwc2l play a role in fat cell differentiation adipogenesis, osteo- and chondrogenic differentiation. Tnfsf14 and Rras2 that controls apoptosis and proliferation in tumors were also enriched.

Interestingly, among the top 200 genes with mutations enriched in Endothelial sample vs Infection control and mESCs control vs Infection control, only 15 were common, which suggests that different genes may be differentially important for the proliferation/survival/fitness of $Flk1^{KI/KI}$ mESCs and endothelial cells (Fig. 34(b)).

The bioinformatic analysis using Enrichr API for the top 500 genes with depleted mutations in the Endothelial sample vs Infection control shows that the depleted genes belonged to the category of core essential processes such as protein translation, nucleic acid metabolism, chromatin regulation, regulation of telomere maintenance and lengthening (Fig. 34(c)).

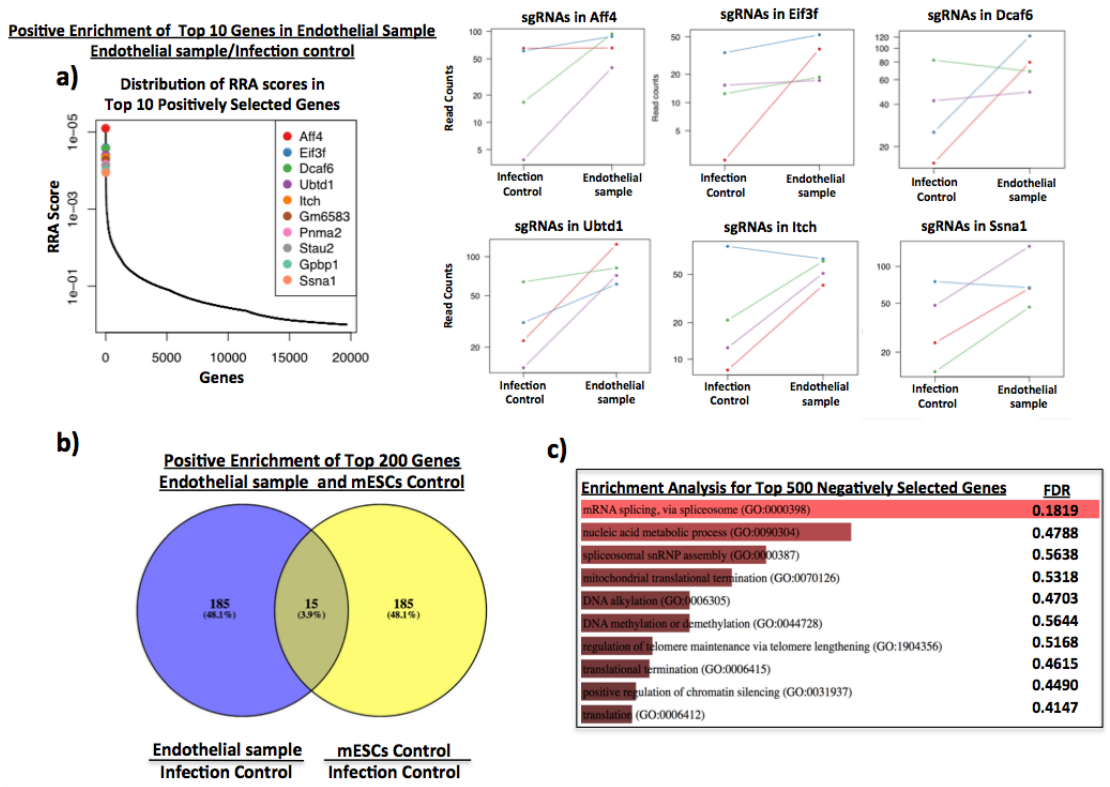


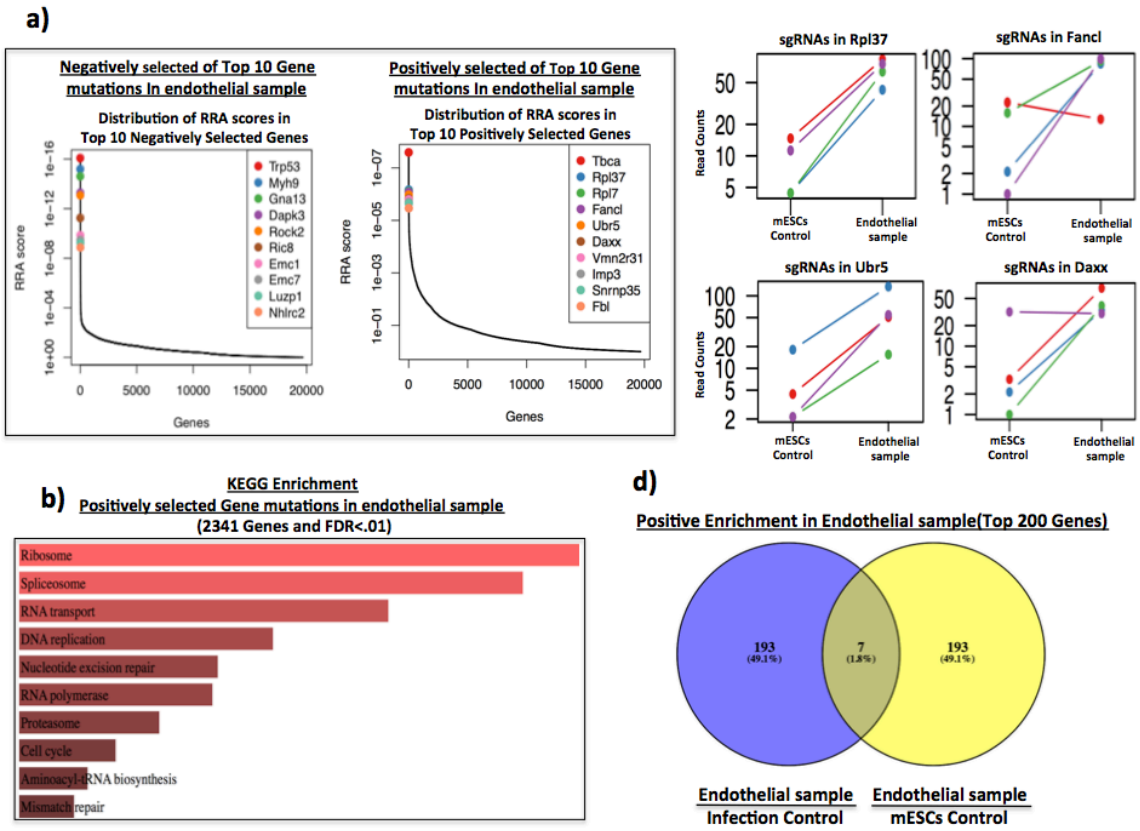
Figure 34: Functional enrichment analysis of the Top Genes Positively/Negatively selected Gene mutations in Endothelial sample: (a) Plot showing top 10 positively selected gene mutations and a few selected one´s with their sgRNA behavior **(b)** Venn diagram showing the common genes in Top 200 positively enriched genes between endothelial sample and mESCs Control **(c)** Functional enrichment analysis of the top 500 negatively selected gene mutations using Enrichr API.

3.7.3. Biological functional analysis of the Enriched and Underrepresented sgRNAs/Genes in the Endothelial sample versus mESCs control

On comparing endothelial sample with mESCs control, we got 2342 positively selected gene mutations at a stringent static threshold of an FDR of <0.01 and 15 negatively selected gene mutations at an FDR of<0.05. The group of negatively selected genes was identical to the ones that were positively selected in mESCs control indicating their strong enrichment and involvement in the mESC biology in our screen (Fig. 35(a); 33(c)). The positively selected gene mutations belong to Ribosome biogenesis, protein expression and degradation of protein, DNA replication and cell cycle. The top 10 gene mutations include Rpl37, a pro-apoptotic gene that binds MDM2 and inhibit MDM2 E3 ligase activity, leading to p53 stabilization and cell cycle arrest (Daftuar et al., 2013). Fancl and Ubr5, ubiquitin ligase proteins targeting specific proteins for ubiquitin-mediated proteolysis thus performing tumor suppressor functions (Meetei et al., 2003;). Daxx, another tumor suppressor, mediates activation of the JNK pathway and apoptosis via MAP3K5 in response to signaling from TNFRSF6 and TGFBR2 (Ueda et al., 2018)(Fig. 35(a,b)). These positively selected mutations belong to the ubiquitin protein ligase activity

regulating ubiquitin-dependent protein catabolic process and mitotic cell cycle through proteasome. The proteasome has been implicated in the control of apoptosis by modulating the levels of both pro- and anti-apoptotic molecules. Studies have shown that the ubiquitin-proteasome pathway mediates the degradation of key apoptotic regulators and dictate the cell survival vs death (Jesenberger et al., 2002). Thus mutation in these genes can give a growth advantage to the Endothelial cells with Flk1 loss of function. Furthermore, many tumor suppressors and oncoproteins regulate cell growth and proliferation through ribosomal biogenesis and protein synthesis (Zhou X et al., 2015) Therefore mutating these tumor suppressive ribosomal proteins can make $Flk1^{KI/KI}$ endothelial cells survive. There was also an enrichment of TNF-alpha NF-kB Signaling Pathway in the enrichment analysis of positively selected genes, which includes various ribosomal proteins and ubiquitin ligases stating the above argument. NF-kappa-B and TNF signaling pathway is activated by multiple stimuli such as inflammatory cytokines, bacterial or viral products, DNA damages or other cellular stresses and regulate cell survival and apoptosis (Fig. 35(b, c)).

