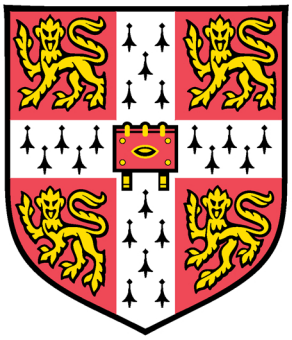
# Role of the cell cycle during human endothelial-to-haematopoietic transition



Giovanni Canu

Jesus College

Cambridge Stem Cell Institute

Department of Surgery

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

September 2018

# DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text.

It is not substantially the same as any that I have submitted or is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted or is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

As stipulated by the Degree Committee of Clinical Medicine, this dissertation does not exceed the 60,000 words in length.

Giovanni Canu, September 2018

*Intelligence without ambition is a bird without wings.*

___ *Salvador Dalí*

# ACKNOWLEDGEMENTS

In the following pages, condensed in a few thousand words, is the chronicle of a long but incredibly satisfying adventure, the final cut of a movie whose shooting has lasted for the past 4 years. Just like every good movie, only a few takes made it to the final version, leaving behind cut scenes made of hard moments which, unavoidably, laid the foundation to the final result. And just like every great movie, many are the people who made all of this possible from behind the scenes. Here I want to turn the spotlight on the backstage, and let the credits roll before the movie starts.

Enormous gratitude goes to my supervisor and mentor Prof Ludovic Vallier, whose support and guidance gave me confidence and allowed me to constantly challenge myself and overcome the difficulties of exploring uncharted lands. He was able to give me the direction, still trusting me with the freedom of following my own ideas. He represents the scientific goals I would only hope to achieve.

Sincere thanks go to Daniel, who backed me from the first steps of this path and endured with me the many hours of meetings and discussions which led to the birth of new ideas and exciting directions, contributing to shaping this dissertation.

The great environment I was blessed to work in was such thanks to the people it was made of. Carola, Will, Rodrigo, Loukia, Kaiser and the rest of the LRM family, past and present, who created the perfect mix of fun and scientific stimulation, entertainment and sometimes even fight. It was never a workplace. Plus, if there is a person without whom none of this would have been possible, this is probably Bob. Thank you Bob, wherever you are now.

Manos and Dr Ana Cvejic were extremely helpful and allowed my project to finally spread its wings, providing the tools to reach new horizons but also the intellectual challenge to look at things from a new perspective.

Infinite gratitude and respect go to the British Heart Foundation, the funding body that believed in me and gave me such a priceless opportunity. With their action, they not only

change the lives of countless patients but give also a higher purpose to incredibly bright minds among whom I was so lucky to be included. My sincere thanks also extend to Katja, who supported me throughout the length of my PhD (and beyond) and to the annual "student review team", first of all Prof Paul Riley for his precious feedbacks and warm encouragements.

Despite never having experienced the pleasure of living on site, Jesus College was for me a second home, always present when I needed. They granted the financial support to attend international conferences which further expanded my scientific perspectives, and provided the rest of a formal dinner or the joy of the best May Balls in town. Few words will always stay with me from my matriculation day. "Once a Jesuan, always a Jesuan".

Special thanks also go to the friends who stayed beside me along this path. The old ones who endured the long distance, Marco and Gianmarco, who shared with me a piece of life I could never forget. And the new ones, Maria, Gianni, Massimo, Elisa, great friends and a small piece of Italy I was able to recreate here.

All of this contributed to make my experience in Cambridge an incredible adventure. Enjoying this historical city, reaching new limits as a scientist and as a person, working and studying in one of the best universities in the world, certainly now and always the first in my heart.

Nevertheless, I believe that what we are, and the goals we reach, are always the product of the influence other people have had on us. Pushes, small or big, that contributed to sending us forwards. Some of these influences are easier to recognise. Mattia was my first mentor in Cambridge. He supported me through the shock of landing in a new country and starting everything anew, believed in me giving me confidence in my capabilities, and pushed me towards the path that led me to write this thesis. Luca was my supervisor in Italy. He gave me my first lab skills, taught me the importance of evaluating critically also things that are taken for granted, and encouraged me to take one of the biggest steps in my life coming to Cambridge.

Other pushes are more difficult to identify, because they run back many years to unsuspected times. Indeed, since I have memory, I have always had my parents support. And they somehow managed to imprint in me the need of always doing more, always wanting more,

being competitive, trying to be the best I can. And doing it while having fun. Most of the things I have, I took from them. They gave me the ambition and the tools to build my own future, the only regret being the distance I had to run for such ambition. This is their gift and their sacrifice. On top of that, they managed to make even better with Enzo, a second improved version, me with a better character. Perhaps the only person I can talk to without talking. If I am here now, a huge push definitely came from my family.

And finally Carlotta, my yesterday, today and tomorrow. My everyday push. Partner in crime on this journey we decided to make of our life. She was part of this adventure well before I even realised. She has celebrated my successes and held up my failures, lighted up the dark days and navigated me through the storm. She who sometimes is the storm. She gave me courage through these years and strength through unconditional love and trust. And most importantly, she has always believed that, whatever the obstacle, together we would have easily overcome it. She is the inspiration for this movie and for those that are yet to come.

# Role of the cell cycle during human endothelial-to-haematopoietic transition

Giovanni Canu

Haematopoietic stem cells (HSC) are responsible for the maintenance of blood homeostasis and are of fundamental importance for the treatment of a variety of life-threatening diseases. However, their scarce availability constitutes a major obstacle. Thus, understanding their embryonic specification is essential for the production of these cells *in vitro* from human pluripotent stem cells (hPSC), but also to develop new therapies. During development, HSCs first arise in the aorta-gonad-mesonephros (AGM) region of the embryo from a population of haemogenic endothelial cells lining the ventral portion of the dorsal aorta which undergo endothelial-to-haematopoietic transition (EHT). This process culminates with the generation of the first HSCs capable of multilineage differentiation and long-term engraftment. Little is known about the molecular mechanisms driving this process, especially in human where the AGM region is not easily accessible *in vivo*.

In this study, I took advantage of hPSCs and single cell transcriptomics to draw an accurate picture of this developmental stage and uncover mechanisms by which the haemogenic endothelium generates early HSCs. Of particular interest, I show that most of the endothelial cells at this stage reside in a quiescent state, with a direct correlation between cell cycle entry and their ability to progress to the haematopoietic fate. Furthermore, I identify CDK4/6 and CDK1 as key regulators affecting this process.

Ultimately, I propose here a direct link between the molecular machineries controlling cell cycle progression and cell fate decision, determining the capability of haemogenic endothelial cells to undertake the haematopoietic fate during EHT.

These results will have a major impact on the improvement of protocols for the production of functional HSCs *in vitro*, but also on the advancement of new therapies based on the culture *ex vivo* of primary HSCs.

# CONTENTS

**PART II - EXPERIMENTAL RESULTS**

| List of abbreviations and acronyms | |
|---|---|
| AGM | aorta-gonad-mesonephros |
| BM-HSC | bone marrow HSC |
| BMP | bone morphogenetic protein |
| CDK | cyclin-dependent kinase |
| CDKi | CDK inhibitors |
| CDK4/6i | inhibition of CDK4/6 by PD0332991 |
| CDK2i | inhibition of CDK2 by Roscovitine |
| CDK1i | inhibition of CDK1 by RO3306 |
| CTRL | control |
| D | day of in vitro differentiation |
| DAPI | 4',6-Diamidino-2-Phenylindole, Dihydrochloride |
| DMSO | dimethyl sulfoxide |
| DSH | dishevelled |
| E | embryonic day |
| EB | embryoid bodies |
| EdU | 5-ethynyl-2´-deoxyuridine |
| EHT | endothelial-to-haematopoietic transition |
| EMP | erythro-myeloid progenitor |
| FACS | fluorescence activated cell sorting |
| FGF | fibroblast growth factor |
| FL-HSC | foetal liver HSC |
| G0 phase | gap 0 phase |
| G1 phase | gap 1 phase |
| G2 phase | gap 2 phase |
| GO | gene ontology |
| HE | haemogenic endothelium |
| HEC | haemogenic endothelial cells |
| hESC | human embryonic stem cells |
| hIPSC | human induced pluripotent stem cells |
| HPC | haematopoietic progenitor cells |
| hPSC | human pluripotent stem cells |
| HSC | haematopoietic stem cells |
| HSPC | haematopoietic stem/progenitor cells |
| IAHC | intra-aortic haematopoietic clusters |
| IGF | insulin-like growth factor |
| IFNγ | interferon gamma |
| LT-HSC | long-term HSC |
| M phase | mitotic phase |
| MAPK | mitogen-activated protein kinase |
| MEF | mitotically inactivated mouse embryonic fibroblasts |

| | |
|---|---|
| NICD | notch intracellular domain |
| Noc | Nocodazole |
| PBMC | peripheral blood mononuclear cells |
| PI3K | phosphatidylinositol-4,5-bisphosphate 3-kinase |
| qPCR | quantitative polymerase chain reaction |
| scRNAseq | single cell RNA sequencing |
| S phase | synthesis phase |
| SEM | standard error of the mean |
| SHH | sonic hedgehog |
| ST-HSC | short-term HSC |
| TGFβ | transforming growth factor beta |
| TNFα | tumor necrosis factor |
| tSNE | t-distributed Stochastic Neighbor Embedding |
| VEC | vascular endothelial cells |
| VEGF | vascular endothelial growth factor |
| WNT | wingless-related integration site |

# LIST OF FIGURES

# LIST OF TABLES

# PART I

# INTRODUCTION

# 1 INTRODUCTION

## 1.1 Summary

The word haematopoiesis comes from the ancient Greek *αἷμα*, for "blood", and *ποιεῖν*, which means "to make". It describes the process by which, in a healthy adult person, approximately $10^{11}$-$10^{12}$ new blood cells are produced daily (Parslow et al., 2001). This high demand of cells is essential for the maintenance of steady-state levels of circulating blood cells and is satisfied by haematopoietic stem cells, a population of adult stem cells residing in the medulla of the bone marrow and characterised by high regenerative capacity.

Prompted by the incredible potential of this population for regenerative medicine, much effort has been spent on trying to achieve their derivation *in vitro* for clinical applications. For this, the main approach has involved the use of human pluripotent stem cells and manipulation of molecular signals in an attempt to recapitulate in a dish the developmental steps leading to the generation of the first haematopoietic stem cell. Signalling pathways controlling embryonic pluripotency and early differentiation, along with molecular mechanisms driving haematopoietic ontogeny and control of their self-renewal capacity, have been extensively studied, and have allowed considerable progress in the generation of blood products *in vitro*.

In this respect, a special role is fulfilled by the cell cycle. This involves a complex network integrating signalling pathways and molecular mechanisms directing proliferation and cell fate decision, along with many other functions, and over the years has emerged as a key player controlling homeostasis and differentiation of haematopoietic stem cells, both during development and in the adult.

## 1.2 Haematopoietic stem cells

Haematopoietic stem cells (HSCs) are a population of self-renewing multipotent cells that in adult mammals reside in the bone marrow. They are functionally defined by their ability to engraft upon transplant into an immunocompetent recipient and to reconstitute the whole haematopoietic system thanks to their capacity to differentiate and produce all the blood cells in the adult body. This fundamental ability is the basis for procedures involving the use of bone marrow, peripheral blood or cord blood for the transplantation of HSCs into patients with life-threatening diseases, a curative treatment for a number of malignant and non-malignant diseases of the haematopoietic system and for some solid tumours, a procedure whose use over recent years has continued to increase both in Europe and in the United States of America (BSBMT, 2013; NHS England, 2015; Passweg et al., 2017; D'Souza et al., 2017).

HSCs are generally considered to be responsible for the production and replenishment of all the blood cells found in the adult body, despite recent evidence suggesting that at least part of the tissue-resident macrophages might instead derive directly from embryonic progenitors and possess self-renewal capacity, therefore being HSC-independent (Ginhoux and Guilliams, 2016; Hoeffel, 2018). Nevertheless, HSCs retain the fundamental role of sustaining lifelong maintenance of the whole haematopoietic system, thanks to their ability to differentiate and replenish mature blood cells which are predominantly short-lived.

This differentiation towards mature blood lineages has been classically described as a process following a strict hierarchical tree-like structure, starting from HSCs at its top and transitioning through well-defined intermediate progenitors belonging to distinct routes of differentiation, characterised by progressive restriction in their fate potential (Figure 1.1A). This view was supported by the implementation of flow cytometry gating strategies allowing for the isolation of distinct populations which were defined by specific sets of surface markers (Venditti et al., 1999; Holyoake et al., 2000; Hao et al., 2001). According to this model, the first branch point was separating the lymphoid route from the erythro-myeloid potential, and was followed by subsequent ramifications of the hierarchy leading to fully committed blood cells.

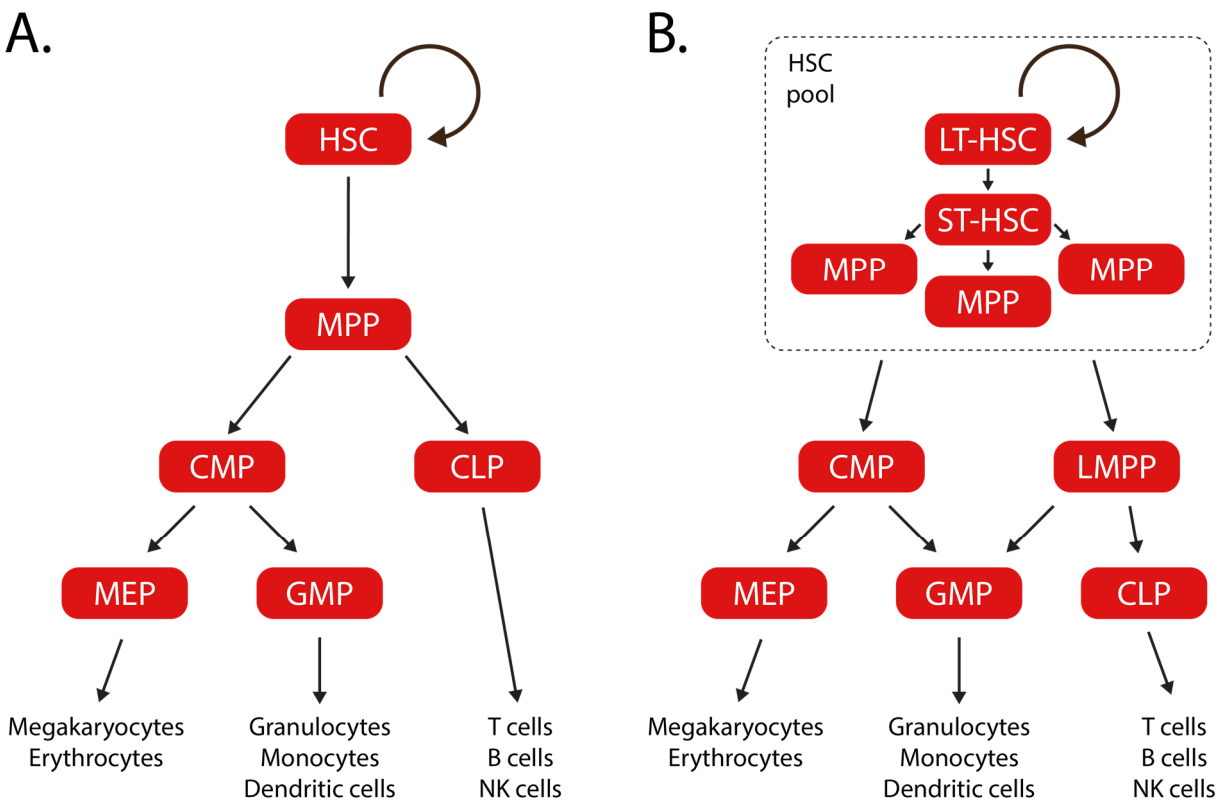



**Figure 1.1 Classical representation of haematopoiesis. (A)** The haematopoietic hierarchy has been classically represented as a strict tree-like structure, with well-defined and separated intermediate stages gradually committing towards differentiated blood cells. **(B)** Progressively, this view has been integrated with new findings suggesting functional heterogeneity in the HSC pool and the existence of crosstalks between branches of differentiation previously thought to be separate (LT: long-term; ST: short-term; HSC: haematopoietic stem cells; MPP: multipotent progenitors; CLP: common lymphoid progenitors; CMP: common myeloid progenitors; GMP: granulocyte-monocyte progenitors; MEP: megakaryocyte-erythrocyte progenitors; LMPP: lymphoid-primed multipotential progenitors).

However, it became clear over time that the reality was more complicated than previously thought. Indeed, a growing number of studies uncovered concepts that challenged the classical model: a crosstalk between different branches, and lymphoid progenitors contributing to the myeloid compartment (Adolfsson et al., 2005; Doulatov et al., 2010); the existence of direct routes of differentiation deriving megakaryocytes directly from HSCs (Sanjuan-Pla et al., 2013); and importantly, the identification of subpopulations of multipotent progenitors able of long-term engraftment but possessing distinct fate propensities (Yamamoto et al., 2013; Pietras et al., 2015; Carrelha et al., 2018). Particularly, the existence of lineage-biased stem cells led to the concept of HSC pool, comprising subpopulations with self-renewal ability but heterogeneous for their differentiation propensity. Furthermore, HSC ability to engraft upon transplant

appeared to be lost progressively, therefore also leading to the distinction between cells able to generate long-term (LT-HSCs) or short-term (ST-HSCs) engraftment (Figure 1.1B).

The advent over recent years of microfluidics technologies have made possible the use of high-throughput single cell gene expression analyses for hundreds of cells, and more recently the transcriptome-wide RNA sequencing for the analysis of even thousands of individual cells, further unveiling the complexity of the haematopoietic tree (Warren et al., 2006; Guo et al., 2013; Nestorowa et al., 2016; Velten et al., 2017). Following these breakthroughs, the haematopoietic stem/progenitor cell (HSPC) compartment appears today to be highly heterogeneous at the transcriptional level, populated of low-primed undifferentiated cells which progressively transition to oligopotent, bipotent or even directly to unipotent progenitor states. This transcriptional heterogeneity is in agreement with the previously mentioned single cell transplantation studies, highlighting the existence of lineage-biased HSCs characterised by distinct differentiation propensities (Sanjuan-Pla et al., 2013; Yamamoto et al., 2013; Carrelha et al., 2018). Additionally, this situation appears to be further complicated by similar functional studies performed on cells derived from distinct human developmental stages, showing that fate propensity of cells in the HSPC compartment, and the differentiation routes they follow, actually change during development, from foetal liver (FL-) to bone marrow (BM-) HSCs (Notta et al., 2015).

Overall, these studies demonstrate that the haematopoietic hierarchy is not composed of clearly separated steps each containing identical progenitors, as previously suggested by the use of well-defined purification strategies. Instead, the classical stepwise differentiation routes emerged to be more similar to a continuous spectrum, along which cells progress in a gradual way from one heterogeneous population of progenitors with a certain lineage bias to another, and with neighbouring routes which are not sharply separated but that can potentially pour into each other, especially at the less committed levels. Therefore, according to this view, the haematopoietic tree would be better represented as a hierarchical but highly flexible continuous system which can easily adapt to meet and satisfy the changing needs of blood demands (Figure 1.2).

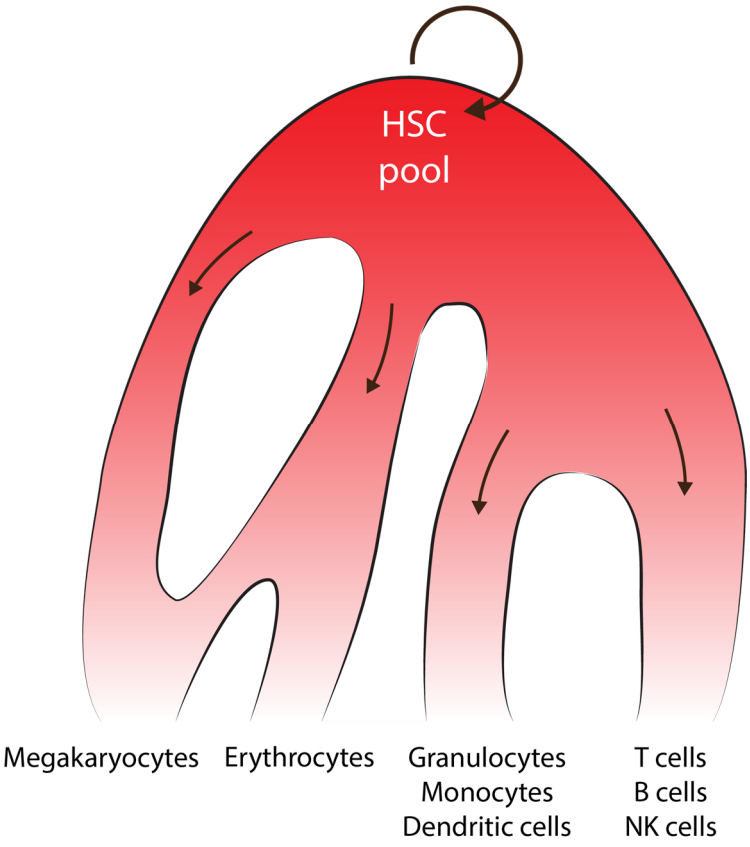



**Figure 1.2 Revised representation of haematopoiesis.** Based on the transcriptomics analysis of individual cells, the haematopoietic hierarchy appears now as a continuous spectrum. Intermediate steps are not well separated and the balance between distinct routes appears to be extremely fluid to adapt to blood demand (adapted from Laurenti and Göttgens, 2018).

## 1.3 Ontogeny of the haematopoietic system

Over the years, the developmental origin of haematopoietic stem cells has been the major focus of many studies, in an attempt to understand, and potentially harness for *in vitro* production, key mechanisms leading to the generation of this population, that as previously mentioned is of fundamental importance for regenerative medicine.

Studies in mouse and a variety of other animal species have shown that haematopoiesis in vertebrate embryos occurs in distinct overlapped waves, conventionally and conveniently divided mainly into primitive and definitive haematopoiesis (Figure 1.3). The first wave, which lasts from embryonic day 7.25 (E7.25) until E9 in mouse, and equivalent to the third week of

gestation in human, occurs in blood islands within the yolk sac and generates the first blood cells of the embryo (Tavian et al., 1999). These mainly contain primitive erythrocytes, derived from mesodermal cells soon after gastrulation through intermediate erythroblast progenitors, and are substantially different from adult red blood cells, being large, nucleated cells which express embryonic haemoglobin proteins. This primitive wave also leads to the generation of low levels of macrophages and megakaryocytes.

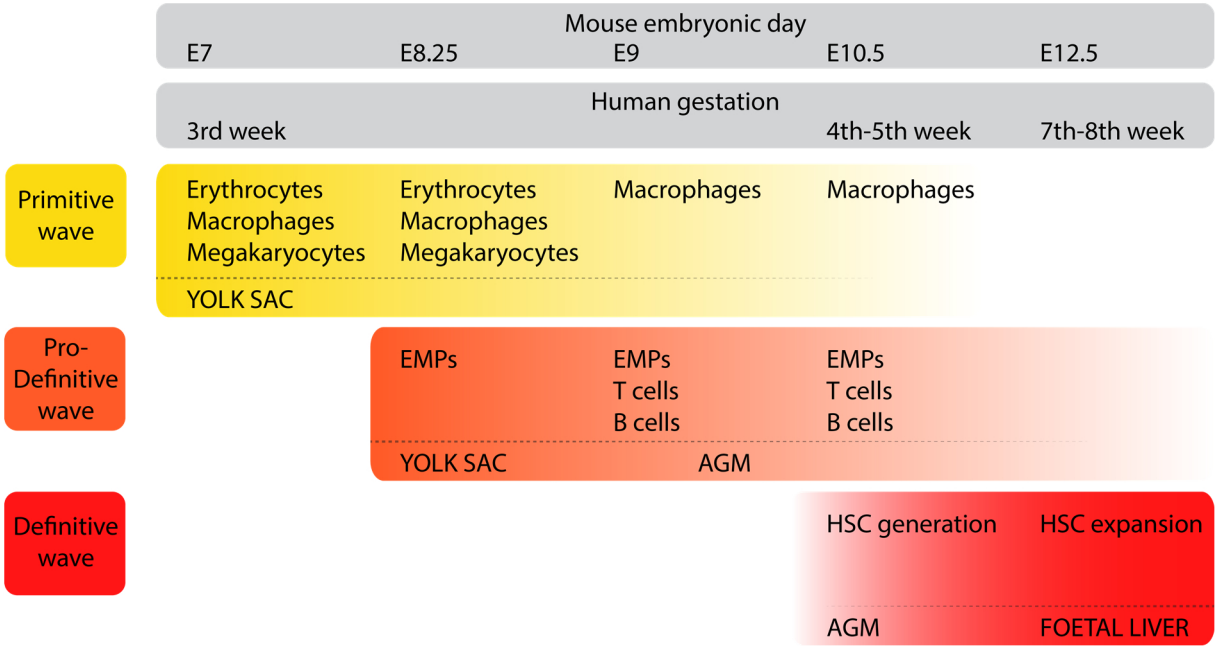



**Figure 1.3 Haematopoietic ontogeny.** During development, multiple haematopoietic waves participate in satisfying the emerging need of blood cells, culminating with the generation in the AGM of the first HSCs. These will then migrate to the foetal liver where they will expand through proliferation and with the contribution of HSCs from secondary haematopoietic sites.

The second definitive wave starts in the yolk sac at E8.25 and produces erythro-myeloid progenitors (EMPs), able to generate myeloid and erythroid cells with adult features and functions, along with some lymphoid cells (Böiers et al., 2013; Palis, 2014; McGrath et al., 2015). This is later followed by the onset of intra-embryonic haematopoiesis in the aorta-gonad-mesonephros (AGM) region of the embryo proper directly from the haemogenic endothelium (HE), a specialised population of endothelial cells (Tavian et al., 1996; Medvinsky and Dzierzak, 1996; Oberlin et al., 2002). The process is marked by the appearance of intra-aortic haematopoietic clusters (IAHCs), groups of cells still physically associated to the haemogenic

endothelial cells (HECs) lining the ventral wall of the dorsal aorta in human, and on a minor extent the dorsal roof in mouse (Yokomizo and Dzierzak, 2010). During the process, the HECs progress to the haematopoietic fate to generate IAHCs through a process called endothelial-to-haematopoietic transition (EHT). According to recent lines of investigation, EMPs generated at this stage are thought to be at the origin of at least part of the tissue-resident macrophages found in the adult (Ginhoux and Guilliams, 2016; Hoeffel, 2018).

This definitive wave culminates at E10.5, corresponding to the fourth-fifth week of gestation in human, with the appearance through EHT of the first self-renewing HSCs, defined by their robust long-term multilineage reconstitution potential (Ivanovs et al., 2011, 2014b). Testimony of their endothelial origin, these cells are characterised in human by the expression of endothelial surface markers like CDH5, along with the typical HSC immunophenotype defined as CD34+/CD90+/CD38-/CD45RA-/CD45+. In addition to the AGM, other secondary sites have been reported to produce HSCs from the HE through EHT later on during development in both mouse and human, such as yolk sac, placenta, vitelline/umbilical arteries, and embryonic head (Ottersbach and Dzierzak, 2005; Robin et al., 2009; Li et al., 2012). Overall, the generation of the first HSCs is sometimes referred to as the beginning of a third haematopoietic wave, characterised by its high regenerative potential.

The first HSCs are generated within IAHCs intermixed with progenitors which are multilineage but devoid of any reconstitution potential. Indeed, the AGM region at this stage is thought to contain only one to three true HSCs (Kumaravelu et al., 2002; Ivanovs et al., 2011). However, it is important to note that at the moment monitoring their origin relies entirely on functional assays, due to the current absence of unique markers defining the first HSCs. For this, they are operationally defined by their engraftment ability upon transplant into adult immunodeficient mice after intravenous injection, which might possibly be not entirely relevant for normal ontogeny. Indeed, the only requirement during development for early HSCs is to migrate to and colonise the foetal liver for further maturation, where they will only later acquire the competence to home to and engraft in the bone marrow. Therefore, the current approach might well underestimate the real number of initial cells produced by the AGM which

will become functional HSCs, especially in human, where this transplantation assay relies on suboptimal interspecies engraftment.

Nevertheless, once generated these first HSCs migrate to the foetal liver where their number has been shown to dramatically increase, both as a consequence of cell proliferation and for the contribution of secondary haematopoietic sites (Kumaravelu et al., 2002). Starting from the seventh-eighth week of human gestation, the liver constitutes the main haematopoietic organ for blood cell production and HSC expansion and maturation until birth, time when the bone marrow progressively takes over this role and becomes the main haematopoietic site. Unfortunately, both these stages remain at the moment largely unexplored, especially in human.

## 1.4 Approaches for *in vitro* generation of haematopoietic cells

As previously mentioned (Chapter 1.2), HSCs constitute the cornerstone for the treatment of a variety of life-threatening diseases. Despite sensible improvements over the years aimed at optimising transplant efficiency to prevent rejection or reducing side effects like graft-versus-host disease, HSC transplantation remains challenging and not always successful. Furthermore, the lack of immunologically matched donors is often a major obstacle. Thus, understating development and biology of HSCs represents an important challenge in the field of regenerative medicine, necessary for the implementation of new strategies to improve the availability and success of such treatments. However, basic studies on HSCs are technically difficult, especially in the human system. Primary HSCs can be obtained from bone marrow, umbilical cord blood or peripheral blood upon mobilisation from the bone marrow, but these sources are relatively rare, while the number of HSCs that can be obtained is limited (Amos and Gordon, 1995; Kekre and Antin, 2014). HSC expansion could provide a potential solution, and considerable progress has been made in defining effective culture conditions for this purpose. However, even the most robust protocols currently available can only achieve a modest expansion of HSCs in culture before losing self-renewing and engraftment ability (Kumar and

Geiger, 2017; Ferreira and Mousavi, 2018). Therefore, the development of alternative systems has become essential, in order to model human haematopoiesis to understand molecular mechanisms driving HSCs. Ideally, this should allow not only to improve culture conditions for the expansion *ex vivo* of primary HSCs derived from donors, but also to eventually produce these cells *in vitro* for their use in autologous transplants or for disease modelling (Vo and Daley, 2015).

In this context, an advantageous alternative is provided by the use of pluripotent stem cells (further discussed in Chapter 1.5). These cells represent the *in vitro* equivalent of the earliest stages of embryonic development, and in the right culture conditions retain the ability to grow almost indefinitely and possess the capacity to differentiate to virtually every cell type in the body. Accordingly, over the years considerable progress has been made in developing methods allowing the differentiation of haematopoietic cells capable of multiple blood cell potential (Kaufman, 2009; Slukvin, 2013). This has been done deploying different strategies, but to date the generation of cells capable of long-term multilineage engraftment still remains a considerable challenge (Figure 1.4).

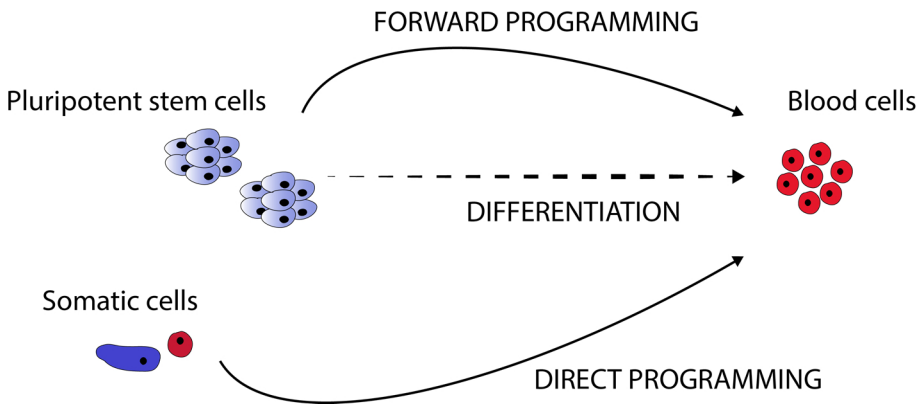



**Figure 1.4 Approaches for the artificial production of blood cells.** The generation of haematopoietic cells *in vitro* has been attempted by different strategies, using differentiation and forward programming of pluripotent stem cells, or direct programming of somatic cells.

The principal approach for the generation of haematopoietic cells *in vitro* using pluripotent cells has involved the sequential administration of small molecules and growth factors in a precise temporal order, in the attempt to recapitulate early events leading to the

generation of these cells during embryonic development. For this, a number of strategies have been developed. The initial use of co-cultures with mouse bone marrow stromal cells or other cell lines, which would ideally provide molecular signals driving haematopoietic differentiation, has been gradually improved towards fully feeder-free systems employing extrinsic manipulation of signals to mimic the *in vivo* environment during development (Kennedy et al., 2007; Niwa et al., 2011; Choi et al., 2012; Sturgeon et al., 2013; Slukvin, 2013; Ramos-Mejía et al., 2014). Overall, these approaches have been based on the knowledge of the molecular signals driving natural development, knowledge that is still currently incomplete, especially in human where early events taking place in the AGM region are not easily accessible. For this, further understanding the early development of the haematopoietic system is of crucial importance, in order to succeed in deriving HSCs and other blood products which are fully functional and suitable for translational applications. On this note, a more comprehensive approach complementing *in vivo* and *in vitro* studies will possibly represent the best strategy to fill the gaps in the current knowledge on human haematopoietic ontogeny.

An alternative strategy has involved bypassing the normal developmental route by directly enforcing the expression of transcription factors leading to the desired cell types, HSCs or other blood products. This approach has been used in pluripotent stem cells (forward programming) and somatic cells (direct programming), and consists in the direct conversion of the source cell type to another through activation of the correct transcriptional machinery (Elcheva et al., 2014; Easterbrook et al., 2016). This method has been proved to be promising, especially for what concerns the production of mature cell types to be used directly for transfusions (Moreau et al., 2016). Nonetheless, a few challenges are still remaining, including immunocompatibility, low-throughput of the process, and full functionality of the product *in vivo*. Importantly, this strategy has been recently employed to successfully convert mouse adult vascular endothelial cells towards the generation of immunocompetent HSPCs (Lis et al., 2017). For this, purified endothelial cells were transduced to obtain the expression of four transcription factors (FOSB, GFI1, RUNX1 and SPI1), followed by co-culture with human umbilical vein endothelial cells modified to express the E4ORF1H gene. This approach was able to convert the isolated endothelial cells to multipotent HSPC-like cells capable of long-term

engraftment. Of note, a similar result was accomplished by using both the methods described above, thus combining hPSC differentiation and direct programming (Sugimura et al., 2017). For this, hPSCs were initially directed through stepwise administration of morphogens to the generation of haemogenic endothelial cells. These cells were then infected to obtain the viral-mediated expression of seven transcription factors (ERG, HOXA5, HOXA9, HOXA10, LCOR, RUNX1 and SPI1) identified by screening of a bigger library. The resulting cells were capable of multilineage engraftment in primary and secondary mouse recipients.

Despite allowing the generation of HSPC-like cells *in vitro*, both these methods still retains features which make them not amenable for clinical application, namely the difficulty of obtaining large-scale production of these cells and the requirement for viral transgene expression. Nevertheless, they undoubtedly sheds lights on the molecular requirements necessary for the successful generation of HSCs *in vitro*.

## 1.5 Pluripotent stem cells as a model of early development and a promise for regenerative medicine

As anticipated in the previous section, human pluripotent stem cells represent a potential alternative solution to the study of molecular mechanisms taking place during haematopoietic development and for the generation *in vitro* of cells for their use in cell therapy, disease modelling and regenerative medicine. In this context, a major breakthrough in the field was the establishment of culture techniques for the expansion and maintenance *ex vivo* of the first pluripotent stem cells, embryonic cells derived from the mouse blastocyst (Evans and Kaufman, 1981; Martin, 1981). This was followed by years of efforts aimed at obtaining similar cells in human, which ultimately led to the derivation of the first human embryonic stem cells (hESCs; Thomson et al., 1998). For this, cells were cultured from the inner cell mass of pre-implantation blastocysts, obtained from surplus embryos for *in vitro* fertilisation procedures. These cells were characterised by self-renewal ability and importantly pluripotency, defined as the capacity to differentiate towards the three embryonic germ layers, namely endoderm, mesoderm and

neuroectoderm, therefore with the potential of generating every cell type in the human body (Figure 1.5). They represented a simplified model for human development, and allowed to study mechanisms otherwise difficult or even impossible to access *in vivo*. Furthermore, the development of protocols for the differentiation of mature cell types with adult phenotype represented a great promise for cell replacement therapy and regenerative medicine.

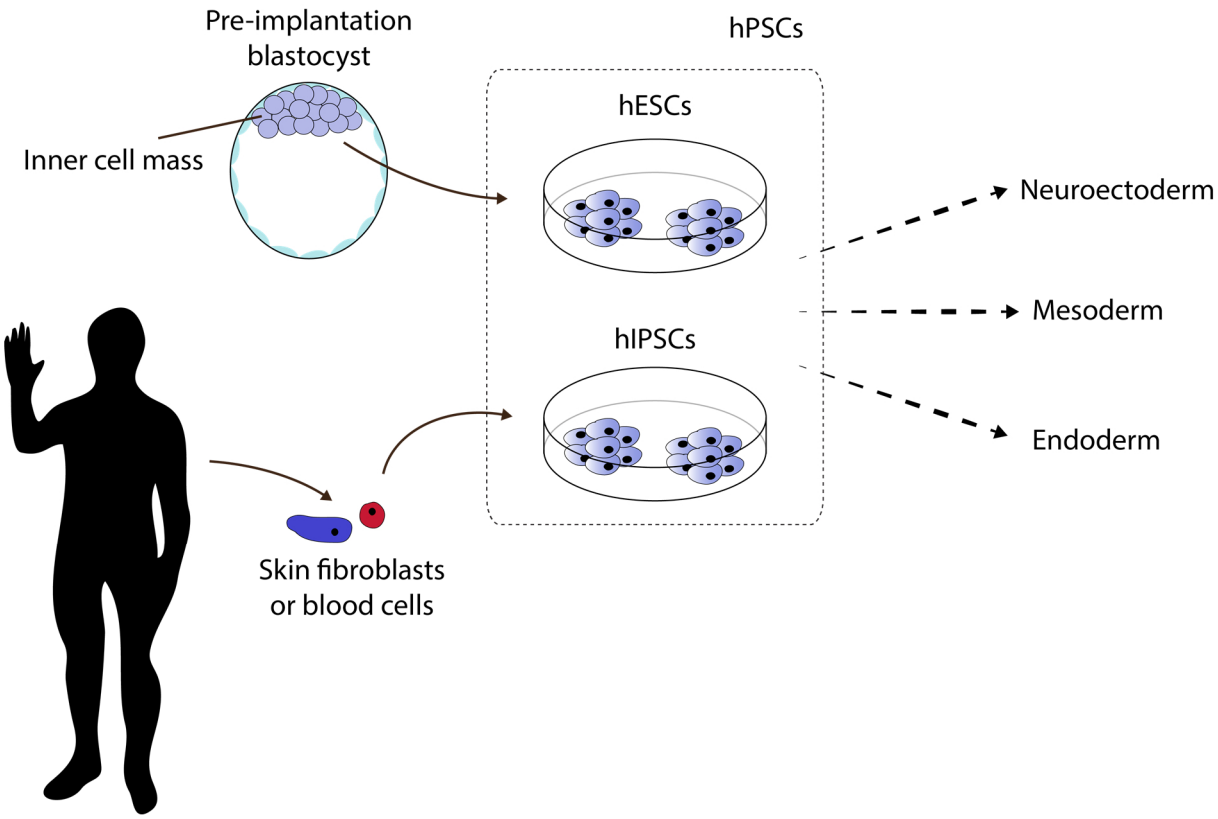



**Figure 1.5 Human pluripotent stem cells.** hESCs derived from the inner cell mass of pre-implantation blastocyst and hIPSCs generated by reprogramming of somatic cells are the *in vitro* equivalent of the earliest stages of human development. They retain self-renewal ability and pluripotency, and can potentially be used to generate every cell type of the human body.