However, we are not sure, if all these positively selected gene mutations are true hits and giving growth advantage to $Flk1^{KI/KI}$ endothelial cells in the endothelial sample. This is because there was a significant down-regulation of these genes in mESCs control population because of their essentiality therefore even a small increase/slow depletion of these genes in endothelial sample cell population will result in a significant positive fold change in endothelial sample. Moreover we observed a very little overlap in the first 200 positively selected gene mutations when the endothelial sample was compared with Infection control and mESCs control references (Fig. 35(d)) Thus it was difficult to extract true hits from the noise of the experimental screen.



c) **Enrichment in GO categories and Wikipathways:**

- **Ribosome biogenesis** (GO:0042254) (FDR:1.221e-73)
- **Positive regulation of ubiquitin protein ligase activity** (GO:1904668) (FDR: 6.790e-18)
- **Negative regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle** (GO:0051436) (FDR: 3.658e-20)
- **Positive regulation of ubiquitin-protein ligase activity involved in regulation of mitotic cell cycle transition** (GO:0051437) (FDR: 1.269e-19)
- **Positive regulation of protein ubiquitination involved in ubiquitin-dependent protein catabolic process** (GO:2000060) (FDR: 1.248e-16)
- **TNF-alpha NF-kB Signaling Pathway** WP246 (FDR: 3.326e-9)
- **Proteasome Degradation** WP519 (FDR: 1.093e-8)

Figure 35: Functional enrichment analysis of the Top Genes Positively/Negatively selected Gene mutations in Endothelial sample: (a) Plot showing top 10 negatively and positively selected gene mutations in endothelial sample. Plots showing sgRNA behavior for the selected positively selected gene mutations; **(b,c)** Functional enrichment analysis of the top 2341 positively selected gene mutations with an FDR<0.01 in different databases using Enrichr API; **(d)** Venn diagram showing the comparisosn of Top 200 positively enriched genes in endothelial sample compared to Infection control and mESCs control references.

Discussion:

The loss of *Flk1* function induces endothelial to mesenchymal transition *in vitro*

The current knowledge on how *Flk1* controls the endothelial and hematopoietic lineages development, or what alternative molecular trajectories and cell fates cells without the *Flk1* function undertake is still incomplete. Shalaby et al., demonstrated that hematopoietic or endothelial lineages fail to develop in *Flk1* mutant embryos using *LacZ* knockin/knockout embryos. Interestingly, the expression of the *Flk1-lacZ* reporter is maintained in these mutant embryos (Shalaby et al., 1995) suggesting that these cells retain at least some endothelial/hematopoietic progenitor characteristics. In contrast, $Flk1^{-/-}$ mutant ESCs can develop endothelial and hematopoietic lineages after in vitro differentiation, but at very low efficiency (Hidaka et al, 1999; Schuh et al., 1999).

Our results show that *in vitro* $Flk1^{KI/KI}$ mESCs can partially bypass the requirement for *Flk1* in endothelial development and form $Flk1^{KI/KI}$ *Flk1-Cerulean+CD31+Icam2+* cells, however with much less efficiency than $Flk1^{KI/WT}$ mESCs. The differentiated $Flk1^{KI/KI}$ *Cerulean+CD31+Icam2+* cells proliferate significantly less and are unable to sprout, consistent with known role of *VEGFR2* and downstream *ERK/MAPK* activity in endothelial proliferation and sprouting (Hidaka, M. *et al.*, 1999 and Schuh *et al.*, 1999). We believe that this $Flk1^{KI/KI}$ *Cerulean+/Icam2+* population is maintained in a relatively inactive endothelial state, but without trans differentiating or undergoing cell death. This is supported by the significantly lower expression of genes related with endothelial migration, protein/MAPK phosphorylation, vasculogenesis and sprouting angiogenesis. It is important to mention that besides *Vegfr2/Flk1*, endothelial progenitor cells also express Vegfr1 and Vegfr3, that eventhough have a much lower VEGF signaling ability, they can maintain some basal level of VEGFA/VEGFC signaling (Augustin et al., 2009). In contrast to this $Flk1^{KI/KI}$ *(C+I+)* population, the $Flk1^{KI/KI}$*(C-I+) population* is formed by a group of cells that lost the expression of the *Flk1* promoter and is becoming more mesenchymal and

differentiating to the cardiac lineage, as evident from the upregulation of epithelial to mesenchymal transition markers and cardiac lineage genes.

Through lineage tracing and cell differentiation studies, it has been shown that Flk1+ mesodermal progenitor cells can give rise to endothelial and blood lineages, cardiac lineages, vascular smooth muscle cells and paraxial mesoderm tissues (Hidaka et al., 1999, Schuh et al., 1999, Shalaby et al., 1997). Our data suggest that upon *Flk1* loss in those progenitors there is a differentiation bias towards the cardiac, vascular smooth muscle cell, and paraxial mesoderm lineages that tend to form mesenchymal and muscle cells. This is in line with other studies showing that in the absence of Etv2, a known regulator of *Flk1* and endothelial development, hemangiogenic mesoderm differentiates to cardiac lineage (Liu *et al*., 2012; Palencia-Desai *et al*., 2011; Rasmussen *et al*., 2011).

Pdgfra is a paraxial mesoderm marker (Sakurai et al., 2006; Darabi et al., 2008). Sakurai et al., suggested that *Flk1+/Pdgfra+* mesoderm, also known as nascent mesoderm, can give rise to *Flk1+Pdfra-* lateral plate mesoderm and *Flk1-/Pdgfra+* paraxial mesoderm (Sakurai et al., 2006). They also showed that cells at these stages are developmentally plastic and interchangeable in culture. Our transcriptome analysis showed that $Flk1^{KI/KI}$*(C-I+)* progenitor cells are also trans-differentiating towards the paraxial $(Flk1^{-}/Pdgfra^{+})$ lineage, suggesting the existence of a bipotent common progenitor. Recently, it has been shown that treatment with an *ERK* inhibitor blocked the progression of *Flk1+/Pdgfra+* cells towards the endothelial lineage and favoured the *Pdgfra+* paraxial mesoderm and cardiac progenitor lineages (Zhao et al., 2017). Since *Flk1* is a strong *VEGF* receptor tyrosine kinase that activates *MAPK/ERK* (Harding, A. etal., 2017), it suggests that *VEGF/Flk1* may be essential to increase *ERK* activity in a subset of progenitors, and that this will give rise to the endothelial/hematopoietic lineages. This is in line with the previous observation in which an elegant whole genome binding study demonstrated the need of *Etv2-Flk1* feed forward mechanism for the expansion of hemangiogenic progenitor *Flk1+/Pdgfra-* thus requiring simultaneous activation of genes specifying endothelial and haematopoietic lineage and Flk1 signaling (Liu et al., 2015). We believe that the absence of this *Etv2-Flk1* feed forward mechanism in $Flk1^{KI/KI}$

progenitor cells is diverting them towards the paraxial *(Flk1^{-}/Pdgfra^{+})* and cardiac mesodermal *(Flk1$^{(Low)}$/Pdgfra^{+})* lineages.

Our work suggest that in the absence of an active *Flk1* signalling input, endothelial cells may either die (these would not be present in our analysis) or undergo EndMT. It is known that endothelial cells experiencing growth factor deprivation or having activation of *ß-catenin*, *Tgfb*, *Bmp* pathways undergo EndMT (Dejana et al., 2017). During this process they acquire a variety of different mesenchymal characteristics and fates (fibroblasts, smooth muscle cells), including secretion of ECM proteins, such as fibronectin and collagen, altered endothelial cell junction organization, loss of cell polarity, and increased cell proliferation and migratory capacity (Kalluri et al., 2009).

Vegfr2 leads to the activation of *MAPK* family of cytoplasmic serine/threonine kinases, *ERK1/2* pathway, while inhibition results in suppression of ERK activation leading to embryonic mortality (Simons *et al.,* 2016). Recently, a study showed that *ERK1/2* loss in endothelial cells, induced the activation of the *Tgfb* signaling pathway and resulted in induction of EndMT in quiescent endothelium (Ricard N et al., 2019). We also observed an increase in EndMT and upregulation of the *Tgfb* and *Bmp* signaling pathways in *Flk1$^{KI/KI}$* cells, particularly in the *Flk1$^{KI/KI}$ (C-I+)* cells and not as much in the *Flk1$^{KI/KI}$ (C+I+)* cells. Therefore, it seems that in our system the EndMT happens mainly in cells that completely loose the expression of Flk1-Cerulean, which are likely cells loosing their endothelial lineage characteristics, eventhough they still express Icam2 (perhaps at lower levels, that are still detectable by FACS).