However, the large-scale production of hESC lines remained a considerable challenge, mainly because of the paucity of human embryos donated for research purposes. Furthermore, the use of hESCs for research and potentially for treatments raised ethical concerns about the use of human embryos (McLaren, 2001). This led for example in the United States of America to an almost complete ban of federal funding for hESC research, which only recently has been

partially lifted (Holden, 2009). Importantly, the derivation of these cells is still illegal in some countries, one example being Italy (Frassoni, 2006; Palazzani, 2011). These concerns and limitations were finally overcome by the development of techniques showing that adult somatic cells could be reprogrammed in both mouse and human to an embryonic-like state by the introduction of specific transcription factors (Takahashi and Yamanaka, 2006; Takahashi et al., 2007). Specifically, the first human induced pluripotent stem cells (hIPSCs) were obtained by viral-mediated overexpression of the four transcription factors OCT4, SOX2, KLF4 and c-MYC in adult fibroblasts obtained through biopsy. Over the years, this approach has been extensively refined in order to improve reprogramming efficiency, allow the process to be carried out on different and more accessible cell types, and also to avoid the integration of transgenes by performing transfection and transient expression of the factors using non-integrating vectors (Hochedlinger and Jaenisch, 2015). Importantly, these cells were shown to share most of the hESC properties and to be capable of self-renewal and differentiation through the three germ layers. For this, both hESCs and hIPSCs are now collectively referred to as human pluripotent stem cells (hPSCs; Figure 1.5).

These cells offer the unprecedented opportunity to model the earliest stages of human development, confirming and complementing studies involving the use of animal models. Additionally, the capacity of these cells to differentiate offers the possibility to deepen our understanding of events taking place later on during development in specific tissues or organs of not immediate access, one example being the onset of definitive haematopoiesis. Furthermore, hIPSCs hold great promise for regenerative medicine. They represent a virtually unlimited source of patient-derived immunologically matched cells, to be used for transplantation, with the perspective of regenerating damaged tissues or whole organs thus avoiding immunosuppression. Moreover, hIPSCs derived from patients with genetic disorders allow disease modelling and potentially the development of treatments based on the study of mechanisms that are not always conserved in animal models.

However, for all these applications, understanding signalling pathways and molecular mechanisms driving self-renewal, differentiation potential and cell fate decision in hPSCs is of fundamental importance.

## 1.6 Signalling pathways controlling pluripotency

The first stages of hPSC research involved considerable efforts aimed at determining conditions that would allow the long-term maintenance of these cells in culture in an undifferentiated state. Initial attempts led to the use of co-culture systems on feeder layers of mitotically inactivated mouse embryonic fibroblasts (MEFs), which were shown to provide factors able to maintain pluripotency. In this system, hPSCs could be detached from the feeder layer and then induced to differentiation as floating clumps of cells in the presence of serum. The lack of pluripotency-inducing factors and the presence of serum were able to induce the formation of 3D structures generating cells of the three germ layers, and were thought to partially approximate the complexity of the early embryo (Itskovitz-Eldor et al., 2000). For this, they were called embryoid bodies (EBs). Despite MEF cultures and EB differentiation represent still today valuable tools for specific applications, stem cell research over the years has been characterised by the attempt at identifying robust feeder-free, chemically defined conditions for the maintenance and differentiation of hPSCs (Wiles and Johansson, 1999; Chen et al., 2011).

These technical improvements had to be accompanied with the identification and an increasingly deeper understanding of signalling pathways controlling pluripotency. Notably, culture conditions that had been described for mouse embryonic stem cells did not prove successful in maintaining pluripotency in feeder-free hESC cultures. It was indeed later demonstrated that the two stem cell models represented distinct stages of embryonic development, corresponding to distinct states of pluripotency and characterised by unique signalling requirements (Pauklin et al., 2011).

### 1.6.1 Activin/NODAL/TGFβ signalling

Ultimately, the feeder-secreted factors responsible for the maintenance of hESC pluripotency in culture were demonstrated to be molecules of the TGFβ superfamily (Vallier et

al., 2004). Indeed, the addition of Activin, NODAL or TGFβ were able not only to promote pluripotency, but also to inhibit the default differentiation pathways leading to neuroectoderm (Vallier et al., 2005). Of note, Activin and NODAL, which bind to the same receptor, are functionally very similar to TGFβ, all of them activating the same downstream effectors SMAD2 and SMAD3 (Figure 1.6). As a result, these are phosphorylated and activated, enter the nucleus and induce transcription of their target genes, which include genes controlling pluripotency and differentiation (Pauklin and Vallier, 2015). One of the fundamental factors controlled by this pathway is NANOG, a master regulator of pluripotency which at the same time interacts with SMAD2/3 and guides their binding to key genomic locations to further reinforce the pluripotency transcriptional network (Brown et al., 2011). SMAD2/3 and NANOG have also been shown to cooperate for the recruitment of methyltransferases that leads to the deposition of epigenetic marks like H3K4me3 on key developmental genes for the maintenance of the pluripotent state (Bertero et al., 2015). Furthermore, SMAD2/3 has been recently shown to participate in an additional layer of complexity in the regulation of pluripotency, by interacting with the METTL3-METTL14-WTAP methyltransferase complex and controlling the m6A modification of a subset of transcripts involved in early cell fate decisions, therefore priming them for rapid downregulation upon commitment and differentiation (Bertero et al., 2018). Overall, the Activin/NODAL/TFGβ pathway appears to be a fundamental driver of pluripotency and a requirement for the maintenance of the undifferentiated state, while their inhibition results in differentiation, preferentially towards the default neuroectoderm lineage.

### 1.6.2 FGF signalling

In addition to this key pathway, FGF2 was also shown to have an important role in hESCs, working synergistically with the Activin/NODAL/TGFβ signalling to maintain the undifferentiated state (Vallier et al., 2005). The interaction of FGF2 with its receptors can result in the activation of four distinct signalling pathways, namely the JAK/STAT, PLCγ, PI3K and MAPK/ERK pathways (Dailey et al., 2005; Lanner and Rossant, 2010). Of these, the MAPK/ERK pathway, and to a lesser extent the PI3K pathway, have been extensively explored. Specifically,

the MAPK/ERK pathway involves the sequential phosphorylation and activation of RAF, MEK and ERK1/2 kinases, which shuttles to the nucleus and activate downstream proteins leading to the induction of target genes (Figure 1.6). Studies have also shown that FGF signalling contributes to the expression of *NANOG* (Greber et al., 2007). Interestingly, FGF alone is not sufficient for the maintenance of pluripotency, and instead it appears to be working synergistically with the Activin/NODAL/TGFβ signalling by increasing the expression of NODAL co-factor *CRIPTO* (Vallier et al., 2005). Furthermore, ERK2 has been reported to interact with multiple genomic loci in hESCs, involved in a number of functions like metabolism, cell cycle progression and pluripotency (Göke et al., 2013). Ultimately, FGF signalling appears to trigger a variety of biological responses, involved in proliferation, cell survival and differentiation of hPSCs. Importantly, Activin/NODAL/TFGβ and FGF2 have been shown to maintain pluripotency also in hIPSCs, confirming these cells to be functionally equivalent to hESCs (Vallier et al., 2009).

Of note, other factors also participate in maintaining pluripotency. The IGF signalling pathway has been shown to be partially redundant with FGF, and activate the PI3K pathway resulting in increased cell proliferation and inhibition of differentiation (Campbell et al., 2012). At the same time the WNT canonical pathway (further discussed in Chapter 1.7.1), acting through the effector β-catenin, is also important for hPSCs, with its role depending on the signalling strength. Indeed, low levels of WNT activation promote the pluripotent state, while stronger stimulation leads to mesendoderm specification (Sumi et al., 2013). Importantly, the undifferentiated state appears to be maintained through an extensive crosstalk between these multiple signalling pathways (Singh et al., 2012).

Remarkably, these molecular requirements closely recapitulate signalling pathways with a key role in the post-implantation mouse embryo, in which NODAL is necessary to maintain the expression of pluripotency markers in the epiblast (Camus et al., 2006). On one side, this further underlines the importance of understanding early development for the correct manipulation of pluripotency and differentiation *in vitro*, and on the other it confirms the potential of hPSCs in allowing fateful modelling of human development.

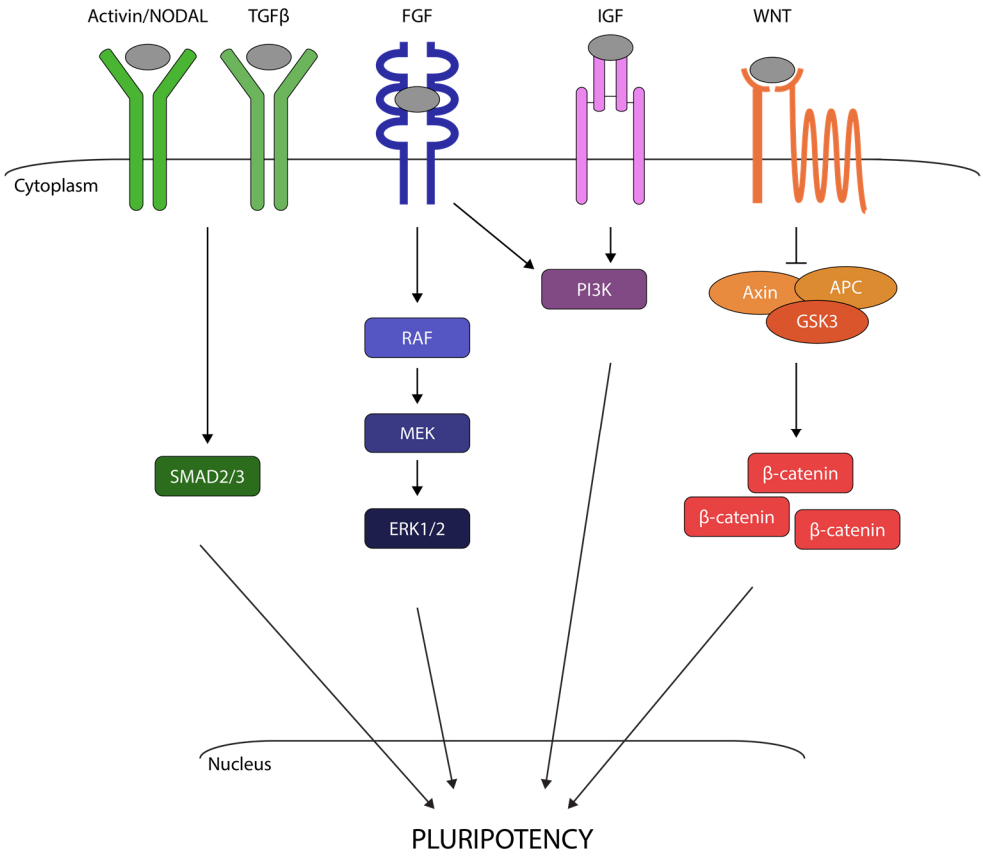



**Figure 1.6 Signalling pathways maintaining pluripotency.** Activin/NODAL/TGFβ and FGF are the main signalling pathways driving pluripotency. These also synergise with other pathways such as IGF and WNT.

## 1.7 Signalling pathways driving haematopoietic differentiation

Initial methods for the *in vitro* differentiation of hPSCs relied on the use of serum in the culture media, whose composition was undefined and included factors which drove commitment at least partially in an uncontrolled way. Over the years, several studies led to the identification of fundamental molecular pathways driving early germ layer specification and commitment to differentiated cell types. The knowledge derived from developmental studies allowed for the progressive development of more refined culture conditions for the differentiation of hPSCs *in vitro* through early germ layers to more differentiated cell types, in an accurately guided manner, whose rationale was the faithful recapitulation of developmental mechanisms (Murry and Keller, 2008). Of note, recapitulating haematopoietic development *in*

*vitro* also requires the tight control of early germ layer induction, specifically towards the generation of the lateral plate mesoderm subtype which developmentally undergoes haematopoietic specification *in vivo* (Dzierzak and Bigas, 2018).

From these developmental studies, it was soon recognised that FGF and most importantly the Activin/NODAL/TGFβ signalling are not only master gatekeepers of pluripotency, but have also a fundamental role during early germ layer specification. This dual action depends on the crosstalk with other key signalling pathways such as BMP and WNT, and notably their relative manipulation *in vitro* drives both early mesendoderm/neuroectoderm specification and mesoderm subtype patterning in hPSCs (Bernardo et al., 2011; Mendjan et al., 2014; Faial et al., 2015). Indeed, the same pathways regulating pluripotency appear to also control a complex process like gastrulation that *in vivo* results from the relative position along the embryo exposing distinct cells to specific levels of agonists and inhibitors, whose expression forms a gradient along the anteroposterior axis (Figure 1.7). This drives the interplay between different pathways, and ultimately determines cell fate decision and differentiation (Tam and Behringer, 1997; Murry and Keller, 2008).

Consequently, the recapitulation of this positional effect is necessary for the control of germ layer induction *in vitro*, which relies on the accurate combination of signalling molecules. As already mentioned, the role of Activin/NODAL/TGFβ and FGF during differentiation depends on their crosstalk with BMP and WNT pathways (Mendjan et al., 2014; Faial et al., 2015). Indeed, by modulating the dosage of these signalling pathways it is possible to obtain mesendoderm and induce different subtypes of mesoderm, characterised by distinct potential to further differentiate towards somatic cell types.

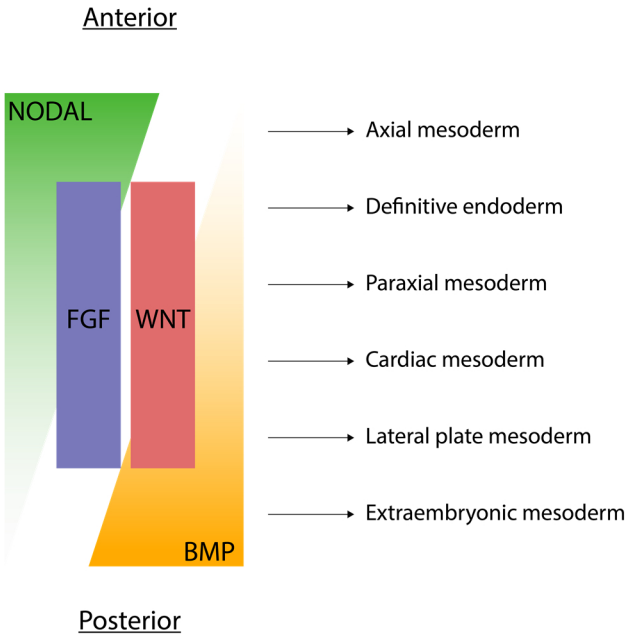



**Figure 1.7 Positional effect during embryonic development.** The relative position along the anteroposterior axis determines the level of activation of distinct signalling pathways and the extent of their crosstalk. This positional effect needs to be fatefully recapitulated *in vitro* for the generation of the correct germ layer able to differentiate towards a certain somatic cell type.

### 1.7.1 WNT signalling

As already mentioned, the WNT canonical pathway leads to the accumulation and nuclear translocation of β-catenin, which functions as a transcriptional co-activator for a number of target genes. For this, WNT binding to its heterodimeric receptor composed of Frizzled and LRP5/6 proteins leads to the disruption of an intracellular protein complex including Axin, APC and GSK3, by causing its translocation to the plasma membrane (Figure 1.6). In WNT absence, the complex phosphorylates and marks for degradation the effector β-catenin. But upon WNT activation, this degradation activity is inhibited, leading to β-catenin accumulation in the cytoplasm and its translocation into the nucleus, where it leads to the activation of its target genes (Clevers and Nusse, 2012).

Beyond this, WNT also signals through non-canonical pathways that do not involve β-catenin. These include the planar cell polarity pathway, which regulates the cytoskeleton, and the calcium pathway, which regulates intracellular calcium levels. Both these pathways rely on the interaction of WNT ligands with the same heterodimeric Frizzled-LRP5/6 receptor, which

intracellularly binds to Dishevelled (DSH; Komiya and Habas, 2008). More specifically, for the planar cell polarity pathway, DSH on one side associates with DAAM1 and activate the protein Profilin, able to bind Actin, and the small G protein Rho which activates ROCK, a major cytoskeleton regulator. On the other side, DSH can signal to RAC1, which activates JNK, leading to its action on target genes. These events result in the restructuring of the cytoskeleton, important, for example, during gastrulation. The other non-canonical pathway is the calcium pathway, in which Frizzled not only activates the intracellular DSH, but also interact with a trimeric G protein. Ultimately, this signalling results in calcium release from the endoplasmic reticulum, which then leads to the activation of proteins like Calcineurin, that in turn activates factors such as NFAT, involved in embryonic ventral patterning, and in the control of cell adhesion, migration and tissue separation.

For *in vitro* directed differentiation of hPSCs, the activation of WNT signalling during an early stage has been reported to be associated with an enrichment in progenitors of the definitive haematopoietic wave. Importantly, this has been achieved through CHIR99021, a small molecule inhibiting GSK3 and therefore leading to β-catenin accumulation, suggesting that WNT canonical pathway is necessary for the generation of lateral plate mesoderm contributing to definitive haematopoiesis (Sturgeon et al., 2014). Furthermore, studies performed in mouse have shown the pathway to be transiently required for HSC specification in the AGM, but not for their subsequent maintenance (Ruiz-Herguido et al., 2012).

### 1.7.2 BMP signalling

BMPs are secreted extracellular matrix-associated proteins of the TGFβ superfamily, important for a wide range of processes during development. An important role in the BMP signalling regulation is played by the secretion of antagonists like Noggin, able to bind BMP members and therefore inhibit the association with their receptor counterparts. Both these groups of molecules are associated with the extracellular matrix in order to limit their effect on neighbouring cells. Upon binding, BMP causes the dimerization of its receptor subunits BMPRI

and BMPRII, leading to intracellular phosphorylation of SMAD1, SMAD5 and SMAD8. These, in association with the cofactor SMAD4, translocate into the nucleus and activate the expression of their target genes (Wang et al., 2014; Brazil et al., 2015). In addition to the canonical BMP/SMAD signalling, other non-canonical roles have been described, including for example the activation of MAPK signalling, or the interaction with PI3K and Rho GTPases. Of note, the balance between canonical and non-canonical functions is likely to depend on the extracellular environment and the cellular context (Wang et al., 2014a).

It is important to note that an extensive crosstalk exists between the BMP signalling and other pathways. Remarkably, differentiation cultures using hPSCs in the presence of different levels of BMP4, Activin, FGF2, the WNT inhibitor CHIR99021 and the PI3K inhibitor Ly294002, have been shown to generate distinct mesendoderm subtypes (Figure 1.7). This is achieved through a finely tuned interaction between key downstream effectors like SMADs, pluripotency regulators like NANOG and early key mesendoderm factors such as BRACHYURY, which collectively directs the earliest stages of embryonic development (Mendjan et al., 2014; Faial et al., 2015).

BMP signalling has also been known to have a role in promoting the generation of haematopoietic cells from the AGM region (Durand et al., 2007). However, its effect seems to be dependent on modulation by other pathways like FGF (Pouget et al., 2014), and its activation appears to be tightly regulated by a temporal and spatial balance between ligands like BMP4 and antagonists like BMPER (McGarvey et al., 2017). BMP activation is indeed needed for the initial HSC specification, but its subsequent inhibition seems to be equally necessary for the successful maturation of functional HSCs (Souilhol et al., 2016). Furthermore, beyond its importance for haematopoietic emergence in the dorsal aorta, its role later during development is still not entirely clear, since foetal liver HSCs show heterogeneity in their ability to respond to BMP (Crisan et al., 2015).

### 1.7.3 Notch signalling

Notch signalling is a highly conserved pathway in which both the receptors and their ligands are usually transmembrane proteins, therefore allowing for short-range communication and requiring physical contact between neighbouring cells. It plays a critical role during development, resulting in a response that is context-dependent and can promote or suppress proliferation, cell death, fate decision and differentiation. A distinguishing hallmark of Notch signalling is its transduction mechanism. Indeed, it relies on the proteolysis of the Notch receptor upon binding to its ligand, which produces an active Notch intracellular domain (NICD). Once released, the NICD fragment shuttles to the nucleus where it binds to other transcription co-factors and activates its target genes (Kopan and Ilagan, 2009).

The Notch pathway is particularly important in angiogenesis and blood development. In HECs undergoing EHT, Notch signalling guides the generation of IAHCs, and this action seems to depend on a NOTCH1-dependent induction of key haematopoietic factors such as RUNX1 and GATA2 (Kumano et al., 2003; Burns et al., 2005; Guiu et al., 2013; Ditadi et al., 2015). Furthermore, it was recently reported that pivotal to the correct induction of the haematopoietic programme is the strength of such signalling. It was indeed shown that in endothelial cells in the mouse dorsal aorta, the distinct arterial and haemogenic programmes rely on the induction by the two Notch ligands DLL4 and JAG1, respectively (Gama-Norton et al., 2015). Both ligands are expressed in the dorsal aorta, but binds to NOTCH1 with a different affinity and activate their signal with different strength. The stronger NOTCH1/DLL4 signalling results in the activation of the arterial programme, but this interaction is outcompeted by JAG1, which has a higher affinity for the receptor. This results in a preferential NOTCH1/JAG1 signalling which possesses weaker strength, and which blocks the arterial programme and activates instead haemogenic specification.

Notably, a strong Notch signalling induction by DLL4 seems to also drive terminal haematopoietic maturation, as shown by studies using co-culture systems with mouse stromal cells overexpressing DLL4 to guide HSPC commitment towards T lymphocytes (Sturgeon et al., 2014).

### 1.7.4 Importance of integrating multiple pathways

The emergence of the first HSCs is a complex process, carefully regulated by the interaction between multiple gradients of molecules on the dorsoventral axis, along with signals coming from the lateral urogenital ridges (Figure 1.8). The aforementioned FGF, WNT and BMP signalling pathways indeed show extensive crosstalk with other pathways. An example is given by an important inducer of HSC emergence in the AGM such as sonic hedgehog (SHH; Gering and Patient, 2005; Peeters et al., 2009). Despite this being produced dorsally from the notochord, the dorsal roof of the aorta has a much lower sensitivity to SHH compared to the ventral portion, explaining the emergence of IAHCs preferentially (in mouse) or exclusively (in human) from this area. Furthermore, the two opposite dorsal SHH and ventral BMP gradients appear to feedback on each other. It was indeed suggested that BMP acts initially in the induction of the haematopoietic programme, and carefully regulates the intensity of SHH signalling by inhibiting its expression. However, BMP signal needs to be turned off to complete HSC emergence, and SHH appears indeed to increase the expression of Noggin which antagonises BMP while at the same time increasing SHH expression itself, overall delineating an important cross-regulation between the two pathways causing a transient BMP activation (Wilkinson et al., 2009; Souilhol et al., 2016).

Another example of signalling pathways interacting to regulate haematopoietic development is given by WNT signalling, that as previously mentioned was shown to be necessary for HSC specification, but not for their maintenance (Ruiz-Herguido et al., 2012). It appears that this temporal regulation might be associated with the action of other signalling, mediated by retinoic acid. This signal was previously suggested to be necessary for haematopoietic specification (Goldie et al., 2008), and it appears that for this effect to take place, WNT signalling needs first to be downregulated. Furthermore, it was suggested that retinoic acid itself might actually inhibit WNT signalling (Chanda et al., 2013). As a result, WNT is inhibited in emerging haematopoietic cells, and its activation remains limited to the

surrounding endothelium, as shown in the mouse AGM by the presence of β-catenin in endothelial cells but not in IAHCs.

Additionally, the Notch signalling has been reported to crosstalk with VEGF, one of the main signals controlling angiogenesis. During vascularisation, endothelial cells growth is stimulated by the action of VEGF signalling, whose gradient guides their migration and proliferation. At the same time, VEGF is able to induce the expression of *DLL4*. The resulting interaction of DLL4 with NOTCH1 counterbalances endothelial growth by inhibiting the expression of VEGF receptors and therefore reducing endothelial cell responsiveness to this growth signal. Coherently, lack of Notch signalling results in uncontrolled endothelial migration and proliferation (Hellström et al., 2007; Leslie et al., 2007; Siekmann and Lawson, 2007). While this mechanism seems to be important during vascularisation, its role in HSC specification is still to be fully elucidated, despite a certain level of interaction taking place during EHT, in which VEGF function appears to be upstream of the Notch pathway (Leung et al., 2013).

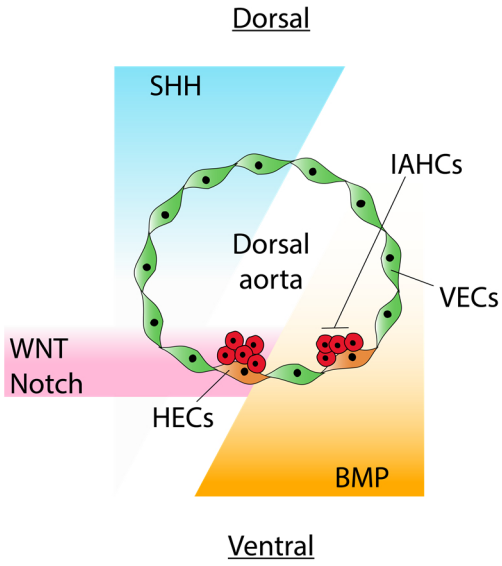



**Figure 1.8 Signalling pathways driving EHT.** The interplay between multiple pathways along the dorsoventral axis determines the emergence of haematopoietic cells in the dorsal aorta (IAHCs: intra-aortic haematopoietic clusters; HECs: haemogenic endothelial cells; VECs: vascular endothelial cells).

Furthermore, interaction with Notch signalling and a role during haematopoietic development was also suggested for inflammatory signals. IFNγ was indeed shown to promote HSC emergence in zebrafish embryos, and the expression of this signal was suggested to be

downstream of Notch (Sawamiphak et al., 2014). Simultaneously, another inflammatory signal like TNFα was also shown to directly upregulate the expression of JAG1, and therefore increase HSC production through promotion of Notch signalling (Espín-Palazón et al., 2014). Interestingly, this last finding is in agreement with the previously mentioned study showing the specific role of the NOTCH1/JAG1 pathway in promoting EHT (Gama-Norton et al., 2015). Further work is needed to fully elucidate the role of inflammatory signals on HSC generation, but it has been suggested from zebrafish studies that these signals might come from primitive macrophages, which in sterile conditions and in the absence of infection contribute to the events that in the developing embryo lead to adult haematopoiesis, therefore delineating a more complex pro-definitive role for the primitive haematopoietic wave.

### 1.7.5 TAL1, RUNX1 and GATA2: master regulators of haematopoiesis

Gradients of signal molecules drive multiple developmental events in different parts of the embryo. Their combination leads to the sequential activation of the lateral plate mesoderm, the endothelial and the haematopoietic programme. This ultimately relies on multiple transcription factors that are involved in HSPC specification, but also in their maintenance and lineage commitment. The diversity of roles for the same factors depends on their highly context-dependent activity, which indeed varies according to the presence of other co-factors guiding the expression of target genes. Multiple transcriptional complexes drive the haematopoietic fate, and although many of the individual players in these complexes have been identified and characterised, their activity and functioning still need to be fully elucidated. Three of the main factors known to drive HSPC specification and with a diversity of roles during haematopoiesis are the master regulators TAL1, RUNX1 and GATA2.

TAL1 is a transcription factor that in combination with VEGF signalling is able to inhibit the cardiac lineage in early mesoderm, specify lateral plate mesoderm towards haematopoietic and endothelial fates, and subsequently consolidate the haematopoietic fate (Org et al., 2015). TAL1 expression is induced by a combination of VEGF and BMP signalling.

It is a strong inducer of HSPC specification, but it is present at every level of the haematopoietic hierarchy, being involved in the onset of haematopoiesis, in HSPC maintenance and in lineage commitment. Its action is indeed strongly context-dependent, and for example can induce HSPC emergence when in combination with GATA2, or erythroid and megakaryocytic differentiation when combined with GATA1 (Shivdasani et al., 1995; Hoang et al., 2016; Vagapova et al., 2018). Interestingly, TAL1 was also shown to control adult HSC entry into the cell cycle. Its expression in mouse is indeed higher in adult LT-HSCs compared to ST-HSCs, and it can keep the stem cell pool in quiescence by activating the expression of genes such as the cell cycle inhibitor *CDKN1A/p21* (Lacombe et al., 2010).

Differently from TAL1, RUNX1 is not required during the primitive haematopoietic wave, but its deletion prevents instead endothelial cells from undergoing EHT and generating IAHCs. This results in a complete loss of definitive blood cells, highlighting RUNX1 specific requirement for the activation of the definitive haematopoietic programme (Okuda et al., 1996; Cai et al., 2000). The observation that RUNX1 is expressed in a small population of endothelial cells where IAHCs will subsequently emerge, therefore anticipating their appearance, gave a first confirmation for the endothelial origin of the definitive haematopoietic wave (North et al., 1999). Overall, RUNX1 appears to be fundamental for EHT and the generation of definitive EMPs and HSPCs, but not for the differentiation of early mesodermal progenitors into primitive haematopoietic cells. Furthermore, its requirement has been suggested to end once the HSPC programme has been specified in the AGM (Chen et al., 2009). Nevertheless, RUNX1 continues to be expressed thereafter, and mutations in its gene are often associated with leukaemia (Ichikawa et al., 2013), suggesting that RUNX1 subsequent role still needs to be fully understood. Indeed, despite RUNX1 having been the focus of many studies, its downstream effectors and the molecular mechanisms of its action still remain largely unknown. One important effect mediated by RUNX1 is the upregulation in the HE of the two transcription factors GFI1 and GFI1B, which play an important role in repressing endothelial genes while RUNX1 activates the haematopoietic programme (Lancrin et al., 2012).

A third master regulator of haematopoietic development is GATA2. This factor is not strictly required for primitive haematopoiesis. Despite being expressed, its absence causes only

a modest effect, and it was indeed shown that at this stage GATA2 functionally overlaps with GATA1, the main factor driving primitive erythropoiesis along with TAL1 (Fujiwara et al., 2004). GATA2 is instead necessary for the specification of definitive haematopoiesis, being expressed in HECs and IAHCs and with its absence preventing EHT (de Pater et al., 2013). The expression of GATA2 was reported to be directly controlled by NOTCH (Robert-Moreno et al., 2005). Importantly, GATA2 is not only required for the generation of HSCs in the AGM, but is also important for their maintenance thereafter. This role requires GATA2 levels to be maintained within stringent levels, as both downregulation and overexpression of this factor are associated with disease and leukaemia (Ostergaard et al., 2011; Luesink et al., 2012; Vicente et al., 2012).

Importantly, a complex crosstalk between RUNX1, TAL1 and GATA2 was shown to be mediated by BMP signalling. Specifically, the BMP effector SMAD1 was reported to upregulate both RUNX1 and GATA2. The latter, in a complex with TAL1 and FLI1, can reinforce this signal by upregulating SMAD1/5, but also upregulate SMAD6 which represses RUNX1, therefore constituting a complex circuit for the temporal and balanced activation of RUNX1 (Oren et al., 2005; Marks-Bluth et al., 2015). Remarkably, RUNX1 was also shown to bind TAL1, causing a global reorganization of its genomic binding pattern, perhaps important for the shift of TAL1 role from the primitive to the definitive haematopoietic programme (Lichtinger et al., 2012). Furthermore, ChIP-sequencing experiments show that RUNX1, GATA2 and TAL1 are often found together in association at genomic sites (Wilson et al., 2010). Overall, these findings delineate a finely tuned crosstalk between these three master regulators of haematopoiesis, whose expression and role are finely controlled, change during development and are highly context-dependent.

## 1.8 Cell cycle

The cell cycle involves all the events leading to DNA duplication and cell division in order to produce two daughter cells, and it is therefore one of the most fundamental molecular

mechanisms. However, its role extends beyond cell growth and division, as it is also associated with other fundamental processes such as DNA repair, morphogenesis and cell fate decision. As such, dysregulations in the control of the cell cycle can lead to aberrant proliferation and cancer. Importantly, cell cycle regulation and key components of its molecular machinery have been proven to drive haematopoietic development and homeostasis.

### 1.8.1 Cell cycle phases

The cell cycle is commonly divided into interphase and mitotic phase. Interphase involves a variety of biochemical processes that prepare the cell for division happening during mitosis. The various stages of interphase are usually not morphologically distinguishable. Nevertheless, it can be further divided into G1 (gap 1), S (synthesis) and G2 (gap 2) phases, each characterised by distinct biochemical processes, and together with the M (mitosis) phase they constitute the most common subdivision in four consecutive cell cycle phases (Schafer, 1998).

The main events characterising cell cycle are DNA replication, taking place during S phase, and cell division, in M phase, both of them complex processes characterised by high fidelity and quality controls to ensure the correct generation of two daughter cells, and with safeguard mechanisms that allow DNA repair upon duplication errors or to get rid of cells with an abnormal chromosomal asset. The remaining gap phases are essentially preparation stages making sure everything is in place for the subsequent events, but also checkpoints integrating different signals to determine cell cycle progression or arrest. Specifically, the G1 phase involves most of the biosynthetic activity of the cell, resumed immediately after the previous division, and consisting of protein synthesis, increase in number of organelles and growth in cell size. This phase includes key control mechanisms making sure that everything is ready for DNA duplication, but also linking the cell cycle state to the external environment. Indeed, these control mechanisms collectively constitute checkpoints, or restriction points, representing moments determining if the cell will proceed into the S phase, remain in G1 until specific conditions apply such as growth signals or nutrients availability, or exit the cell cycle and enter

a longer reversible quiescent state also known as G0. In this context, it is important to note that adult HSCs are well known to mostly reside in G0, and to only re-enter the cell cycle when needed in order to satisfy blood demand. These checkpoints are of fundamental importance, as once the cell enters the S phase, it is usually committed for cell division. As such, the duration of G1 is highly variable, and dependent on developmental stage, cell type and external conditions. The G2 phase occurs instead immediately after DNA duplication and corresponds to a period of rapid growth, protein synthesis and early microtubule reorganization, events that are necessary for the subsequent mitosis.

### 1.8.2 Cyclins and CDKs

Cell cycle progression needs to be tightly regulated, on one side to detect and repair DNA lesions and ensure that daughter cells inherit the correct genetic material; on the other to prevent uncontrolled proliferation which would disrupt tissue homeostasis and ultimately lead to cancer. Two classes of molecules are responsible for this regulation, functioning as both checkpoints for the progression into the subsequent cell cycle phase, and integrating the complexity of the signalling pathways instructing cell proliferation. These regulators are cyclins and cyclin-dependent kinases (CDKs), and include several members that function as a complex in specific combinations during distinct cell cycle phases. In these complexes, the CDKs constitute the effector subunits, having phosphorylation activity upon binding to their specific cyclins, which are instead the regulatory subunits. Specific cyclin-CDK complexes are activated at distinct cell cycle phases. The CDKs are usually constitutively expressed, while cyclin levels change in a periodic way and, upon expression, they bind to their CDK partner and allow cell cycle progression, for example inducing the expression of target genes (Figure 1.9). The specific downstream effect depends on the cyclin-CDK combination (Schafer, 1998).

Signals promoting cell cycle progression such as growth factors, induce the expression of D cyclins during G1 which bind to CDK4 and CDK6. The cyclin D-CDK4/6 complex phosphorylates and inactivates the transcriptional repressors of the Rb family, consisting of

RB1, RBL1/p107 and RBL2/p130. These proteins restrict cell cycle entry and maintain the cell in G1 by binding and repressing E2F, thus preventing the activation of genes involved in cell cycle progression. Upon phosphorylation by cyclin D-CDK4/6, the Rb members are partially inactivated and E2F activates the transcription of genes necessary for the transition to the next phase, like cyclin Es and cyclin As. As a result, cell cycle progresses through G1. In late G1, additional phosphorylation by the now active cyclin E-CDK2 complex further inactivates the Rb-mediated inhibition of E2F, resulting in a complete G1 exit and entry into the S phase. At this stage, cyclin A-CDK2 is the main active complex, which induces the expression of proteins involved in DNA replication and drives the further transition through the S phase. Once accumulated, cyclin As also bind to CDK1, and the cyclin A-CDK1 complex starts the G2/M transition. During G2 there is also expression of cyclin Bs, whose peak happens slightly later compared to cyclin As. The resulting cyclin B-CDK1 complex further promotes the G2/M transition, subsequently causing the breakdown of the nuclear envelope and promoting mitosis during M phase (Poon, 2016).

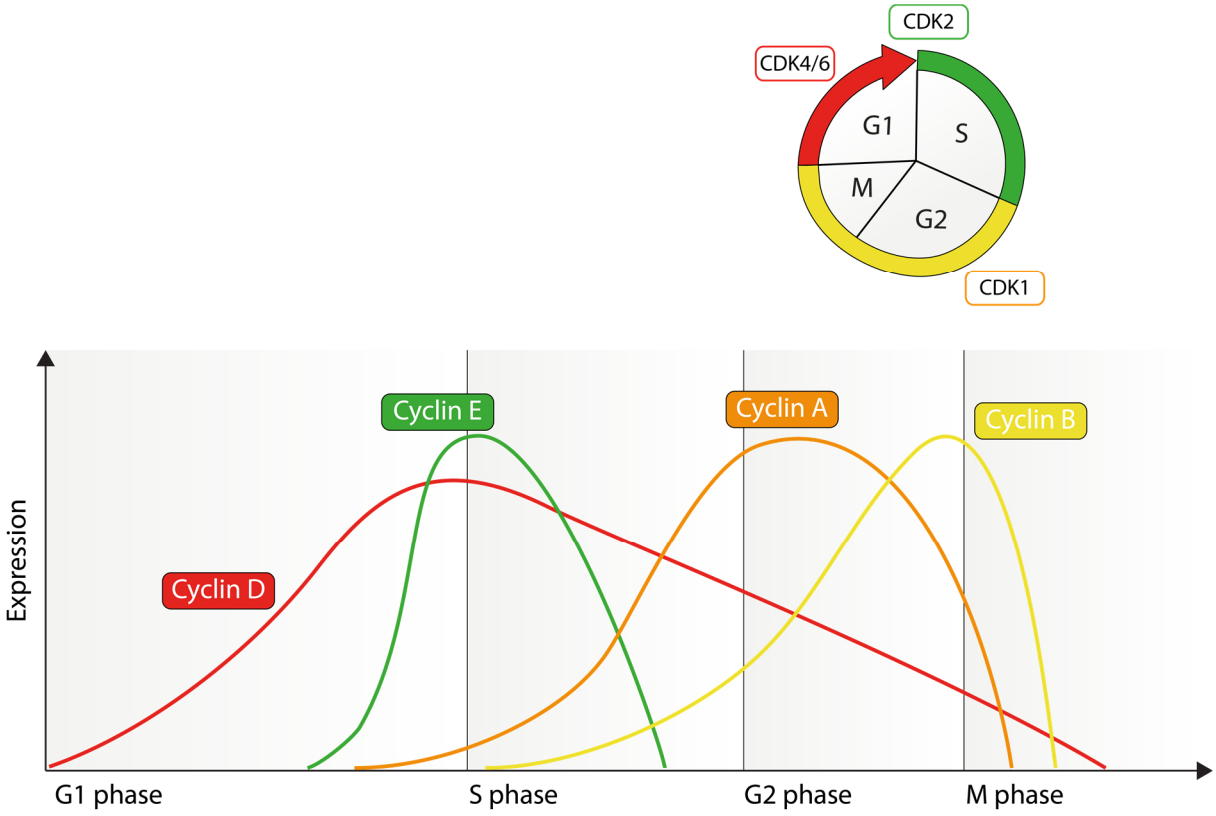



**Figure 1.9 Cell cycle regulation.** The dynamic expression of regulatory cyclins is responsible for the timely activation of specific CDKs and for the progression through successive cell cycle phases.

The machinery directing cell cycle progression is tightly regulated, not only by the temporal upregulation of cyclins but also by other factors which can inhibit cyclin-CDK complexes, especially during G1, when the commitment to enter cell cycle is established. Beyond the aforementioned Rb proteins, many other factors contribute to this tight regulation. One example are the members of the INK4 family, that include p16, p15, p18 and p19, and whose role is to block CDK4/6 and thus restrict cells in G1. Other key CDK inhibitors are the proteins of the CIP/KIP family p21, p27 and p57, which can inhibit both CDK4/6 in G1 and CDK2 in S phase. These and other factors contribute to drawing the complex regulatory landscape that determines cell cycle progression or arrest (Pietras et al., 2011).

Remarkably, increasing evidence over the years has shown a variety of non-canonical functions for both cyclins and CDKs. Direct kinase-independent transcriptional roles were for example reported for G1 cyclins and CDKs. Indeed, Cyclin D1 can directly bind to and inhibit p300 and CBP, therefore inhibiting their acetyltransferase activities and repressing various promoters (McMahon et al., 1999). Importantly, during haematopoietic differentiation both cyclin D3 and CDK6 can instead individually interact with RUNX1 and reduce its ability to bind to DNA and activate target genes. This activity does not require binding to their partner or a kinase activity, and was shown to block myeloid differentiation (Peterson et al., 2005; Fujimoto et al., 2007). Furthermore, CDK6 can also induce the transcription of VEGF, therefore promoting angiogenesis (Kollmann et al., 2013). Additionally, these regulators have been reported to have key roles in other fundamental processes. In hPSCs, the cyclin D-CDK4/6 complex can direct cell fate decision, and was indeed shown to phosphorylate SMAD2/3 preventing their entry into the nucleus and the activation of target genes driving mesendoderm differentiation (Pauklin and Vallier, 2013). Cyclin D1 was reported to direct DNA repair by binding chromatin on irradiation-induced double-strand breaks and recruiting RAD51, therefore promoting homologous recombination (Jirawatnotai et al., 2011). Cyclin B1-CDK1 was instead reported to localise to the matrix of mitochondria and to phosphorylate various mitochondrial proteins, including components of the respiratory chain, leading to an increased mitochondrial respiration. This was hypothesised to allow the cell to sense and respond to the increased energy demand upon G2/M transition (Wang et al., 2014b). Finally, Cyclin E/A-CDK2 complexes play a major role in tumorigenesis, by phosphorylating MYC and allowing it

to suppress RAS-induced senescence upon damage, thus promoting tumour formation (Hydbring et al., 2010).

These are just a few examples of the non-canonical functions of cyclins and CDKs that have emerged over recent years, including transcription, DNA damage repair, metabolism, immune response, cell fate decision and differentiation (Hydbring et al., 2016). Ultimately, this represents a tight link between essential cellular mechanisms and cell cycle progression, and adds an additional layer of control to these processes. Indeed, cyclins and CDKs have emerged not only as regulators of cell cycle progression, but also as part of multiple and complex regulatory mechanisms.

### 1.8.3 Cell cycle and haematopoiesis

The earliest stages of embryonic development are characterised by short cell cycle, limited G1 and G2 phases and quick cell division, a feature observed also in their *in vitro* pluripotent stem cell counterpart. Studies in murine embryos and in mouse and human pluripotent stem cells indicate that once cells start to commit during gastrulation, they acquire a longer cell cycle, most notably indicated by an elongated G1 phase, and their proliferation rate decreases. This remodelling of the cell cycle profile has been linked to the exit from pluripotency, with the G1 phase representing a time window during which most of the cell fate decision is undertaken (Pauklin et al., 2016; Boward et al., 2016).

Several genetic studies using mouse models have contributed over the years to shed light on the role of distinct cell cycle regulators during development, and importantly highlighted the importance of cyclins and CDKs during haematopoiesis. Remarkably, many of the aforementioned factors possess considerable functional redundancy, as shown by the modest effects observed when individual regulators are disrupted. Indeed, lack of either of the cyclin D members is associated only with minimal haematopoietic defects during development, the major impact observed with cyclin D3 affecting T lymphocyte production (Sicinska et al., 2003). Mutations in cyclin D2 and D3 appear to be more relevant in the adult, as shown by studies

associating these genes with an increased incidence of lymphoma (Fantl et al., 1995; Metcalf et al., 2010). However, the most striking effect is observed in mice deficient in all the three D cyclins, which die during late embryogenesis due to heart abnormalities and severe haematopoietic defects, characterised by significant reduction in the number of peripheral red blood cells. Importantly, these mice show a strong reduction in the numbers of HSPCs found in their foetal liver, and these fewer cells appear to be enriched in the G1 phase of the cell cycle and unable to provide even short-term reconstitution upon transplantation (Kozar et al., 2004). Similarly, individual loss of CDK4 and CDK6 produces viable embryos, and only CDK6 is associated with slight haematopoietic impairment. However, CDK4/6 double knockout mice are characterised by late embryonic lethality associated with severe anaemia and defects in foetal haematopoiesis, phenotypically very similar to the lack of cyclin D1/D2/D3 (Malumbres et al., 2004). As expected, this phenotype is also confirmed by the triple knockout of all the three interphase kinases CDK4/6/2. Surprisingly, the severe haematopoietic impairment causing embryonic lethality also indicates that the cell cycle is able to progress and support proliferation and morphogenesis at least until this late stage, when the lethal haematopoietic phenotype becomes evident. This suggests that the remaining CDK1 alone is able to bind interphase cyclins and thus support the whole cell cycle in the absence of CDK4, CDK6 and CDK2, at least for early development (Santamaría et al., 2007a). This is confirmed by experiments showing that instead mice devoid of CDK1 fail to complete more than few divisions and die very early at the blastocyst stage (Diril et al., 2012). Overall this suggests that CDK1 is the principal regulator of the cell cycle, essential for its progression, and suggests that the remaining regulators, during development and possibly in the adult, might be more important for an accurate fine-tuning of this regulation and a balanced cell cycle progression, in addition to cell cycle independent, cell type-specific functions that would explain the haematopoietic phenotypes observed.

As previously mentioned, adult HSCs are mostly quiescent and reside in the G0 phase, albeit ready for a rapid re-entry into the cell-cycle to respond to haematopoietic demand. This state is thought to be necessary for HSCs to persist for a lifetime, carefully balancing quiescence, proliferation and commitment towards differentiation in a tightly regulated manner. This ensures blood homeostasis while avoiding HSCs depletion and the accumulation of DNA damage. The specialised adult bone marrow niche is responsible for the maintenance of

quiescence, which is believed to contribute to HSC longevity and preserve their function at least in part by minimising stresses derived by DNA replication and cellular metabolism (Eliasson and Jönsson, 2010). Almost all the foetal liver HSCs (FL-HSCs) are actively cycling, possibly to expand the stem cell pool and at the same time satisfy blood demand during foetal development, and only later in the foetal liver and more prominently after seeding the bone marrow (BM-HSCs), they gradually acquire the quiescent phenotype (Bowie et al., 2006). It is currently unclear if the acquisition of quiescence is necessary to complete their maturation towards an adult phenotype. Remarkably, both for FL- and BM-HSCs, only the fraction of cells found in G0/G1 is able to engraft and repopulate a host upon transplant, while cells in S/G2/M are temporarily devoid of this property. However, upon completing the cell cycle and reaching G1, these cells reacquire engraftment ability (Bowie et al., 2006). A similar finding was obtained in human, using cord blood HSCs, although in this case cells were maintained *ex vivo* for 5 days before transplant, which might affect their self-renewal ability (Glimm et al., 2000). Nevertheless, these studies delineate an important link between cell cycle state and HSC regenerative potential.

Importantly, the exit from quiescence is a tightly regulated process. In this respect, long-term (LT-) and short-term (ST-) HSCs show different division kinetics which depend on their cell cycle machinery. Despite both equally residing in quiescence, ST-HSCs show immediate re-entry into the cell cycle upon mitogenic signal, while LT-HSCs show a G0 exit delayed by 5-6 hours. This is caused by the absence of CDK6 in LT-HSCs, which needs to be expressed before cell cycle entry can be engaged (Laurenti et al., 2015). Overall this possibly represents a safeguard control mechanism for preserving the HSC pool, on one side ensuring that LT-HSCs only exit G0 upon sustained signals, on the other maintaining a population of ST-HSCs already primed to quickly enter the cell cycle in order to address the haematopoietic demand.

## 1.9 Objectives

As delineated in the previous sections, haematopoietic stem cells are of fundamental importance for the treatment of a variety of life-threatening diseases. However, their scarce availability constitutes a major obstacle, and a number of approaches have been deployed for their production *in vitro*, efforts which still remain unsuccessful. For that, understanding key molecular mechanisms driving their generation during development is essential. However, despite substantial progress that has been made over the years in this direction, our knowledge still remains uncomplete.

The aim of this dissertation is to contribute to this knowledge by gaining insights into the fundamental mechanisms driving the onset of early haematopoiesis and that *in vivo* lead to the specification and emergence of the first haematopoietic stem cell. Towards this end, the objective of this study is to answer the following questions:

- How does the endothelial-to-haematopoietic transition proceeds and what are the dynamics between the cell types involved?
- What are the molecular mechanisms still unknown but fundamental for this transition?
- What is the relationship between the cell cycle state and haematopoietic specification, and can we modulate key cell cycle regulators in order to control the generation and maturation of haematopoietic cells *in vitro*?

For this, I took advantage of multiple approaches that will be discussed in the following chapters.

# PART II

# EXPERIMENTAL RESULTS

# 2 OPTIMISATION AND CHARACTERISATION OF A MODEL SYSTEM TO STUDY HUMAN HAEMATOPOIETIC SPECIFICATION

## 2.1 Summary

Understanding the molecular mechanisms controlling specification and generation of HSPCs during early human development is challenging for obvious technical and ethical reasons. Indeed, developmental processes like the onset of the foetal definitive haematopoietic wave taking place in the yolk sac and the AGM are not easily accessible *in vivo* in human. Thus, the development of an *in vitro* model for HSPC production would address this major challenge. Human pluripotent stem cells (hPSCs) represent an advantageous platform to achieve this objective. Indeed, hPSCs provide a stable and unlimited source of cells, avoiding the effects of the genetic diversity of primary cells which often complicate the interpretation of molecular mechanisms. Furthermore, differentiation of hPSCs offers an unprecedented opportunity to study mechanisms controlling natural paths of development *in vitro*. Consequently, I decided to take advantage of a recently published protocol for differentiating hPSCs into HPCs. This method recapitulates the early stages of human development for the production of progenitors showing characteristics of definitive haematopoiesis (Kennedy et al., 2012; Sturgeon et al., 2014). This chapter describes the optimisation of this protocol and the characterisation of the resulting haematopoietic cells.

## 2.2 Establishment and optimization of the first stages of differentiation

The first step of this protocol consists in differentiating hPSCs as embryoid bodies (EB) in the presence of a cocktail of growth factors and small molecules inducing mesoderm

patterning and directing endothelial and haematopoietic specification, thus following developmental events taking place in the early embryo (Figure 2.1).

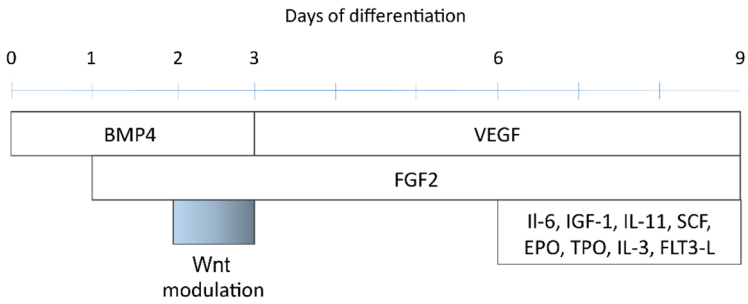



**Figure 2.1 Schematic of the differentiation system.** In this model, during the first stage of differentiation EBs are directed for 9 days with a combination of growth factors and small molecules.

More specifically, this approach aims at generating a mixed population of progenitors after 8 to 10 days characterized by the expression of the CD34 and CD43 surface markers. At this stage CD43 marks a more primitive population mainly capable of erythroid potential, while CD34+/CD43- cells represent instead progenitors of definitive foetal haematopoiesis (Kennedy et al., 2012). Importantly, an early stimulation of WNT signalling is able to enrich for these definitive progenitors (Sturgeon et al., 2014).

In order to maximise the generation of this population, my first step was the optimisation of the mesoderm stage. As previously discussed (1.7), a correct recapitulation of the earlier stages of development is critical, especially for what concerns the induction and subsequent patterning of the specific mesoderm subtype capable of further differentiating towards an endothelial and then a haematopoietic fate. For this, I focused on modulating BMP4 and FGF2 signalling pathways during the earliest days of differentiation, with the aim of achieving the highest number of CD34+/CD43- cells before further characterisation. Specifically, I first tested different concentrations of BMP4 during mesoderm specification, corresponding to the first 3 days of the differentiation protocol. This was initially done in the presence of low concentrations of FGF2 (2 ng/ml) and the small molecule CHIR99021, inhibitor of GSK3 and thus indirect activator of the WNT canonical pathway (applied as shown in Figure 2.1).

The subsequent differentiation was then performed using concentrations of growth factors as previously described (Sturgeon et al., 2014), carried on until day 9 (D9) and then tested by flow cytometry for the expression of the surface markers CD34 and CD43. These experiments revealed that different concentrations of BMP4 caused only a minor change in the expression of CD34, but an important effect on the number of CD43+ cells. Based on this result, I determined that the best dose of BMP4 for mesoderm induction and patterning was 10 ng/ml (Figure 2.2).

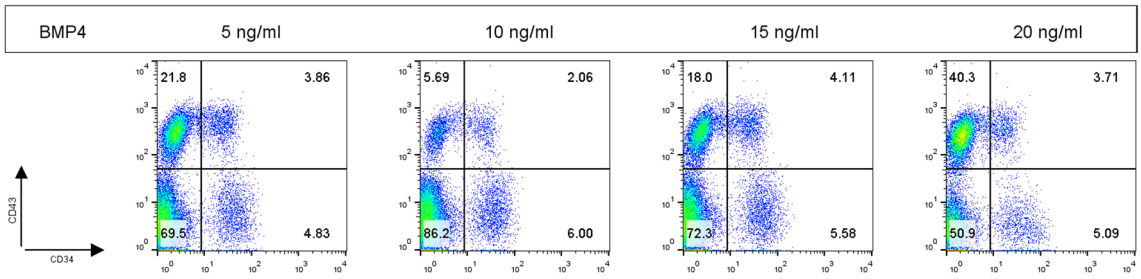



**Figure 2.2 Optimisation of BMP4 concentration during mesoderm specification.** Differentiation was performed using low levels of FGF2 (2 ng/ml) and testing different concentrations of BMP4 during mesoderm induction and patterning. It was carried on until D9 following the standard protocol, and expression of CD34 and CD43 was assessed by flow cytometry. n=1 experiment.

Subsequently, I repeated these experiments using the selected dose of BMP4, and testing for different concentrations of FGF2. In this case, the number of CD34+ cells was affected, with only small changes for CD43 (Figure 2.3).

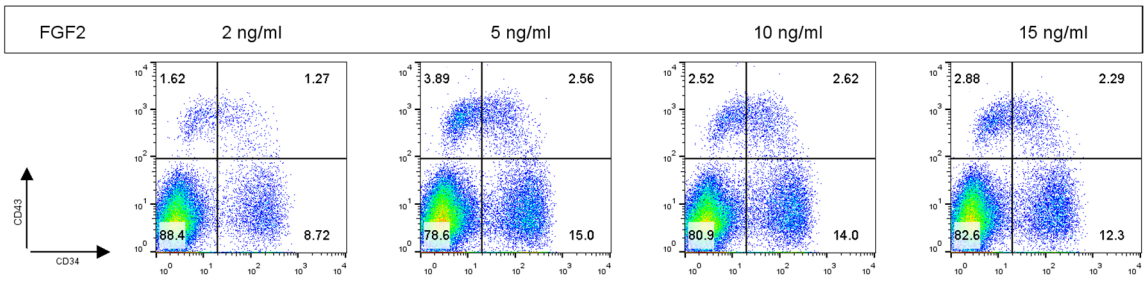



**Figure 2.3 Optimisation of FGF2 concentration during mesoderm specification.** Differentiation was performed using the best concentration of BMP4 as determined from the previous experiment (10 ng/ml) and testing different concentrations of FGF2 during mesoderm induction and patterning. It was carried on until D9 following the standard protocol, and expression of CD34 and CD43 was assessed by flow cytometry. Graphs are representative of technical duplicates from n=1 experiment.

Ultimately, this approach allowed me to select 10 ng/ml of BMP4 and 5 ng/ml of FGF2 as the best combination to achieve induction and patterning of the correct mesoderm subtype competent for the generation of CD34+/CD43- cells.

As briefly mentioned, the CD43 compartment at this early stage represents a population of primitive embryonic progenitors mostly capable of differentiation into erythroid cells. On the other hand, the CD34+/CD43- population is not supposed to have haematopoietic colony forming capability, but instead to represent progenitors with the potential of further differentiating towards a definitive foetal phenotype upon co-culture on OP9 mouse bone marrow stromal cells (Kennedy et al., 2012; Sturgeon et al., 2014). To test this hypothesis, hPSCs were induced to differentiate into mesoderm using the selected conditions and then further differentiated until D9. The resulting cells were then sorted based on their expression of CD34 and CD43 into three populations (CD34+/CD43-, CD34-/CD43+, CD34+/CD43+) and then further differentiated in a colony-forming unit (CFU) assay, in order to assess their potential for differentiation towards the erythroid and myeloid lineages (Figure 2.4). The assay confirmed that at this stage only CD43+ cells were capable of producing haematopoietic colonies. More specifically, CD34-/CD43+ cells were able to give almost exclusively erythroid colonies, while CD34+/CD43+ cells contained a much lower number of colony-forming cells, although of mixed identity. Most importantly, CD34+/CD43- cells were confirmed to not be able to generate any colonies at this stage of differentiation.

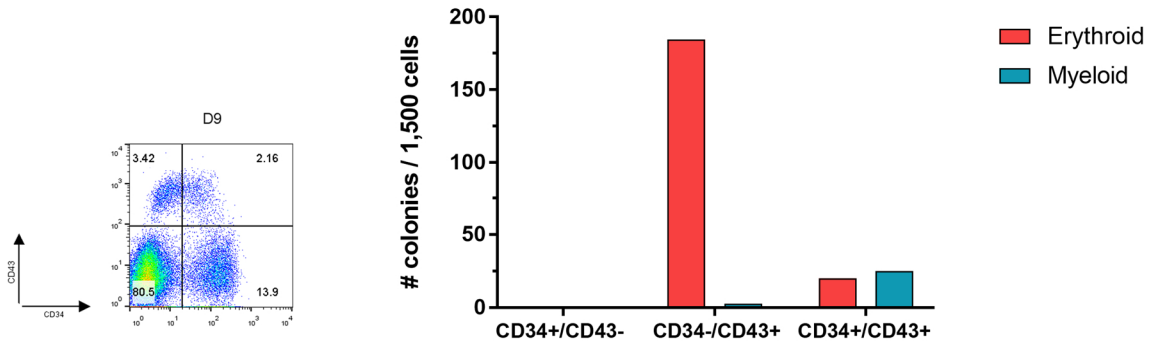



**Figure 2.4 Maturation potential of D9 cells in a CFU assay.** Cells were differentiated until D9 and distinct populations were FACS-sorted based on their expression of CD34 and CD43 surface markers. For each sorted population, 1,500 cells were used for a CFU assay to confirm their maturation potential. n=1 experiment.

During the establishment and optimisation of these first steps of differentiation, an updated version of the protocol became available (Ditadi et al., 2015). The first stage of this revised protocol had undergone only minor changes, still relying on an EB system for mesoderm induction and patterning, and for endothelial specification. However, a major improvement was included in the second stage by avoiding the use of feeder cells. Indeed, cells sorted as CD34+/CD43- were directly plated on Matrigel-coated plates and cultured as a 2D monolayer to undergo an endothelial-to-haematopoietic transition similar to events taking place *in vivo* during foetal development. This improved approach had the important and attractive feature of being entirely feeder-free and relying on the use of well-defined culture conditions, therefore eliminating the requirement for OP9 mouse bone marrow stromal cells. For these reasons, I decided to continue with this revised version of the differentiation protocol.

As briefly mentioned, in addition to this important improvement, few alterations also appeared to have been made to the first stage of the protocol, specifically a reduction in the number of cytokines applied after D5, a slight change in the timing of their administration and an anticipation to D8 for the sorting of CD34+/CD43- cells (Figure 2.5).

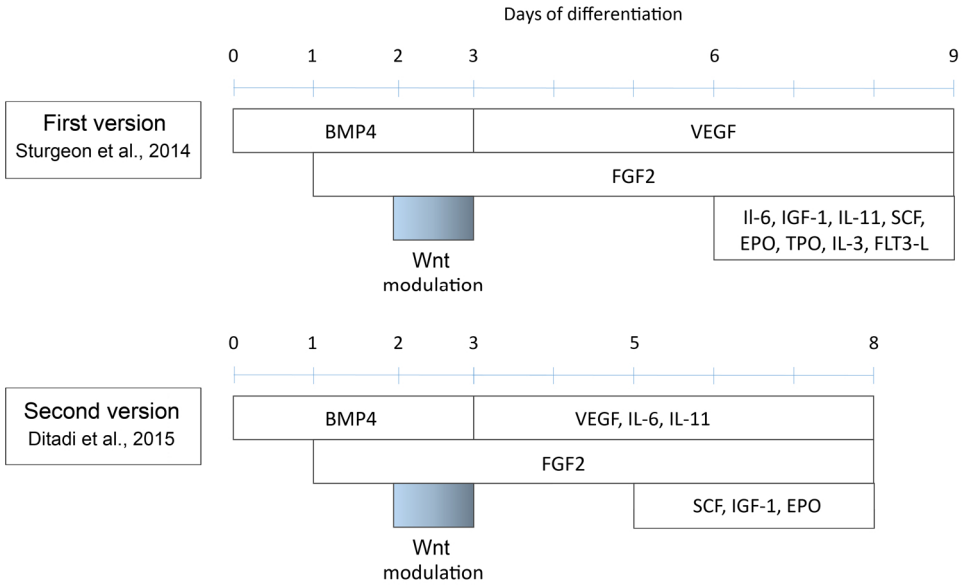



**Figure 2.5 Comparison between the first stage of the two protocols.** Here a schematic of the first stage of differentiation is shown, comparing the two versions.

By comparing the two methods using mesoderm induction and patterning conditions as previously selected, I was able to show that the revised protocol generates a higher number of CD34+/CD43- cells at D8, one day earlier compared to the previous version (Figure 2.6). Therefore, I decided to use the new version of the protocol for the production of the intermediate CD34+/CD43- population and to test the second stage of differentiation modelling the endothelial-to-haematopoietic transition.

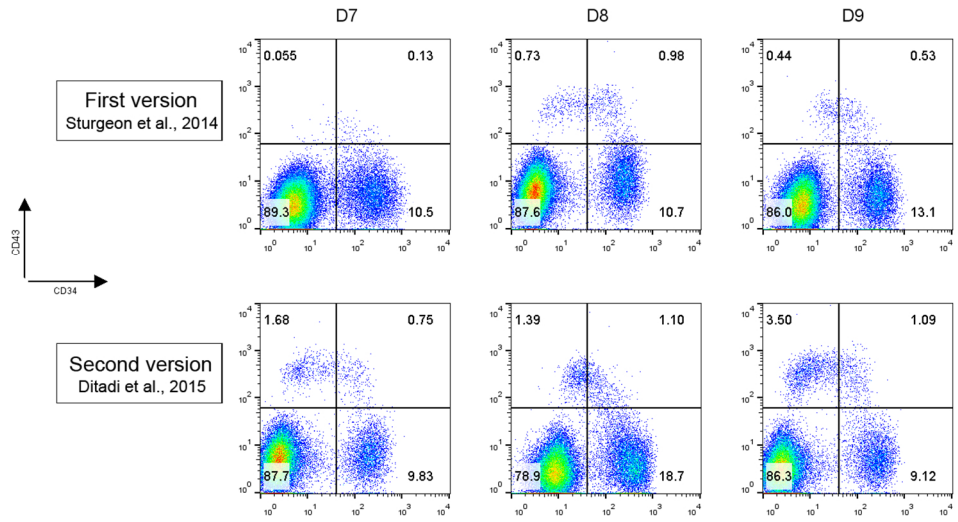



**Figure 2.6 Comparison between the two protocols for the generation of CD34+/CD43- cells.** Differentiation was performed using the two systems described, and assessing the expression of CD34 and CD43 by flow cytometry between D7 and D9 of differentiation. This approach allowed me to select the best protocol and determine the trend of expression for the markers analysed. Graphs are representative of technical duplicates from n=1 experiment.

## 2.3 Characterisation of the endothelial-to-haematopoietic transition stage

In order to use this *in vitro* model to uncover mechanisms controlling early haematopoiesis, after a first optimisation of the system my efforts were then focused on characterising the generation of haematopoietic progenitor cells (HPCs). Ultimately, my aim was to identify the exact time window during which HPCs are first generated from haemogenic endothelial cells (HECs) *in vitro*.

For this, CD34+/CD43- cells sorted at D8 were re-aggregated overnight into small clumps and seeded the day after onto Matrigel-coated plates, marking the first day of endothelial-to-haematopoietic transition culture (EHT D1). This allowed for the cells to attach and grow to form a layer of endothelial cells at EHT D3, from which during the next 2 days by EHT D5 clusters of round cells were produced, resembling intra-aortic haematopoietic clusters (IAHCs) generated *in vivo* (Figure 2.7). These clusters appeared to keep attachment to the underlying endothelium, and on the subsequent days they showed an increase in their size and most importantly the generation of new cells which started to be released as single cells floating in the culture medium, with their numbers increasing dramatically on the following days.

This process appeared to morphologically recapitulate developmental events taking place during EHT in the foetal yolk sac and AGM region, corresponding to the release from the haemogenic endothelium (HE) of the first HSPCs which then migrate to the foetal liver for further maturation and expansion.

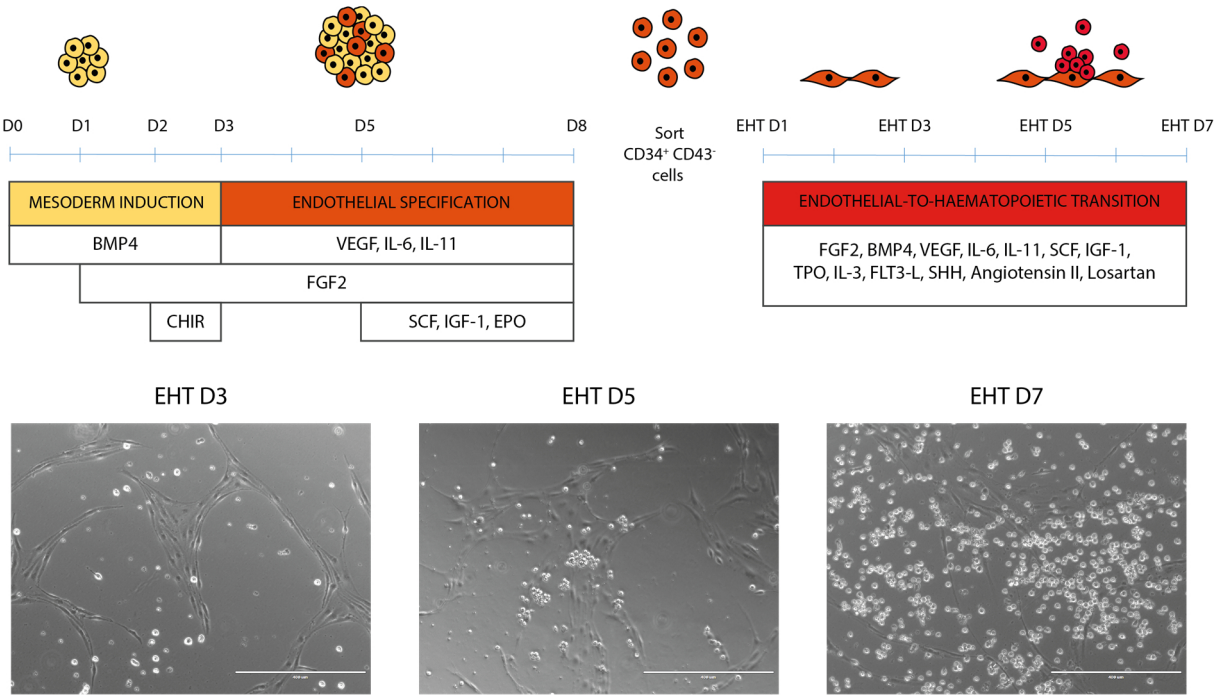



**Figure 2.7 Culture system for *in vitro* modelling of EHT.** The protocol allows modeling the generation of IAHC from the HE, which produces and releases cells in the culture media. Scale bar is 400 µm.

The endothelial identity of cells used for this second stage of differentiation was confirmed by the co-expression of CDH5. Indeed, this endothelial marker appeared to be

expressed on the majority of the CD34+ cells sorted at D8 (Figure 2.8A). Importantly, during the EHT culture the number of cells expressing this marker rapidly decreased, concomitant with an increase in the number of cells expressing the pan-haematopoietic marker CD43, first surface marker reported to appear on HSPCs generated both *in vitro* and *in vivo* (Vodyanik et al., 2006; Ivanovs et al., 2014a). The inverse correlation for the expression of these markers illustrated the transition from an endothelial to a haematopoietic cell identity (Figure 2.8B).

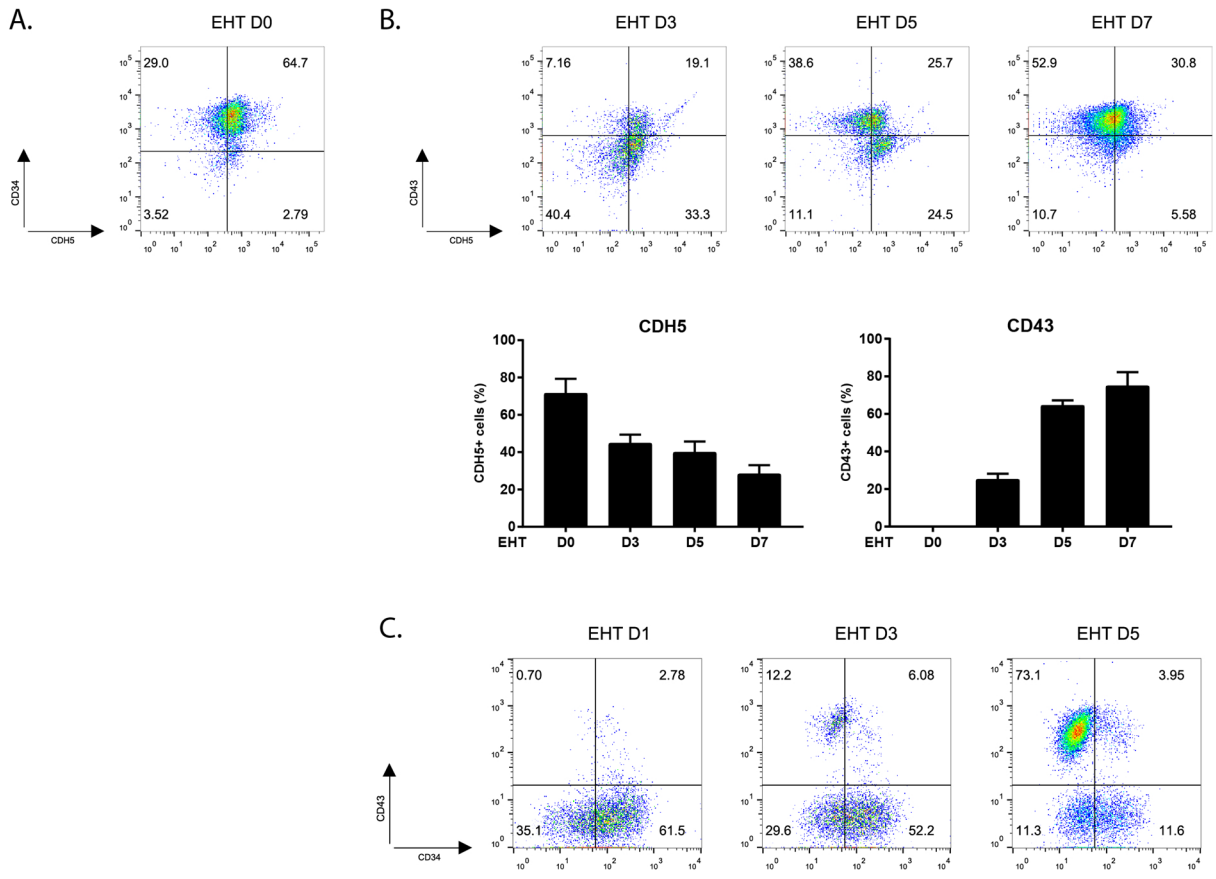



**Figure 2.8 Characterisation of EHT by flow cytometry. (A)** CD34+/CD43- cells were sorted at D8, marking the initial day of EHT culture (EHT D0). Most of these cells also co-expressed the endothelial marker CDH5. **(B)** On the subsequent days, EHT was characterised by the acquisition of CD43 marking the first haematopoietic cells generated, concomitant with the loss of CDH5. Results in (A) and (B) represent n=3 independent experiments, with error bars ±SEM. **(C)** The first haematopoietic cells generated during EHT were CD34+/CD43+. Subsequently, most of these cells lost the expression of CD34. n=1 experiment.

This transition was also evident by looking at the expression of key marker genes by quantitative polymerase chain reaction (qPCR; Figure 2.9). The expression of endothelial markers, such as *VWF*, *PECAM1* and *CDH5*, showed an inverse progression over time

compared to haematopoietic differentiation markers, such as *SPI1*, *GATA1* and *KLF1*. It was also interesting to observe the trend of expression for key markers and transcription factors usually associated with HSC specification and function, like *RUNX1*, *MEIS1*, *GATA2*, *GFI1* and *MYB*. Indeed, all these markers appeared to show a peak of expression at EHT D3, corresponding with the first morphological appearance of the haematopoietic clusters in culture. The observation that these key HSC factors showed the highest expression at EHT D3 and decreased afterwards, along with the fact that differentiation markers appeared to increase over time, suggested that EHT in these culture conditions was a very transitory process and that once generated, HPCs were either unstable and directly proceeded towards more differentiated cell types, or that they quickly became outnumbered by more mature cell types.

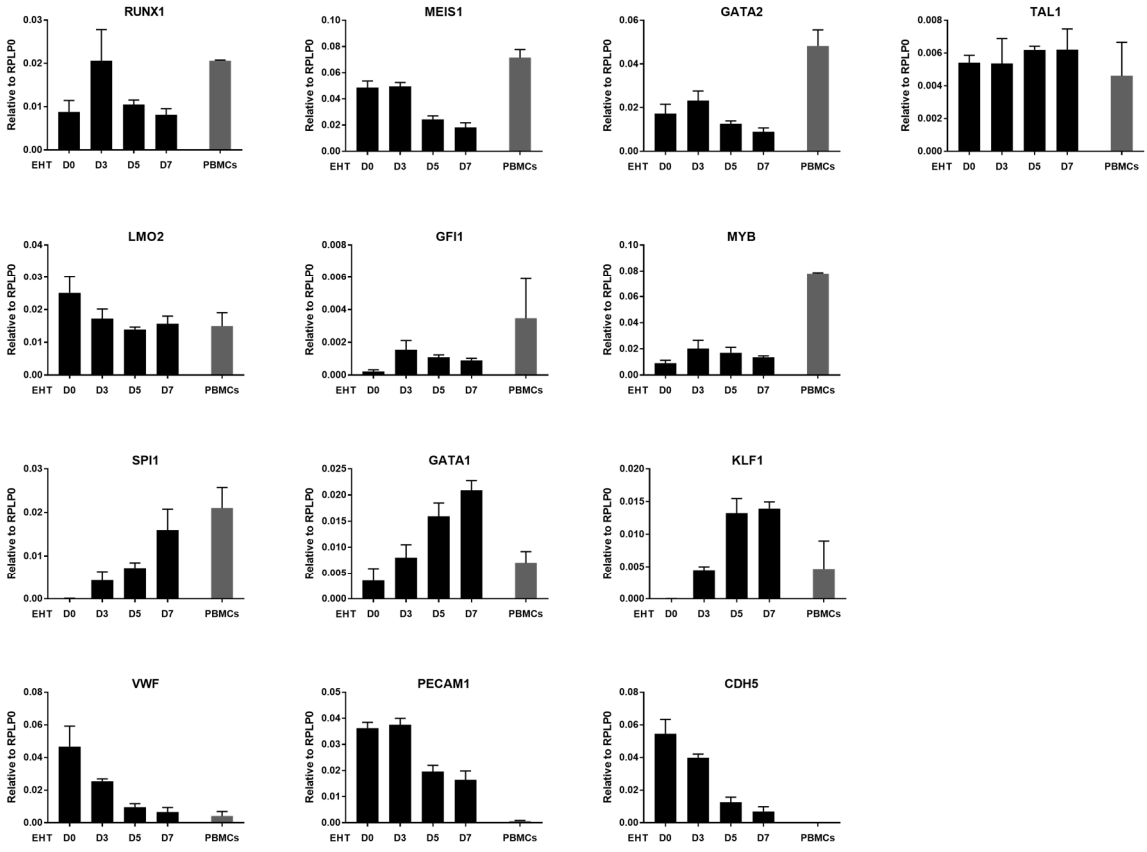



**Figure 2.9 Characterisation of gene expression during EHT.** The expression of key haematopoietic and endothelial genes was measured during EHT by qPCR. CD34+ peripheral blood mononuclear cells (PBMCs) from primary samples are shown as control. Results for EHT cultures represent n=3 independent experiments, with error bars ±SEM. Results for PBMCs represent n=2 independent samples, with error bars ±SEM.

In order to confirm if this culture system was able to generate functional HPCs capable of maturation into different blood cell types, cells generated during EHT were tested for their capability to terminally differentiate. For this, cells were collected at EHT D5 and used for a CFU assay. By the end of the assay, I could observe the production of different colonies of erythroid and myeloid lineages, confirming that the *in vitro* generated HPCs were functional and able to further differentiate. Most importantly I could also confirm the presence of colonies of mixed phenotype, representative of multipotent haematopoietic progenitors (Figure 2.10A,B). These terminally differentiated cells were collected and further analysed by qPCR for the expression of haemoglobin genes. Using this approach, I was able to show that the resulting red cells expressed the foetal γ globin encoded by the *HBG* genes, with lower levels of expression of the embryonic isoform ε globin, encoded by *HBE1* (Figure 2.10C). This result was consistent with previous studies showing a similar pattern occurring during development, when the main site of haematopoiesis is shifting from the AGM region towards the foetal liver and the first haemoglobin isoform switch is taking place with the γ globin becoming predominant (Sankaran and Orkin, 2013). This suggested that HPCs generated in these culture conditions recapitulated the onset of foetal definitive haematopoiesis, as opposed to embryonic primitive haematopoiesis. At the same time, lack of expression of the adult isoform β globin encoded by *HBB* was not surprising, considering that *in vivo* this second switch from γ to β globin only happens shortly after birth.