Recently it was proposed the existence of a common *AplnR+* progenitor for mesenchymal and endothelial cells, termed as mesenchymoangioblast (MB), in human pleuripotent stem cells (hPSC) cultures (Vodyanik et al., 2010). This *AplnR+* MB progenitor form after two days of hPSC differentiation and proceeds through endothelial/angioblastic cell intermediates and subsequently gives rise to mesenchymal stem cells and its lineages through EndMT. We found an upregulation of the *AplnR* gene expression in *Flk1$^{KI/KI}$* cells and this may reveal an enrichment of

mesenchymal stem cells. However, since we did bulk RNAseq, and not single-cell RNAseq, we cannot confirm this. This requires further investigation.

The loss of Flk1 function induces endothelial apoptosis *in vivo*

It has been shown that E7.5 embryos deficient in *Flk1* receptor (*Flk1*$^{-/-}$) possess hematopoietic progenitors however E8.5 *Flk1*$^{-/-}$ embryos are deficient in such cells (Schuh et al., 1999). Because *Flk-1*$^{-/-}$ mESCs can differentiate and give rise to endothelial and hematopoietic progenitors when differentiated on OP9 stromal cells (Hidaka et al, 1999; Schuh et al., 1999) and as shown by us in the previous section with Flk1$^{KI/KI}$ cells, we wondered if *Flk1*$^{KI/KI}$ embryos could have those cells. However, we did not detect Icam2+ endothelial progenitor cells in *Flk1*$^{KI/KI}$ E9.5 embryos. Therefore we decided to conditionally deleted Flk1 with Tie2-Cre mice in endothelial progenitor cells. Tie2-Cre is expressed after the commitment of mesoderm into angioblasts and hemogenic endothelium, a little later than the start of Flk1 expression (Lancerin, C et al., 2009; Goode, D. K. et al., 2016). We thought that this would allow the early progenitor population to differentiate and expand giving us more endothelial cells for our analysis. Indeed we have detected Icam2+ ECs in *Flk1*$^{KI/flox}$ *Tie2-Cre iSuRe-Cre* embryos, but in very low amounts. This is certainly due to the important role of *VEGF/Flk1* signaling in expansion of both arterial, capillary and venous endothelial cells, once they are differentiated (Kataoka et al., 2011; Rasmussen et al., 2013; Wythe et al., 2013).

Surprisingly, we have found that the conditional *Flk1-null Flk1-Cerulean+ Icam2+* ECs express endothelial and hematopoietic genes at comparable levels to *Flk1* heterozygous cells, mantaining their endothelial status. Therefore these cells can differentiate independently of the *Flk1* function but upregulate molecular pathways involved in apoptosis and downregulate pathways related to cellular proliferation and protein biosynthesis. *In vitro Flk1*$^{KI/KI}$ differentiated endothelial cells displayed impaired proliferation, but not apoptosis. This suggests that *in vivo*, the existent

growth factors and environmental cues, are not suficient to induce alternative fates in *Flk1*$^{KI/flox}$ *Tie2-Cre iSuRe-Cre* cells, and they instead undergo apoptosis.

The apoptosis of *Flk1- null* endothelial cells *in vivo* is in line with the recent study in zebrafish where they have shown that early inhibition of Vegf signaling in embryo led to apoptosis in endothelial cells due to reduction of Etv2 (Chetty, S. A. Et al., 2017).

In contrast to the severe endothelial defects after the loss of Flk1, it has also been shown that *Flk1*$^{-/-}$ ESCs and *Flk1*$^{-/-}$ E7.5 embryos are able to generate blood precursors (Hidaka et al, 1999; Schuh et al., 1999). Indeed we also observed an enrichment of a differentiating megakaryocyte (myeloid) signature in the *Flk1 null* mutant ECs *in vivo* suggesting they are able to differentiate into primitive haematopoietic precursors.

Genome wide loss-of-function genetic screen identifies mutations in regulators of apoptosis and cell-cycle that can favor endothelial expansion or survival in the absence of the *Flk1* function

We conducted a whole genome genetic screen based on CRISPR/Cas9 (Doench et al., 2016) to see if there are any unidentified genes that functionally interact with *Flk1* thereby rescuing its severe loss of function phentype and giving growth advantage to the endothelial progenitors with full loss of *Flk1* function.

Whole genome screens are very complex in their nature, and we had to carefully optimize several parameters before performing the final experiment. This involved a lot of trial and error, to achieve the desired infection rates and cellular amounts. In the case of our experiment, the genetic screen was significantly more difficult than usual due to the strong impact of *Flk1* loss-of-function in the development of the desired endothelial cells.

At the end, and even with lower than optimal number of cells, we were able to conduct our whole genome genetic screen in a relatively successful manner. Due to

problems in lentiviral library amplification or sequencing, we obtained less reads than initially desirable or expected. Still it was enough to perform our analysis. For positive selection screens, library coverage and read depth can be reduced to 50 to 100-fold cellular representation since only a small cell population is expected to survive (Chan et al., 2019).

The comparison of our *Flk1*$^{KI/KI}$ endothelial sample with the respective infection and mESCs controls revealed an enrichment of mutations in genes controlling apoptosis. These genes code for proteins involved in proteasome, mitochondria and ribosomes function, and have been shown to modulate pro or anti-apoptosis regulators (Gupta et al., 2018; Friedman et al., 2004; Wang et al., 2001; Zhou X., 2015; Golomb L, et al., 2014; Eberhard et al., 2013). It is interesting to see that mutations in genes involved in apoptosis will give these growth stunted endothelial population some competitive advantage, not necessarily through proliferation, but rather survival.

The pooled genomic screen relies on hits that elicit a strong differential growth or survival of cells in the period after library infection. The faster the cells divide and the more cells obtained, the easier is to quantify changes in the distribution of the mutations. That is likely why the essential known genes, important for cellular proliferation, were well identified in the mESCs control population and not in the endothelial population screen. In mESCs we could identify known regulators of stem cell biology (Hart et al., 2015). Perharps we should have extended the growth and differentiation of *Flk1*$^{KI/KI}$ endothelial cells for more days to obtain a larger list of differentially represented mutations, since these cells proliferate very slowly, even with mutations that partially rescue the *Flk1* loss of function.

Whole genome screens are a powerful strategy to identify in a unbiased way new genes or mechanisms that are relevant for a given biological process, in this case for survival and proliferation of cells in the absence of the *Flk1* function. However, these assays only reveal potential hits that need to be later validated by other assays. During the period of this thesis work, I did not have time to follow up with the validation of the hits and prove that indeed the loss of some of them rescues the function of *Flk1* during endothelial differentiation, expansion or survival. This

analysis may confirm that some of the identified genes are indeed capable of modulating or partially substituting the role of *Flk1* in endothelial biology.

Conclusions

1) The new *Flk1-H2B-Cerulean-2A-CreERT2* knock-in reporter enables the separation of two distinct cell populations; cells with a clear endothelial phenotype (Cerulean+ Icam2+) and cells with a mixed, hematopoietic or mesenchymal phenotype $Flk1^{KI/WT}$(Cerulean- Icam2+).

2) Full loss of *Flk1* function *in vitro* is compatible with endothelial-hematopoietic differentiation and proliferation, but not endothelial sprouting.

3) Full loss of *Flk1 in vitro* induces endothelial to mesenchymal transition.

4) Full loss of *Flk1 in vitro* induces alternative cell fates, such as cardiac and mesenchymal lineages.

5) *In vivo* endothelial cells with full loss of *Flk1* function do not undergo alternative developmental pathways and instead upregulated pathways related to apoptosis.

6) Genome wide loss of function genetic screens identified new genes involved in the regulation of cell proliferation, differentiation and apoptosis of Flk1 null cells.

7) Our results may shed light on molecular mechanisms that bypass the requirement of *Flk1* for endothelial expansion and survival.

Conclusiones:

1) El nuevo reportero knock-in Flk1-H2B-Cerulean-2A-CreERT2 permite la separación de dos poblaciones celulares distintas; células con un fenotipo endotelial claro (Cerulean + Icam2 +) y células con un fenotipo mixto, hematopoyético o mesenquimatoso Flk1KI / WT (Cerulean-Icam2 +).

2) La pérdida total de la función Flk1 *in vitro* es compatible con la diferenciación y proliferación endotelial-hematopoyética, pero no con el brote endotelial.

3) La pérdida total de *Flk1 in vitro* induce la transición endotelial a mesenquimal.

4) La pérdida total de *Flk1 in vitro* induce destinos celulares alternativos, como los linajes cardíacos y mesenquimales.

5) Las células endoteliales *in vivo* con pérdida total de la función Flk1 no experimentan vías de desarrollo alternativas y en su lugar aumentan las vías relacionadas con la apoptosis.

6) Los ensayosde pérdida de función identificaron nuevos genes involucrados en la regulación de la proliferación celular, diferenciación y apoptosis de las células sin *Flk1*.