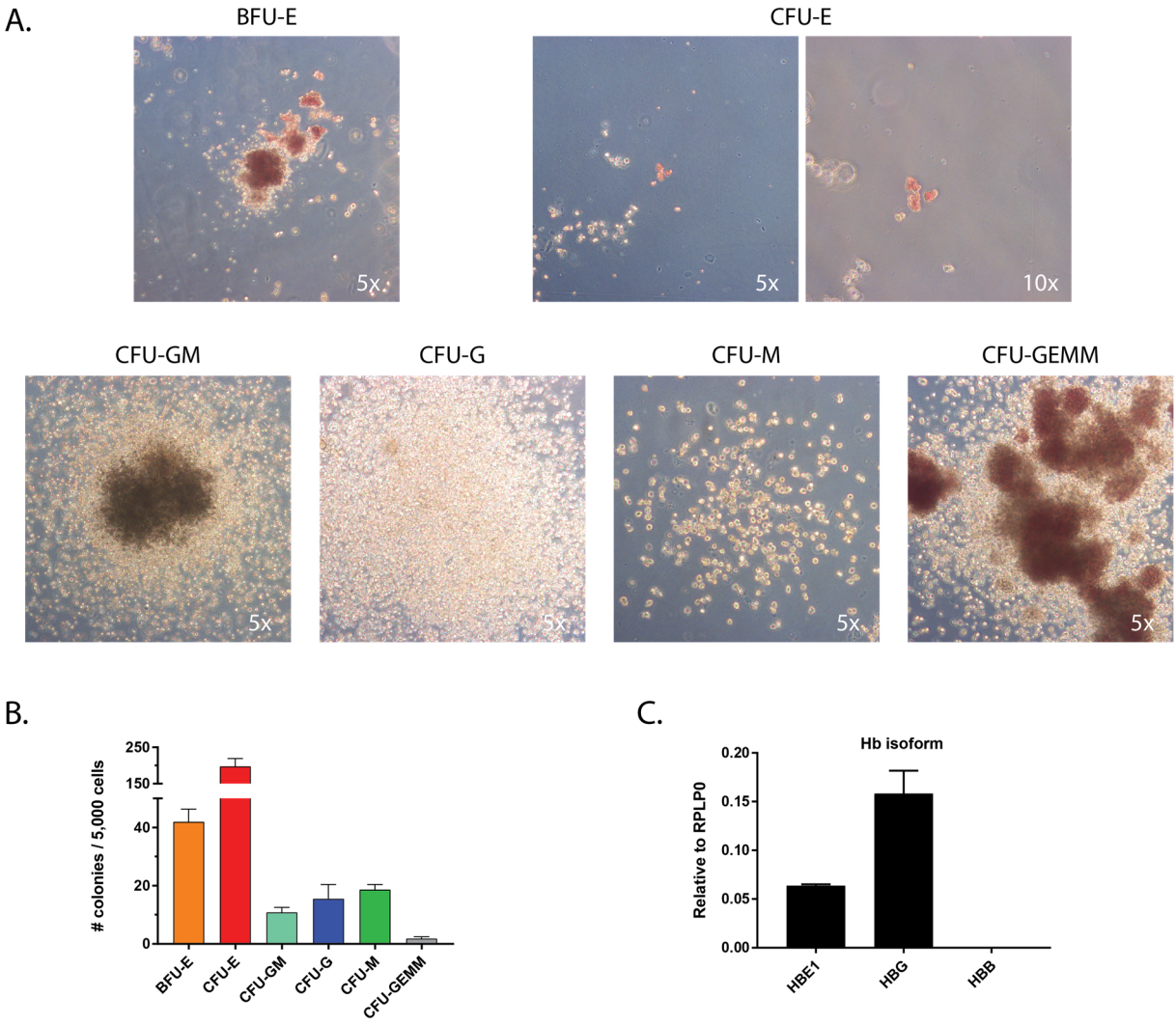



**Figure 2.10 Terminal differentiation of HPCs generated *in vitro*. (A, B)** Cells at EHT D5 were collected and used for a CFU assay. Progenitors generated in these culture conditions were capable of terminal maturation to cells of different blood lineages. **(C)** qPCR data showed that the main β-like globin isoform expressed by red cells generated at the end of the CFU assay was the foetal γ globin. Results represent n=3 independent experiments, with error bars ±SEM

## 2.4 Discussion

In this chapter, I described the optimisation and characterisation of a model system for the recapitulation of early developmental events taking place during foetal haematopoiesis and leading to the generation of HSPCs from the haemogenic endothelium. My strategy was based on the use of hPSCs as an unlimited source of cells to be used for this *in vitro* differentiation, with the aim of uncovering fundamental molecular mechanisms driving the onset of definitive haematopoiesis in human, thus overcoming the difficulty of studying this process *in vivo* due

to limited accessibility. My first efforts were directed towards the optimization of the culture conditions used to generate the intermediate population of CD34+/CD43- cells. Indeed, the exact molecular requirements for the induction of pluripotent stem cells to differentiate show variation across distinct cell lines (Cahan and Daley, 2013). This is caused by genetic and epigenetic heterogeneity, which causes cell line variability in their response to distinct signalling pathways, ultimately resulting in lineage biases. For this reason, I decided to focus on the earliest days of the protocol corresponding to early germ layer development. This is certainly a key step, given that the specification of the correct germ layer is essential for the generation of any specific cell type *in vitro*. Importantly, enrichment for cells representing primitive or definitive haematopoiesis was shown to be mediated by signalling pathways at the mesoderm stage (Slukvin, 2013). Indeed, the two haematopoietic waves appear to derive from distinct mesoderm subtypes, and modulation of WNT signalling is able to control this early cell fate decision (Sturgeon et al., 2014). Of note, a similar role was previously shown for the activation or inhibition of the Activin/NODAL/TGFβ pathway at the same stage, leading to an enrichment of primitive or definitive haematopoiesis, respectively (Kennedy et al., 2012). Thus, one possibility is that redundant mechanisms might control similar differentiation outcomes. Alternatively, a crosstalk between the two signalling pathways might be in place for the regulation of downstream mechanisms directing differentiation. Therefore, a combined modulation of the two signalling could possibly allow for a more accurate control over this early cell fate decision event. For my experiments I performed the activation of the WNT canonical pathway and optimised BMP and FGF signalling during mesoderm induction and patterning. With this approach, I was able to increase the efficiency of generation of CD34+/CD43- cells.

I confirmed this population to represent endothelial cells able to transition to the haematopoietic fate and generate cells resembling the IAHCs produced *in vivo* during development. Importantly, I showed that these cells have the ability to further differentiate in a CFU assay, and I confirmed that the resulting red cells mainly express foetal γ-globin, expressed during the definitive haematopoietic wave. This is also in agreement with previous findings showing this model to possess the potential to generate T lymphocytes (Ditadi et al., 2015), suggested to be a hallmark of definitive haematopoiesis and initially produced during

development from the yolk sac and the AGM region (Böiers et al., 2013; McGrath et al., 2015). In agreement with this, the current protocol was recently suggested to produce cells which are very close to an HSC state. Indeed, haemogenic endothelial cells produced with this differentiation system, but not other types of endothelium such as human umbilical vein endothelial cells, acquired multilineage engraftment potential when induced to ectopically overexpress seven specific transcription factors important for HSC specification and function (Sugimura et al., 2017). Furthermore, the acquisition of reconstitution potential with this approach was not successful when cells were induced prior to EHT D3, suggesting that the onset of EHT is essential for the establishment of the correct molecular landscape necessary for the induction of an HSC-like state. Importantly, this is in agreement with my gene expression data suggesting that the peak of expression of factors considered to be fundamental for HSPC specification takes place at EHT D3. Potentially, this time point represents the onset of transcriptional programmes leading to the first HPCs *in vitro*, and the stage when the induction to an HSC-like state becomes possible. Of note, the morphological appearance of haematopoietic clusters takes place between EHT D3 and EHT D5. Based on these observations, I hypothesised this specific stage to constitute a short and transitory time window during which HPCs are generated *in vitro*. For this reason I selected this narrow interval for my subsequent experiments.

# 3 ANALYSIS OF IN VITRO HUMAN HAEMATOPOIETIC DEVELOPMENT USING SINGLE CELL TRANSCRIPTOMICS

## 3.1 Summary

In the previous chapter, I have shown that the generation of HPCs *in vitro* involved endothelial cells transitioning toward a haematopoietic cell identity, similarly to the EHT happening *in vivo*. This is likely to be a progressive and gradual transition rather than a sudden single step leap, composed of continuous consecutive states. Some cells will be engaged in EHT, and some others will have already completed the transition and potentially will be free to further differentiate. This also implies that the EHT culture is a very heterogeneous system, with different cell types and states coexisting and often sharing similar characteristics, especially cell surface markers. Thus, isolation of different cell states by conventional cell sorting strategies could be difficult if not impossible, rendering the study of molecular mechanisms directing the transition very challenging. To overcome these limitations, I decided to take advantage of single cell RNA sequencing (scRNAseq), which allows the study of the transcriptional landscape of individual cells. As anticipated in the previous chapter, there is a specific time window during which clusters similar to the IAHCs are produced *in vitro* and when the HSPCs transcriptional programme appears to be activated. Therefore, cells involved in the transition at this particular stage were isolated and further characterised at the single cell level.

## 3.2 Identification of distinct cell types involved in the endothelial-to-haematopoietic transition

For the experiment, hPSCs were differentiated as described in Chapter 2.3, and cells were collected at EHT D3 and EHT D5 for scRNAseq. The bioinformatic analyses of the transcriptional data reported in this dissertation were performed with the help of Dr Emmanouil I. Athanasiadis from Dr Ana Cvejic's group.

Briefly, based on transcriptional data and after removal of cells that did not pass quality control, we performed cell clustering in the 3D t-distributed Stochastic Neighbor Embedding (tSNE) space. With this approach we originally identified multiple clusters of cells, and upon calculation of the average expression level of the top 20 marker genes per cluster, we were able to merge highly correlated clusters based on transcriptional similarity, ending up with the final identification of 4 distinct populations of cells in our samples (Figure 3.1A).

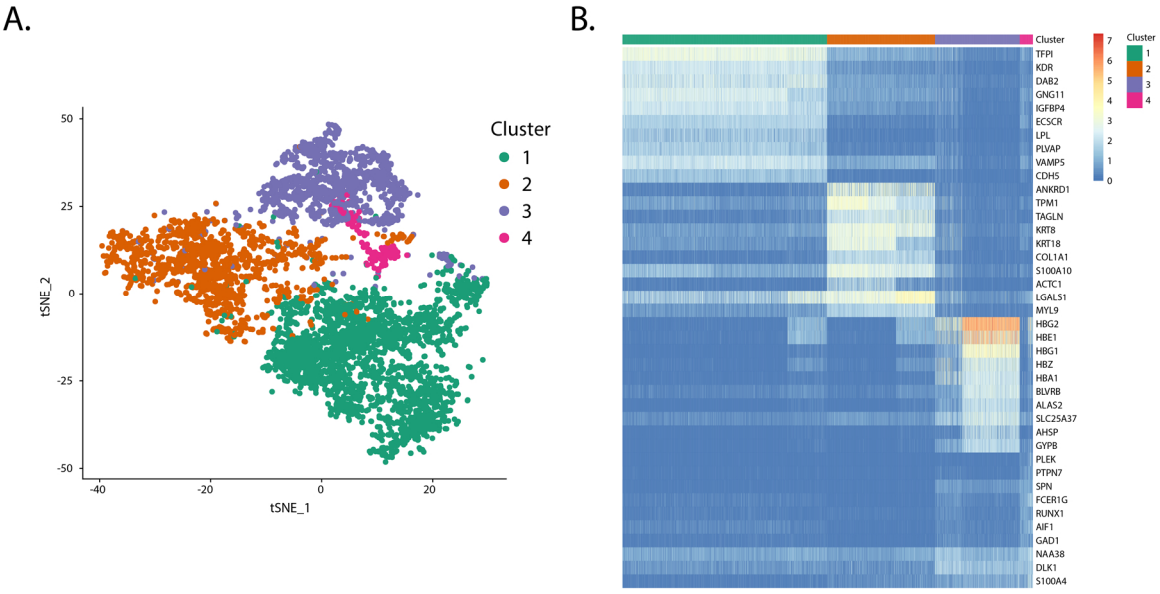



**Figure 3.1 Cell clustering in the 3D tSNE space. (A)** Based on transcriptional identity, we were able to identify 4 clusters in the tSNE space. **(B)** The top 10 marker genes are shown in the heatmap for each cluster. For this analysis, EHT D3 and EHT D5 samples were combined.

We then calculated the marker genes of the resulting clusters (Figure 3.1B) and monitored the expression of known lineage genes to assign cell identity. Based on these results, one group of cells was labelled as endothelial cells, as it was shown to express endothelial

markers including *CDH5*, *KDR*, *PECAM1*, *ECSCR* and *ENG* (Figure 3.2). It was interesting to note that this cluster also expressed low levels of *CD43*, along with intermediate or low levels of key transcription factors involved in HSC specification, such as *RUNX1*, *GATA2*, *TAL1* and *SOX18* (Figure 3.3). Additionally, at least part of this cluster also expressed genes like *SOX17*, known to mark HECs surrounding newly generated IAHCs (Lizama et al., 2015), or *HOXA7*, *HOXA9* and *HOXA10*, suggested to characterize the onset of definitive haematopoiesis and the generation of the first HSCs *in vivo* (Ng et al., 2016; Dou et al., 2016). These observations were therefore consistent with an HE profile.

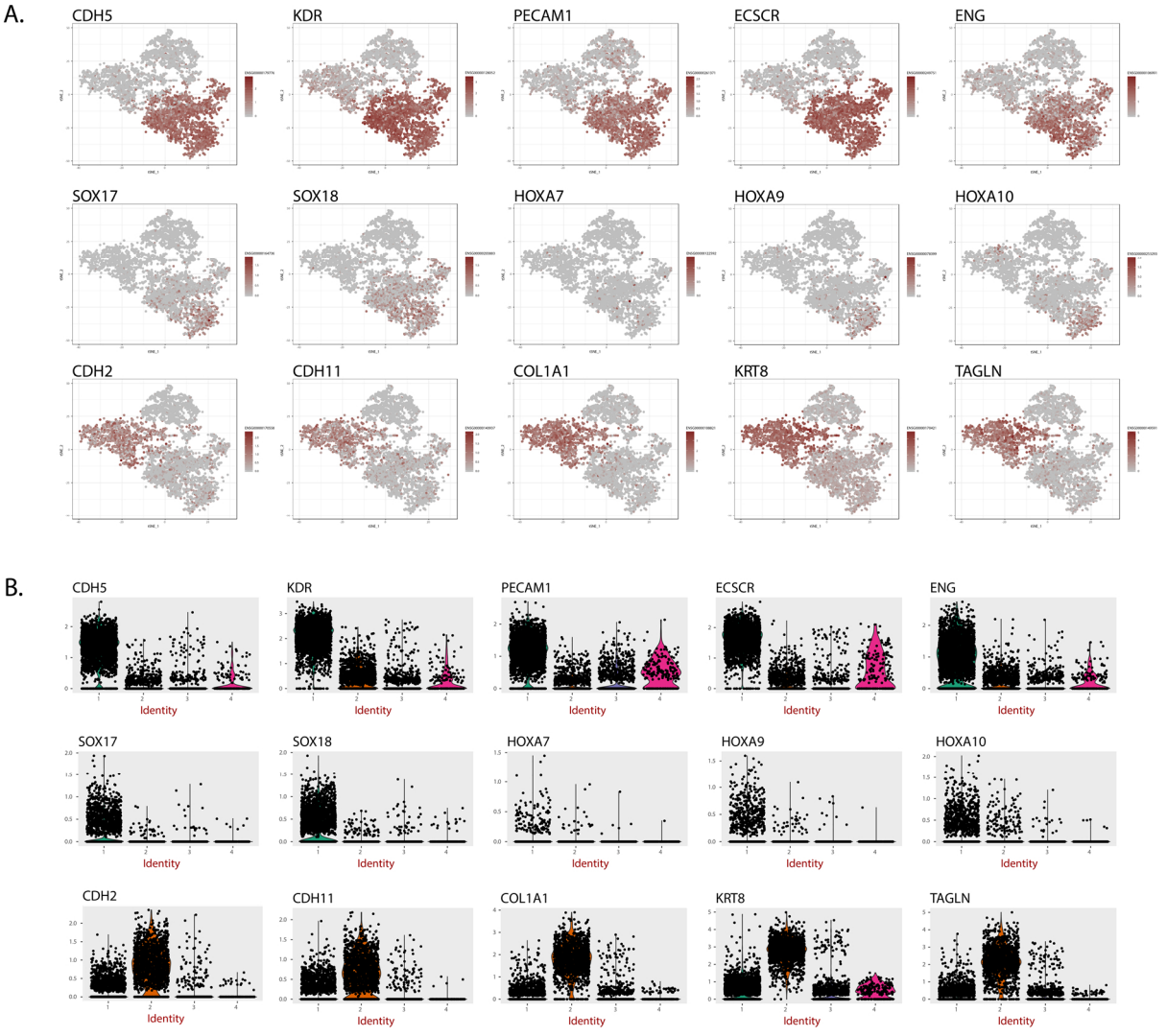



**Figure 3.2 Endothelial and mesenchymal lineage markers.** Based on the expression of *CDH5*, *KDR*, *PECAM1*, *ECSCR* and *ENG*, cluster 1 was labelled as endothelial cells. Importantly, this cluster also contained genes marking the HE like *SOX17*, or genes important for specification of definitive HSPCs, like *SOX18*, *HOXA7*, *HOXA9*, *HOXA10*. Based on the expression of genes such as *CDH2*, *CDH11*, *COL1A1*, *KRT8* and *TAGLN*, cluster 2 instead was labelled as mesenchymal cells. Expression of key lineage markers refers to EHT D3 and EHT D5 data combined, and are shown as **(A)** tSNE plots and **(B)** violin plots. Numbers on the x axis in violin plots (1 - 4) correspond to cluster numbers as in Figure 3.1.

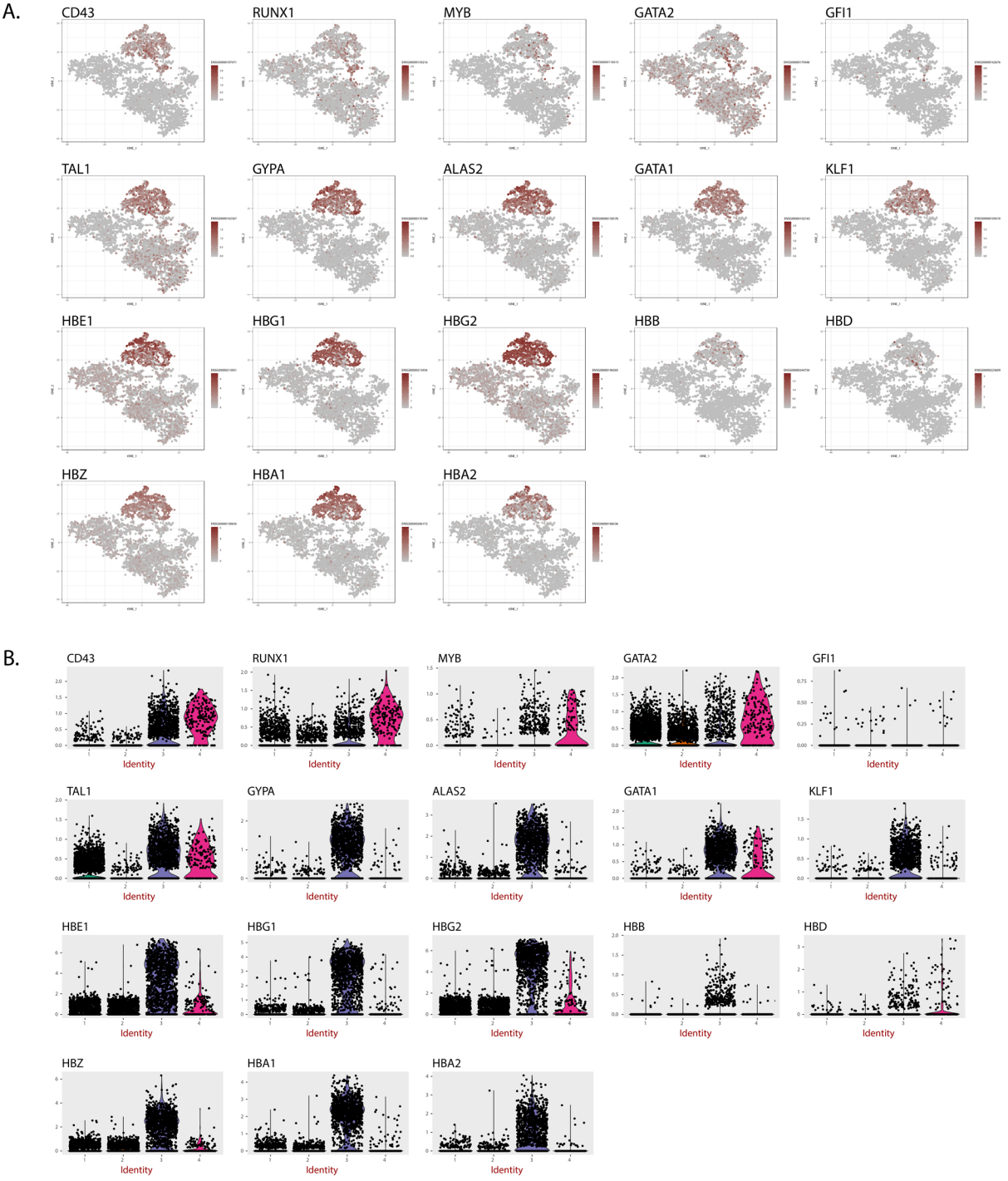



**Figure 3.3 HSPC and erythroid lineage markers.** Cluster 3 and 4 both appeared to express haematopoietic markers such as *CD43*, *RUNX1*, *MYB*, *GATA2*, *GFI1* and *TAL1*. Cluster 3 appeared to be particularly enriched in erythroid genes, such as *GYPA*, *ALAS2*, *GATA1*, *KLF1* and several haemoglobin isoforms, while cluster 4 showed a much lower expression for these genes, while also showing low expression of other blood lineage genes (Figure 3.4). For this and for the lack of mature blood cell morphologies (Figure 3.5), cluster 3 was labelled as erythroid progenitors, while cluster 4 as HPCs. Lower levels of key haematopoietic markers, co-expressed with endothelial genes, were also found on cluster 1, consistent with an HE identity. Expression of key lineage markers refers to EHT D3 and EHT D5 data combined, and are shown as **(A)** tSNE plots and **(B)** violin plots. Numbers on the x axis in violin plots (1 - 4) correspond to cluster numbers as in Figure 3.1.

Two clusters appeared instead to be positive for haematopoietic markers such as *RUNX1*, *CD43* and *TAL1* (Figure 3.3). Cluster 3 was particularly enriched for factors known to characterise the erythroid lineage, like *GYPA*, *ALAS2*, *GATA1*, *KLF1* and different haemoglobin isoforms. However, the May–Grünwald-Giemsa stain I performed with samples from EHT D5 showed the absence of cells with clear erythroid or other mature blood cell morphologies (Figure 3.5). Instead, the histological morphology was similar and consistent with a previous study where the same staining was performed on EHT samples from mouse E11.5 AGM (Taoudi et al., 2005). Prompted by these observations, I hypothesised these cells to represent progenitors committed to the erythroid lineage but not yet showing the typical differentiated morphology or high levels of lineage genes. Cluster 4 on the other hand was the more interesting of the two haematopoietic groups of cells. Indeed, it expressed low levels of myeloid markers such as *SPI1* and *CD33* (Figure 3.4), but it was also weakly positive for erythroid factors such as *GATA1, KLF1*, and even very low levels of haemoglobins. Small portions of this cluster showed expression of megakaryocyte markers as *ITGA2B* and *GP1BA*, or lymphoid genes as *CD3* and *RUNX3* at very low levels. Most importantly, the cluster mainly expressed transcription factors involved in HSPC specification and function, namely *RUNX1*, *MYB*, *GFI1* and *GATA2* (Figure 3.3). All these observations led me to hypothesise that this cluster might represent the first population of HPCs, possibly multipotent progenitors with the ability to differentiate towards different blood cell types, as suggested by the heterogeneous gene expression profile.

Finally, a fourth group of cells did not appear to express high levels of either endothelial or haematopoietic markers. Instead, it expressed factors involved in processes like wound healing and muscle tissue morphogenesis, and it was characterized by genes like *CDH2*, *CDH11*, *COL1A1*, *KRT8* and *TAGLN* (Figure 3.2). Thus, I labelled the cluster as mesenchymal cells. Importantly, the identification of this population highlighted the power of the analysis, which allowed to reveal a previously unexpected cell type.

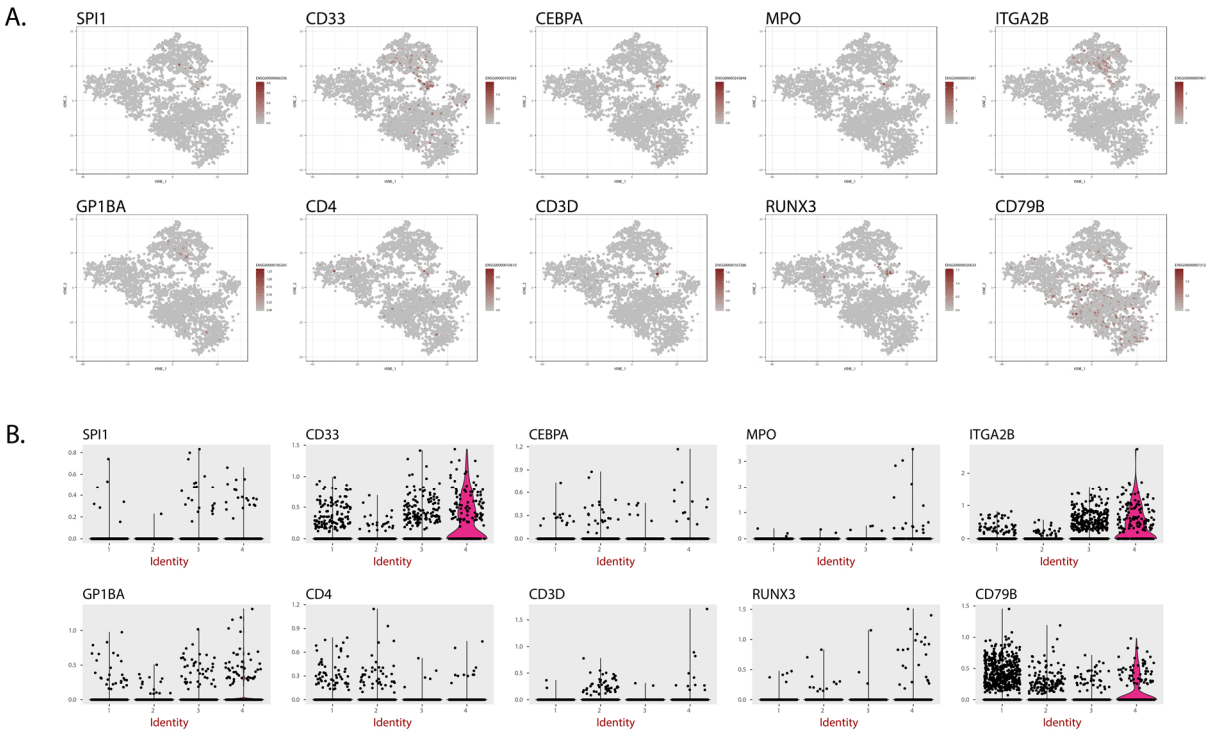



**Figure 3.4 Myeloid, megakaryocytic and lymphoid lineage markers.** Cluster 4 was shown to have low expression levels for myeloid genes like *SPI1* or *CD33*, megakaryocytic markers like *ITGA2B* and *GP1BA*, or very low levels of lymphoid genes like *CD3D* and *RUNX3*. Combined with low levels of erythroid markers and with the higher expression of genes involved in HSPC specification (Figure 3.3), this suggested the cluster to represent multipotent HPCs. Expression of key lineage markers refers to EHT D3 and EHT D5 data combined, and are shown as **(A)** tSNE plots and **(B)** violin plots. Numbers on the x axis in violin plots (1 - 4) correspond to cluster numbers as in Figure 3.1.

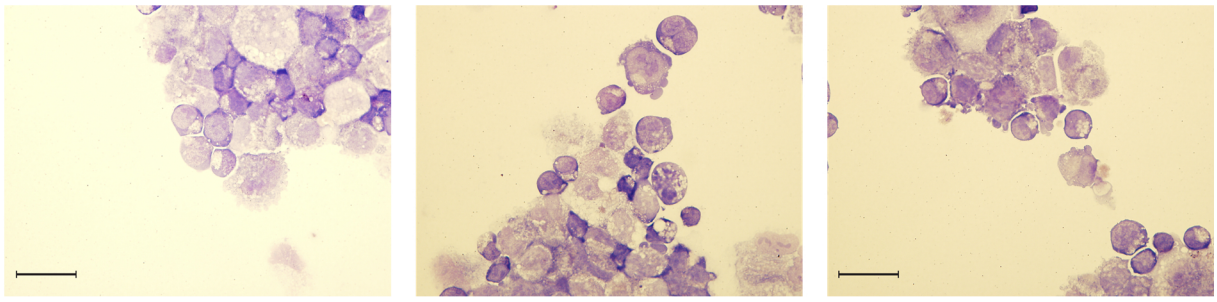


**Figure 3.5 Histological staining of cells at EHT D5.** Cells were collected at EHT D5, prepared on microscopy slides by cytospin centrifugation and stained with May–Grünwald-Giemsa method. The assay revealed the absence of obvious morphologies indicating the presence of terminally differentiated blood cells in culture, but consistency with previous studies of mouse HE during EHT from E11.5 AGM. Scale bar is 15 µm.

To confirm these hypotheses, I used the scRNAseq dataset to identify surface markers with the aim of designing a simple but effective sorting strategy which would allow me to isolate

or enrich for distinct populations and therefore confirm the assigned identities in a functional experiment. This was of fundamental importance for the subsequent steps of my investigation, and it was indeed crucial for the correct interpretation of molecular mechanisms uncovered by this analysis. My primary focus was the isolation of cluster 4, which I hypothesised to be the first population of multipotent HPCs generated *in vitro*. Using this approach, I ultimately selected the surface markers CDH5, CD43 and CD44 (Figure 3.6).

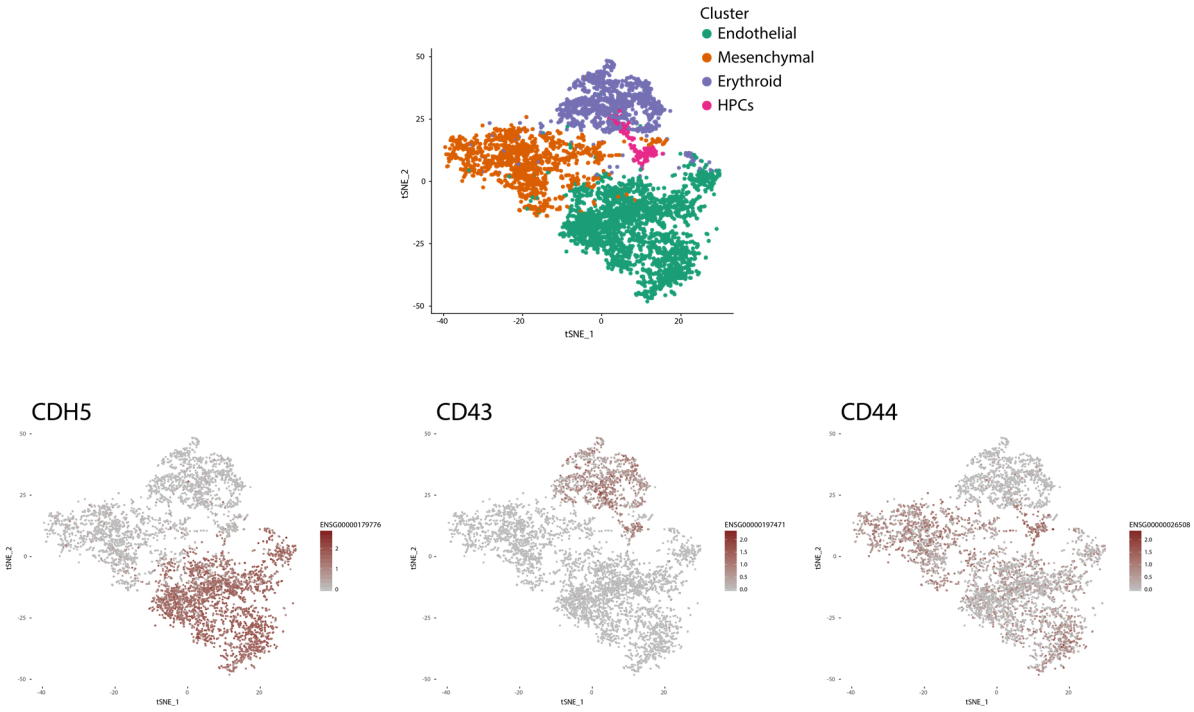



**Figure 3.6 Markers used for the isolation of distinct populations.** Based on scRNAseq data, CDH5, CD43 and CD44 were chosen to design a sorting strategy for the isolation of the clusters identified.

CDH5 was mainly expressed in cluster 1, which I labelled as endothelial cells. Importantly, it appeared to also be very weakly expressed on portions of cluster 3 and 4. As already mentioned, CD43 is a pan-haematopoietic marker and it was shown to be expressed on cluster 3 and 4, which based on the markers shown above I hypothesised to be erythroid progenitors and multipotent HPCs, respectively. CD44 was instead expressed prevalently on cluster 2, previously labelled as mesenchymal cells, as well as on cluster 4 and, at lower levels, on cluster 1 (Figure 3.6). Interestingly, CD44 was previously shown to be expressed on a plethora of cell types *in vivo*, but most importantly it was suggested to be expressed on IAHCs emerging during EHT both in human and mouse (Watt et al., 2000; Ohata et al., 2009).

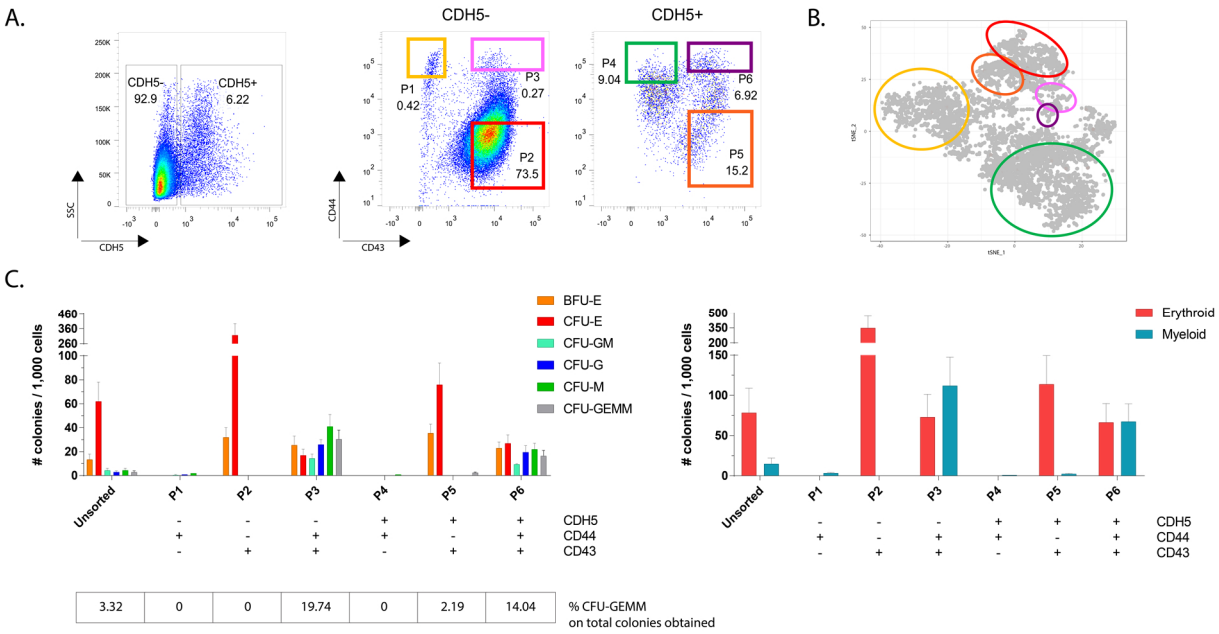



**Figure 3.7 Sorting and functional validation of distinct clusters. (A)** Using this strategy, six populations were FACS-sorted at EHT D5 for further functional validation. **(B)** The six sorted populations corresponded to distinct fractions of cells in the tSNE plot, based on the expression of the selected markers. **(C)** After isolation, 1,000 cells per population were tested in a CFU assay, and the colonies obtained were scored based on their morphology and plotted individually (left) or based on their erythroid or myeloid identity (right). The percentage of multilineage CFU-GEMM colonies on the total number of colonies obtained is also shown. Results represent n=2 independent experiments, with error bars ±SEM.

Using this combination of surface markers, I was able to sort at EHT D5 six distinct populations, which I then tested in a CFU assay in order to assess their differentiation potential (Figure 3.7) Importantly, CDH5-/CD44+/CD43- (P1) and CDH5+/CD44+/CD43- (P4) cells, corresponding to populations previously labelled as mesenchymal and endothelial cells respectively, did not show significant levels of blood colonies, confirming their lack of haematopoietic potential at this stage. On the other hand, CDH5-/CD44-/CD43+ (P2) and CDH5+/CD44-/CD43+ (P5) cells were exclusively able to generate colonies representative of the erythroid lineage, and more specifically CDH5-/CD44-/CD43+ cells had a higher enrichment in the number of mature CFU-E colonies, thereby suggesting a higher level of commitment. Based on these results, I speculated these two populations to represent distinct portions of cluster 3, labelled as erythroid progenitors. Part of this cluster is indeed weakly positive for CDH5 in our transcriptomics data and could represent a less committed type of progenitor, as indicated by a different number of CFU-E generated. Finally, CDH5-/CD44+/CD43+ (P3) and CDH5+/CD44+/CD43+ (P6) were able to generate all the different

types of colonies. As for cluster 3, the two populations potentially represented distinct portions of the same cluster 4, given that only part of this cluster was shown to be positive in the transcriptomic analysis for CDH5, which once again seemed to mark the transition from an early to a late progenitor state. These two populations not only had the ability to differentiate into all the different types of colonies, but they also showed a limited number of CFU-E colonies, suggesting a lower degree of commitment. These colonies are indeed derived from mature erythroid progenitors which are also very proliferative and therefore usually predominant even in unsorted samples. Most importantly, these two populations showed an important enrichment in the amount of the CFU-GEMM type of colony, representative of multipotent progenitors and not produced by the remaining 4 populations sorted. Specifically, 19.74% and 14.04% of all the blood colonies produced in the assay by CDH5-/CD44+/CD43+ and CDH5+/CD44+/CD43+ cells, respectively, were CFU-GEMM, against only a 3.32% generated in the unsorted sample. This confirmed that the two populations were indeed enriched in multipotent progenitors as hypothesised based on transcriptomics data, and highlighted the effectiveness of the sorting strategy.

All these findings together confirmed the cluster identities previously assigned based on gene expression. Cluster 1 and 2 are non-haematopoietic groups of cells likely to represent endothelial and mesenchymal cells, respectively, based on their gene expression profiles. The remaining clusters are haematopoietic populations. Specifically, cluster 3 contains progenitors that can only produce erythroid colonies, while cluster 4 is able to generate multiple types of erythroid and myeloid colonies and most importantly is highly enriched with multilineage potential, and thereby likely represents multipotent HPCs. Additionally, my data suggested the possibility of using CDH5 to mark distinct states of maturation for haematopoietic progenitors, with more mature cells likely to be devoid of the endothelial marker.

## 3.3 Analysis of EHT progression by scRNAseq

Having confirmed their identity by functional validation, it was interesting to see how the 4 clusters progressed over time from EHT D3 to EHT D5. This time window represented the most likely culmination of the transition from the endothelial to haematopoietic cell fate and thus was chosen for the study of this developmental process. For this analysis the sequencing data, previously combined for the identification of the distinct populations, was now used for a comparison between the two time points.

From this comparison it was evident that the major change from EHT D3 to EHT D5 involved a reduction in the endothelial cluster and an increase in the haematopoietic compartment, especially for the erythroid progenitors (Figure 3.8A). The cluster of multipotent HPCs only showed a small increase, possibly indicating that once generated these cells have a low proliferation rate. Alternatively, they might represent a transitory population which quickly commits and is immediately replaced by new cells. It was interesting to note how the increase in the haematopoietic compartment at the expense of the endothelial cluster seemed to directly represent the EHT process and the progression from one cell type to the other. Importantly, the mesenchymal cluster did not appear to be affected.

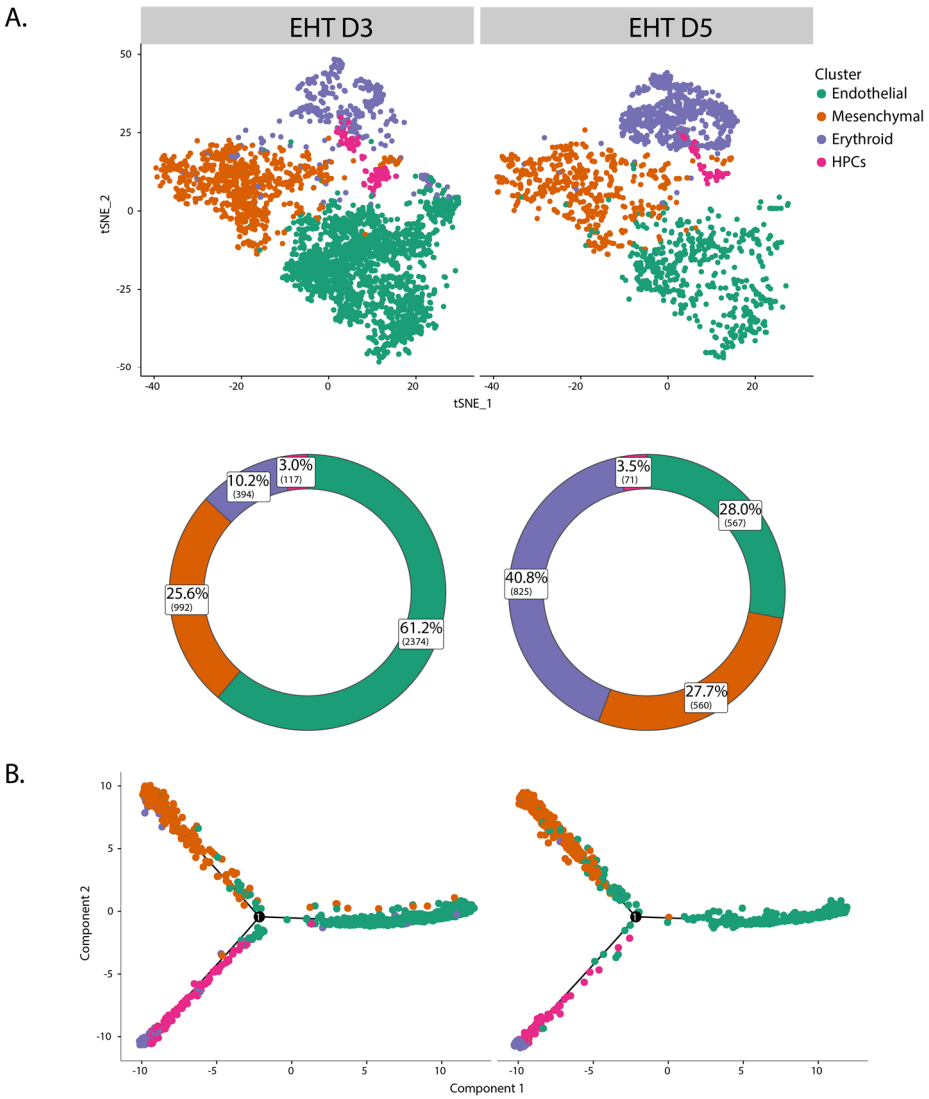



**Figure 3.8 Cluster progression during EHT. (A)** Comparing the two time points, the major change appeared to be a reduction in the endothelial cluster and an enrichment in the haematopoietic compartment, especially for the erythroid progenitors. **(B)** In the pseudotime trajectories, it was also possible to see that the branch point, likely to be occupied by cells in the transition process, was less populated at EHT D5, while the haematopoietic branch showed cells progressing further towards the end point. Overall these trajectories gave a graphical representation of EHT based on transcriptional data.

We were also able to use our single cell transcriptomics data to further elucidate mechanisms controlling EHT. For that, we generated pseudotime trajectories depicting the transition between different cell states using the package Monocle2 (Trapnell et al., 2014; Qiu et al., 2017a; Qiu et al., 2017b). In brief, the analysis took advantage of the fact that individual cells engaged in a progressive process are usually asynchronous for their progression along that process. This means that at any time distinct cells will be at slightly different stages and therefore form a continuum from point A to point B in that specific process, with their gene expression reflecting such variability. Monocle allows to take a multitude of variable genes across the whole

sample into consideration and use their expression to deconstruct this asynchrony and align individual cells in a pseudotemporal order. With this approach, we were able to use our data to generate pseudotime trajectories (Figure 3.8B). Confirming our expectations, the analysis appeared to depict a transition from the endothelial cluster to the haematopoietic compartment containing both erythroid progenitors and multipotent HPCs, in agreement with the identities previously assigned. In addition, the analysis suggested the possibility that the mesenchymal cluster might also be directly derived from endothelial cells. Furthermore, a comparison between pseudotimes for the two time points revealed that during this time most of the cells that were engaged in the transition successfully completed their commitment, as shown by the fact that the branch point, likely to be populated by cells undergoing the transition at EHT D3, was instead poorly occupied at EHT D5, when cells appeared to show higher density at the end points of the trajectory.

Overall, the analysis allowed to gain important insights on the populations participating in the transition and to monitor their evolution during this key developmental stage.

## 3.4 Discussion

In this chapter, I took advantage of novel technologies for single cell transcriptomics to study the EHT process leading to the generation of early HPCs. This technique, combined with the *in vitro* system previously described, allowed me to define with unprecedented detail the transcriptional signature of cells involved in the transition. This allowed me to characterise the endothelial cells undergoing EHT with single cell precision, and to determine that two distinct populations of haematopoietic cells were present in this culture system. A first group of cells had an erythroid transcriptional identity and was confirmed to produce exclusively erythroid colonies. A second smaller fraction of cells expressed instead low levels of different lineage markers, but also transcription factors involved in HSPCs specification and thus resembling multipotent progenitors. Importantly, the latter population was functionally confirmed to

possess multilineage differentiation potential, and therefore to likely constitute the first example of multipotent progenitors generated in this system.

The comparison of single cell transcriptomics data for two consecutive time points during the EHT process allowed me to study for the first time the evolution of this transition. This approach revealed that the number of multipotent HPCs increased only slightly during the time considered, an observation which might be explained by different hypotheses. One possibility is that, once generated, these cells have a low proliferation rate and possibly, over the 48 hours analysed in the study, maintain their number stable through asymmetric division, therefore contributing to the erythroid population. Alternatively, most of the HPCs might quickly commit and contribute to the erythroid cluster, while at the same time being replaced by cells freshly generated from the HE, with the total number of multipotent HPCs remaining almost constant over time, at least for the short time window taken into consideration. This would suggest the HPC population to be extremely transitory, characteristics perhaps exacerbated by culture conditions not optimised for the stabilisation of the stem/progenitor state. A third scenario could be represented by a synthesis of the two hypothesis. Part of the HPCs might indeed quickly commit and contribute to the erythroid cluster, and be replaced by new cells generated both as a result of HPC asymmetric cell division and of endothelial cells transitioning to the haematopoietic fate. Interestingly, in all these scenarios the erythroid commitment appears to be the dominant differentiation programme, while other fate choices are expressed when more permissive conditions become available, such as in a CFU assay. This might suggest that further optimisation of the EHT culture conditions is necessary to prevent the generated progenitors from rapidly committing. Indeed, the strong decrease in the number of endothelial cells corresponding to the increase in the erythroid cluster suggests that the intermediate HPC state is extremely transitory in these culture conditions.

An additional cluster identified in the analysis was constituted by mesenchymal cells, previously unexpected in this *in vitro* system. This group was characterised by the expression of genes involved in smooth muscle tissue morphogenesis, wound healing and matrix deposition, and I originally hypothesised it to be an undesired by-product of the differentiation protocol, possibly caused by culture conditions which should be further optimised. Indeed, it

could be possible that endothelial and mesenchymal cells have a common origin from mesodermal precursors, and that a suboptimal differentiation fails to prevent the generation of mesenchymal cells. Alternatively, they might instead represent a cell type contributing and necessary for the formation of the correct developmental niche from which HSPCs are generated *in vivo* in the dorsal aorta and recapitulated *in vitro* in this model system. Indeed, mesenchymal cells have been previously reported to originate in the AGM prior to the HSPC emergence and to participate to the haematopoietic niche, possibly with a supportive role similar to what is observed in the bone marrow (Wang et al., 2008). It is therefore possible that these cells are also required for the generation of functional haematopoietic cells *in vitro*. Additionally, the generation of pseudotime trajectories, beyond providing an important confirmation for the EHT as a progression from an endothelial to a haematopoietic fate, revealed the potential developmental origin of mesenchymal cells. Indeed, the analysis suggested that the mesenchymal cluster might be directly derived from endothelial cells, potentially through an endothelial-to-mesenchymal transition (EndoMT) which was previously described *in vivo* both in development and disease (Ten Dijke et al., 2012; Chen et al., 2015; Good et al., 2015; Zhong et al., 2018; Man et al., 2018). This could represent a biological process overlapped with EHT and perhaps necessary for the emergence of IAHCs. Additionally, EndoMT and EHT might also represent two divergent directions originating from a common endothelial progenitor. As such, this might constitute a cell fate decision event, with the balance between the two transitions possibly controlled by similar mechanisms to determine if endothelial cells will undertake the haematopoietic or mesenchymal fate. Of note, our data showed the mesenchymal cluster to remain almost constant between the two time points. This could be caused by culture conditions optimised for the promotion of EHT, and it is possible that modulation of specific signalling pathways might change this behaviour and increase EndoMT. In this regards, TGFβ was reported to promote EndoMT (Wermuth et al., 2016), and it would be interesting to test if this signalling could also play a role in this *in vitro* system.

An important result achieved with the experiments presented was the design of a simple and effective sorting strategy that allowed me to isolate the distinct populations and functionally test their blood differentiation capacity. Using this approach, I was able to corroborate the non-

haematopoietic identity of the endothelial and mesenchymal clusters. Furthermore, I could confirm the restricted differentiation capability of the erythroid population and the wider multilineage potential of the HPC cluster. Of note, it is also possible that the HPC cluster is composed of multiple progenitors with different fate propensities, as previously discussed for the HSPC compartment *in vivo* (Chapter 1.2). Indeed, the multilineage potential was observed at the population level, and single cell functional experiments will be necessary to further test the differentiation capability of individual cells within the cluster. Nevertheless, the possibility of isolating the first population of multipotent progenitors generated *in vitro* based on the expression of the three markers CDH5, CD43 and CD44 was an important achievement, and paves the way to future studies specifically addressing this population. Interestingly, it was previously suggested that CD43 is the first surface marker to appear during the generation of HSPCs, both *in vitro* and *in vivo* (Vodyanik et al., 2006*;* Ivanovs et al., 2014a). Importantly, here I have shown that the co-expression of CD43 and CD44 specifically marks a population possessing multilineage potential. In agreement with these results, CD44 was previously reported to be expressed on IAHCs in the AGM region, both in human and mouse (Watt et al., 2000; Ohata et al., 2009), and to be involved in the migration of foetal HSCs and in their homing to the bone marrow (Cao et al., 2016). Our data also revealed that the mesenchymal cluster, similarly to HPCs, express high levels of CD44. This would be in agreement with previous studies reporting the appearance of mesenchymal cells in the AGM and their presence at subsequent haematopoietic sites, including foetal liver and bone marrow, along with the embryonic circulation, suggesting that at least some of these cells might migrate across different haematopoietic sites during development (Mendes et al., 2005). Therefore, CD44 could possibly be involved in the migration of both HSPCs and mesenchymal cells, and allow their homing to common sites during haematopoietic development. Additionally, I proposed that the expression of CDH5 possibly marks earlier subpopulations within the two haematopoietic clusters. Both CDH5+ and CDH5- HPCs have similar differentiation potential, with the latter population possessing higher multilineage capacity, possibly indicating a higher degree of maturation. Indeed, their endothelial origin would be in agreement with the earliest progenitors expressing the marker. Additionally, this would confirm previous studies showing that HSCs emerging in the AGM express CDH5, which is downregulated upon maturation (Ivanovs et al.,

2014a). Similarly, I showed that erythroid progenitors devoid of CDH5 produce a higher number of mature CFU-E colonies, and therefore appear to represent more committed progenitors. The expression of CDH5 on less mature erythroid cells suggests that they are possibly derived from CDH5+ HPCs.

In conclusion, the combination of scRNAseq and *in vitro* differentiation of hPSCs gave me a detailed picture of developmental events which would be otherwise difficult to access in human. Furthermore, this approach allowed me to solve the complexity of this heterogeneous system and to identify, characterise and isolate the populations involved in the foetal haematopoietic specification.

# 4 ROLE OF CELL CYCLE PROGRESSION DURING HAEMATOPOIETIC SPECIFICATION

## 4.1 Summary

A key observation generated by our single cell analyses is the potential difference in proliferation between the distinct populations generated. In agreement with this, from morphological observations cell cycle appeared to be dynamic during the second stage of the differentiation protocol, as cells in this culture system initially seemed to show very little amount of proliferation between EHT D0 and EHT D2, while their proliferation increased after EHT D3 when clusters of haematopoietic cells started to appear, similar to the IAHCs found *in vivo* during development. Subsequently, single cells released in culture from these clusters after EHT D5 showed an extremely high proliferation rate. A dynamic regulation of the cell cycle machinery was also suggested by qPCR data, showing that cell cycle regulators were differentially expressed over time during EHT. To quantify all these observations, we decided to use our single cell transcriptomic data to examine the cell cycle state of the different populations involved in EHT. This approach suggested a role for cell cycle regulation in the generation of early HSPCs, which was further confirmed through functional validations. Overall, these results suggest an additional layer of complexity in the regulation of early haematopoiesis.

## 4.2 Populations involved in EHT show distinct cell cycle states

As previously shown, qPCR analyses of gene expression during the second stage of differentiation was able to give a first representation of EHT, with the expression of endothelial and haematopoietic genes showing an inverse correlation during the transition. Similarly, the

expression of key genes controlling cell cycle was in agreement with morphological observations, suggesting a dynamic behaviour for proliferation and cell cycle regulation (Figure 4.1). Specifically, I could note the transcriptional upregulation of factors involved in S and G2/M phases and indicative of active cell cycle, including *cyclin E1*, *cyclin B1*, *CDK2* and *CDK1*, and at the same time the transcriptional downregulation of proteins known to act as inhibitors of cell cycle progression, such as *RB1*, *RBL2/p130* and *CDKN1C/p57*. Altogether, this data suggested a progressive increase in the number of cells entering the cell cycle, with a timing that corresponded to EHT progression.

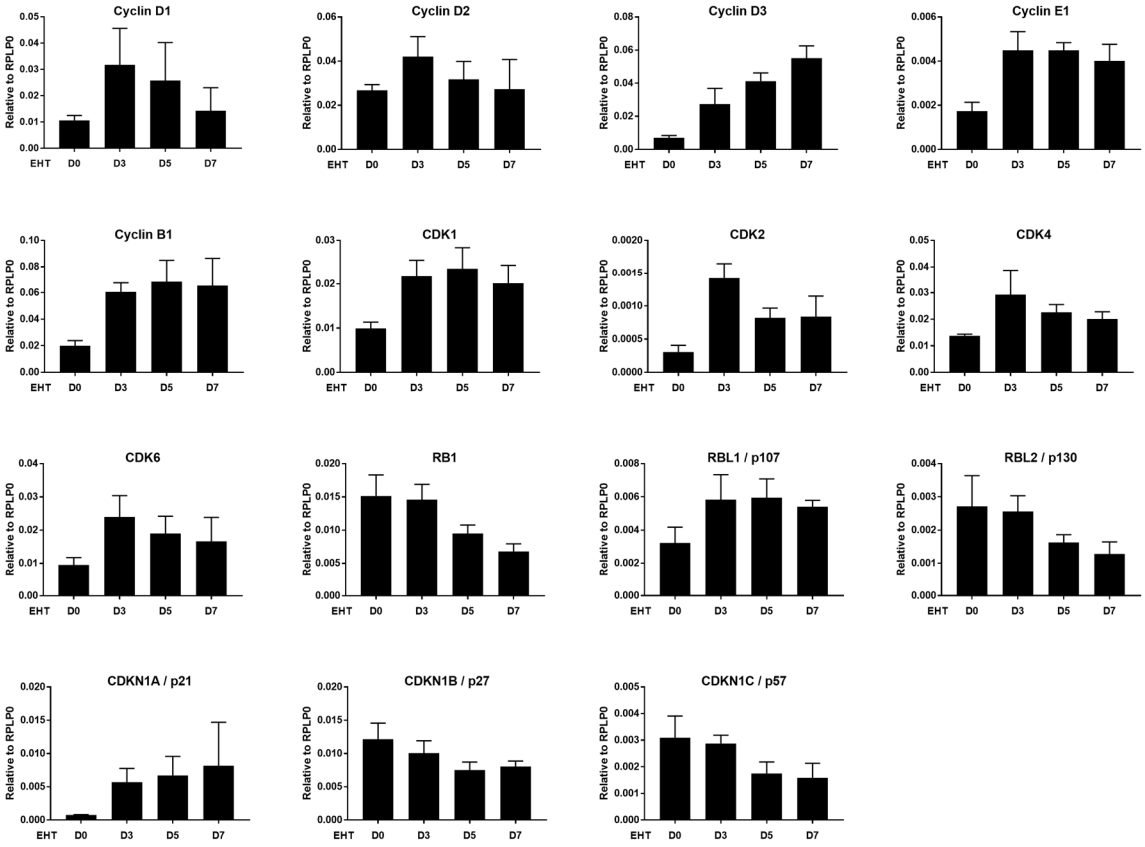



**Figure 4.1 Expression of cell cycle regulators during EHT.** The expression of key cell cycle genes was measured during EHT by qPCR. Results represent n=3 independent experiments, with error bars ±SEM.

Prompted by these observations, we decided to interrogate our scRNAseq dataset and to investigate the cell cycle state of the previously identified populations. By calculating cell cycle phase scores based on canonical conserved markers, we were able to infer the cell cycle phase of individual cells in each cluster (Butler et al., 2018). The results showed a heterogeneous but

cell type-specific distribution in distinct cell cycle phases across different clusters (Figure 4.2). Endothelial cells were particularly enriched in G1, with almost 90% of the cells found in this phase. Mesenchymal cells and HPCs appeared to have a more active cell cycle, with HPCs still fairly enriched in G1. On the other hand, erythroid progenitors showed a completely opposite behaviour compared to endothelial cells, being characterized by an incredibly active cell cycle with only 2.5% of the cells in G1.

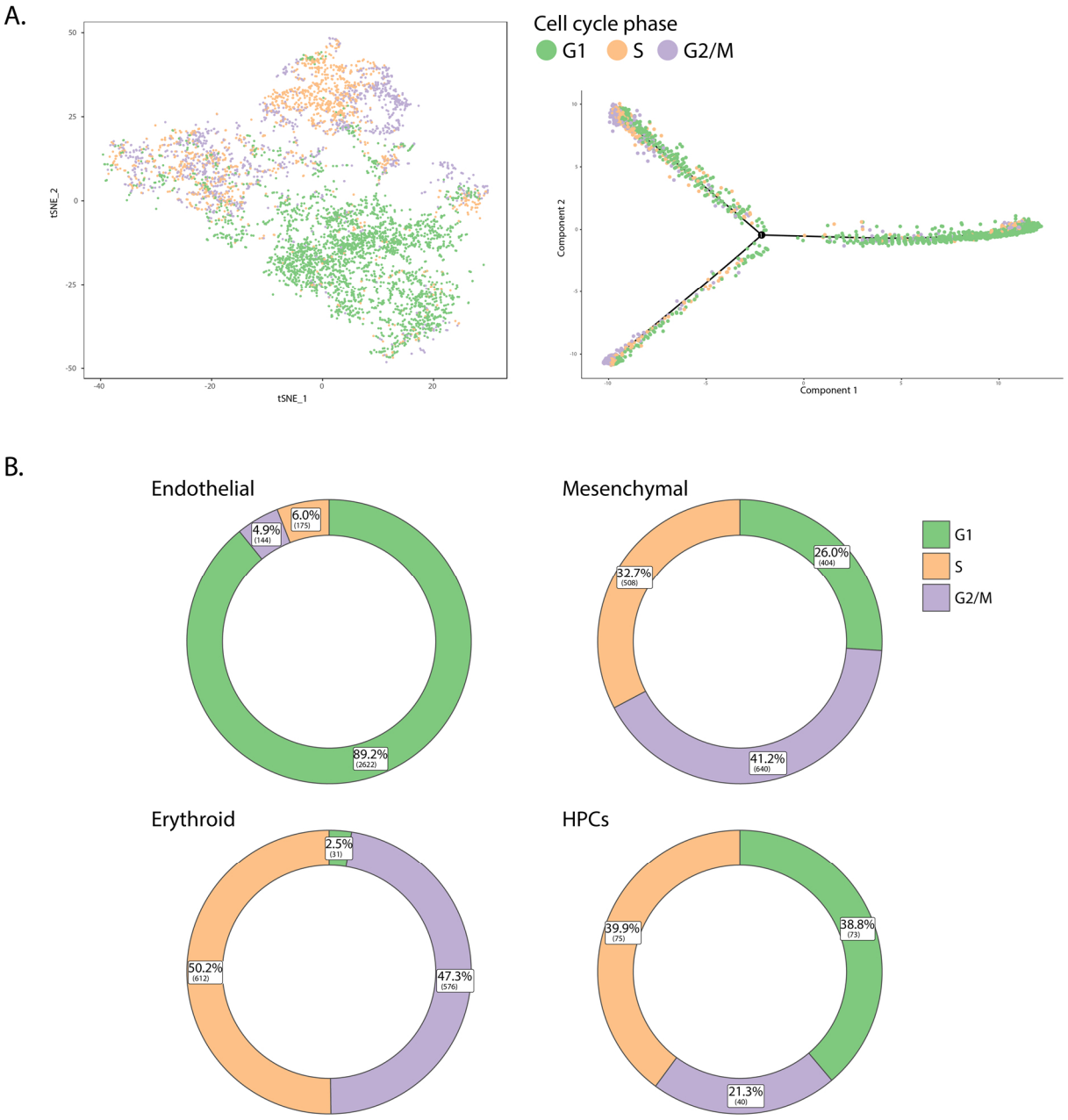



**Figure 4.2 Cell cycle analysis of populations involved in EHT. (A)** Cell cycle scores for individual cells were calculated based on the expression of canonical marker genes, and cells were painted to graphically portray their cell cycle phase, both in the tSNE space and in the pseudotime trajectory. **(B)** The analysis revealed that distinct clusters are characterised by very distinct cell cycle distributions.

The analysis confirmed our previous observations and also suggested a potential connection between cell identity and cell cycle state. HECs generated in this system appeared either to be a population of quiescent cells or to be characterized by a low proliferation rate, and their transition to the haematopoietic compartment appeared to be tightly connected with the acquisition of an active cell cycle state. Furthermore, once generated, HPCs appeared to be actively cycling, a property which became drastically more prominent in the erythroid progenitors.

These differences were further explored by comparing distinct clusters for the expression of specific regulators controlling the cell cycle state (Figure 4.3). With this approach we could see that compared to other clusters, endothelial cells were enriched for the expression of genes belonging to the Rb and CIP/KIP families, such as *RB1*, *RBL2/p130*, *CDKN1A/p21* and *CDKN1C/p57*. These factors act by inhibiting G1 progression and G1/S transition, as previously discussed (Chapter 1.8.2). Overall the expression of these genes was consistent with an enrichment of the cells in G1 and a slow proliferation or quiescent phenotype. In contrast to this, both the haematopoietic clusters showed a lower expression for these cell cycle inhibitors, and especially the erythroid progenitors appeared to be enriched for factors like *cyclin A2*, usually expressed in dividing cells, or *MKI67*, a marker of proliferation.

It was also interesting to note a differential expression for cyclin D genes, necessary for the activation of CDK4/6 and progression through the G1 phase. Endothelial cells were indeed enriched for *cyclin D1* and *D2*, while both the haematopoietic clusters showed a preferential expression of *cyclin D3*, possibly suggesting a divergent requirement for cyclin D genes, although it is still unclear how the differential expression of specific cyclin D members might affect quiescence and progression through G1 phase.

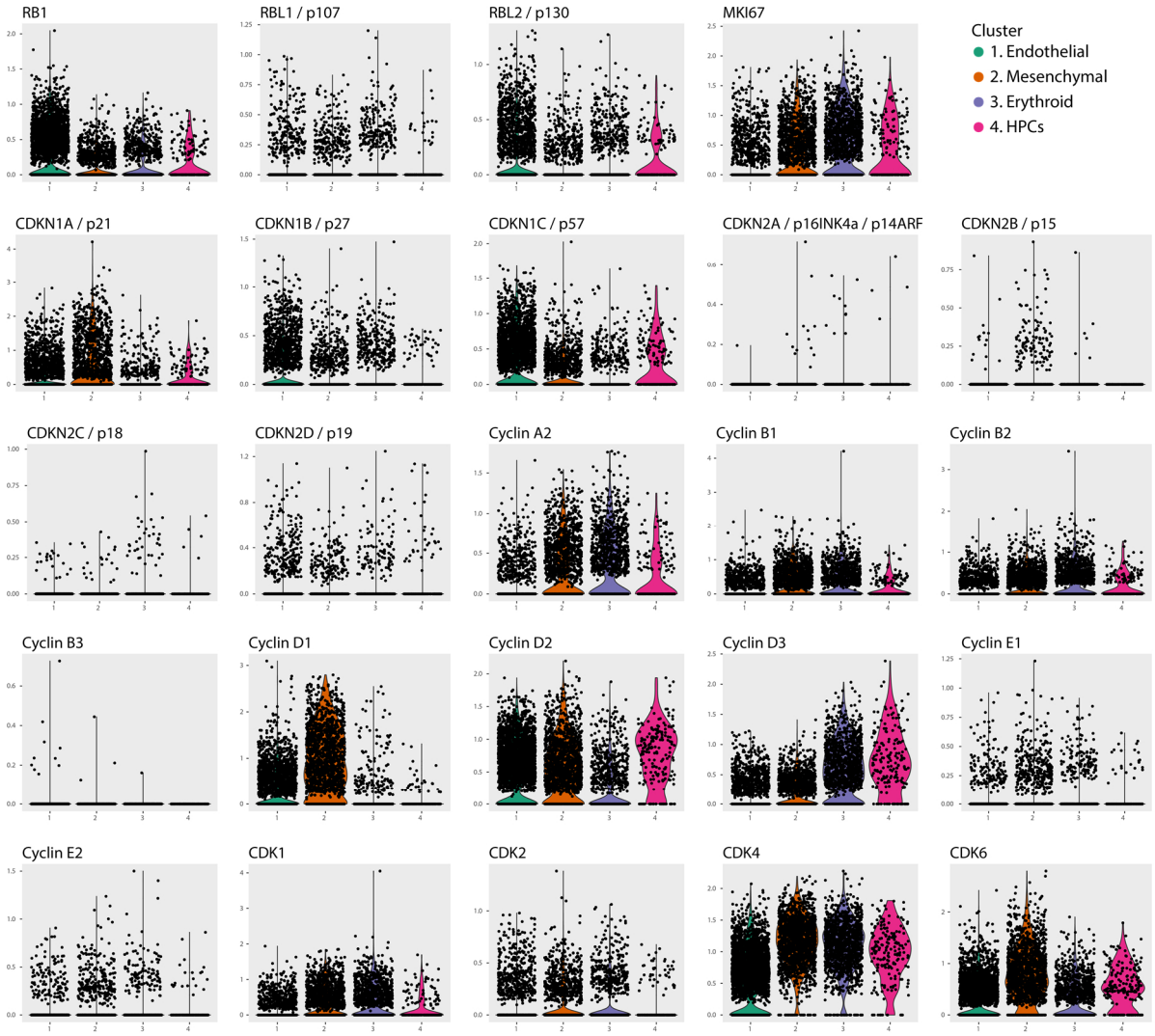



**Figure 4.3 Expression of cell cycle regulator genes on distinct clusters.** The expression of key cell cycle regulators was assessed in distinct clusters based on scRNAseq data.

To further confirm these results, I performed cell cycle profile analyses during EHT. For this, I collected samples at different time points during differentiation and performed staining for 5-ethynyl-2´-deoxyuridine (EdU) and DNA content. EdU is a nucleoside analogue to thymidine and is incorporated into DNA during active DNA synthesis, in this way effectively marking cells in S phase. DNA content instead allows discrimination between cells in G1 and G2/M. By combining this approach with staining for the CD43 surface marker, I was able to visualize by flow cytometry the cell cycle profile during EHT of cells in the haematopoietic and non-haematopoietic compartments (Figure 4.4).

The experiment confirmed that the emerging haematopoietic population was characterized by an active cell cycle profile, opposed to the non-haematopoietic fraction always containing a large majority of cells in the G1 phase. This was in agreement with the single cell transcriptomics data, and suggested a potential connection between cell cycle state and transition from endothelial to haematopoietic identity. Ultimately, this led me to the hypothesis that endothelial cells with haemogenic potential might reside in a quiescent or almost quiescent state, and that their engagement into haematopoietic commitment could be associated with, and perhaps require, cell cycle re-entry and progression.

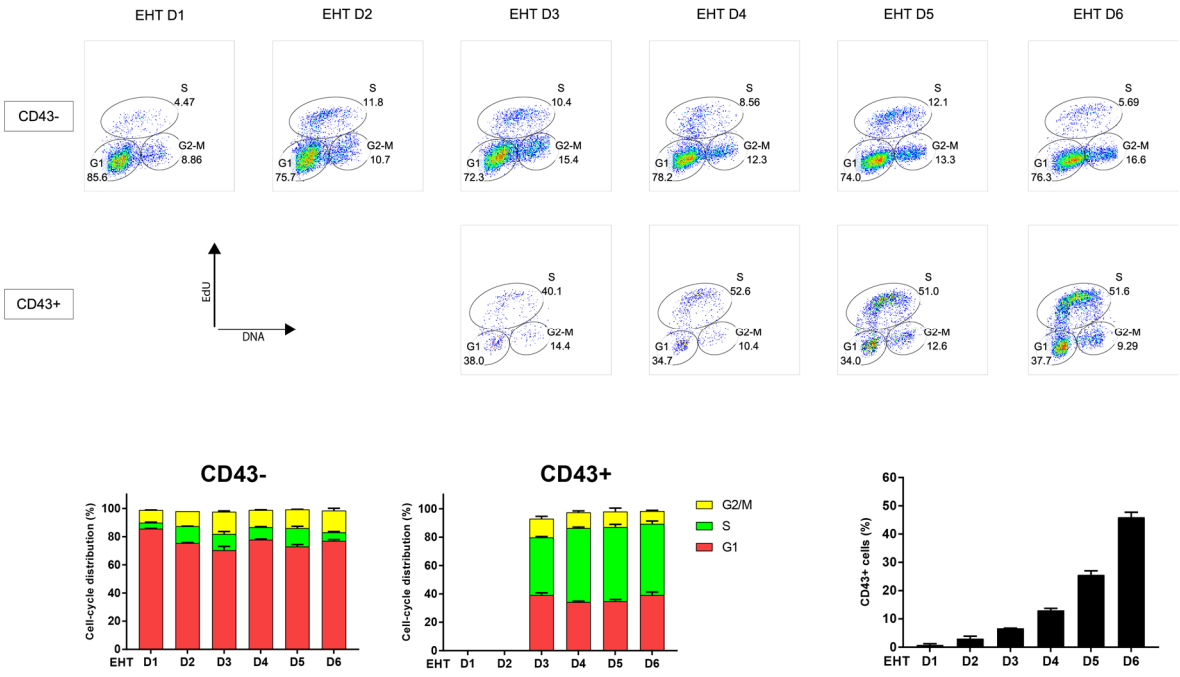



**Figure 4.4 Cell cycle state during EHT.** Using EdU staining assay, cell cycle state for CD43- and CD43+ fractions was monitored during EHT, confirming previous data from scRNAseq. Results represent technical duplicates from n=1 experiment, with error bars ±SEM.

## 4.3 Cell cycle progression is necessary for the specification of functional HPCs

Following this data, I decided to investigate whether cell cycle progression was not only associated with, but also necessary for, EHT. For this, I used the small molecule Nocodazole, an antimitotic agent that acts by binding to β-tubulin and therefore disrupting the polymerization of microtubules. This leads to an inhibition of microtubule dynamics, and ultimately to disruption of mitotic spindle function and arrest of the cell cycle in the G2/M phase (Blajeski et al., 2002). I started a treatment with 0.1 μg/mL Nocodazole at EHT D3 and collected the cells after 48 hours at EHT D5. The cell cycle profile was once again determined through EdU staining, showing an expected enrichment for cells in G2/M following the treatment. Surprisingly, the assay also revealed that a significant fraction of cells were still found in G1, probably representing cells that during the 48 hours treatment never entered the cell cycle, therefore not reaching the G2/M phase to be affected by Nocodazole (Figure 4.5A). Considering data previously shown, this population of cells which appeared to be blocked in G1 likely represented endothelial cells. And given the length of the treatment with Nocodazole, this suggested that these cells remained in G1 for at least 48 hours, therefore likely defining a quiescent population.

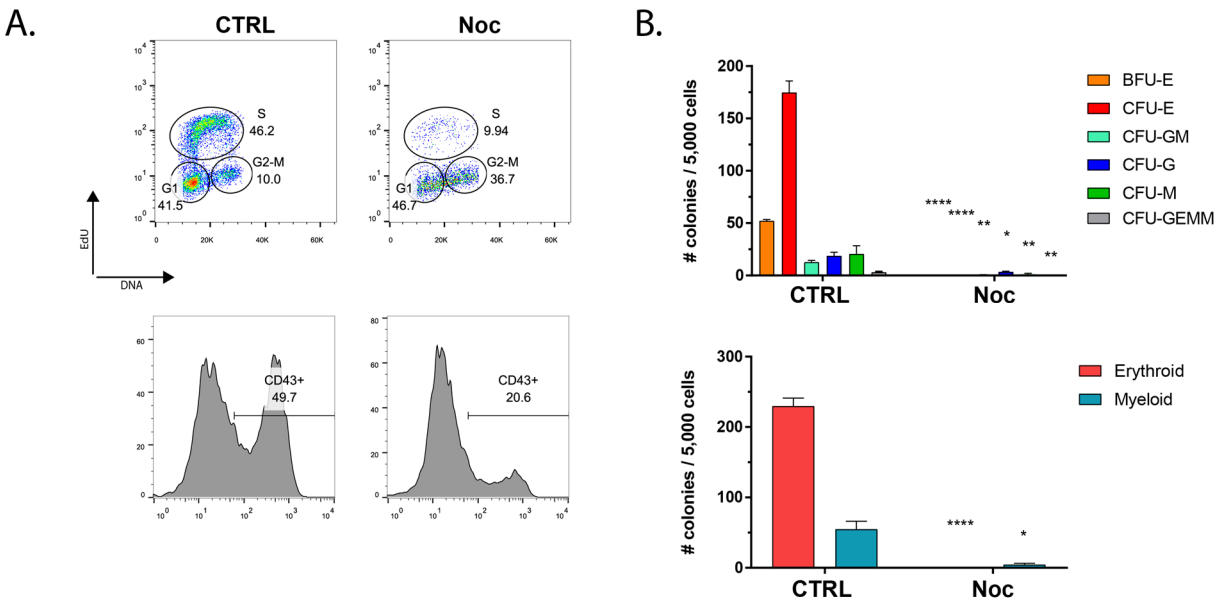



**Figure 4.5 Cell cycle progression is necessary for HPCs specification.** Cells were differentiated as previously described, and at EHT D3 media was supplemented either with 0.1% DMSO (CTRL) or with 0.1 µg/mL Nocodazole (Noc). After 48h at EHT D5, cells were either **(A)** processed for EdU and CD43 staining, or **(B)** alternatively they were washed to remove Nocodazole and further cultured in a CFU assay. Results represent n=3 independent experiments, with error bars ±SEM, and statistical significance compared to control as *P<0.5, **P<0.01, ***P<0.001, ****P<0.0001 by unpaired t-test.

Additionally, the experiment showed a decrease in CD43+ cells, suggesting a reduction `in the number of haematopoietic cells generated as a consequence of the treatment. As previously noticed, scRNAseq data and cell cycle profile analysis had shown that haematopoietic cells, especially erythroid progenitors, had an active cell cycle. Therefore, one possibility for the reduced number of CD43+ cells was that Nocodazole had simply blocked proliferation of haematopoietic cells once produced, without affecting their generation. In order to test this possibility, samples were collected after treatment and washed to remove Nocodazole, therefore releasing them from the cell cycle block. These cells were then used for a CFU assay to assess their capacity to terminally differentiate. If the effect of the treatment was limited to cell proliferation, the assay would have shown haematopoietic colonies, although with reduced numbers. Instead, the experiment showed that treated cells barely had any blood colony forming potential despite the presence of CD43+ cells, revealing that blocking cell cycle progression did not only affect proliferation of progenitors already generated, but the EHT process itself and the effective specification of these progenitors (Figure 4.5B). As a result, these cells had not acquired full haematopoietic competence and capacity to further differentiate into committed lineages. Based on these results, it is possible that the cell cycle state might control

key mechanisms necessary for the acquisition of full haematopoietic functionality, and that a normal cell cycle progression is fundamental and necessary to correctly complete EHT and for the generation of functional HPCs.

## 4.4 Dissection of endothelial subpopulations involved in EHT

Transcriptomic data and functional experiments confirmed a role for cell cycle progression in EHT. It was clear though that the endothelial cluster itself was heterogeneous, as already indicated by the non-uniform expression of key marker genes or transcription factors necessary for HSPC specification (Figure 3.2, Figure 3.3). Of note, the pseudotime trajectories also showed that these cells constitute a continuum from the endothelial state along their progression to the branch point and beyond to the mesenchymal and haematopoietic lineages (Figure 3.8). Therefore, an intriguing possibility was that endothelial cells found at different points along the trajectory might represent distinct endothelial states, with differences between these states potentially revealing important information about the transition. This would allow to distinguish endothelial cells engaged in EHT and committed to the haematopoietic fate from cells retaining a pure endothelial identity.

For this analysis, we used the pseudotime previously generated but we focused uniquely on endothelial cells, removing the remaining 3 clusters. As a result, endothelial cells were divided into 4 subclusters based on their relative position along the trajectory (Figure 4.6). With this approach, we were able to compare the expression of key marker genes across these different endothelial subpopulations. In confirmation of the pseudotime trajectory as representative of a progressive exit from the endothelial state, we could observe dynamic expression of endothelial markers such as *CDH5* or *PECAM1*. It was also interesting to note that subcluster D not only showed lower expression of endothelial genes, but also the concomitant expression of markers like *KRT8* and *TAGLN*. This was consistent with their allocation to the portion of the pseudotime trajectory previously shown to correspond to the mesenchymal cluster (Figure 3.8), suggesting these cells to represent endothelial cells

transitioning to the mesenchymal fate. Importantly, a different behaviour was observed on subcluster C, which also showed downregulation of endothelial markers, but this time not coincidental with the expression of haematopoietic markers, which were either only weakly upregulated, like for *MYB*, or at the same level compared to subcluster A, as was the case for *RUNX1*. This observation could suggest that in HECs undergoing EHT, downregulation of endothelial factors and loss of endothelial identity needs to happen before the upregulation of haematopoietic factors can take place, shedding light on a possible temporal sequence of events which *in vivo* leads to the specification of early HSCs. However, additional experiments aimed at further elucidating these temporal changes taking place during EHT will be necessary, as previous studies have instead shown that RUNX1 is responsible for the upregulation of GFI1 and GFI1B, which in turn cause the repression of endothelial genes (Lancrin et al., 2012).

Based on my previous results showing a distinct cell cycle state for endothelial cells and the haematopoietic compartment, I decided to investigate whether these distinct endothelial subpopulations might also show a differential enrichment for specific cell cycle regulators (Figure 4.7). This analysis revealed that cells in subcluster A were characterized by a higher expression of cell cycle inhibitors of the Rb and CIP/KIP families, similarly to what we had previously shown for the totality of endothelial cells compared to other cell types. On the other hand, the group of endothelial cells in subcluster C, supposedly the most advanced along haematopoietic commitment, had lower levels of these regulators, at the same time without any sign of upregulation of proliferation markers such as *MKI67* or *cyclin A2*. Importantly, this might indicate that the downregulation of cell cycle inhibitors is the first event required for EHT, preceding the upregulation of proliferation markers and cell cycle entry.