7) Nuestros resultados pueden arrojar luz sobre los mecanismos moleculares que permiten la expansión y supervivencia de las células endoteliales sin *Flk1*.

References

Adams, W.J., Zhang, Y., Cloutier, J., Kuchimanchi, P., Newton, G., Sehrawat, S., Aird, W.C., Mayadas, T.N., Luscinskas, F.W., & García-Cardeña, G. 2013. Functional vascular endothelium derived from human induced pluripotent stem cells. *Stem Cell Reports, 1*(2), 105–13. doi: 10.1016/j.stemcr.2013.06.007.

Ahmed, T., Tsuji-Tamura, K., & Ogawa, M. 2016. CXCR4 signaling negatively modulates the bipotential state of hemogenic endothelial cells derived from embryonic stem cells by attenuating the endothelial potential. *Stem Cells, 34*(12), 2814–2824. doi: 10.1002/stem.2441.

Bahary, N., Goishi, K., Stuckenholz, C., Weber, G., Leblanc, J., Schafer, C.A, Berman, S.S, Klagsbrun, M., & Zon, L.I. 2007. Duplicate VegfA genes and orthologues of the KDR receptor tyrosine kinase family mediate vascular development in the zebrafish. *Blood, 110*(10), 3627–36. doi: 10.1182/blood-2006-04-016378.

Benedito, R., Rocha, S.F., Woeste, M., Zamykal, M., Radtke, F., Casanovas, O., Duarte, A., Pytowski, B., & Adams, R. H. 2012. Notch-dependent VEGFR3 upregulation allows angiogenesis without VEGF-VEGFR2 signalling. *Nature*, *484*, 110-4. doi: 10.1038/nature10908.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y., & Traver, D. 2010. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature*, *464*(7285), 108–111. doi: 10.1038/nature08738.

Boisset, J.C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., & Robin, C. 2010. *In vivo* imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature*, *464*(7285), 116–120. doi: 10.1038/nature08764.

Carmeliet, P., Ferreira, V., Breier, G., Pollefeyt, S., Kieckens, L., Gertsenstein, M., Fahriq, M., Vandenhoeck, A., Harpal, K., Eberhardt, C., Declercg, C., Pawling, J., Moons, L., Collen, D., Risau, W., & Nagy, A. 1996. Abnormal blood vessel

development and lethality in embryos lacking a single VEGF allele. *Nature*, *380*(6573), 435–439. doi: 10.1038/380435a0.

Chan, S.S., Shi, X., Toyama, A., Arpke, R.W., Dandapat, A., Iacovino, M., Kang, J., Le, G., Hagen, H.R., Garry, D. J., & Kyba, M. 2013. Mesp1 patterns mesoderm into cardiac, hematopoietic, or skeletal myogenic progenitors in a context-dependent manner. *Cell Stem Cell*, *12*(5), 587–601. doi: 10.1016/j.stem.2013.03.004.

Chetty, S. C., Rost, M. S., Enriquez, J. R., Schumacher, J. A., Baltrunaite, K., Rossi, A., Stainier, D.Y. R., & Sumanas, S. 2017. Vegf signaling promotes vascular endothelial differentiation by modulating *etv2* expression. *Dev. Biol., 424*(2), 147-161. doi: 10.1016/j.ydbio.2017.03.005.

Choi, K. 2002. The hemangioblast: a common progenitor of hematopoietic and endothelial cells. *J Hematother Stem Cell Res., 11*(1), 91–101. doi: 10.1089/152581602753448568.

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., & Keller, G. 1998. A common precursor for hematopoietic and endothelial cells. *Development, 125*(4), 725–732.

Chung, Y.S., Zhang, W.J., Arentson, E., Kingsley, P.D., Palis, J., & Choi, K. (2002). Lineage analysis of the hemangioblast as defined by FLK1 and SCL expression. *Development, 129*(23), 5511–20. doi:10.1242/dev.00149.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., & Zhang, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science, 339*(6121), 819-23. doi: 10.1126/science.1231143.

Covassin, L.D., Villefranc, J.A., Kacergis, M.C., Weinstein, B.M., & Lawson, N.D. 2006. Distinct genetic interactions between multiple Vegf receptors are required for development of different blood vessel types in zebrafish. *Proc. Natl. Acad. Sci. U. S. A., 103*(17), 6554–9. doi: 10.1073/pnas.0506886103.

Daftuar, L., Zhu, Y., Jacq, X., & Prives, C. 2013. Ribosomal Proteins RPL37, RPS15 and RPS20 Regulate the Mdm2-p53-MdmX Network . *PLoS One,8*(7), e68667. doi: 10.1371/journal.pone.0068667.

Darabi, R., Gehlbach, K., Bachoo, R.M., Kamath, S., Osawa, M., Kamm, K. E., Kyba, M., & Perlingeiro, R.C. 2008. Functional skeletal muscle regeneration from differentiating embryonic stem cells. *Nat. Med., 14*(2), 134–143. doi: 10.1038/nm1705.

de Bruijn, M.F., Speck, N.A., Peeters, M.C., & Dzierzak, E. 2000. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J., 19*(11), 2465–74. doi: 10.1093/emboj/19.11.2465.

de Peppo, G.M., Svensson, S., Lennerås, M., Synnergren, J., Stenberg, J., Strehl, R., Hyllner, J., Thomsen, P., & Karlsson, C. 2010. Human embryonic mesodermal progenitors highly resemble human mesenchymal stem cells and display high potential for tissue engineering applications. *Tissue Eng. Part A., 16*(7), 2161–82. doi: 10.1089/ten.TEA.2009.0629.

De Val, S., & Black, B.L. 2009. Transcriptional control of endothelial cell development. *Dev Cell, 16*(2), 180–195. doi: 10.1016/j.devcel.2009.01.014.

De Val, S., Chi, N.C., Meadows, S.M., Minovitsky, S., Anderson, J.P., Harris, I.S., Ehlers, M.L., Agarwal, P., Visel, A., Xu, S.M., Pennacchio, L.A., Dubchak, I., Krieg, P.A., Stainier, D.Y., & Black, B.L. 2008. Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell, 135*(6), 1053–1064. doi: 10.1016/j.cell.2008.10.049.

Dejana, E., Hirschi, K.K., & Simons, M. 2017. The molecular basis of endothelial cell plasticity. *Nat Commun., 8*, 14361. doi: 10.1038/ncomms14361.

Ding, G., Tanaka, Y., Hayashi, M., Nishikawa, S.I., & Kataoka, H. 2013. PDGF receptor alphaþ mesoderm contributes to endothelial and hematopoietic cells in mice. *Dev. Dyn.,* 242(3), 254–268. doi: 10.1002/dvdy.23923.

Ditadi, A., Sturgeon, C.M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A.D., Azzola, L., Ng, E.S., Stanley, E.G., French, D.L., Cheng, X., Gadue, P., Speck, N.A., Elefanty, A.G., & Keller, G. 2015. Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat. Cell Biol., 17*(5), 580–591. doi: 10.1038/ncb3161.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H.W., Listgarten, J., & Root, D.E. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol., 34*(2), 184-191. doi: 10.1038/nbt.3437.

Doss, M.X., Gaspar, J.A., Winkler, J., Hescheler, J., Schulz. H., & Sachinidis, A. 2012. Specific gene signatures and pathways in mesodermal cells and their derivatives derived from embryonic stem cells. *Stem Cell Rev. Rep., 8*(1), 43–54. doi: 10.1007/s12015-011-9263-5.

Eberhard, R., Stergiou, L., Hofmann, E.R., Hofmann, J., Haenni, S., Teo, Y., Furger, A., & Hengartner , M.O. 2013. Ribosome Synthesis and MAPK Activity Modulate Ionizing Radiation-Induced Germ Cell Apoptosis in *Caenorhabditis elegans. PLoS Genet., 9*(11), e1003943. doi: 10.1371/journal.pgen.1003943.

Eilken, H.M., Nishikawa, S., & Schroeder, T. 2009. Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature, 457*(7231), 896–900. doi: 10.1038/nature07760.

Eliades, A., Wareing, S., Marinopoulou, E., Fadlullah, M.Z.H., Patel, R., Grabarek, J.B., Plusa, B., Lacaud, G., & Kouskoff, V. 2016. The hemogenic competence of endothelial progenitors is restricted by Runx1 silencing during embryonic development. *Cell Reports, 15*(10), 2185–2199. doi: 10.1016/j.celrep.2016.05.001.

Ema, M., Faloon, P., Zhang, W.J., Hirashima, M., Reid, T., Stanford, W.L., Orkin, S., Choi, K., & Rossant, J. 2003. Combinatorial effects of Flk1 and Tal1 on vascular and hematopoietic development in the mouse. *Genes Dev,. 17*(3), 380–93. doi: 10.1101/gad.1049803.

Ema, M., Takahashi, S., & Rossant, J. 2006. Deletion of the selection cassette, but not cis-acting elements, in targeted Flk1-lacZ allele reveals Flk1 expression in multipotent mesodermal progenitors. *Blood, 107(1),* 111–7. doi: 10.1182/blood-2005-05-1970.