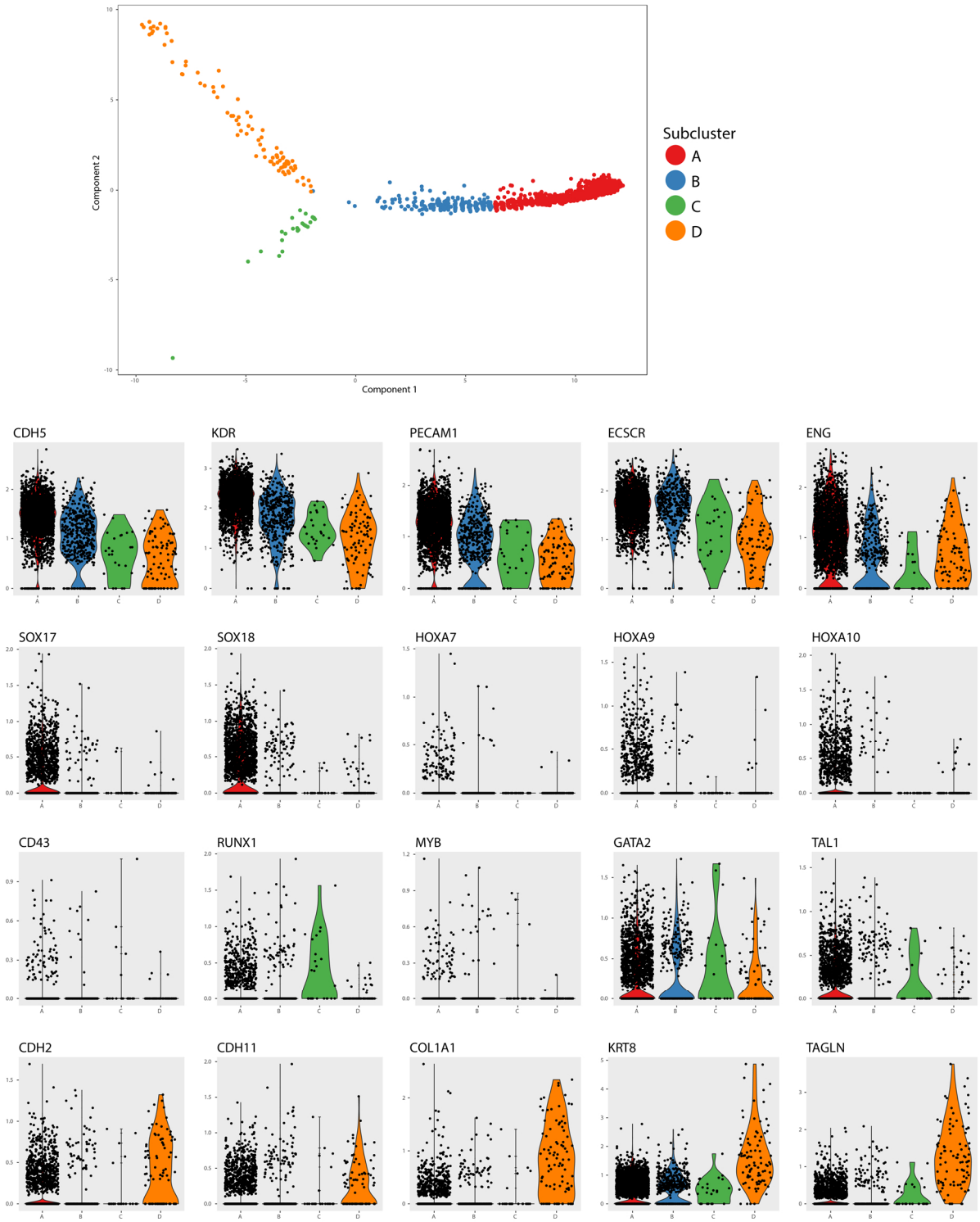



**Figure 4.6 Analysis of endothelial subclusters.** The endothelial cluster was further divided, and subclusters were characterised for the expression of endothelial and mesenchymal markers, along with transcription factors necessary for early HSPC specification.

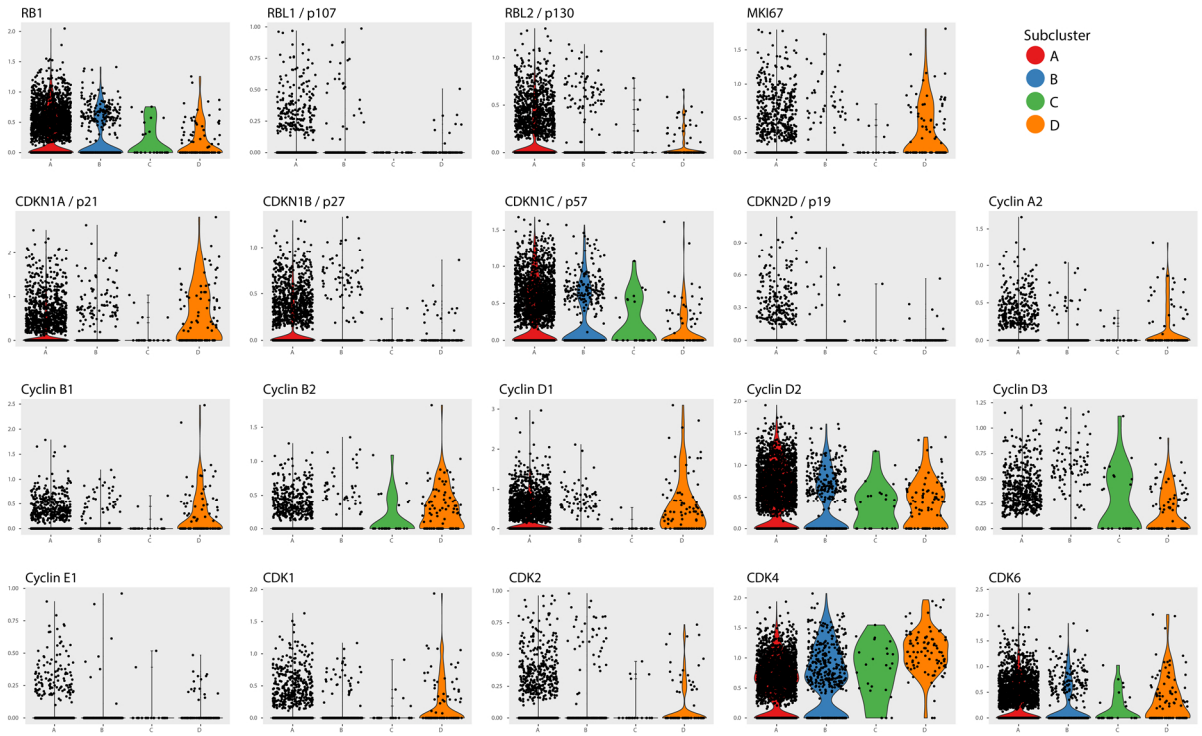



**Figure 4.7 Expression of cell cycle regulator genes in endothelial subclusters.** The expression of key cell cycle regulators was assessed in distinct endothelial subgroups.

## 4.5 Discussion

In this chapter, I showed that the dynamic regulation of the cell cycle progression is associated with the control of haematopoietic specification. Indeed, prompted by morphological observations and qPCR results, a deeper analysis of the cell cycle state was performed by interrogating our single cell transcriptomics data. The analysis revealed how the distinct cell types previously identified were also characterized by peculiar cell cycle states, the two extremes being quiescent endothelial cells and extremely proliferating erythroid progenitors. This *in silico* analysis was functionally confirmed and further expanded by experiments in which blocking cell cycle progression appeared to reduce the number of CD43+ cells. One obvious explanation for this reduction upon cell cycle arrest would be a lack of proliferation in the haematopoietic fraction, especially considering that the erythroid progenitors had a high proliferative cell cycle profile and their number quickly expanded over

time, as previously shown (Figure 3.8). However, proliferation could not entirely explain the phenotype observed. Indeed, upon treatment with Nocodazole, cells displayed an almost complete lack of blood colony forming potential, suggesting that blocking cell cycle progression during EHT might impair the generation of functional haematopoietic cells. Therefore, the CD43+ cells produced in the presence of the cell cycle blocker appeared to be qualitatively different, and possibly represented progenitors that emerged from EHT but failed to acquire the capability to further differentiate. This would suggest that key molecular mechanisms necessary for the correct specification of functional progenitors are cell cycle dependent. In this regard, our group has previously shown that hPSCs acquire responsiveness to distinct differentiation signals during specific phases of the cell cycle, through a post-translational modulation of SMAD2/3 transcriptional activity controlled by cyclin D-CDK4/6 (Pauklin and Vallier, 2013). Similarly, it is possible that during EHT the cells need to transition through specific cell cycle phases to activate mechanisms that are required for the acquisition of haematopoietic functionality. Of note, my data showed that distinct groups of cells had different requirement for cyclin D members, with endothelial cells enriched for cyclin D1/D2 and haematopoietic clusters preferentially expressing cyclin D3. This might suggest a role for cyclin D-CDK4/6 in controlling developmental mechanisms during EHT. In this regard, further investigations might reveal that distinct cyclin D proteins possibly modulate this activity through stage- or cell type-specific mechanisms.

In agreement with my results showing EHT to be associated with the acquisition of an active cell cycle profile, it was recently reported that the earliest haematopoietic progenitors generated in the mouse AGM are slowly cycling cells, and that their maturation is associated with the acquisition of a more active cell cycle (Batsivari et al., 2017). Furthermore, the anatomical position of these cells within IAHCs appears to be correlated with their cell cycle state, with slowly cycling cells frequently found at the base of the cluster in association with the underlying endothelium and rapidly cycling cells located at more apical positions. Our results agree with this observation and further expand it. Indeed, my data suggested that endothelial cells are quiescent and enter the cell cycle upon commitment to the haematopoietic fate, and additionally I showed that cell cycle progression might be necessary for this transition, possibly revealing a tight link between the molecular machineries controlling cell cycle state and cell fate

decision during EHT. In this regards, I also showed that endothelial cells are enriched for the expression of cell cycle inhibitors, such as members of the CIP/KIP family. These regulators might actively prevent endothelial cells from entering the cell cycle and undergoing EHT, and only those cells in which they are successfully downregulated might be able to start the transition. Importantly, these regulators have been reported to be modulated by TGFβ signalling and induce cell cycle exit of LT-HSCs after their migration to the foetal liver (Hur et al., 2016). Therefore, the inhibition of this signalling pathway *in vitro* could instead allow the downregulation of these cell cycle inhibitors and potentially increase the number of endothelial cells undergoing EHT, thus improving the efficiency of differentiation. However, the accurate temporal regulation of the pathway might be essential, as an excessive activation of the cell cycle could potentially make the HPC state even more transitory and rapidly cause their further commitment.

In support to the hypothesis of the cell cycle activation as a molecular requirement for EHT, further dissection of the endothelial cluster revealed that endothelial cells at the beginning of the pseudotime trajectory were enriched for the expression of these cell cycle inhibitors, consistent with a quiescent state or very slow cell cycle progression. On the other hand, endothelial cells that localised at the onset of the haematopoietic branch on the trajectory, and thus likely to represent cells actively engaged in EHT, were revealed to have downregulated such inhibitors but to not yet express proliferation markers. These cells could potentially represent an intermediate state, likely very transitory, in which endothelial markers have started to be transcriptionally downregulated, despite a lack of haematopoietic genes upregulation, possibly temporally subsequent. Based on these results, I hypothesise that haemogenic endothelial cells are normally kept in a quiescent state by the expression of cell cycle inhibitors. The first requirement for the onset of EHT would then be the removal of these molecular breaks to allow cell cycle entry and progression. This would be necessary for endothelial cells to start the transition to the haematopoietic fate, associated with the transcriptional downregulation of endothelial genes and acquisition of haematopoietic transcriptional identity.

In conclusion, cell cycle progression appears to be necessary for endothelial cells to start EHT and for the acquisition of full functionality and blood forming capacity. Mechanistically,

one possibility is that cell cycle progression is required for the cells to become responsive to specific signals and for the activation of differentiation mechanisms. In this context, understanding signalling pathways driving the cell cycle entry could help improve the generation of haematopoietic progenitors *in vitro*.

# 5 CYCLIN-CDK COMPLEXES LINK CELL CYCLE REGULATION TO HAEMATOPOIETIC DIFFERENTIATION

## 5.1 Summary

In the previous chapter, I unveiled how cell cycle and haematopoietic specification are tightly co-regulated. Cell cycle entry of HECs appeared to be limited to cells undergoing EHT and committing to blood lineages, while cell cycle progression was shown to be essential for the correct generation of haematopoietic cells. Based on these findings, I decided to further explore the connection between cell cycle regulation and haematopoietic commitment. For that, I focused on the role of CDK proteins during EHT. CDKs are the effector subunits of key regulators of the cell cycle progression that act during checkpoints in G1/S and G2/M upon binding to their regulatory cyclins. Importantly, CDKs have also been shown to possess a variety of non-canonical functions and to be important for other processes such as differentiation. In pancreatic progenitors, lengthening of the G1 phase was shown to be necessary for the complete induction of NEUROG3, essential for endocrine development. Upon cell cycle entry, the G1/S kinases CDK4/6 and CDK2 were suggested to phosphorylate NEUROG3, causing its subsequent degradation and ensuring that endocrine differentiation only occurs in non-dividing cells (Krentz et al., 2017). CDK5 was reported to have a neuroprotective role by promoting cell cycle arrest in neurons, but also to regulate through its kinase activity other events during development of the nervous system, including neuronal migration and axon and dendrite development (Zhang et al., 2010; Su and Tsai, 2011). Furthermore, CDK4/6 can direct early cell fate decisions in hPSCs by controlling the transcriptional activity of key transcription factors such as SMAD2/3, therefore determining cell fate propensity towards mesendoderm or neuroectoderm (Pauklin and Vallier, 2013). Based on my results revealing a link between cell cycle progression and haematopoietic differentiation, CDKs might have a similar role during

EHT in controlling specification of early HSPCs during development. Therefore, I decided to investigate the role of these regulators.

## 5.2 Functional effect of CDK inhibitors during EHT

As previously discussed (Chapter 1.8.2), CDKs are master regulators of the cell cycle which are activated by cyclins specifically expressed at different phases. Importantly, this specific expression allows for the temporary activation of cyclin-CDK complexes at distinct cell cycle phases. To study the role of these regulators during EHT, I took advantage of three small molecules, well-characterised inhibitors of cyclin-CDK complexes (CDKi) which bind to distinct CDK proteins and block their phosphorylation activity (McInnes, 2008; Figure 5.1).

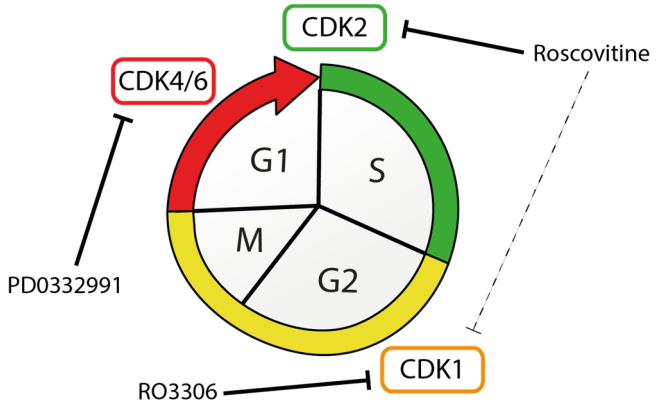



**Figure 5.1 Inhibition of cyclin-CDK function by small molecules.** For the study I have used the small molecules PD0332991, Roscovitine and RO3306, inhibitors of CDK4/6, CDK2 and CDK1, respectively.

PD0332991 inhibits CDK4 and CDK6 (Finn et al., 2009; Rocca et al., 2014), both of which form complexes with cyclin Ds and phosphorylate Rb proteins to enable progression through G1 phase. RO3306 inhibits CDK1 (Vassilev, 2006; Prevo et al., 2018) which interact with cyclin Bs and cyclin As for the progression through G2/M. The main Roscovitine target is CDK2 (Vella et al., 2016), which upon binding to cyclin Es and cyclin As leads to G1/S transition and progression through the S phase. This molecule has also been reported to inhibit CDK1, although with lower affinity (Meijer et al., 1997; Knockaert et al., 2000).

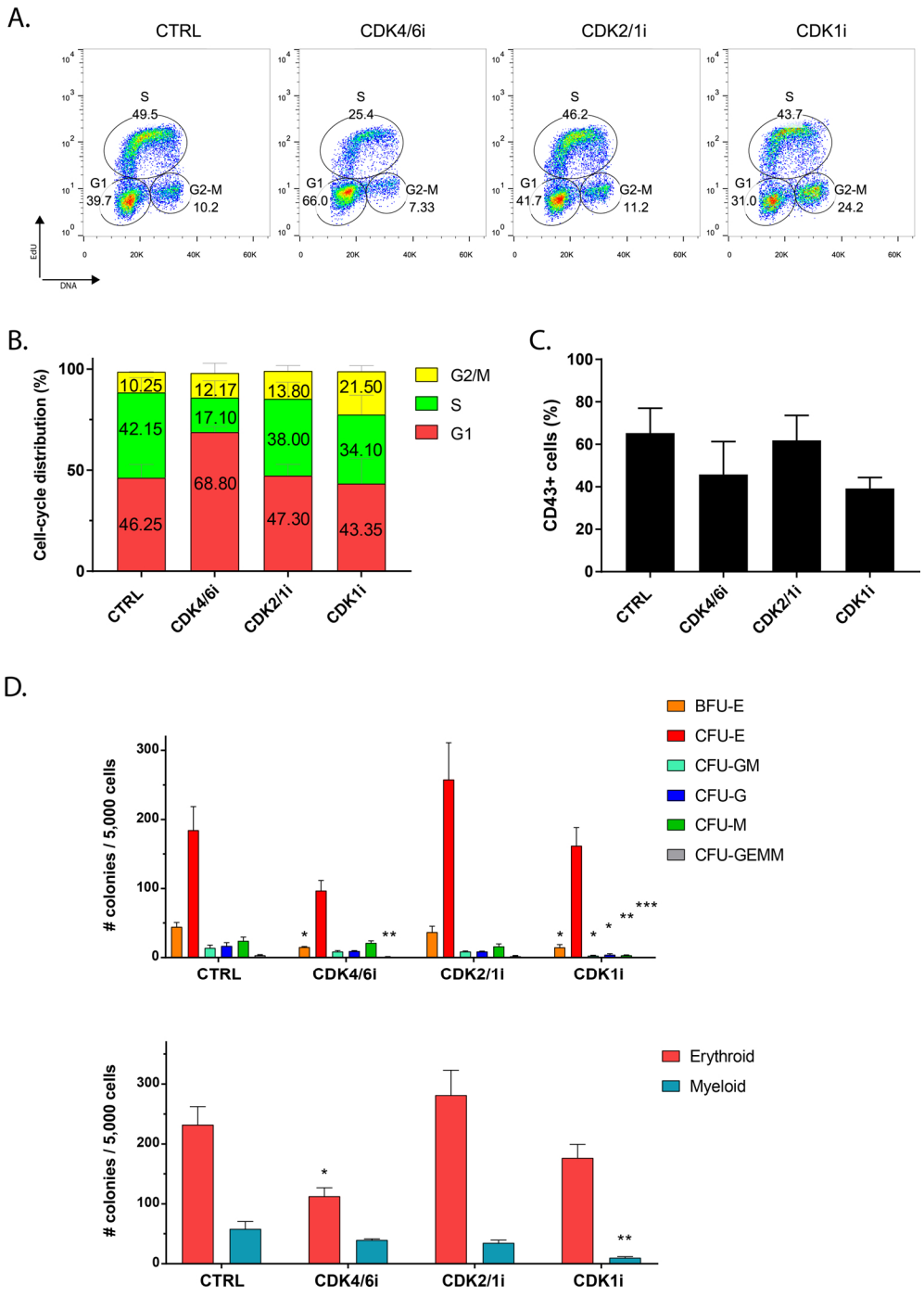



**Figure 5.2 Effect of CDK inhibitors during EHT.** Samples were treated with either of the three small molecules or with 0.1% DMSO (CTRL) for 48 hours between EHT D3 and EHT D5. **(A, B)** As shown by EdU staining, the treatments caused an expected change in the cell cycle profile: enrichment of cells in G1 phase for the CDK4/6 inhibitor PD0332991 (1 µM); very limited increase in both G1 and G2/M for the CDK2/1 inhibitor Roscovitine (4 µM); increase in G2/M phase for the CDK1 inhibitor RO3306 (10 µM). **(C)** This was associated with a decrease in the number of CD43+ cells for samples treated with CDK4/6i and even more with CDK1i, as assessed by flow cytometry. **(D)** Most importantly, the treatment affected the blood colony forming potential upon CFU assay. CDK4/6i caused a significant decrease in the number of erythroid colonies, particularly of the BFU-E type. CDK2/1i caused an increase, although not significant, in the number of mature CFU-E erythroid colonies. Most importantly, CDK1i caused the greatest effect, with significant disruptions in both the erythroid and myeloid colonies. Of note, both CDK4/6i and CDK1i had an important effect on the number of CFU-GEMM colonies derived from multipotent progenitors. Results in (B) and (C) represent n=2 independent experiments, one of which is shown in (A), and error bars are ±SEM. Results in (D) represent n=3 independent experiments, with error bars ±SEM, and statistical significance compared to control as *P<0.5, **P<0.01, ***P<0.001 by one-way ANOVA.

As for previous experiments, cells were differentiated until EHT D3 and then grown in the presence of one of the inhibitors for 48 hours. At EHT D5 cells were collected for further analyses. Specifically, I assessed their cell cycle profile by EdU staining, the number of CD43+ cells produced by flow cytometry and, upon removal of the inhibitors, their blood colony forming capacity by CFU assay. The results revealed interesting effects which appeared to be specific for distinct treatments (Figure 5.2).

The inhibition of CDK4/6 (CDK4/6i) caused an expected enrichment of cells in the G1 phase, a reduction in the number of CD43+ cells, and an overall reduction in the number of colonies generated in CFU assay, especially of the erythroid lineage. This effect would be consistent with my previous results showing that cell cycle entry is necessary for correct EHT. In this context, the inhibition of CDK4/6 would possibly limit endothelial cells exiting G1 phase or delay their transition to the S phase, therefore preventing them from engaging EHT and reducing the number of haematopoietic progenitors generated.

Treatment with Roscovitine did not cause major disruptions in the cell cycle profile, as shown by a limited increase in the number of cells found in the G1 and G2/M phases. Nevertheless, the fact that this small increase was observed in these two phases seemed to confirm that this molecule affected both CDK2 and CDK1 (CDK2/1i). This limited change in cell cycle profile was associated with a very small decrease in the number of CD43+ cells and with an increase, although not significant, in the number of CFU-E colonies, representative of more mature erythroid progenitors, possibly suggesting an increased commitment towards this lineage.

Finally, the inhibition of CDK1 (CDK1i) blocked the cell cycle in the G2/M fraction, decreased the CD43 compartment and reduced the number of BFU-E and myeloid colonies. Furthermore, both the inhibition of CDK1 and CDK4/6 caused a depletion in the number of multilineage CFU-GEMM colonies. My previous data (Figure 3.7) showed that myeloid and multilineage colonies were specifically derived from the small cluster labelled as multipotent HPCs in the scRNAseq tSNE plot. Therefore, this would suggest the two molecules to possibly affect generation, self-renewal and/or commitment capacity of this specific population, with the inhibition of CDK1 being the most detrimental treatment.

In order to further investigate the effect of the CDK inhibitors on EHT, I compared the expression of key marker genes by qPCR at EHT D5 after a 48 hours treatment (Figure 5.3, Figure 5.4). This approach showed only mild differences, the strongest being achieved with the inhibition of CDK4/6. Specifically, I could observe a small increase in the expression of the endothelial markers *VWF*, *PECAM1* and *CDH5* upon this treatment, and at the same time a limited decrease in HSC genes like *MYB* and *RUNX1*. This would be consistent with a delayed G1 exit causing reduced EHT transition and thus an enrichment in endothelial cells. In agreement with this hypothesis, I could also observe an increase in the expression of *cyclin D1* and *D2*, which I had previously shown through scRNAseq to be more expressed on endothelial and mesenchymal cells, while haematopoietic cells were relying more on *cyclin D3* (Figure 4.3).

However, these differences were very limited and gene expression measured on a population level by qPCR could not reveal important effects induced by inhibition of CDK regulators. This observation could be explained by the heterogeneous nature of the culture system. Indeed, different cell types participating in the transition could be differentially affected by the treatments and contribute in distinct ways to the phenotypes. For this reason, I decided to perform scRNAseq to capture differences between intermixed cell types and therefore possibly isolate and discriminate lineage-specific effects.

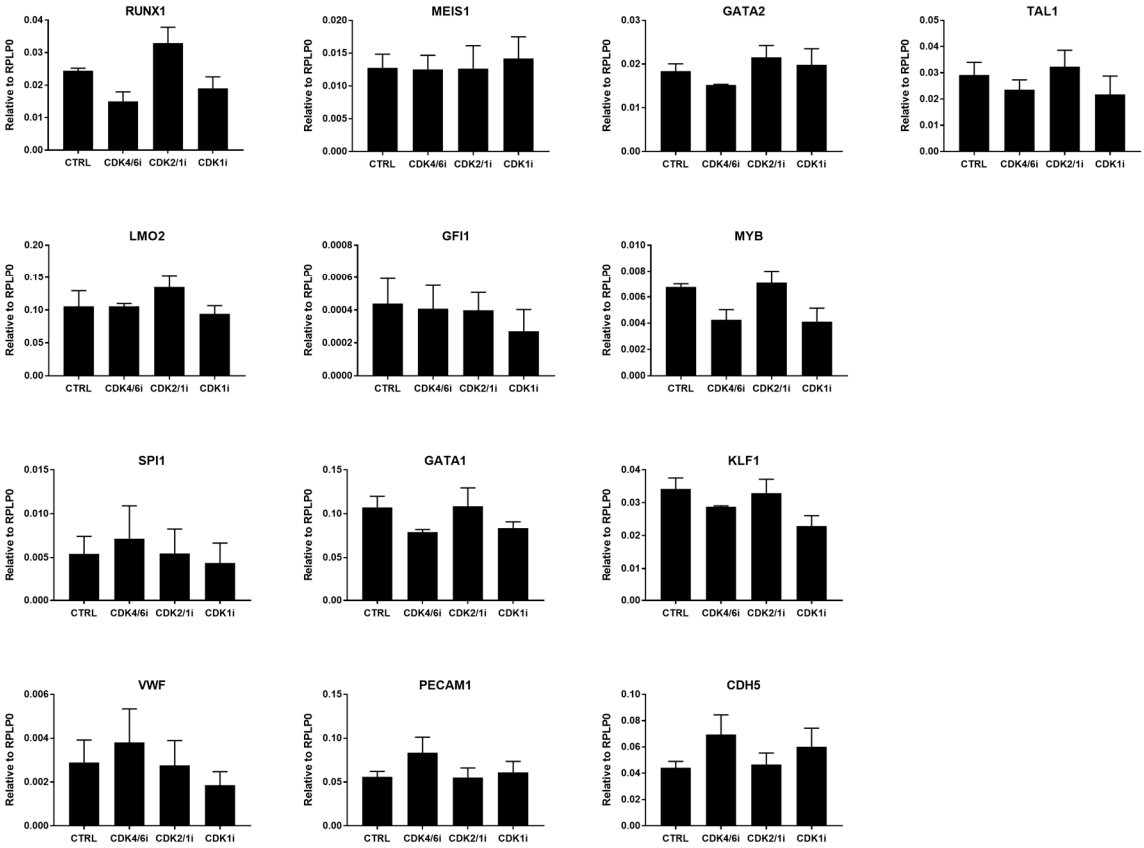



**Figure 5.3 Expression of key marker genes by qPCR upon treatment with CDK inhibitors.** The expression of key haematopoietic and endothelial genes upon treatment with CDKi was measured at EHT D5 by qPCR. Results represent n=2 independent experiments, with error bars ±SEM.

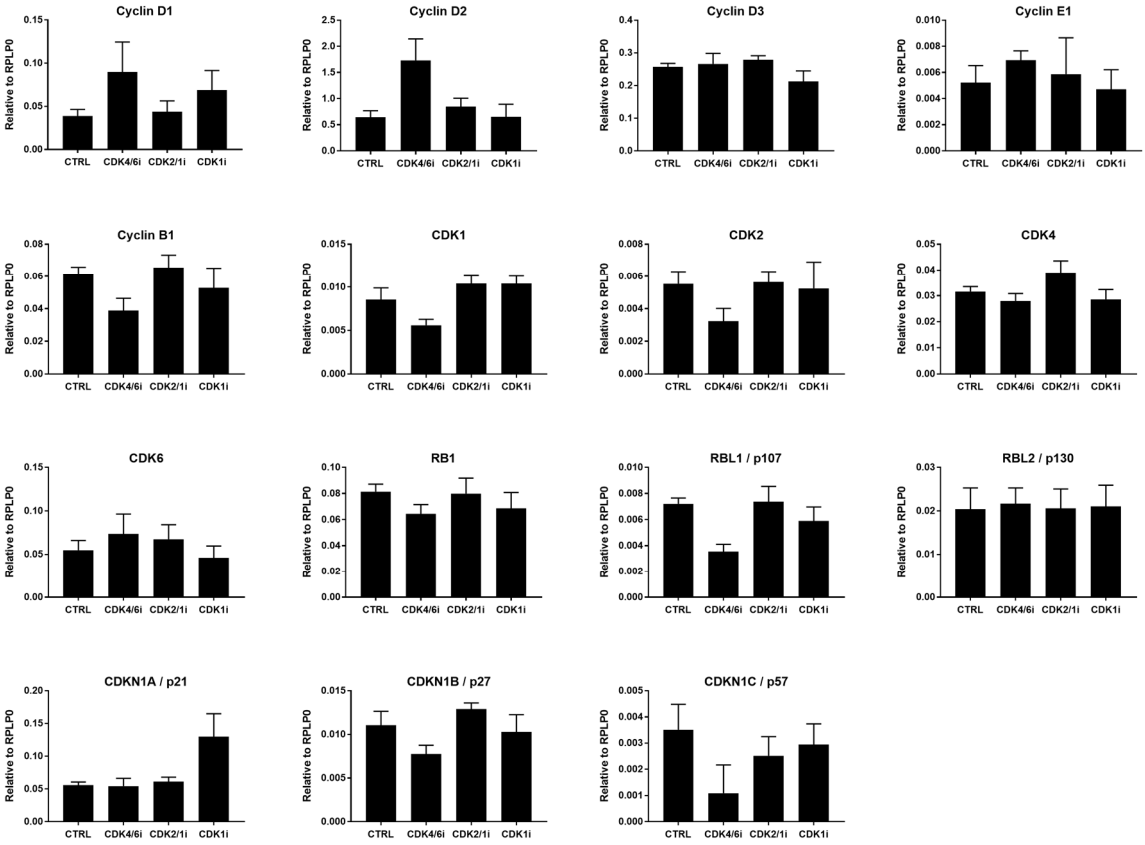



**Figure 5.4 Expression of cell cycle regulator genes by qPCR upon treatment with CDK inhibitors.** The expression of key cell cycle regulators upon treatment with CDKi was measured at EHT D5 by qPCR. Results represent n=2 independent experiments, with error bars ±SEM.

## 5.3 Single cell transcriptomics reveal disruptions in cluster organisation and cell cycle state upon inhibition of CDKs

Cells were grown for 48 hours starting at EHT D3 in the presence of specific CDK inhibitors or in the presence of 0.1% DMSO. The resulting cells collected at EHT D5 were then processed for scRNAseq. Data was analysed as previously described (Chapter 3.2), combining the four samples and clustering in the 3D tSNE space (Figure 5.5). By combining and analysing the samples together, it was possible to detect the 4 clusters previously identified and compare each cell population across different conditions. Using this approach, I could observe that the relative population frequencies were specifically affected by distinct treatments (Figure 5.6).

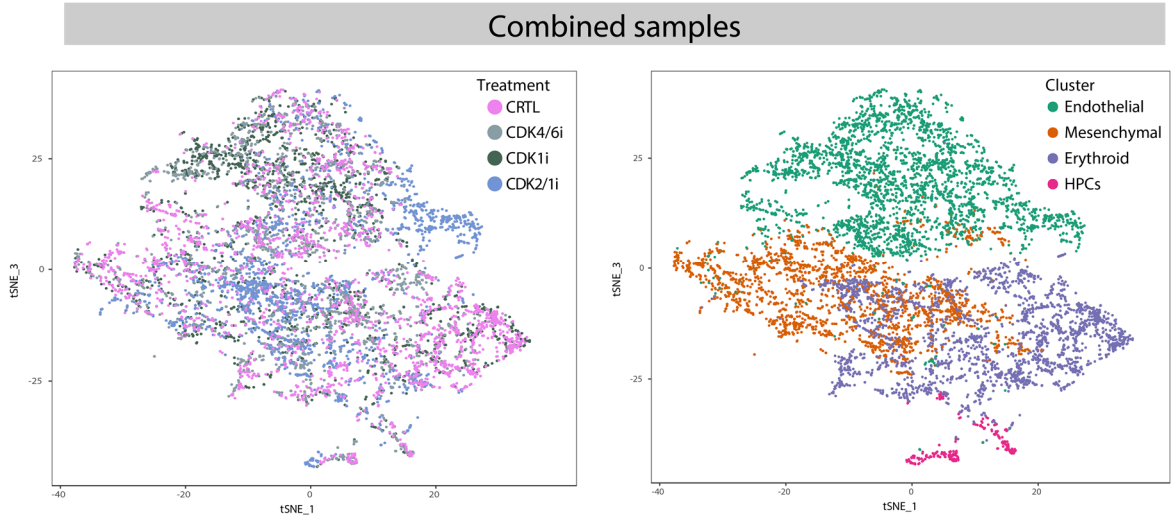



**Figure 5.5 Clustering in the tSNE space of samples upon different treatments.** Samples treated with either of the three small molecules or with DMSO were processed for scRNAseq. Data for the four samples was combined, and distinct clusters in the tSNE space were identified based on transcriptional identity.

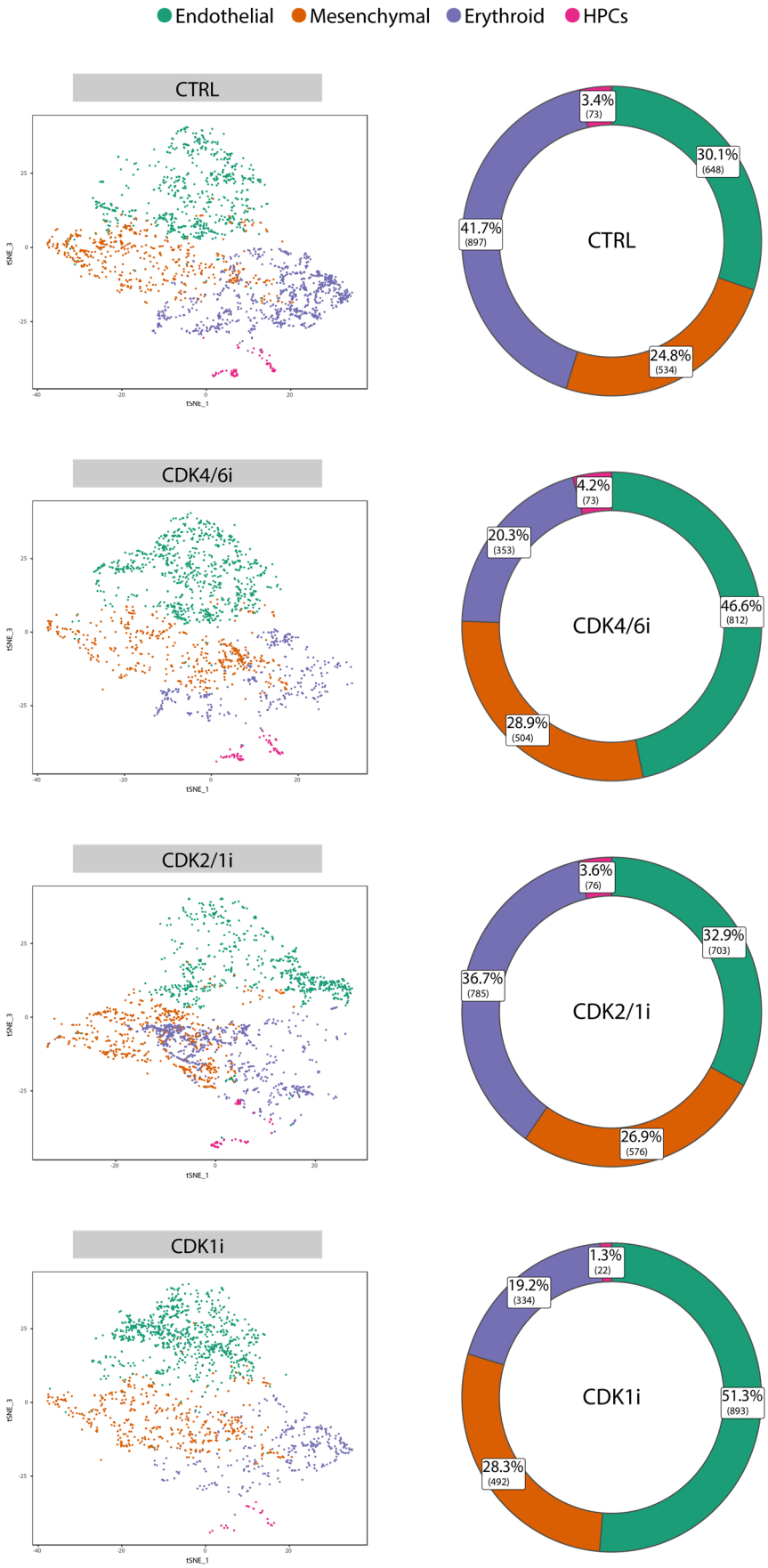



**Figure 5.6 Effect of CDK inhibitors on the tSNE generated clusters.** The analysis allowed to observe the effect of CDKi during EHT with single cell precision and to directly compare equivalent clusters across samples. The treatments appeared to affect the distribution of the populations previously described.

These results showed enrichment in endothelial cells upon CDK4/6 inhibition. This change was associated with a decrease in the haematopoietic compartment, specifically for the erythroid progenitors, consistent with previous CFU results (Figure 5.2D). Interestingly, the population of HPCs did not appear to be reduced but instead slightly increased. This was not consistent with the observed decrease in the CFU assay in the number of multilineage colonies, known from previous data to derive exclusively from this cluster (Figure 3.7). This suggested that beyond affecting the proliferation of erythroid progenitors, the inhibition of CDK4/6 might be affecting the generation of functional HPCs or their ability to further differentiate.

The CDK2/1 inhibitor did not cause major differences in the number of cells belonging to each cluster compared to control, consistent with the fact that at the functional level, except for a limited increase in the number of CFU-E colonies, this treatment was the least disruptive.

On the other hand, cells grown in the presence of the CDK1 inhibitor were strongly affected. Indeed, I could observe an increased number of endothelial cells concomitant with a decrease in the haematopoietic compartment. Moreover, this treatment was the most detrimental for the HPC cluster. This was in agreement with previous results showing that the treatment caused a reduced number of myeloid and multilineage colonies, which are derived from this population (Figure 5.2D, Figure 3.7). Thus, these results appeared to suggest an important role for CDK1 in these early multipotent progenitors, possibly affecting their generation or self-renewal.

Of note, none of the treatments caused major changes in the size of the mesenchymal cluster, possibly indicating that the cell cycle regulation is specifically important for EHT but not for EndoMT.

Additionally, the transcriptomics data was analysed to examine the cell cycle distribution of individual cell clusters in each sample (Figure 5.7). One possibility was indeed that distinct cell types would be affected with different specificity by CDK inhibition. As previously shown, we determined the cell cycle state of individual cells based on the expression of canonical marker genes, and compared corresponding clusters across different conditions.

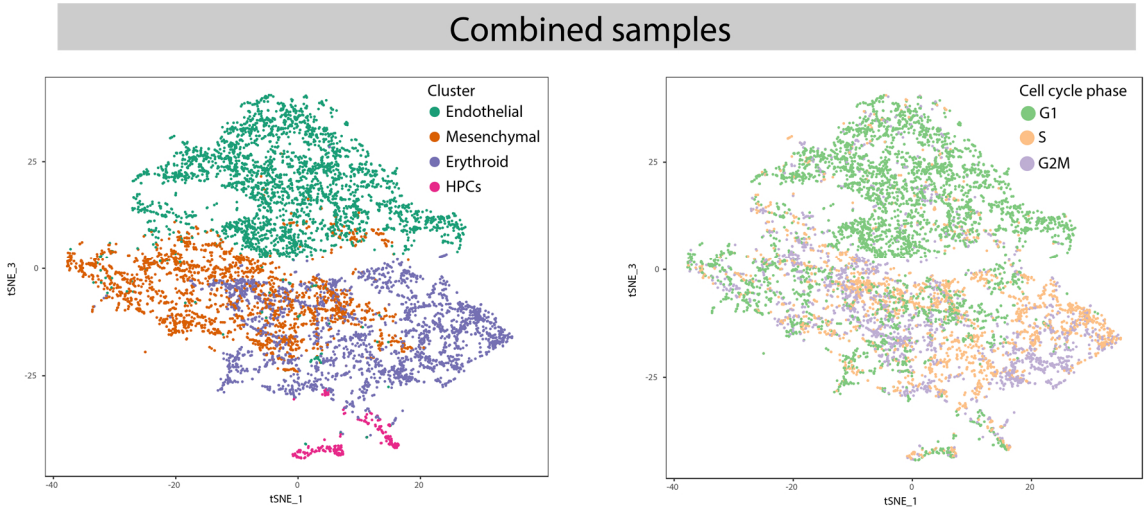



**Figure 5.7 Cell cycle profile of distinct clusters in the combined dataset.** Cell cycle marker genes identified in the scRNAseq dataset were used to calculate cell cycle scores.

As expected and in agreement with EdU staining (Figure 5.2A,B), the inhibition of CDK4/6 caused an enrichment of cells in G1 (Figure 5.8). More precisely, the treatment increased the number of endothelial cells found in G1 to almost 97%. Thus, it appeared that the inhibition of CDK4/6 could prevent or delay their cell cycle entry and possibly prevent them from undergoing EHT. This would explain the observed increase in the endothelial fraction compared to control (Figure 5.6). Importantly, an increase in the number of cells in the G1 phase was also shown for both the HPC and erythroid clusters, the latter previously shown to be decreased in its size upon treatment. These profiles and the relative cluster frequencies might suggest that endothelial cells produce erythroid progenitors through a highly transitory HPC population, and that blocking CDK4/6 affects this process.

Overall, these results explain the reduction in the number of CD43+ cells and of colonies produced in the CFU assay, especially of the erythroid lineage (Figure 5.2). At the same time, multilineage CFU-GEMM colonies were shown to derive exclusively from HPCs, whose cluster size was not depleted by the treatment. Therefore, this would suggest that the reduced number of this type of colony was entirely dependent on an altered capacity to differentiate as a consequence of CDK4/6 inhibition, rather than decreased HPC production.

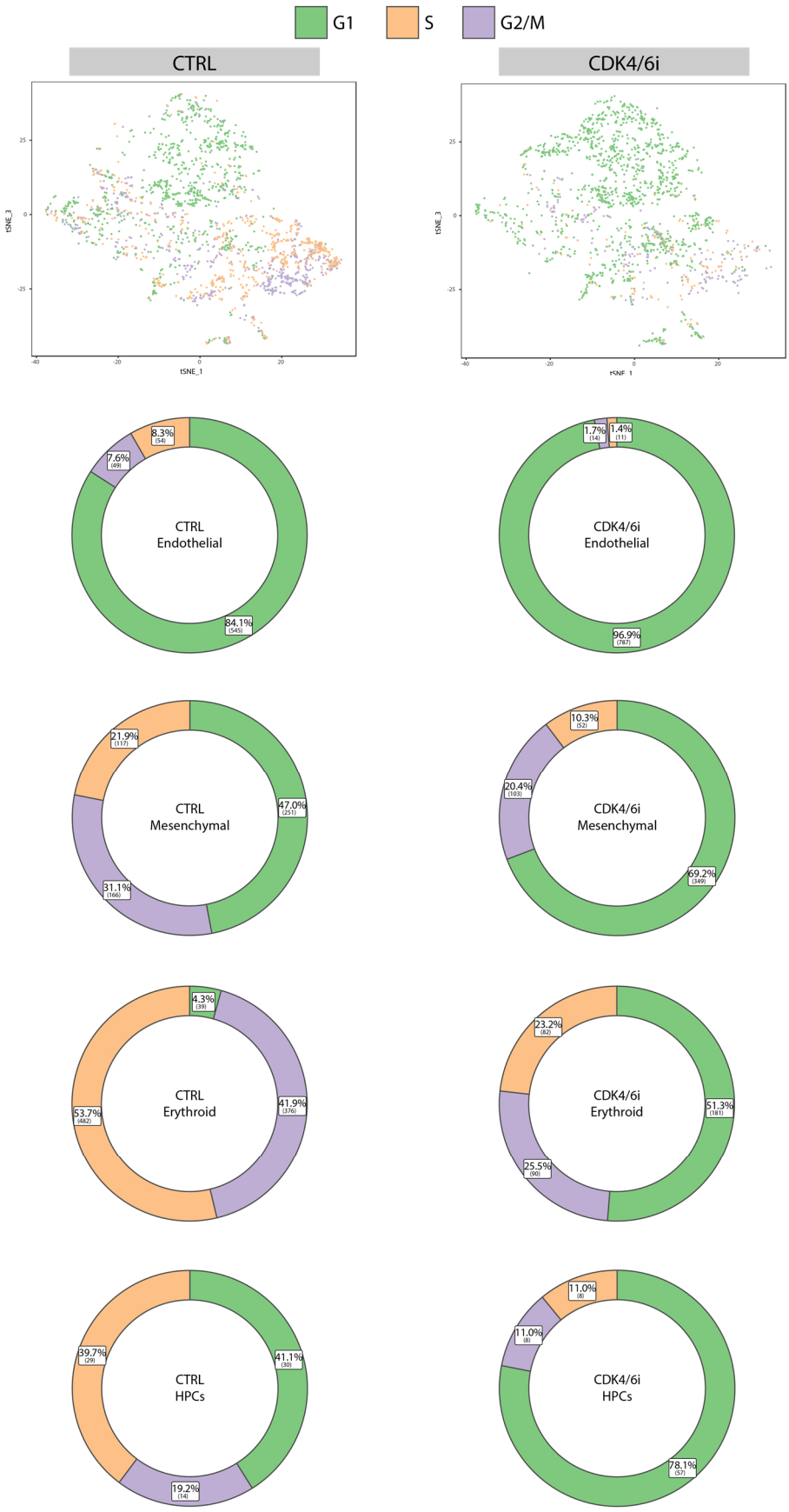



**Figure 5.8 CDK4/6i cell cycle profiles.** The cell cycle profile of individual clusters upon inhibition of CDK4/6 was calculated based on transcriptomics data and compared to the control condition.

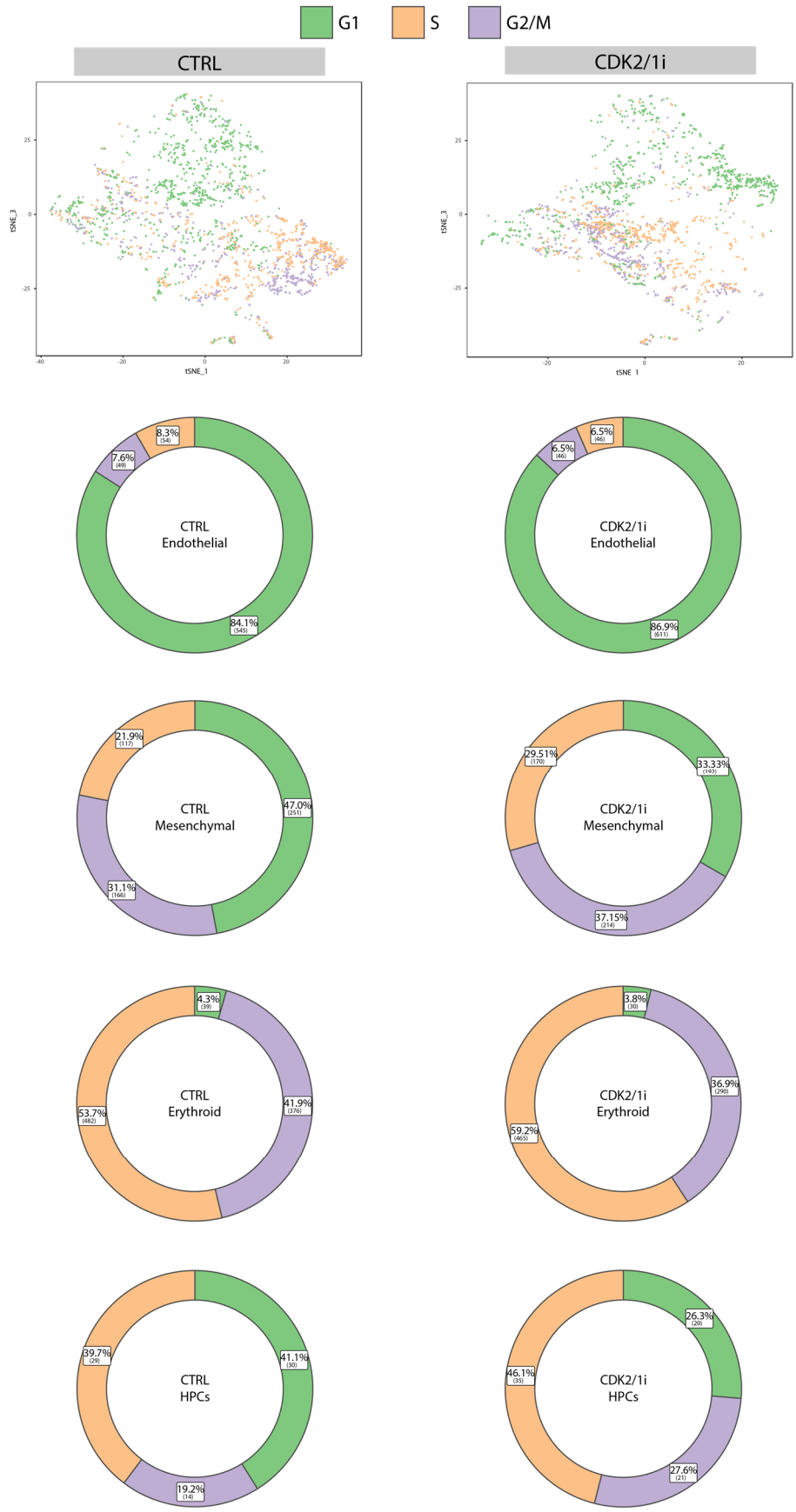



**Figure 5.9 CDK2/1i cell cycle profiles.** The cell cycle profile of individual clusters upon inhibition of CDK2/1 was calculated based on transcriptomics data and compared to the control condition.

In contrast to this, the inhibition of CDK2/1 caused overall a less intense effect and the cell cycle profile of the endothelial cluster was essentially unaffected (Figure 5.9). The main differences appeared on the multipotent HPC cluster and, on a lesser extent, the erythroid progenitors. A limited enrichment in S phase at the expenses of G2/M was observed in the erythroid cells suggesting a specific effect of the inhibitor on CDK2 for this particular population. On the other hand, the treatment increased both S and G2/M phases in the HPC cluster, suggesting that in this specific population the inhibitor could affect both CDK2 and CDK1. Nevertheless, considering these and previous results, the effect of this treatment appeared to be overall very limited, both at the transcriptional and functional level.

Finally, the inhibition of CDK1 caused an expected enrichment in G2/M for the endothelial cluster (Figure 5.10). Given the key function of CDK1 in G2/M, it appeared that the cells blocked in this phase were endothelial cells engaged in cell cycle progression, possibly representing cells undergoing EHT, as previously discussed (Chapter 4.5). Accordingly, upon CDK1 inhibition I observed a reduction in the number of multipotent HPCs, whose number was uniquely affected by this treatment. Nevertheless, I previously assessed the presence of HPCs already at EHT D3 (Figure 3.8), the time when the CDK1 inhibitor was applied. This would suggest that beyond a possible influence over the generation of new HPCs from haemogenic endothelial cells, the treatment might also affect previously generated progenitors, possibly altering their survival or self-renewal capacity, for example promoting their further differentiation and exhaustion. This might confirm that the HPC population is normally very transitory and that their further differentiation is associated with cell cycle progression. Consequently, the inhibition in G2/M might promote their commitment. Indeed, the analysis showed that the remaining HPCs were enriched in the G1 phase, possibly suggesting that these cells need to progress through the cell cycle to differentiate and that instead they retain an undifferentiated state in G1. This would be in agreement with previous results showing that the HPC cluster is instead preserved and slightly enriched when CDK4/6 is inhibited (Figure 5.8), although in this case their ability to further differentiate was affected. Furthermore, the cell cycle of the erythroid progenitors was surprisingly not enriched in G2/M phase, possibly suggesting functional redundancy in this specific cell type.


**Figure 5.10 CDK1i cell cycle profiles.** The cell cycle profile of individual clusters upon inhibition of CDK1 was calculated based on transcriptomics data and compared to the control condition.

## 5.4 CDK inhibition differentially affects gene expression of distinct clusters

As previously mentioned, cell cycle regulators have been shown to have multiple roles beyond their control over cell cycle progression (Hydbring et al., 2016) and to direct differentiation and cell fate decision in multiple systems (Zhang et al., 2010; Su and Tsai, 2011; Pauklin and Vallier, 2013; Krentz et al., 2017). Therefore, it is possible that the CDK role in specification and further differentiation of *in vitro* generated haematopoietic progenitors might be related not uniquely to their function in cell cycle progression but also associated with secondary roles performed by these regulators. Ultimately, the inhibition of CDK function might possibly affect signalling pathways or other molecular mechanisms taking place during EHT.

To investigate this possibility, I compared differentially expressed genes for each cluster across distinct treatments. For this analysis, only genes showing a differential expression greater than 0.5 $\log_2$ fold changes were considered, and I used the Enrichr resource to perform gene ontology (GO) analysis and determine biological processes involving such differentially expressed genes (Kuleshov et al., 2016). This approach showed small changes in specific clusters in the expression of a variety of genes.

Upon inhibition of CDK4/6, each cluster compared to control showed downregulation of genes for ribosomal subunits or enrichment in GO terms involved in mitochondrial functioning, possibly consistent with a reduced translational and metabolic activity as a consequence of a slowed cell cycle progression (Figure 5.11, Table 5.1 to Table 5.6). At the same time, the haematopoietic clusters showed a limited upregulation of genes associated with membrane reorganization and exocytosis. Of note, the multipotent HPC cluster upon treatment showed a small increase in few haemoglobin genes, possibly suggesting a priming toward the erythroid lineage.

On the other hand, the very limited functional effect on differentiation and cell cycle profile observed for samples treated with the CDK2/1 inhibitor was reflected by their

transcriptional signature, as almost all the clusters showed no or very little change (Figure 5.12, Table 5.7). The only relevant effect was observed on the HPC cluster, showing upregulation of erythroid genes such as *ALAS2* and haemoglobins, along with the platelet factors *PPBP* and *PF4*. This would be in agreement with the increase in mature erythroid CFU-E colonies found in the CFU assay, although that increase was not statistically significant (Figure 5.2).

Finally, the inhibition of CDK1 caused downregulation of ribosomal subunits and of genes enriched in GO terms associated with mitochondrial functions, together with the downregulation in HPCs of factors involved in gene expression and mRNA processing (Figure 5.13, Table 5.8 to Table 5.14). This could be possibly associated once again with a reduced cellular activity caused by the enrichment in G2/M and delayed cell cycle progression. In agreement with this effect on the cell cycle, the upregulation of genes associated with mitosis, cytoskeleton rearrangement and G2/M transition was observed in the endothelial population. In addition, the cluster of multipotent HPCs showed upregulation of the erythroid *ALAS2* and haemoglobin genes, with downregulation of the myeloid factor *MPO*, in agreement with the lack of myeloid potential observed in the CFU assay. Interestingly, inhibition of CDK1 also caused in this cluster the upregulation of *MALAT1*, *NEAT1* and *TXNIP*, previously reported to crosstalk with the MAPK signalling in endothelial cells and mediate the cellular response to stress conditions like oxidative stress (Li et al., 2009; Liu et al., 2014; Fuschi et al., 2017; Tian et al., 2017). Some of these factors were also upregulated in erythroid (*MALAT1* and *NEAT1*) and endothelial cells (*NEAT1*). Overall, the upregulation of these genes, along with the downregulation of genes associated with mitochondrial functions, potentially suggested that the inhibition of CDK1 promote the cell response to stress.

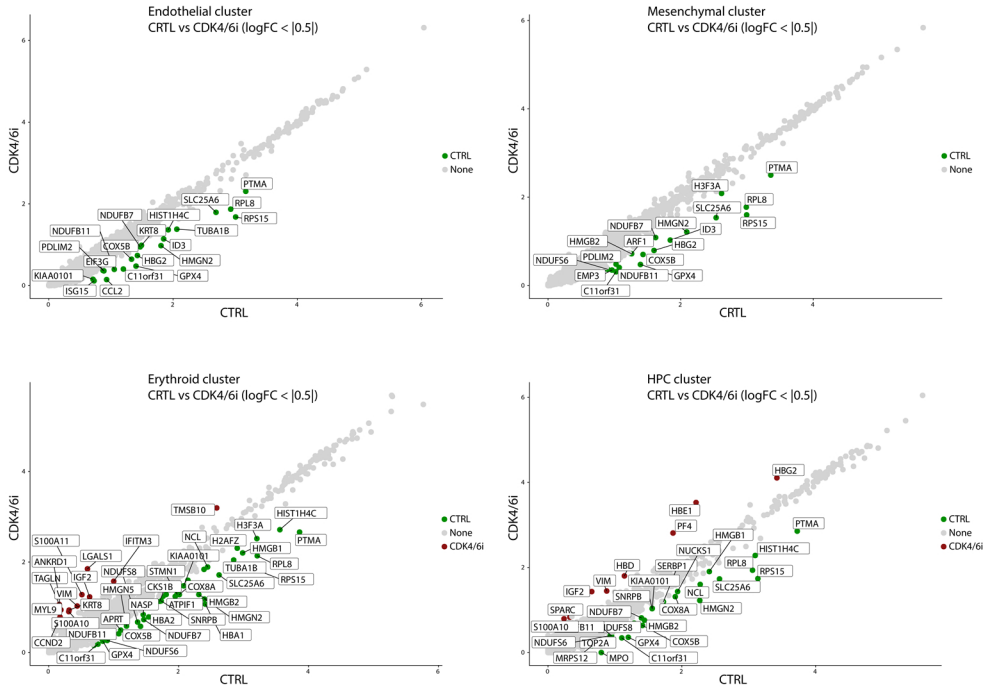



**Figure 5.11 CDK4/6i differentially expressed genes per cluster.** Differentially expressed genes between treatment and control are shown in these regression plots divided by cluster. For each gene and on each sample, axes are $\log_2$ of the average counts. Differentially expressed genes are highlighted. Genes falling into the diagonal are not considered as differentially expressed. Cut-off used is $\log_2$ fold change = 0.5.

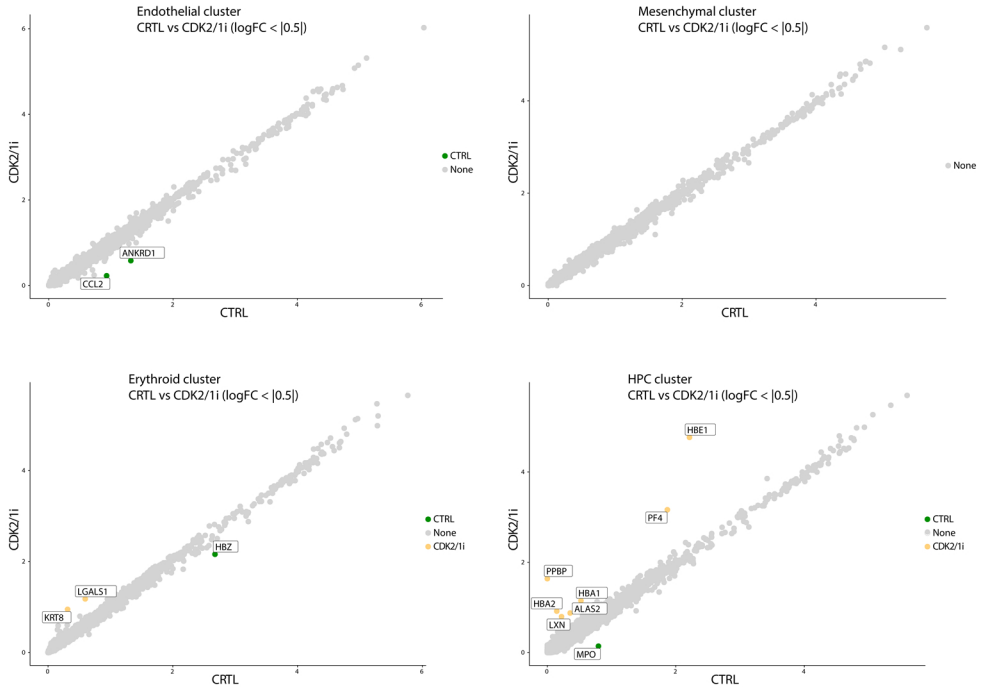



**Figure 5.12 CDK2/1i differentially expressed genes per cluster.** Differentially expressed genes between treatment and control are shown in these regression plots divided by cluster. For each gene and on each sample, axes are $\log_2$ of the average counts. Differentially expressed genes are highlighted. Genes falling into the diagonal are not considered as differentially expressed. Cut-off used is $\log_2$ fold change = 0.5.

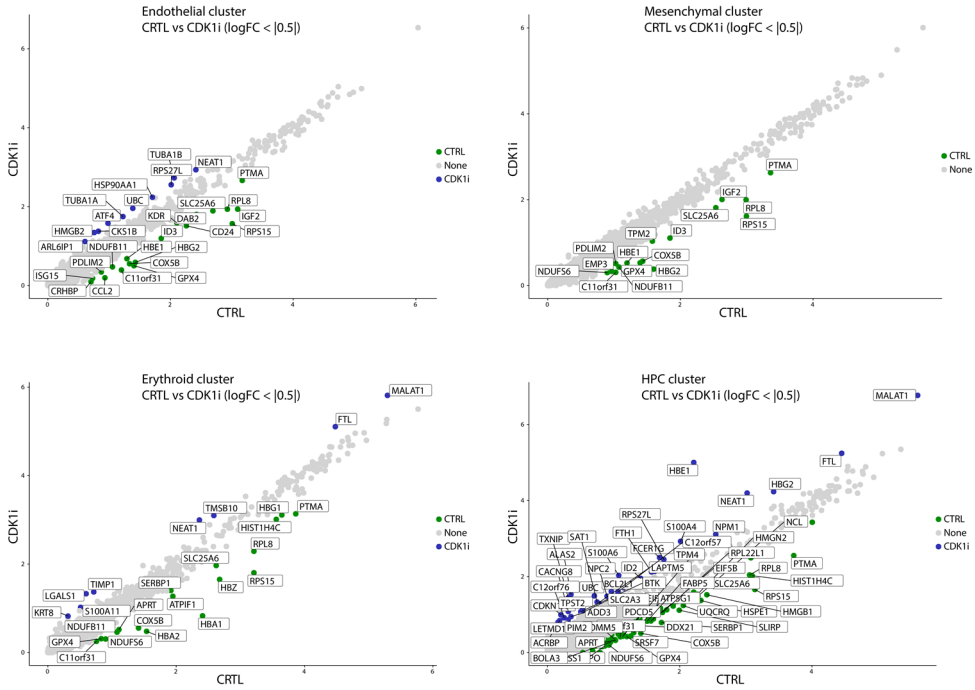



**Figure 5.13 CDK1i differentially expressed genes per cluster.** Differentially expressed genes between treatment and control are shown in these regression plots divided by cluster. For each gene and on each sample, axes are $\log_2$ of the average counts. Differentially expressed genes are highlighted. Genes falling into the diagonal are not considered as differentially expressed. Cut-off used is $\log_2$ fold change = 0.5.

## 5.5 Discussion

In this chapter I examined the role during EHT of key CDK proteins, master regulators of the cell cycle also involved in a variety of non-canonical functions, as previously discussed (Chapters 1.8.2 and 5.1). For this, I took advantage of three small molecules, well-characterised inhibitors which bind to the CDKs and block their function. With this approach, and by combining functional experiments and scRNAseq analyses, I was able to study the role of these regulators in the distinct populations participating in the transition.

Based on these results, CDK4/6 appears important for endothelial cell entry into the cell cycle and commitment to the haematopoietic fate. Specifically, this regulator is necessary for G1 progression and transition to the S phase, and blocking its activity causes an expected enrichment in G1 for both the endothelial and haematopoietic compartments. In this context, the reduction in the number of erythroid progenitors might be explained in part by decreased proliferation. However, this is also associated with a concomitant increase in the number of endothelial cells, most likely as a result of their reduced transition to the haematopoietic fate. On the other hand, the number of multipotent HPCs is not affected and is instead slightly increased. Overall, these results suggest that endothelial cells undergoing EHT normally generate a population of multipotent HPCs which, in the current culture conditions, is highly transitory and quickly commit to the erythroid fate (Figure 5.14A). In this scenario, the activity of CDK4/6 appears to be required for both HPC generation and for erythroid differentiation and proliferation. This hypothesis is in agreement with the observed effects upon inhibition of CDK4/6, consisting in a reduced number of CD43+ cells, fewer erythroid colonies generated in the CFU assay, and decrease in the erythroid cluster concomitant with an increase in the endothelial population. Furthermore, this would be consistent with previous studies showing that the erythroid commitment is reinforced during the S phase by the repression of the myeloid transcriptional programme (Pop et al., 2010). In agreement with this, my results show that preventing cells from transitioning to the S phase causes a reduction in the erythroid lineage, therefore suggesting that a similar mechanism might take place in the generation of early erythroid progenitors during development. Accordingly, despite the size of the HPC cluster not

being affected, a reduction in the number of multilineage CFU-GEMM colonies is observed upon inhibition of CDK4/6. This might confirm that CDK4/6 activity and G1/S transition are necessary for erythroid, but not for myeloid, commitment, and that a longer G1 phase results in restricted commitment and therefore fewer multilineage colonies. To address this hypothesis, it will be necessary to isolate the HPC population and specifically test their potential upon inhibition of CDK4/6 by single cell differentiation experiments. This would also allow examination of the capacity of these progenitors to generate erythroid colonies upon CDK inhibition, currently masked by the presence in culture of erythroid cells generated before the treatment, and therefore to discriminate between HPC differentiation and erythroid proliferation. Importantly, the depletion of CDK4/6 activity does not affect the size of the multipotent HPC cluster and these cells do not accumulate despite their reduced differentiation, suggesting that beyond being necessary for their differentiation, CDK4/6 is also essential for endothelial cells to undergo EHT and generate more HPCs. Of note, it was previously reported that CDK6 can bind to RUNX1 and interfere with its transcriptional activity (Fujimoto et al., 2007). RUNX1 is an essential regulator of EHT, which synergises with other key factors such as TAL1 and GATA2. It is necessary for the induction of the definitive transcriptional programme and its requirement in the AGM has been suggested to end after HSPC specification, as previously discussed (Chapter 1.7.5). Thus, it is possible that an important role for the cell cycle machinery is to modulate the timely activation of the haematopoietic transcriptional network, and in this context the inactivation of RUNX1 by CDK6 might be necessary for the completion of the haematopoietic specification. In agreement with this hypothesis, based on my results CDK4/6 activity, which is necessary for cell cycle entry and G1/S progression, also appears to be important during EHT for both HPC specification and further differentiation to the erythroid fate (Figure 5.14B).

Similarly, CDK1 also appears to have an important role in endothelial cells during EHT. In agreement with its function in cell cycle progression, endothelial cells display an enrichment in the G2/M phase upon inhibition of CDK1. This results in an increased endothelial population, suggesting that CDK1 is necessary for their progression to the haematopoietic fate. Indeed, this corresponds to a reduction in the haematopoietic compartment, as also indicated by the lower number of CD43+ cells. However, the decrease is not limited to the erythroid

cluster as seen for CDK4/6, and instead involves also the population of multipotent HPCs, in agreement with the fewer myeloid and multilineage colonies observed in the CFU assay. This potentially highlights an essential role for CDK1 during EHT, with its activity being important for the generation of multipotent HPCs. In addition, the reduction in the number of HPCs might also depend on a role of CDK1 on their commitment, so that blocking HPC progression once the cell cycle has been engaged might be not able to prevent their further differentiation. Indeed, it is possible that the undifferentiated state of this transitory population can only be preserved in the G1 phase. In this scenario, CDK1 would be essential for the completion of the cell cycle, thus allowing the progenitors to return to G1, possibly to achieve balanced asymmetric cell division. In agreement with this, the number of HPCs appear to be decreased also compared to pre-treatment levels, and the fewer HPCs that remain following inhibition of CDK1 are counterintuitively enriched in G1, and possibly represent cells that did not enter the cell cycle and were not affected by the treatment. Accordingly, and as discussed above, the inhibition of CDK4/6 causes enrichment of these cells in G1, thus preserving the cluster size, although impairing their multilineage differentiation potential. Together, these results suggest that HPCs differentiate upon cell cycle entry, likely following a default differentiation programme towards the erythroid compartment, and that the population is exhausted when G1 re-entry is prevented. Thus, CDK1 activity appears to have a key role during EHT, being potentially essential for both the specification of multipotent HPCs and the prevention of their uncontrolled differentiation leading to exhaustion (Figure 5.14C). Further experiments including the isolation and single cell differentiation of the HPCs upon treatment will be able to confirm if this model is correct. Of note, the cell cycle state of the erythroid progenitors does not seem to be affected by the inhibition of CDK1. This regulator was previously reported to be absolutely fundamental during mouse embryonic development, with other CDKs not able to drive proliferation beyond the blastocyst stage (Santamaría et al., 2007b). However, other CDKs can replace CDK1 in the adult, as shown by the use of conditional knockout mice revealing that liver regeneration is not impaired by loss of CDK1 (Diril et al., 2012). Thus, a similar functional redundancy with other CDKs replacing CDK1 might take place in the erythroid cluster, allowing a normal cell cycle progression.

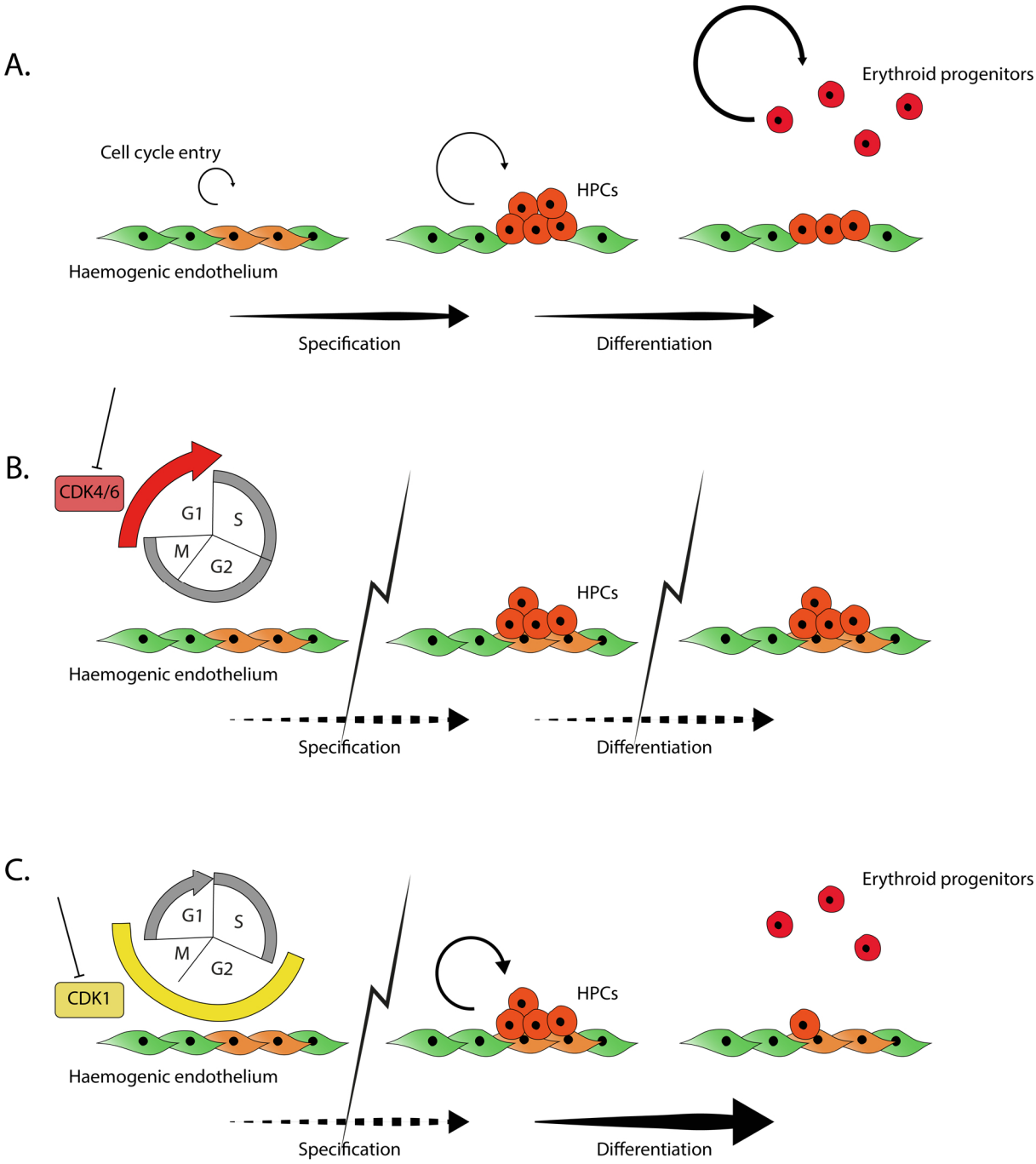



**Figure 5.14 Proposed model for the role of CDK4/6 and CDK1 during EHT. (A)** Endothelial cells transitioning to the haematopoietic fate are associated with cell cycle entry. This results in the specification of HPCs, a transitory population which in culture quickly commits to produce highly proliferative erythroid progenitors. **(B)** CDK4/6 is necessary for both specification and further differentiation of HPCs. Its inhibition results in the endothelial cells blocked in G1 and not undergoing EHT, and with the already generated HPCs also blocked in G1 and unable to further differentiate. The consequence is an enrichment in endothelial cells, a decrease in erythroid progenitors, and the preservation of the HPC population associated with depletion of their multilineage potential, possibly skewed towards the myeloid lineage. **(C)** CDK1 is necessary for HPC specification and balanced differentiation. Its inhibition blocks EHT, while the commitment of already generated HPCs is not blocked, resulting in their further differentiation and quick exhaustion. This leads to enrichment in endothelial cells, decrease in erythroid progenitors, and HPCs depletion.

Functional redundancy might also explain the nonessential requirement observed for CDK2 during EHT. It was previously shown that CDK2 knockout mice are viable, despite germ cell development in these animals being affected (Berthet et al., 2003). Similarly, CDK2 does not appear to be required for endothelial cells to undergo EHT, possibly replaced by other CDKs, and its inhibition causes in erythroid cells only a very limited enrichment in S phase. Of note, the treatment instead increases both S and G2/M phases in the HPC cluster, suggesting that in this particular population both CDK2 and CDK1 might be affected by the inhibitor. This might explain the little increase in the number of CFU-E colonies generated in the CFU assay, in agreement with what discussed for the role of CDK1 and with previous studies suggesting that the erythroid differentiation programme is reinforced during the S phase (Pop et al., 2010). However, this treatment does not seem to block CDK1 very efficiently, while CDK2 might be functionally replaced by other CDKs, overall explaining the limited effect observed.

Additionally, the analysis of genes differentially expressed across different conditions shows that each treatment causes, to a different extent, upregulation of erythroid genes in the HPC cluster. This suggests that perturbations of the cell cycle progression might interfere with the maintenance of an undifferentiated state and that the erythroid fate constitutes a default differentiation programme in this culture conditions. In this context, despite the transcriptional upregulation, CDK4/6 activation and G1/S transition appear to be necessary for the fulfilment of this programme, while CDK1 and G1 re-entry seems to be essential for the prevention of uncontrolled differentiation. An additional role for CDK1 in HPCs might also involve a crosstalk with the MAPK signalling and the cellular response to stress conditions, as shown by the upregulation of *MALAT1, NEAT1* and *TXNIP* (Li et al., 2009; Liu et al., 2014; Fuschi et al., 2017; Tian et al., 2017). This will need to be further explored, for example with the use of agonists and antagonists for the modulation of the MAPK pathway. Overall, a balanced control of the cell cycle progression appears to be fundamental for the HPC population.

Finally, the mesenchymal cluster does not appear to be affected by perturbations of the cell cycle machinery, despite changes in its cell cycle profile. This would suggest that the production of mesenchymal cells occurs without cell cycle progression, and therefore that EHT and EndoMT are controlled through very different mechanisms, with the cell cycle machinery

regulating the generation of haematopoietic, but not mesenchymal, cells. Alternatively, it is also possible that EndoMT takes place earlier than EHT and is mostly complete at the time examined for this study. Indeed, my results reveal that the mesenchymal cluster remains constant during the interval analysed, with only a slight increase from EHT D3 to EHT D5 (Figure 3.8). In addition, a deeper analysis of the endothelial population showed that cells engaged in EndoMT have already started upregulating mesenchymal genes, compared to the fewer cells engaged in EHT in which haematopoietic genes have not yet been induced, despite downregulation of endothelial markers (Figure 4.6) This might indicate that EndoMT takes place earlier than EHT and is already in an advanced state at the time considered for my experiments. Therefore, my results would only show that the cell cycle does not play a role in mesenchymal cell homeostasis. If this is the case, further experiments focusing on earlier time points will be necessary to study the role of cell cycle regulation in EndoMT leading to mesenchymal cell generation. Of note, EndoMT has been previously shown to be induced by TGFβ signalling, leading to tissue fibrosis (Wermuth et al., 2016). TGFβ is also known to cause cell cycle arrest by inducing the expression of regulators such as p21 and the inactivation of G1/S cyclin-CDK complexes (Datto et al., 1995; Hocevar and Howe, 1998; Yoo et al., 1999). Therefore, it is possible that EndoMT does not require cell cycle progression and that cell cycle arrest promotes the generation of mesenchymal cells. Thus, the *in vitro* model used for this dissertation could allow studying the production of these cells and potentially reveal novel mechanisms controlling EndoMT, relevant for development and disease. My results suggest that EndoMT and EHT might represent two alternative cell fate choices undertaken by endothelial cells during development. Intriguingly, an early cell cycle state modulation, possibly regulated by TGFβ, might control this choice, with cell cycle arrest promoting EndoMT and cell cycle progression essential for EHT. Further experiments focusing on earlier endothelial cells will be necessary to test this hypothesis. Importantly, the gating strategy developed in this study will allow to quickly monitor any effect on the balance between EHT and EndoMT.

In conclusion, cell cycle entry and progression, associated with the timely activation of CDK4/6 during the G1 phase and CDK1 in G2/M, appear to be essential for endothelial cells to undergo EHT and generate HPCs. Once generated, these progenitors engage cell cycle more

actively, and quickly differentiate, with the activity of CDK4/6 necessary for cell cycle progression and differentiation, and the activity of CDK1 important for G1 re-entry to maintain the HPC state and avoid uncontrolled commitment. Overall, a complex interplay between the molecular machineries controlling cell cycle progression and cell fate decision appears to be fundamental during EHT, and potentially plays additional roles in processes such as EndoMT in determining endothelial cell plasticity and differentiation dynamics.