Faloon, P., Arentson, E., Kazarov, A., Deng, C.X., Porcher, C., Orkin, S., & Choi, K. 2000. Basic fibroblast growth factor positively regulates hematopoietic development. *Development, 127*(9), 1931–1941.

Fehling, H. J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., & Kouskoff, V. 2003. Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development, 130*(17), 4217–4227. doi:10.1242/dev.00589.

Ferdous, A., Caprioli, A., Iacovino, M., Martin, C.M., Morris, J., Richardson, J.A., Latif, S., Hammer, R.E., Harvey, R.P., Olson, E.N., Kyba, M., & Garry, D.J. 2009. Nkx2-5 transactivates the Ets-related protein 71 gene and specifies an endothelial/endocardial fate in the developing embryo. *Proc. Natl. Acad. Sci. U. S. A., 106*(3), 814–19. doi: 10.1073/pnas.0807583106.

Ferrara, N., Carver-Moore, K., Chen, H., Dowd, M., Lu, L., O'Shea, K.S., Powell-Braxton, L., Hillan, K.J., & Moore, M.W.1996. Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature, 380*(6573), 439–442. doi:10.1038/380439a0.

Flamme, I., Frölich, T., & Risau, W. 1997. Molecular mechanisms of vasculogenesis and embryonic angiogenesis. *J. Cell. Physiol., 173*(2), 206–10. doi: 10.1002/(SICI)1097-4652(199711)173:2<206::AID-JCP22>3.0.CO;2-C.

Friedman, J., & Xue, D. 2004. To Live or Die by the Sword: The Regulation of Apoptosis by the Proteasome. *Dev Cell., 6*(4), 460-1.

Fujimoto, H., & Yanagisawa, K.O. 1984. Defects in the archenteron of mouse embryos homozygous for the T-mutation. *Differentiation, 25*(1-3), 44–47. doi: 10.1111/j.1432-0436.1984.tb01336.x.

Galic, Z., Kitchen, S.G., Kacena, A., Subramanian, A., Burke, B., Cortado, R., & Zack, J.A.

2006. T lineage differentiation from human embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A., 103*(31), 11742–7. doi: 10.1073/pnas.0604244103.

Gilbert SF. Developmental Biology. 6th edition. Sunderland (MA): Sinauer Associates; 2000a. Paraxial Mesoderm: The Somites and Their Derivatives.

Gilbert SF. Developmental Biology. 6th edition. Sunderland (MA): Sinauer Associates; 2000b.

Gluecksohn-Schoenheimer S. 1944. The Development of Normal and Homozygous Brachy (T/T) Mouse Embryos in the Extraembryonic Coelom of the Chick. *Proc Natl Acad Sci U S A., 30*(6), 134–140. doi: 10.1073/pnas.30.6.134.

Golomb, L., Volarevic, S., & Oren′ M. 2014. p53 and ribosome biogenesis stress: The essentials. *FEBS Lett., 588*(16), 2571-9. doi: 10.1016/j.febslet.2014.04.014.

Goode, D. K., Obier, N., Vijayabaskar, M.S., Lie-A-Ling, M., Lilly, A.J., Hannah, R., Lichtinger, M., Batta, K., Florkowska, M., Patel, R., Challinor, M., Wallace, K., Gilmour, J., Assi, S.A., Cauchy, P., Hoogenkamp, M., Westhead, D.R., Lacaud, G., Kouskoff, V., Gottgens, B., & Bonifer, C. 2016. Dynamic gene regulatory networks drive hematopoietic specification and differentiation. *Dev. Cell, 36*(5), 572–587. doi: 10.1016/j.devcel.2016.01.024

Gupta, I., Singh, K., Varshney, N.K., & Khan, S. 2018. Delineating Crosstalk Mechanisms of the Ubiquitin Proteasome System That Regulate Apoptosis. *Front. Cell Dev. Biol., 6*, 11. doi: 10.3389/fcell.2018.00011.

Habeck, H., Odenthal , J., Walderich, B., Maischein, H., & Schulte-Merker, S. 2002. Analysis of a zebrafish VEGF receptor mutant reveals specific disruption of angiogenesis. *Curr. Biol., 12*(16), 1405–12. doi: 10.1016/s0960-9822(02)01044-8.

Harding, A., Cortez-Toledo, E., Magner, N.L., Beegle, J.R., Coleal-Bergum, D.P., Hao, D., Wang, A., Nolta, J.A., & Zhou, P. 2017. Highly efficient differentiation of endothelial cells from pluripotent stem cells requires the MAPK and the PI3K pathways. *Stem Cells*, *35*(4), 909–19. doi: 10.1002/stem.2577.

Hart, A., Melet, F., Grossfeld, P., Chien , K., Jones, C., Tunnacliffe, A., Favier, R., & Bernstein, A. 2000. Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. *Immunity, 13*(2), 167–77. doi: 10.1016/s1074-7613(00)00017-0.

Hidaka, M., Stanford, W.L., & Bernstein, A. 1999. Conditional requirement for the Flk-1 receptor in the in vitro generation of early hematopoietic cells. *Proc. Natl. Acad. Sci. U. S. A.*, *96*(13):7370–7375.doi: 10.1073/pnas.96.13.7370.

Hirata, H., Kawamata, S., Murakami, Y., Inoue, K., Nagahashi, A., Tosaka, M., Yoshimura, N., Miyamoto, Y., Iwasaki, H., Asahara, T., & Sawa, Y. 2007. Co-expression of platelet-derived growth factor receptor alpha and fetal liver kinase 1 enhances cardiogenic potential in embryonic stem cell differentiation *in vitro*. J*. Biosci. Bioeng., 103*(5):412–419. doi:10.1263/jbb.103.412.

Hirschi, K.K. 2012. Hemogenic endothelium during development and beyond. *Blood*, *119*(21), 4823–4827. doi: 10.1182/blood-2011-12-353466.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., & Melton, D.A. 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol., 26*(7), 795–7. doi: 10.1038/nbt1418.

Huber, T.L., Kouskoff, V., Fehling, H.J., Palis, J., & Keller, G. 2004. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature*, *432*(7017), 625–630. doi: 10.1038/nature03122.

Irion, S., Clarke, R.L., Luche, H., Kim, I., Morrison, S.J., Fehling, H.J., & Keller, G.M. 2010. Temporal specification of blood progenitors from mouse embryonic stem cells and induced pluripotent stem cells. *Development, 137*(17), 2829–39. doi: 10.1242/dev.042119.

Ishitobi, H., Wakamatsu, A., Liu, F., Azami, T., Hamada, M., Matsumoto, K., Kataoka, H., Kobayashi, M., Choi, K., Nishikawa, S., Takahashi, S., & Ema, M. 2011. Molecular basis for Flk1 expression in hemato-cardiovascular progenitors in the mouse. *Development, 138*(24), 5357–5368. doi: 10.1242/dev.065565.

Ismailoglu, I., Yeamans, G., Daley, G.Q., Perlingeiro, R.C., & Kyba, M. 2008. Mesodermal patterning activity of SCL. *Exp Hematol, 36*(12), 1593–1603. doi: 10.1016/j.exphem.2008.07.005.

Itatani, Y., Kawada, K., Yamamoto, T., & Sakai, Y. 2018. Resistance to Anti-Angiogenic Therapy in Cancer-Alterations to Anti-VEGF Pathway. *Int. J. Mol. Sci.*, *19***(4),** pii: E1232. doi: 10.3390/ijms19041232.

Jaffredo, T., Gautier, R., Eichmann, A., & Dieterlen-Lievre, F. 1998. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development*, 125(22), 4575–4583.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., & Charpentier, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science, 337*(6096), 816-21. doi: 10.1126/science.1225829.

Kalluri, R., & Weinberg, R. A. 2009. The basics of epithelial-mesenchymal transition. *J. Clin. Invest., 119*(6)*,* 1420–1428 . doi: 10.1172/JCI39104.

Kataoka, H., Hayashi, M., Nakagawa, R., Tanaka, Y., Izumi, N., Nishikawa, S., Jakt, M.L., Tarui, H., & Nishikawa, S. 2011. Etv2/ER71 induces vascular mesoderm from Flk1+PDGFRα+ primitive mesoderm. *Blood, 118*(26), 6975–6986. doi: 10.1182/blood-2011-05-352658.

Kataoka, H., Takakura, N., Nishikawa, S., Tsuchida, K., Kodama, H., Kunisada, T., Risau, W., Kita, T., & Nishikawa, S.I. 1997. Expressions of PDGF receptor alpha, c-Kit and Flk1 genes clustering in mouse chromosome 5 define distinct subsets of nascent mesodermal cells. *Dev. Growth Differ., 39*(6), 729–740. .

Kim, I., Saunders, T.L., & Morrison, S.J. 2007. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell, 130*(3), 470–83. doi: 10.1016/j.cell.2007.06.011.

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., & Tam, P.P. 1999. The orderly allocation of mesodermal cells to the extraembryonic structures and the antero- posterior axis during gastrulation of the mouse embryo*. Development,*

126(21), 4691–4701.

Kispert, A., Herrmann, B.G. 1994. Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Dev. Biol., 161*(1), 179–193. doi:10.1006/dbio.1994.1019.

Kissa, K., & Herbomel, P. 2010. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature*, *464*(7285), 112–115. doi: 10.1038/nature08761.

Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera Mdel, C., & Yusa, K. 2014. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol, 32*(3), 267-73. doi: 10.1038/nbt.2800.

Kouskoff,V., Lacaud, G., Schwantz, S., Fehling, H.J., & Keller, G. 2005. Sequential development of hematopoietic and cardiac mesoderm during embryonic stem cell differentiation. *Proc. Natl. Acad. Sci. U.S.A., 102*(37), 13170–13175. doi:10.1073/pnas.0501672102.

Koyano-Nakagawa, N., Kweon, J., Iacovino, M., Shi, X., Rasmussen, T.L., Borges, L., Zirbes, K.M., Li, T., Perlingeiro, R.C., Kyba, M., & Garry, D.J. 2012. Etv2 is expressed in the yolk sac hematopoietic and endothelial progenitors and regulates Lmo2 gene expression. *Stem Cells, 30*(8), 1611–1623. doi: 10.1002/stem.1131.

Koyano-Nakagawa, N., Shi, X., Rasmussen, T.L., Das, S., Walter, C.A., & Garry, D.J. 2015. Feedback mechanisms regulate Ets variant 2 (Etv2) gene expression and hematoendothelial lineages. *J. Biol. Chem*., 290(47), 28107–28119. doi: 10.1074/jbc.M115.662197.

Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., Zuyev, S., Ansell, J., & Medvinsky, A. 2002. Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development, 129*(21), 4891–9.

Kyba, M., Perlingeiro, R.C.R., & Daley, G.Q. 2002. HoxB4 confers definitive lymphoid-

myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell, 109*(1), 29–37. doi:10.1016/s0092-8674(02)00680-3.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V., & Lacaud, G. 2009.The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature*, *457*(7231), 892–895. doi: 10.1038/nature07679.

Landry, J-R., Kinston, S., Knezevic, K., de Bruijn, M.F.T.R, Wilson, N., Nottingham, W.T., Peitz, M., Edenhofer, F., Pimanda, J.E., Ottersbach, K., & Gottgens, B. 2008. Runx genes are direct targets of Scl/Tal1 in the yolk sac and fetal liver. *Blood, 111*(6), 3005–3014. doi:10.1182/blood-2007-07-098830.

Lee, D., Park, C., Lee, H., Lugus, J.J., Kim, S.H., Arentson, E., Chung, Y.S., Gomez, G., Kyba, M., Lin, S., Janknecht, R., Lim, D.S., & Choi, K. 2008. ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell Stem Cell, 2*(5), 497–507. doi: 10.1016/j.stem.2008.03.008.

Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., & Liu, X.S. 2014. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology*, *15*(12), 554. doi: 10.1186/s13059-014-0554-4.

Liao, E.C., Paw, B.H., Oates, A.C., Pratt, S.J., Postlethwait, J.H., & Zon, L.I. 1998. SCL/Tal-1 transcription factor acts downstream of cloche to specify hematopoietic and vascular progenitors in zebrafish. *Genes Dev., 12*(5), 621–6. doi:10.1101/gad.12.5.621.

Liu, F., Bhang, S.H., Arentson, E., Sawada, A., Kim, C.K., Kang, I., Yu, J., Sakurai, N., Kim, S.H., Yoo, J.J., Kim, P., Pahng, S.H., Xia, Y., Solnica-Krezel, L., & Choi, K. 2013. Enhanced hemangioblast generation and improved vascular repair and regeneration from embryonic stem cells by defined transcription factors. *Stem Cell Rep., 1*(2), 166–82. doi: 10.1016/j.stemcr.2013.06.005.

Liu, F., Kang, I., Park, C., Chang, L.W., Wang, W., Lee, D., Lim, D.S., Vittet, D., Nerbonne, J.M., & Choi, K. 2012. ER71 specifies Flk-1þ hemangiogenic mesoderm by inhibiting cardiac mesoderm and Wnt signaling. *Blood*, *119*(14), 3295–3305. doi:

10.1182/blood-2012-01-403766.

Liu, F., Li, D., Yu, Y.Y., Kang, I., Cha, M.J., Kim, J.Y., Park, C., Watson, D.K., Wang, T., & Choi, K. 2015. Induction of hematopoietic and endothelial cell program orchestrated by ETS transcription factor ER71/ETV2. *EMBO Rep., 16*(5)*,* 654–669. doi: 10.15252/embr.201439939.

Lugus, J.J., Chung, Y.S., Mills, J.C., Kim, S.I., Grass, J., Kyba, M., Doherty, J.M., Bresnick, E.H., & Choi, K. 2007. GATA2 functions at multiple steps in hemangioblast development and differentiation. *Development, 134*(2), 393–405. doi: 10.1242/dev.02731.

Ma, Z., Chen, C., Tang, P., Zhang, H., Yue, J., & Yu, Z. 2017. BNIP3 induces apoptosis and protective autophagy under hypoxia in esophageal squamous cell carcinoma cell lines: BNIP3 regulates cell death. *Dis Esophagus*, *30*(9), 1-8. doi: 10.1093/dote/dox059.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, .JE., & Church, G.M. 2013. RNA-guided human genome engineering via Cas9. *Science, 339*(6121), 823-6. doi: 10.1126/science.1232033.

Margariti, A., Winkler , B., Karamariti, E., Zampetaki, A., Tsai, T. N., Baban, D., Ragoussis, J., Huang, Y., Han, J-DJ., Zeng, L., Hu, Y., & Xu, Q. 2012. Direct reprogramming of fibroblasts into endothelial cells capable of angiogenesis and reendothelialization in tissue-engineered vessels. *Proc. Natl. Acad. Sci. U. S. A., 109* (34), 13793–8. doi: 10.1073/pnas.1205526109.

Meetei, A.R., de Winter, J.P., Medhurst, A.L., Wallisch, M., Waisfisz, Q., van de Vrugt, H.J., Oostra, A.B., Yan, Z., Ling, C., Bishop, C.E., Hoatlin, M.E., Joenje, H., & Wang, W. 2003. A novel ubiquitin ligase is deficient in Fanconi anemia. *Nat Genet., 35*(2), 165-70. doi: 10.1038/ng1241.

Mikkola, H.K.A, Fujiwara, Y., Schlaeger, T.M., Traver, D., & Orkin, S.H. 2003. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood,*

101(2), 508–16. doi:10.1182/blood-2002-06-1699.

Millauer, B., Wizigmann-Voos, S., Schnürch, H., Martinez, R., Møller, N.P., Risau, W., & Ullrich, A. 1993. High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of vasculogenesis and angiogenesis. *Cell*, *72*(6), 835–846. doi:10.1016/0092-8674(93)90573-9.

Morita, R., Suzuki, M., Kasahara, H., Shimizu, N., Shichita, T., Sekiya, T., Kimura, A., Sasaki, K., Yasukawa, H., & Yoshimura, A. 2015. ETS transcription factor ETV2 directly converts human fibroblasts into functional endothelial cells. *Proc. Natl. Acad. Sci. U. S. A., 112*(1), 160–5. doi: 10.1073/pnas.1413234112.

Motoike, T., Markham, D.W., Rossant, J., & Sato, T.N. 2003. Evidence for novel fate of Flk1+ progenitor: contribution to muscle lineage. *Genesis, 35*(3):153–159. doi: 10.1002/gene.10175.

Murray, P.D.F. 1932. The development *in vitro* of the blood of the early chick embryo. *Proc R Soc Lond B Biol Sci., 111*(773), 497–521. doi: 10.1098/rspb.1932.0070.

Nakano, H., Liu, X., Arshi, A., Nakashima, Y., van Handel, B., Sasidharan, R., Harmon, A.W., Shin, J.H., Schwartz, R.J., Conway, S.J., Harvey, R.P., Pashmforoush, M., Mikkola, H.K., & Nakano, A. 2013. Haemogenic endocardium contributes to transient definitive haematopoiesis. *Nat. Comms. 4*, 1564. doi: 10.1038/ncomms2569.

Nishikawa, S.I., Nishikawa, S., Hirashima, M., Matsuyoshi, N., & Kodama, H. 1998. Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. *Development, 125*(9), 1747–1757.

Niu, Z., Iyer, D., Conway, S.J., Martin, J.F., Ivey, K., Srivastava, D., Nordheim, A., Schwartz, R.J. 2008. Serum response factor orchestrates nascent sarcomerogenesis and silences the biomineralization gene program in the heart. *Proc. Natl. Acad. Sci. U.*

S. A., 105(46), 17824–9. doi: 10.1073/pnas.0805491105.

Padron-Barthe, L., Temino, S., Villa Del Campo, C., Carramolino, L., Isern, J., & Torres, M. 2014. Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo. *Blood, 124*(16), 2523–2532. doi: 10.1182/blood-2013-12-545939.

Pimanda, J. E., Ottersbach, K., Knezevic, K., Kinston, S., Chan, W.Y., Wilson, N.K., Landry, J.R., Wood, A.D., Kolb-Kokocinski, A., Green, A.R., Tannahill, D., Lacaud, G., Kouskoff, V., & Gottgens, B.2007. Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc. Natl Acad. Sci. USA., 104*(45), 17692–17697. doi: 10.1073/pnas.0707045104.

Poole, T.J., & Coffin, J.D. 1989. Vasculogenesis and angiogenesis: two distinct morphogenetic mechanisms establish embryonic vascular pattern. *J. Exp. Zool., 251*(2)*,* 224–31. doi: 10.1002/jez.1402510210.

Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F.W., & Orkin, S.H. (1996). The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell, 86*(1)*,* 47–57. doi: 10.1016/s0092-8674(00)80076-8.

Rasmussen, T.L., Kweon, J., Diekmann, M.A., Belema-Bedada, F., Song, Q., Bowlin, K., Shi, X., Ferdous, A., Li, T., Kyba, M., Metzger, J.M., Koyano-Nakagawa, N., & Garry, D.J. 2011. ER71 directs mesodermal fate decisions during embryogenesis. *Development, 138*(21), 4801–4812. doi: 10.1242/dev.070912.

Rasmussen, T.L., Martin, C.M., Walter, C.A., Shi, X., Perlingeiro, R., Koyano-Nakagawa, N., & Garry, D.J. 2013. Etv2 rescues Flk1 mutant embryoid bodies. *Genesis, 51*(7), 471–480. doi: 10.1002/dvg.22396.

Rasmussen, T.L., Shi, X., Wallis, A., Kweon, J., Zirbes, K.M., Koyano-Nakagawa, N., & Garry, D.J. 2012. VEGF/Flk1 signaling cascade transactivates Etv2 gene expression. *PLoS One, 7*(11), e50103. doi: 10.1371/journal.pone.0050103.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., & Bongso, A. 2000. Embryonic stem cell lines from human blastocysts: somatic differentiation *in vitro*. *Nat.*

Biotechnol. 18(4), 399–404. doi: 10.1038/74447.

Rhodes, K.E., Gekas, C., Wang, Y., Lux, C.T., Francis, C.S., Chan, D.N., Conway, S., Orkin, S.H., Yoder, M.C., & Mikkola, H.K.A. 2008. The emergence of hematopoietic stem cells is initiated in the placental vasculature in the absence of circulation. *Cell Stem Cell, 2*(3), 252–63. doi: 10.1016/j.stem.2008.01.001.

Ricard, N., Scott, R.P., Booth, C.J., Velazquez, H., Cilfone, N.A., Baylon, J.L., Gulcher, J.R., Quaggin, S.E., Chittenden, T.W., & Simons, M. 2019. Endothelial ERK1/2 signaling maintains integrity of the quiescent endothelium. *Journal of Experimental Medicine*, *216* (8), 1874. doi: 10.1084/jem.20182151.

Risau, W. 1997. Mechanisms of angiogenesis. *Nature, 386*(6626)*,* 671-4. doi: 10.1038/386671a0.

Robb, L., Lyons, I., Li, R., Hartley, L., Köntgen, F., Harvey, R.P., Metcalf, D., & Begley, CG. 1995. Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proc. Natl. Acad. Sci. U. S. A., 92*(15), 7075–9. doi:10.1073/pnas.92.15.7075.

Sabin, F.R. 1920. Studies on the origin of blood vessels and of red corpuscles as seen in the living blastoderm of the chick during the second day of incubation. *Contributions to Embryology, 9*, 213–262.

Saga, Y., Hata, N., Kobayashi, S., Magnuson, T., Seldin, M.F., & Taketo, M.M. 1996. MesP1: a novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. *Development, 122*(9), 2769–2778.

Saga, Y., Miyagawa-Tomita, S., Takagi, A., Kitajima, S., Miyazaki, J., & Inoue, T. 1999. Mesp1 is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development, 126*(15), 3437–3447.

Saiti, D., & Lacham-Kaplan, O. 2007. Mouse germ cell development *in-vivo* and *in-vitro*. *Biomarker Insights*, *2*, 241–252.

Sakurai, H., Era, T., Jakt, L.M., Okada, M., Nakai, S., Nishikawa, S., & Nishikawa, S. 2006. *In vitro* modeling of paraxial and lateral mesoderm differentiation reveals early

reversibility. *Stem Cells, 24*(3), 575–586. doi:10.1634/stemcells.2005-0256.

Schuh, A. C., Faloon, P., Hu, Q. L., Bhimani, M., & Choi, K. 1999. *In vitro* hematopoietic and endothelial potential of flk-1(-/-) embryonic stem cells and embryos. *Proc. Natl Acad. Sci. USA., 96*(5), 2159–2164. doi: 10.1073/pnas.96.5.2159.

Shalaby, F., Ho, J., Stanford, W.L., Fischer, K.D., Schuh, A.C., Schwartz, L., Bernstein, A., Rossant, J. 1997. A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis. *Cell*. *89*(6):981–990. doi: 10.1016/s0092-8674(00)80283-4.

Shalaby, F., Rossant, J., Yamaguchi, T.P., Gertsenstein, M., Wu, X.F., Breitman, M.L., Schuh, A.C. 1995. Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature, 376*(6535), 62–6. doi: 10.1038/376062a0.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., Zhang, F. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science, 343*(6166), 84-87. doi: 10.1126/science.1247005.

Shi, X., Richard, J., Zirbes, K.M., Gong, W., Lin, G., Kyba, M., Thomson, J.A., Koyano-Nakagawa, N., & Garry, D.J. 2014. Cooperative interaction of Etv2 and Gata2 regulates the development of endothelial and hematopoietic lineages. *Developmental Biology, 389* (2), 208–218. doi: 10.1016/j.ydbio.2014.02.018.

Shi, X., Zirbes, K.M., Rasmussen, T.L., Ferdous, A., Garry, M.G., Koyano-Nakagawa, N., & Garry, D.J. 2015.The transcription factor Mesp1 interacts with cAMP-responsive element binding protein 1 (Creb1) and coactivates Ets variant 2 (Etv2) gene expression. *J. Biol. Chem., 290*(15), 9614–9625. doi: 10.1074/jbc.M114.614628.

Shibuya, M., & Claesson-Welsh, L. 2006. Signal transduction by VEGF receptors in regulation of angiogenesis and lymphangiogenesis. *Exp Cell Res., 312*(5), 549-60. doi: 10.1016/j.yexcr.2005.11.012.

Shivdasani, R.A., Mayer, E.L., & Orkin, S.H. 1995. Absence of blood formation in mice lacking the T- cell leukaemia oncoprotein tal-1/SCL. *Nature, 373*(6513), 432–4. doi: 10.1038/373432a0.

Simons, M., Gordon, E., & Claesson-Welsh, L. 2016. Mechanisms and regulation of endothelial VEGF receptor signalling. *Nat. Rev. Mol. Cell Biol., 17, 611–625.* doi:10.1038/nrm.2016.87.

Stainier, D.Y., Weinstein, B.M., Detrich, H.W. 3rd, Zon, L.I., & Fishman, M.C. 1995. Cloche, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. *Development, 121*(10), 3141–50.

Stevens, L.C. 1960. Embryonic potency of embryoid bodies derived from a transplantable testicular teratoma of the mouse. *Dev. Biol. 2*, 285–97. doi:10.1016/0012-1606(60)90010-5.

Sumanas, S., Gomez, G., Zhao, Y., Park, C., Choi, K., & Lin, S. 2008. Interplay among Etsrp/ER71, Scl, and Alk8 signaling controls endothelial and myeloid cell formation. *Blood, 111*(9), 4500–10. doi: 10.1182/blood-2007-09-110569.

Swiers, G., Baumann, C., O'rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K.D., Yoshimoto, M., Yoder, M.C., Frampton, J., Schroeder, T., Enver, T., Gottgens, B., & de Bruijn, M.F.T.R. 2013. Early dynamic fate changes in haemogenic endo- thelium characterized at the single-cell level. *Nat Comms, 4*, 2924. doi: 10.1038/ncomms3924.

Takahashi, K., & Yamanaka, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell, 126*(4), 663–76. doi:10.1016/j.cell.2006.07.024.

Takakura, N., Yoshida, H., Ogura, Y., Kataoka, H., Nishikawa, S., & Nishikawa, S. 1997. PDGFR alpha expression during mouse embryogenesis: immunolocalization analyzed by whole-mount immunohistostaining using the monoclonal anti-mouse PDGFR alpha antibody APA5. *J. Histochem. Cytochem., 45*(6), 883–893. doi: 10.1177/002215549704500613.

Tornack, J., Seiler, K., Grützkau, A., Grun, J.R., Onodera, M., Melchers, F., & Tsuneto, M. 2013. Ectopic Runx1 expression rescues Tal-1-deficiency in the generation of

primitive and definitive hematopoiesis," *PLoS One, 8*(7), e70116. doi: 10.1371/journal.pone.0070116.

Tsai, F.Y., Keller, G., Kuo, F.C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F.W., & Orkin, S.H. 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature, 371*(6494), 221–226. doi: 10.1038/371221a0.

Ueda, H., Akiyama, Y., Shimada, S., Mogushi, K., Serizawa, M., Matsumura, S., Mitsunori, Y., Aihara, A., Ban, D., Ochiai, T., Kudo, A., Tanabe, M., Tanaka, S. 2018. Tumor suppressor functions of DAXX through histone H3.3/H3K9me3 pathway in pancreatic NETs. *Endocr. Relat. Cancer, 25*(6), 619-631. doi: 10.1530/ERC-17-0328.

Ueno, H., & Weissman, I.L. 2006. Clonal analysis of mouse development reveals a polyclonal origin for yolk sac blood islands. *Dev Cell, 11*(4), 519–533. doi: 10.1016/j.devcel.2006.08.001.

Visvader, J.E., Fujiwara, Y., & Orkin, S.H. 1998. Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development. *Genes & Development*, *12*(4), 473–479. doi: 10.1101/gad.12.4.473.

Vodyanik, M.A., Yu, J., Zhang, X., Tian, S., Stewart, R., Thomson, J.A., & Slukvin, I.I. 2010. A mesoderm-derived precursor for mesenchymal stem and endothelial cells. *Cell Stem Cell, 7*(6), 718–729. doi: 10.1016/j.stem.2010.11.011.

Wang, Y., Miao, X., Liu, F., Li, F., Liu, Q., Sun, J., & Cai, L. 2014. Dysregulation of histone acetyltransferases and deacetylases in cardiovascular diseases. *Oxid. Med. Cell Longev.*, 2014, 641979. doi: 10.1155/2014/641979.

Wang, Y., Yates, F., Naveiras, O., Ernst, P., & Daley, G.Q. 2005. Embryonic stem cell-derived hematopoietic stem cells. *Proc. Natl. Acad. Sci. U. S. A., 102*(52), 19081–6. doi:10.1073/pnas.0506127102.

Wareing, S., Eliades, A., Lacaud, G., & Kouskoff ,V. 2012. ETV2 expression marks blood and endothelium precursors, including hemogenic endothelium, at the onset of blood development. *Dev Dyn, 241*(9), 1454–1464. doi: 10.1002/dvdy.23825.

Wareing, S., Mazan, A., Pearson, S., Gottgens, B., Lacaud, G., & Kouskoff, V. 2012. The Flk1-Cre- mediated deletion of ETV2 defines its narrow temporal requirement during embry- onic hematopoietic development. *Stem Cells, 30*(7), 1521–1531. doi: 10.1002/stem.1115.

Wesemann, D.R., Portuguese, A.J., Magee, J.M., Gallagher, M.P., Zhou, X., Panchakshari, R.A., & Alt. F.W. 2012. Reprogramming IgH isotype-switched B cells to functional-grade induced pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A. 109*(34), 13745–50. doi: 10.1073/pnas.1210286109.

Wilkinson, D.G., Bhatt, S., & Herrmann, B.G. 1990. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature, 343(6259)*, 657–659. doi: 10.1038/343657a0.

Wilson, V., Manson, L., Skarnes, W.C., & Beddington, R.S. 1995. The T gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development*, *121*(3), 877–886.

Xiaodong Wang, X. 2001. The expanding role of mitochondria in apoptosis. *Genes Dev., 15*(22), 2922-33.

Xu, C.X., Lee, T.J., Sakurai, N., Krchma, K., Liu, F., Li, D., Wang, T., & Choi, K. 2017. ETV2/ER71 regulates hematopoietic regeneration by promoting hematopoietic stem cell proliferation. *The Journal of Experimental Medicine*, *214*(6), 1643–1653. doi: 10.1084/jem.20160923.

Xu, J., Liu, X., Jiang, Y., Chu, L., Hao, H., Liua, Z., Verfaillie, C., Zweier, J., Gupta, K., & Liu, Z. 2008. MAPK/ERK signalling mediates VEGF-induced bone marrow stem cell differentiation into endothelial cell. *J. Cell Mol. Med., 12*(6a), 2395–406. doi: 10.1111/j.1582-4934.2008.00266.x.

Yamaguchi, T.P., Dumont, D.J., Conlon, R.A., Breitman, M.L., Rossant, J. 1993. flk-1, an flt-related receptor tyrosine kinase is an early marker for endothelial cell precursors. *Development, 118*(2), 489–498.

Yamamizu, K., Matsunaga, T., Katayama, S., Kataoka, H., Takayama, N., Eto, K., Nishikawa, S., & Yamashita, J.K. 2012. PKA/CREB signaling triggers initiation of endothelial and hematopoietic cell differentiation via Etv2 induction. *Stem Cells, 30*(4), 687–696. doi: 10.1002/stem.1041.

Yamashita, J', Itoh, H., Hirashima, M., Ogawa, M., Nishikawa, S., Yurugi, T., Naito, M., Nakao, K., & Nishikawa, S. 2000. Flk1-positive cells derived from embryonic stem cells serve as vascular progenitors. *Nature*, *408*(6808), 92–6. doi: 10.1038/35040568.

Zape, J.P., & Zovein, A.C. 2011. Hemogenic endothelium: origins, regulation, and implications for vascular biology. *Semin Cell Dev Biol*., *22*(9), 1036–1047. doi: 10.1016/j.semcdb.2011.10.003.

Zhao, H., & Choi, K. 2017. A CRISPR screen identifies genes controlling *Etv2* threshold expression in murine hemangiogenic fate commitment. *Nature Communications, 8,* 541. doi: 10.1038/s41467-017-00667-5.

Zhen, F., Lan, Y., Yan, B., Zhang, W., & Wen, Z. 2013. Hemogenic endothelium specification and hematopoietic stem cell maintenance employ distinct Scl isoforms. *Development, 140*(19)*,* 3977–3985. doi: 10.1242/dev.097071.

Zhou, H., Chen, X., Chen, L., Zhou, X., Zheng, G., Zhang, H., Huang, W., & Cai, J. 2014. Anti-fibrosis effect of scutellarin via inhibition of endothelial-mesenchymal transition on isoprenaline-induced myocardial fibrosis in rats. *Molecules, 19*, 15611–15623. doi: 10.3390/molecules191015611.

Zhou, X., Liao, W-J., Liao, J-M., Liao, P., & Lu, H. 2015. Ribosomal proteins: functions beyond the ribosome. *J Mol Cell Biol., 7*(2), 92-104. doi: 10.1093/jmcb/mjv014.

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S.,

Zanetta, L., Dejana, E., Gasson, J.C., Tallquist, M.D., & Iruela-Arispe, M.L. 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell*, *3*(6), 625–636. doi: 10.1016/j.stem.2008.09.018.

Acknowledgement

It is a moment of reminiscence that I take as an opportunity to sum up all these years that went by in the process of this work called Ph.D. These years have sketched a lasting impression on my mind and will determine my attitude in any future endeavour that I plunge myself in. No journey is undertaken alone and my journey is no different. How can I not thank these wonderful people who have not only been my constant companions but have made this journey all the more worthwhile.

Life has graciously blessed me with unwavering positivity of my mentor and guide ***Dr. Rui Benedito,*** *who always helped me in many ways rather than one. I wholeheartedly express my deepest gratitude towards him for constantly being there for motivating me to realise my full potential. His inquisitive approach, deep focus, astute knowledge and patient listening abilities have facilitated me in treading my research path with the utmost zeal. To me more than a guide he has been a source of positive vibes with a charismatic personality and a caring attitude. I would profoundly thank him for being my mentor and inspiring me and challenging me to be the best version of myself.*

He has always inspired me with his scientific prowess, meticulous approach and expertise at the eleventh hour. The patience with which he clarified my doubts and his constructive criticism has immensely helped me bring this work to completion. In spite of his hectic schedule, he made sure to spare his precious time for fruitful discussions and valuable suggestions. I would like to thank him for providing me all the necessary facilities required for the execution of my research and for giving me the opportunity of a phenomenal learning experience.

I thank all the faculty members at CNIC for their wise counsel, support and help. I sincerely thank the entire non-research staff at CNIC. I grab the opportunity to thank my previous and current lab mates at CNIC for giving me memorable moments and being a constant source of help.

Last but not the least I would thank the universe and its positive cosmic energies which united for the completion of my Ph.D. I thank everyone who in some way or the other has been a part of the kaleidoscope of my life and helped me finish this doctoral work.

Oh! And lest I forget, I would like to thank my beautiful, amazing, stunning and supremely heavenly wife, Aakrati Agarwal, my wonderful mother, my annoying Flatmates who gave me a pain in the ass on various occasions and reminded me of being alive and yes, the cat that frequents my house to lap up the milk (you see, I am a nature lover)!!

Mayank Bansal