## 5.6 Annexe: Gene Ontology analyses for differentially expressed genes

**Table 5.1 GO terms for CDK4/6i downregulated genes in the endothelial cluster.**

| Endothelial cluster \| CDK4/6i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| ribosomal small subunit export from nucleus (GO:0000056) | 0.00698 | 0.05914 | -4.17 | 20.7 |
| mitochondrial ATP synthesis coupled proton transport (GO:0042776) | 0.02178 | 0.07542 | -4.96 | 18.99 |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 0.001 | 0.04224 | -2.73 | 18.88 |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 8.3E-05 | 0.006288 | -1.9 | 17.87 |
| viral process (GO:0016032) | 6.1E-05 | 0.006288 | -1.7 | 16.45 |
| response to type I interferon (GO:0034340) | 0.00797 | 0.05914 | -2.97 | 14.36 |
| positive regulation of apoptotic cell clearance (GO:2000427) | 0.00797 | 0.05914 | -2.78 | 13.42 |
| ribosomal subunit export from nucleus (GO:0000054) | 0.01293 | 0.06483 | -3.08 | 13.38 |
| regulation of glial cell apoptotic process (GO:0034350) | 0.00698 | 0.05914 | -2.65 | 13.17 |
| positive regulation of nitric-oxide synthase biosynthetic process (GO:0051770) | 0.00996 | 0.06483 | -2.79 | 12.84 |

**Table 5.2 GO terms for CDK4/6i downregulated genes in the mesenchymal cluster.**

| Mesenchymal cluster \| CDK4/6i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 0.000009682 | 0.000568 | -2.74 | 31.58 |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 9.31E-07 | 0.0001222 | -1.9 | 26.42 |
| ribosomal small subunit export from nucleus (GO:0000056) | 0.006284 | 0.04096 | -4.16 | 21.09 |
| mitochondrial ATP synthesis coupled proton transport (GO:0042776) | 0.01962 | 0.05571 | -4.94 | 19.43 |
| respiratory electron transport chain (GO:0022904) | 0.000001389 | 0.0001222 | -1.36 | 18.33 |
| mitochondrial respiratory chain complex I biogenesis (GO:0097031) | 0.00002582 | 0.0007574 | -1.58 | 16.68 |
| NADH dehydrogenase complex assembly (GO:0010257) | 0.00002582 | 0.0007574 | -1.57 | 16.64 |
| mitochondrial respiratory chain complex assembly (GO:0033108) | 0.00008825 | 0.002219 | -1.78 | 16.58 |
| V(D)J recombination (GO:0033151) | 0.007179 | 0.04211 | -3.26 | 16.12 |
| mitochondrial respiratory chain complex I assembly (GO:0032981) | 0.00002582 | 0.0007574 | -1.41 | 14.9 |

**Table 5.3 GO terms for CDK4/6i downregulated genes in the erythroid cluster.**

| Erythroid cluster \| CDK4/6i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 2.92E-10 | 1.25E-07 | -1.9 | 41.77 |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 0.000002073 | 0.0002228 | -2.73 | 35.77 |
| V(D)J recombination (GO:0033151) | 0.000103 | 0.00492 | -3.28 | 30.13 |
| respiratory electron transport chain (GO:0022904) | 5.92E-10 | 1.27E-07 | -1.36 | 28.88 |
| mitochondrial respiratory chain complex assembly (GO:0033108) | 0.000001286 | 0.0001843 | -1.78 | 24.16 |
| DNA topological change (GO:0006265) | 0.0001651 | 0.007099 | -2.47 | 21.54 |
| DNA geometric change (GO:0032392) | 0.00001531 | 0.0008227 | -1.87 | 20.72 |
| DNA ligation involved in DNA repair (GO:0051103) | 0.0002851 | 0.009431 | -2.5 | 20.44 |
| mitochondrial respiratory chain complex I biogenesis (GO:0097031) | 0.000007671 | 0.0004712 | -1.58 | 18.59 |
| NADH dehydrogenase complex assembly (GO:0010257) | 0.000007671 | 0.0004712 | -1.57 | 18.54 |

**Table 5.4 GO terms for CDK4/6i upregulated genes in the erythroid cluster.**

| Erythroid cluster \| CDK4/6i upregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| membrane raft assembly (GO:0001765) | 0.004193 | 0.05535 | -3.89 | 21.31 |
| regulation of viral entry into host cell (GO:0046596) | 0.0001149 | 0.02343 | -1.89 | 17.14 |
| regulation of muscle system process (GO:0090257) | 0.005985 | 0.05535 | -2.94 | 15.05 |
| membrane raft organization (GO:0031579) | 0.006582 | 0.05535 | -2.96 | 14.86 |
| genetic imprinting (GO:0071514) | 0.005388 | 0.05535 | -2.57 | 13.41 |
| response to muscle stretch (GO:0035994) | 0.005985 | 0.05535 | -2.57 | 13.14 |
| regulation of glycogen (starch) synthase activity (GO:2000465) | 0.004791 | 0.05535 | -2.44 | 13.05 |
| negative regulation of DNA metabolic process (GO:0051053) | 0.0003079 | 0.0314 | -1.52 | 12.28 |
| regulation of apoptotic process (GO:0042981) | 0.001047 | 0.05535 | -1.79 | 12.27 |
| positive regulation of G1/S transition of mitotic cell cycle (GO:1900087) | 0.01016 | 0.05535 | -2.66 | 12.22 |

**Table 5.5 GO terms for CDK4/6i downregulated genes in the HPC cluster.**

| HPC cluster \| CDK4/6i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 2.75E-07 | 0.00001807 | -2.73 | 41.3 |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 6.70E-10 | 1.62E-07 | -1.9 | 40.19 |
| DNA topological change (GO:0006265) | 1.81E-07 | 0.00001588 | -2.48 | 38.54 |
| V(D)J recombination (GO:0033151) | 0.00003847 | 0.0008432 | -3.28 | 33.34 |
| respiratory electron transport chain (GO:0022904) | 1.23E-09 | 1.62E-07 | -1.36 | 27.89 |
| DNA ligation (GO:0006266) | 0.000001453 | 0.00004245 | -2.07 | 27.79 |
| apoptotic nuclear changes (GO:0030262) | 0.000003022 | 0.00007949 | -1.91 | 24.29 |
| DNA ligation involved in DNA repair (GO:0051103) | 0.0001068 | 0.00216 | -2.5 | 22.89 |
| mitochondrial respiratory chain complex I biogenesis (GO:0097031) | 0.000001028 | 0.00003862 | -1.58 | 21.76 |
| NADH dehydrogenase complex assembly (GO:0010257) | 0.000001028 | 0.00003862 | -1.57 | 21.7 |

**Table 5.6 GO terms for CDK4/6i upregulated genes in the HPC cluster.**

| HPC cluster \| CDK4/6i upregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| regulated exocytosis (GO:0045055) | 0.00002208 | 0.001921 | -2.3 | 24.7 |
| membrane raft assembly (GO:0001765) | 0.002797 | 0.03577 | -3.89 | 22.88 |
| defense response to protozoan (GO:0042832) | 0.003595 | 0.03577 | -2.98 | 16.76 |
| platelet degranulation (GO:0002576) | 0.00001305 | 0.001921 | -1.44 | 16.15 |
| membrane raft organization (GO:0031579) | 0.004392 | 0.03577 | -2.96 | 16.05 |
| regulation of MHC class II biosynthetic process (GO:0045346) | 0.003595 | 0.03577 | -2.73 | 15.38 |
| positive regulation of cAMP-mediated signaling (GO:0043950) | 0.003595 | 0.03577 | -2.7 | 15.2 |
| positive regulation of macrophage differentiation (GO:0045651) | 0.003595 | 0.03577 | -2.59 | 14.56 |
| genetic imprinting (GO:0071514) | 0.003595 | 0.03577 | -2.57 | 14.45 |
| regulation of glycogen (starch) synthase activity (GO:2000465) | 0.003196 | 0.03577 | -2.44 | 14.04 |

**Table 5.7 GO terms for CDK2/1i upregulated genes in the HPC cluster.**

| HPC cluster \| CDK2/1i upregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| oxygen homeostasis (GO:0032364) | 0.002448 | 0.01284 | -4.71 | 28.29 |
| gas homeostasis (GO:0033483) | 0.002797 | 0.01284 | -4.18 | 24.56 |
| positive regulation of neutrophil chemotaxis (GO:0090023) | 0.00003138 | 0.0008552 | -2.2 | 22.84 |
| positive regulation of granulocyte chemotaxis (GO:0071624) | 0.0000367 | 0.0008552 | -1.97 | 20.14 |
| positive regulation of neutrophil migration (GO:1902624) | 0.00003399 | 0.0008552 | -1.95 | 20.01 |
| regulation of leukocyte chemotaxis (GO:0002688) | 0.00004859 | 0.0008552 | -2 | 19.83 |
| oxygen transport (GO:0015671) | 0.002797 | 0.01284 | -3.32 | 19.53 |
| regulation of neutrophil chemotaxis (GO:0090022) | 0.00004546 | 0.0008552 | -1.77 | 17.74 |
| defense response to protozoan (GO:0042832) | 0.003146 | 0.01284 | -2.97 | 17.11 |
| positive regulation of leukocyte migration (GO:0002687) | 0.00006214 | 0.0009113 | -1.65 | 16.01 |

**Table 5.8 GO terms for CDK1i downregulated genes in the endothelial cluster.**

| Endothelial cluster \| CDK1i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| positive regulation of nitric-oxide synthase biosynthetic process (GO:0051770) | 0.0000383 | 0.01255 | -2.82 | 28.65 |
| regulation of nitric-oxide synthase biosynthetic process (GO:0051769) | 0.00006628 | 0.01255 | -2.58 | 24.83 |
| ribosomal small subunit export from nucleus (GO:0000056) | 0.006632 | 0.05774 | -4.15 | 20.81 |
| mitochondrial ATP synthesis coupled proton transport (GO:0042776) | 0.0207 | 0.05998 | -4.86 | 18.84 |
| positive regulation of activated T cell proliferation (GO:0042104) | 0.0001608 | 0.01267 | -2.15 | 18.76 |
| positive regulation of cell migration by vascular endothelial growth factor signaling pathway (GO:0038089) | 0.007576 | 0.05774 | -3.47 | 16.95 |
| regulation of positive chemotaxis (GO:0050926) | 0.008519 | 0.05774 | -3.55 | 16.93 |
| protein kinase B signaling (GO:0043491) | 0.0005284 | 0.02313 | -1.97 | 14.87 |
| response to type I interferon (GO:0034340) | 0.007576 | 0.05774 | -2.95 | 14.4 |
| positive regulation of apoptotic cell clearance (GO:2000427) | 0.007576 | 0.05774 | -2.76 | 13.47 |

**Table 5.9 GO terms for CDK1i upregulated genes in the endothelial cluster.**

| Endothelial cluster \| CDK1i upregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| cytoskeleton-dependent intracellular transport (GO:0030705) | 0.00007845 | 0.006058 | -2.09 | 19.73 |
| regulation of transcription from RNA polymerase II promoter in response to oxidative stress (GO:0043619) | 0.004492 | 0.04395 | -3.6 | 19.45 |
| cell cycle G2/M phase transition (GO:0044839) | 0.00002973 | 0.005055 | -1.82 | 18.98 |
| V(D)J recombination (GO:0033151) | 0.003994 | 0.04395 | -3.27 | 18.05 |
| regulation of mitotic cell cycle phase transition (GO:1901990) | 0.00008908 | 0.006058 | -1.91 | 17.85 |
| outer mitochondrial membrane organization (GO:0007008) | 0.003994 | 0.04395 | -3.23 | 17.84 |
| regulation of G2/M transition of mitotic cell cycle (GO:0010389) | 0.00004774 | 0.005411 | -1.77 | 17.61 |
| negative regulation of translation in response to stress (GO:0032055) | 0.003495 | 0.04395 | -3.09 | 17.48 |
| regulation of translational initiation in response to stress (GO:0043558) | 0.003495 | 0.04395 | -2.84 | 16.07 |
| regulation of nuclease activity (GO:0032069) | 0.003495 | 0.04395 | -2.83 | 15.99 |

**Table 5.10 GO terms for CDK1i downregulated genes in the mesenchymal cluster.**

| Mesenchymal cluster \| CDK1i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| ribosomal small subunit export from nucleus (GO:0000056) | 0.005587 | 0.0459 | -4.17 | 21.63 |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 0.0006352 | 0.0305 | -2.74 | 20.14 |
| mitochondrial ATP synthesis coupled proton transport (GO:0042776) | 0.01746 | 0.05732 | -4.96 | 20.08 |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 0.00004128 | 0.004197 | -1.9 | 19.21 |
| bleb assembly (GO:0032060) | 0.006383 | 0.0459 | -3 | 15.15 |
| ribosomal subunit export from nucleus (GO:0000054) | 0.01035 | 0.05579 | -3.08 | 14.08 |
| respiratory electron transport chain (GO:0022904) | 0.00005559 | 0.004197 | -1.36 | 13.32 |
| genetic imprinting (GO:0071514) | 0.007178 | 0.04927 | -2.56 | 12.64 |
| regulation of glycogen (starch) synthase activity (GO:2000465) | 0.006383 | 0.0459 | -2.43 | 12.27 |
| protein targeting to ER (GO:0045047) | 0.002727 | 0.03956 | -2.08 | 12.26 |

**Table 5.11 GO terms for CDK1i downregulated genes in the erythroid cluster.**

| Erythroid cluster \| CDK1i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| regulation of protein targeting (GO:1903533) | 0.005936 | 0.03332 | -4.26 | 21.86 |
| ribosomal small subunit export from nucleus (GO:0000056) | 0.005936 | 0.03332 | -4.17 | 21.37 |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 0.0007188 | 0.02581 | -2.74 | 19.8 |
| mitochondrial ATP synthesis coupled proton transport (GO:0042776) | 0.01854 | 0.04191 | -4.96 | 19.77 |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 0.00004997 | 0.003801 | -1.9 | 18.84 |
| folic acid-containing compound biosynthetic process (GO:0009396) | 0.005936 | 0.03332 | -3.45 | 17.69 |
| oxygen transport (GO:0015671) | 0.006781 | 0.03332 | -3.32 | 16.57 |
| positive regulation of autophagy of mitochondrion (GO:1903599) | 0.006781 | 0.03332 | -3 | 15 |
| ribosomal subunit export from nucleus (GO:0000054) | 0.011 | 0.04056 | -3.08 | 13.89 |
| pteridine-containing compound biosynthetic process (GO:0042559) | 0.006781 | 0.03332 | -2.68 | 13.41 |

**Table 5.12 GO terms for CDK1i upregulated genes in the erythroid cluster.**

| Erythroid cluster \| CDK1i upregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| negative regulation of membrane protein ectodomain proteolysis (GO:0051045) | 0.003196 | 0.04232 | -2.66 | 15.26 |
| negative regulation of reproductive process (GO:2000242) | 0.004392 | 0.04232 | -2.54 | 13.81 |
| sequestering of actin monomers (GO:0042989) | 0.004392 | 0.04232 | -2.43 | 13.17 |
| regulation of integrin-mediated signaling pathway (GO:2001044) | 0.004791 | 0.04232 | -2.3 | 12.27 |
| regulation of trophoblast cell migration (GO:1901163) | 0.004791 | 0.04232 | -2.24 | 11.96 |
| negative regulation of DNA replication (GO:0008156) | 0.006781 | 0.04303 | -2.2 | 10.96 |
| positive regulation of viral entry into host cell (GO:0046598) | 0.003595 | 0.04232 | -1.85 | 10.44 |
| transition metal ion homeostasis (GO:0055076) | 0.009958 | 0.04303 | -2.23 | 10.26 |
| cellular iron ion homeostasis (GO:0006879) | 0.02258 | 0.0497 | -2.61 | 9.91 |
| negative regulation of actin filament polymerization (GO:0030837) | 0.009165 | 0.04303 | -1.83 | 8.57 |

**Table 5.13 GO terms for CDK1i downregulated genes in the HPC cluster.**

| HPC cluster \| CDK1i downregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| mitochondrial ATP synthesis coupled electron transport (GO:0042775) | 8.30E-11 | 4.67E-08 | -1.9 | 44.16 |
| mitochondrial electron transport, NADH to ubiquinone (GO:0006120) | 0.000001458 | 0.0001368 | -2.73 | 36.68 |
| respiratory electron transport chain (GO:0022904) | 2.06E-10 | 5.79E-08 | -1.36 | 30.32 |
| mitochondrial ATP synthesis coupled proton transport (GO:0042776) | 0.003724 | 0.03913 | -4.95 | 27.71 |
| mitochondrial electron transport, cytochrome c to oxygen (GO:0006123) | 0.0001003 | 0.002822 | -2.57 | 23.62 |
| mRNA splicing, via spliceosome (GO:0000398) | 0.000001398 | 0.0001368 | -1.75 | 23.54 |
| histone mRNA metabolic process (GO:0008334) | 0.0001484 | 0.003632 | -2.29 | 20.16 |
| gene expression (GO:0010467) | 1.29E-07 | 0.00002419 | -1.26 | 19.94 |
| establishment of protein localization to mitochondrion (GO:0072655) | 0.00005491 | 0.001932 | -2.02 | 19.86 |
| nucleocytoplasmic transport (GO:0006913) | 0.00004293 | 0.001859 | -1.95 | 19.59 |

**Table 5.14 GO terms for CDK1i upregulated genes in the HPC cluster.**

| HPC cluster \| CDK1i upregulated genes \| GO Biological Process | | | | |
|---|---|---|---|---|
| Name | p-value | adjusted p-value | z score | combined score |
| cellular iron ion homeostasis (GO:0006879) | 0.000003262 | 0.0008498 | -2.65 | 33.52 |
| iron ion homeostasis (GO:0055072) | 9.99E-08 | 0.00005206 | -1.57 | 25.36 |
| oxygen homeostasis (GO:0032364) | 0.01253 | 0.09524 | -4.65 | 20.36 |
| positive regulation of cell cycle arrest (GO:0071158) | 0.00001468 | 0.001912 | -1.71 | 19.01 |
| ribosomal small subunit export from nucleus (GO:0000056) | 0.01253 | 0.09524 | -4.12 | 18.06 |
| gas homeostasis (GO:0033483) | 0.01431 | 0.09524 | -4.11 | 17.47 |
| neutrophil degranulation (GO:0043312) | 0.0001957 | 0.01014 | -2.01 | 17.14 |
| cellular transition metal ion homeostasis (GO:0046916) | 0.00002402 | 0.002503 | -1.59 | 16.88 |
| neutrophil mediated immunity (GO:0002446) | 0.000214 | 0.01014 | -1.93 | 16.31 |
| DNA damage response, signal transduction by p53 class mediator (GO:0030330) | 0.00001468 | 0.001912 | -1.42 | 15.84 |

# PART III

# CONCLUSIONS

# 6 FUTURE PERSPECTIVES AND CONCLUSIONS

## 6.1 Future directions

The results discussed in this dissertation contributed to expand our knowledge about the specification of haematopoietic progenitors taking place during development. At the same time, they also paved the way for future experiments for the further exploration of these mechanisms and to new lines of investigation extending beyond haematopoietic development.

## 6.2 Molecular networks linking cell cycle and cell fate decision

My results suggest that the cell cycle progression could be a requirement for EHT, and I have proposed distinct roles for CDK4/6 and CDK1 during haematopoietic specification and differentiation. However, the molecular network linking cell fate decision and cell cycle regulation still remains to be elucidated.

As previously discussed (Chapter 5.5), CDK4/6 might bind to RUNX1 and participate in the regulation of its transcriptional activity. Of note, cyclin D3 was also suggested to have a similar role (Peterson et al., 2005). This would be in agreement with my results showing a cell type-specific expression for distinct cyclin D isoforms, with cyclin D1/D2 expressed in endothelial and mesenchymal cells and cyclin D3 preferentially expressed in the haematopoietic compartment (Figure 4.3). This potentially suggests that RUNX1 activity might be cell cycle regulated by its interaction in the G1 phase with cyclin D3 and CDK6. However, RUNX1 is also part of a complex transcriptional network involving other key players such as GATA2 and TAL1 (Chapter 1.7.5). Therefore, if the cell cycle dependent regulation of RUNX1 was confirmed to take place during EHT, this could possibly extend to other factors controlling the haematopoietic transcriptional programme. In this regards, the *in vitro* model presented in this

dissertation offers the unprecedented opportunity to further explore these hypotheses and potentially identify effector proteins interacting with key cell cycle regulators. For that, it will be interesting to test the interaction of these and other cell cycle regulators with key haematopoietic transcription factors controlling EHT, for example with the use of techniques such as the proximity ligation assay. Furthermore, the combination of these experiments with the FUCCI reporter system might potentially confirm if this interaction takes place during specific cell cycle phases.

On the other hand, the analysis of genes differentially expressed upon inhibition of CDK1 revealed a potential role for this factor in the regulation of the cellular response to oxidative stress, suggested to be regulated by MALAT1, NEAT1 and TXNIP and mediated by the MAPK signalling pathway (Li et al., 2009; Liu et al., 2014; Fuschi et al., 2017; Tian et al., 2017). Of note, cyclin B1-CDK1 was previously shown to localise to the mitochondrial matrix and to phosphorylate proteins including components of the respiratory chain, thus increasing mitochondrial respiration (Wang et al., 2014b). These findings suggest that CDK1 could potentially modulate the oxidative metabolism and at the same time a MAPK-mediated response to oxidative stress. Further experiments will allow exploring the role of this mechanism during haematopoietic specification and differentiation, for example performing a metabolomics analysis upon CDK1 inhibition or using agonists and antagonist of the MAPK signalling to study the role of this pathway during distinct stages of EHT.

Overall, these experiments will allow to further elucidate the molecular network linking cell cycle state and haematopoietic specification.

## 6.3 Manipulation of the cell cycle machinery for the improved production of haematopoietic cells

The work presented in this dissertation shows that the haemogenic endothelial cells seem to be in a quiescent state, characterised by the expression of cell cycle inhibitors of the Rb and CIP/KIP families, and as previously hypothesised (Chapter 4.5), the removal of these molecular

breaks might be the first necessary event for the onset of EHT. Thus, inhibiting the activity of these cell cycle inhibitors might potentially induce a larger fraction of endothelial cells to transition towards the haematopoietic fate. This could be done with the use of small molecule inhibitors or knockdown mediated by short hairpin RNAs. Based on my results, I anticipate that the inhibition of these cell cycle inhibitors might increase the number of endothelial cells entering and progressing through the cell cycle, and consequently the number of cells undergoing EHT. This approach might be used to screen multiple regulators and identify those with the higher impact on differentiation. Furthermore, multiple signalling pathways might be tested for their capacity to control these cell cycle regulators and promote cell cycle entry and EHT. Overall this approach might allow the identification of signalling pathways controlling cell cycle and able to improve the efficiency of differentiation.

On the other hand, my results suggest that uncontrolled cell cycle progression would lead to exhaustion of the HPC population. Thus, an accurate control over these mechanisms would be likely required, possibly involving a sequential activation and inhibition of these signalling pathways to achieve an initial cell cycle activation to promote EHT and a subsequent cell cycle exit to preserve the HPC population. Indeed, a similar regulation appears to take place *in vivo*, when early HSPCs are generated in the AGM and migrate to the foetal liver where they are gradually induced to exit the cell cycle (Hur et al., 2016).

Of note, an important candidate signal would be TGFβ, known to induce the expression of cell cycle inhibitors such as p21 and promote cell cycle arrest (Datto et al., 1995; Hocevar and Howe, 1998; Yoo et al., 1999). In agreement with my hypothesis, TGFβ signalling was reported to inhibit EHT (Vargel et al., 2016). Indeed, based on my results I speculate that an initial inhibition of the pathway might be necessary for cell cycle activation and induction of EHT, followed by its activation to promote cell cycle exit, preserve the HPC state and possibly induce maturation.

## 6.4 Further characterisation and maturation of multipotent HPCs

The simple but effective sorting strategy developed in this study allows the isolation of the first population of haematopoietic progenitors generated *in vitro*. Further characterization of this population will be necessary to determine if these cells are indeed multipotent or if they represent instead heterogeneous progenitors capable of multilineage differentiation only at the population level. For that, single cell differentiation experiments will be essential. The isolation of this population will also be necessary to confirm the hypothesised roles of CDK4/6 and CDK1 in differentiation. Furthermore, this will allow to further expand their characterisation and to potentially identify signalling pathways that control their commitment or self-renewal.

Of note, our group has developed protocols for the differentiation of hPSCs towards liver cell types such as hepatocytes or cholangiocytes, and systems for their co-culture in 3D environments based on Matrigel or collagen. Importantly, these cells are characterised by a foetal phenotype, and appear to increase their maturation upon 3D co-culture (data not published). Therefore, a novel line of investigation would be represented by the isolation of HPCs generated upon EHT for their 3D co-culture with *in vitro* differentiated hepatic cell types, in an attempt to model HSPC migration to the foetal liver taking place during development upon their generation in the AGM, and therefore recapitulate their developmental niche and possibly induce maturation.

## 6.5 Regulation of endothelial-to-mesenchymal transition

Finally, my results revealed an additional population generated in this *in vitro* system, constituted by mesenchymal cells possibly generated from the endothelium through EndoMT. This population does not appear to be affected by perturbations of the cell cycle, although it is possible that EndoMT is a process temporally preceding EHT and that the treatments used in

this study need to be anticipated to produce an effect. Indeed, early cell cycle arrest might promote mesenchymal differentiation, which would be consistent with the reported role of TGFβ in both promoting EndoMT and inducing cell cycle arrest, as previously discussed (Chapter 5.5), and in agreement with its inhibitory effect on EHT (Vargel et al., 2016). Thus, the *in vitro* differentiation here characterised could possibly be adapted to model this additional process, relevant for both development and disease, and it appears that EndoMT and EHT might represent two alternative cell fate choices. Importantly, using the same gating strategy developed in this work for the isolation of HPCs, it could be possible to isolate the mesenchymal population for further characterisation. This would allow to functionally test these cells for their ability to differentiate, for example towards the generation of cell types like osteoblast cells. Ultimately, this would open multiple lines of investigation, involving the characterisation of the mesenchymal population and the identification of novel molecular mechanisms important for diseases such as fibrosis.

## 6.6 Conclusions

This dissertation contributes to elucidate molecular mechanisms driving a complex developmental event such as the endothelial-to-haematopoietic transition, and to the understanding of dynamics and mechanisms driving early haematopoietic specification and commitment. The use of human pluripotent stem cells and the optimisation of an *in vitro* differentiation system to model human haematopoietic ontogeny, in combination with high-throughput single cell RNA sequencing, allowed me to draw an accurate picture of a developmental stage which is otherwise difficult to access in human *in vivo*. This approach allowed me to explore the complexity of this heterogeneous system and to identify and characterise with unprecedented detail the populations involved. I was able to determine that such populations are characterised by distinct and specific cell cycle states, which I suggested to be determinant in the capability of the endothelial cells to undergo the transition towards the haematopoietic fate. Indeed, cell cycle progression appears to have an important role in this process, and I hypothesise that this might be due to the activation of transcriptional

programmes which are cell cycle regulated, possibly controlled by the timely activation of CDK regulators at specific cell cycle phases. I have indeed proposed a model for the role of CDK4/6 and CDK1 in the specification of haematopoietic progenitors and their further differentiation, although the specific molecular network modulated by these regulators still needs to be explored. For that, the isolation of the first population capable of multilineage differentiation will substantially increase our understanding of these mechanisms.

Collectively, the results presented in this dissertation contribute to the knowledge currently available in the field, and pave the way to future studies and exciting developments in regenerative medicine, such as the *in vitro* generation of cells capable of self-renewal and multilineage engraftment or the design of more effective culture conditions for the *ex vivo* expansion of primary haematopoietic stem cells.

# PART IV

# MATERIAL AND METHODS

# 7 MATERIAL AND METHODS

## 7.1 Culture of hPSCs

The hIPSC lines A1AT-RR (Yusa et al., 2011) and FSPS13B were maintained as previously described (Chen et al., 2011b) on plates coated with 10 μg/ml Vitronectin (Stem Cell Technologies) and cultured in E6 media supplemented with 2 ng/ml TGFβ (R&D) and 25 ng/ml FGF2 (from Dr. Marko Hynoven, Department of Biochemistry, University of Cambridge). Cells were maintained at 37 °C and 5% CO2, and passaged every 5-6 days by dissociation with 0.5 mM EDTA (ThermoFisher Scientific). For coating plates, 10 μg/ml Vitronectin in PBS (ThermoFisher Scientific) was applied for at least 1 hour at room temperature.

## 7.2 Haematopoietic differentiation

Differentiation was performed at 5% $O_2$ by adapting a protocol previously described (Ditadi et al., 2015). Briefly, a serum-free differentiation (SFD) medium was used, consisting of 75% Iscove's modified Dulbecco's medium (ThermoFisher Scientific) and 25% Ham's F12 medium (ThermoFisher Scientific) supplemented with 1% N2 (ThermoFisher Scientific), 0.5% B27 (ThermoFisher Scientific), 0.05% BSA (Sigma-Aldrich), 1 mM ascorbic acid (2-phospho-L-ascorbic acid trisodium salt, Sigma-Aldrich), $4.5x10^{-4}$ monothioglycerol (Sigma-Aldrich), 2 mM L-glutamine (ThermoFisher Scientific), 150 μg/ml transferrin (Sigma-Aldrich) and 10 ng/ml penicillin/streptomycin (ThermoFisher Scientific). Undifferentiated cells were dissociated using 0.5 mM EDTA and small aggregates were resuspended in SFD medium supplemented with 10 ng/ml BMP4 (R&D) and cultured as EBs on non-treated plates (Starlab). After 24 hours, an equivalent volume of SFD was added directly on top to not perturb the small EBs, supplemented with final concentrations of 10 ng/ml BMP4 and 5 ng/ml FGF2. At day 2, developing EBs were collected and washed. For this, they were centrifuged 3 minutes at 100 g,

gently resuspended in SFD supplemented with 10 ng/ml BMP4, 5 ng/ml FGF2 and 3 μM CHIR99021 (Tocris), and seeded back on the plate. After 24 hours, EBs were again collected and washed, this time by centrifuging at 300 g for 3 minutes, and resuspended in StemPro-34 SFM (ThermoFisher Scientific) supplemented with 1 mM ascorbic acid, $4.5x10^{-4}$ monothioglycerol, 2 mM L-glutamine, 150 μg/ml transferrin, 10 ng/ml penicillin/streptomycin (from now on referred to as complete SP34) containing 5 ng/ml FGF2, 15 ng/ml VEGF (Peprotech), 10 ng/ml IL-6 (R&D) and 5 ng/ml IL-11 (R&D), and cultured for 48 hours. At day 5, EBs were again collected, washed and resuspended in complete SP34 with 5 ng/ml FGF2, 15 ng/ml VEGF, 10 ng/ml IL-6, 5 ng/ml IL-11, 50 ng/ml SCF (R&D), 5 ng/ml IGF-1 (R&D), 2 U/ml EPO (R&D) and cultured until day 8. At this stage, EBs were collected and prepared for sorting. For this, they were incubated for 20 minutes at 37 °C with collagenase solution, composed of Advanced DMEM/F12 (ThermoFisher Scientific), 20% KSR (ThermoFisher Scientific), 1% L-glutamine and 1 mg/ml collagenase IV (ThermoFisher Scientific), followed by 3 minutes incubation with TrypLE (ThermoFisher Scientific). Single cells were then stained for flow cytometry and the CD34+/CD43- fraction was sorted and used for the second stage of differentiation. For this, the population was resuspended at a density of $10^6$ cells/ml in complete SP34 with 10 ng/ml BMP4, 5 ng/ml FGF2, 5 ng/ml VEGF, 10 ng/ml IL-6, 5 ng/ml IL-11, 100 ng/ml SCF, 25 ng/ml IGF-1, 30 ng/ml TPO (Peprotech), 30 ng/ml IL-3 (Peprotech), 10 ng/ml Flt3-L (R&D), 20 ng/ml SHH (R&D), 10 ng/ml angiotensin II (Sigma-Aldritch) and 100 μM losartan potassium (R&D). Cells were transferred to a non-treated round-bottom 96 well plate (Corning), 200 μl/well (corresponding to $2x10^5$ cells/well), centrifuged 3 minutes at 300 g, and incubated overnight to allow the cells to re-aggregate. On the following day, marking EHT D1, the small aggregates were plated on Matrigel (Corning). For this, they were gently transferred to thin-layer Matrigel-coated wells, with a density of $2x10^5$ cells/well in a 24 well plate, and cultured for additional 2-4 days using the same media, replaced every 2 days. For coating plates, Matrigel was diluted into cold medium with a concentration of 35 μg/cm$^2$ of surface to be coated, and applied overnight at 37 °C. For treatments with cell cycle inhibitors, 0.1 μg/mL Nocodazole (Sigma-Aldrich), 1 μM PD0332991 (Tocris), 4 μM Roscovitine (Sigma-Aldrich), 10 μM RO3306 (Sigma-Aldrich) or 0.1% DMSO were added at EHT D3 for 48 hours.

## 7.3 Isolation of CD34+ peripheral blood mononuclear cells

Under sterile conditions, peripheral blood was diluted with room temperature PBS supplemented with 1 M trisodium citrate and 20% human serum albumin. 2 volumes of blood dilution were transferred in 50 ml tubes on a layer of 1 volume of Ficoll-Paque (Sigma-Aldrich). After centrifugation for 15 minutes at 800 g, the resulting layer of mononuclear cells was carefully removed, transferred to a new tube, further diluted using the same buffer and centrifuged for 6 minutes at 600 g. The resulting supernatant was removed, the pellet was resuspended in cold PBS supplemented with 0.5 M EDTA and 20% human serum albumin, and centrifuged again for 6 minutes at 600 g. Cells were then processed for CD34+ enrichment using labelling with magnetic beads. Briefly, cells were resuspended in the same PBS/EDTA/human serum albumin buffer, added with 50 µl/$10^8$ cells of FcR blocking reagent (Miltenyi Biotec) and 50 µl/$10^8$ cells of CD34 magnetic beads (Miltenyi Biotec) and incubated for 30 minutes at 4 °C. After incubation, cells were washed using the same buffer and processed with AutoMACS Pro Separator to enrich for CD34+ peripheral blood mononuclear cells.

## 7.4 May–Grünwald-Giemsa staining

Cells were resuspended in culture media and concentrated by cytospin centrifugation at 700 g for 5 minutes onto SuperFrostPlus slides (ThermoFisher Scientific) using a Shandon Cytospin 3 cytocentrifuge. Slides were fixed for 3 minutes in cold methanol and stained with May–Grünwald Giemsa (Sigma). Images were captured using a Leica DM5000b microscope in conjunction with a ×63 oil-immersion lens and an Olympus DP72 camera

## 7.5 Flow cytometry

Cells were dissociated into single cells using TrypLE for 3 minutes at 37 °C, washed with 0.1% BSA-PBS and either fixed with 1% paraformaldehyde and kept at 4 °C for maximum 1 week or immediately stained for flow cytometry. For the staining, after a wash with PBS, cells were blocked with 10% donkey serum (Bio-Rad) for 30 minutes at room temperature. Cells were then stained with the relevant conjugated antibodies (Table 7.1) diluted in PBS for 1 hour at room temperature, protected from light. Following two washes with PBS, cells were analysed using the Cyan ADP flow cytometer, or sorted on the BD Influx cell sorter. Data was analysed using the FlowJo VX software.

**Table 7.1 Antibodies used for flow cytometry**

| Target | Dilution | Supplier | Clone |
|---|---|---|---|
| CD34 | 1:100 | Biolegend | 581 |
| CD43 | 1:25 | BD Biosciences | 1G10 |
| CDH5 | 1:50 | Biolegend | BV9 |
| CD44 | 1:50 | Biolegend | BJ18 |

## 7.6 RNA extraction, cDNA synthesis and qPCR

Total RNA was extracted using the GenElute Mammalian Total RNA Miniprep Kit (Sigma-Aldrich) and the On-Column DNase I Digestion set (Sigma-Aldrich) according to the manufacturer's instructions. RNA was reverse-transcribed using 250 ng random primers (Promega), 0.5 mM dNTPs (Promega), 20 U RNAseOUT (Invitrogen), 0.01 M DTT (Invitrogen) and 25 U of SuperScript II (Invitrogen). For the qPCR reaction, the resulting cDNA was diluted 30-fold. Quantitative PCR mixtures were prepared using the KAPA SYBR FAST qPCR Master Mix Kit (Kapa Biosystems), 4.2 µl of cDNA and 200 nM of each of the forward and reverse primers (Table 7.2). Technical duplicates of the samples were run on 384 well plates using the QuantStudio 12K Flex Real-Time PCR System machine and results

analysed using the delta cycle threshold (ΔCt) method. Expression values were normalized to the housekeeping gene RPLP0.

**Table 7.2 Primers used for qPCR**

| Gene | Forward Primer (5'-3') | Reverse Primer (5'-3') |
|---|---|---|
| *CDH5* | TGGCCAGCTGGTCCTGCAGAT | TGCCCGTGCGACTTGGCATC |
| *CDK1* | GGAAGGGGTTCCTAGTACTGC | AAGCACATCCTGAAGACTGACT |
| *CDK2* | CTCCAGGGCCTAGCTTTCTG | CCGGCGAGTCACCTCATGG |
| *CDK4* | TGAGGGGGCCTCTCTAGCTT | CAAGGGAGACCCTCACGCC |
| *CDK6* | ACAGAGCACCCGAAGTCTTG | GGGAGTCCAATCACGTCCAA |
| *CDKN1A / p21* | GGCAGACCAGCATGACAGAT | GATGTAGAGCGGGCCTTTGA |
| *CDKN1B / p27* | TAATTGGGGCTCCGGCTAAC | GAAGAATCGTCGGTTGCAGGT |
| *CDKN1c / p57* | GCTGCGGTGAGCCAATTTAG | AACAAAACCGAACGCTGCTC |
| *CYCLIN B1* | CGCCTGAGCCTATTTTGGTTG | AGTGACTTCCCGACCCAGTA |
| *CYCLIN D1* | GCTGTGCATCTACACCGACA | AAATGAACTTCACATCTGTGGCA |
| *CYCLIN D2* | GCCACCGACTTTAAGTTTGC | CGGTACTGCTGCAGGCTATT |
| *CYCLIN D3* | TGTGCTACAGATTATACCTTTGCC | GCTTCGATCTGCTCCTGACA |
| *CYCLIN E1* | GACGGGGAGCTCAAAACTGA | TCGGGCTTTGTCCAGCAAAT |
| *GATA1* | CTACACCAGGTGAACCGGC | CTTTTCCAGATGCCTTGCGG |
| *GATA2* | CTGTTCAGAAGGCCGGGAG | AATTTGCACAACAGGTGCCG |
| *GFI1* | ATCCACACTGGTGAGAAGCC | GCTGCCCTCTGTAGTGTTGT |
| *LMO2* | CAGAACATTGGGGACCGCTA | GTCTTGCCCAAAAAGCCTGA |
| *KLF1* | CCGAGGAAGAGGAGGCTTGAG | GGAAGTCATCCTGTGTGTCCG |
| *MEIS1* | GCGCAAAGGTACGACGATCT | GGTACTGATGCGAGTGCAGA |
| *MYB* | GCTACTGCCTGGACGAACTG | GTTGTTAACAGTGGGCTGGC |
| *PECAM1* | CAGGCGCCGGGAGAAGTGAC | CGTCCAGTCCGGCAGGCTCT |
| *RB1* | CTGTGGATGGAGTATTGGGAGG | TCTCATCTAGGTCAACTGCTGC |
| *RBL1 / p107* | AGCAGAGGAGGATTCCTTGCAG | GGGCACATAATCGCATTGGC |
| *RBL2 / p130* | GCTACACGCTGGAGGGAAAT | TCCTTCCACTGTCCCTTTGC |
| *RPLP0* | GGCGTCCTCGTGGAAGTGAC | GCCTTGCGCATCATGGTGTT |
| *RUNX1* | CATCGCTTTCAAGGTGGTGG | CATGGCTGCGGTAGCATTTC |
| *SPI1* | CCCCACGACCGTCCAG | GTAATGGTCGCTATGGCTCTCC |
| *TAL1* | TACTGATGGTCCCCACACCA | CCAGGCGGAGGATCTCATTC |
| *VWF* | TTACGTGGGTGGGAACATGG | TCTGTGGTGACTGTGCCATC |

## 7.7 CFU assay

The assay was performed using the MethoCult H4435 Enriched medium (STEMCELL Technologies), following the manufacturer's instructions. Briefly, cells were added in a tube with the medium, mixed by vortex, plated on non-treated 35 mm culture dishes (Corning) and incubated for 14 days at 37 °C, 5% CO2 and 5% O2. Different types of colony were recognised based on morphology, counted and, when relevant, collected for RNA extraction.

## 7.8 Cell cycle profile analysis

Cell cycle profile analysis was performed using the Click-iT EdU Alexa Fluor 488 Flow Cytometry Assay Kit (ThermoFisher Scientific) according to the manufacturer's instructions. Briefly, cultured cells were incubated at 37 °C with 10 μM EdU for 1 hour and harvested after dissociation with TrypLE. After 3 washes with 0.1% BSA-PBS, cells were fixed with 1% paraformaldehyde for 15 minutes at room temperature and washed three more times with 0.1% BSA-PBS. Cells were then permeabilised for 15 minutes with saponin-based permeabilisation/wash buffer and incubated with the Click-iT reaction cocktail for 30 minutes protected from light. Cells were washed once with saponin-based permeabilisation/wash buffer, stained for DNA content using DAPI (ThermoFisher Scientific) and analysed on the Cyan ADP flow cytometer and FlowJo VX software.

## 7.9 Statistical analysis

Statistical analyses were performed using the GraphPad Prism 7 software. The type of statistical analysis performed and the number of replicates used in each experiment are described in the figure legends. For the comparison of two or multiple groups, unpaired t-test

or one-way ANOVA test was performed, respectively. Significance in each analysis is represented as *P<0.5, **P<0.01, ***P<0.001, ****P<0.0001.

## 7.10 Single cell RNA sequencing

The bioinformatic analyses of the transcriptional data reported in this dissertation were performed with the help of Dr Emmanouil I. Athanasiadis from Dr Ana Cvejic's group.

All the methods were adopted as previously described (Butler et al., 2018).

### 7.10.1 Single cell RNA processing

Following dissociation, cells were resuspended at a concentration of 1,500 cells/μl in ice-cold SP34 medium, complete with cytokines as for EHT culture. Libraries were constructed using Chromium Controller and Chromium Single Cell 3' Library & Gel Bead Kit (10x Genomics) according to the manufacturer's instructions for the recovery of 2,000 cells for each sample. Briefly, the cellular suspension was added to the master mix containing nuclease-free water, RT Reagent Mix, RT Primer, Additive A and RT Enzyme Mix. Master mix with cells was transferred to the wells in row 1 on the Chromium Single Cell A Chip (10x Genomics). Single Cell 3' Gel Beads were transferred in row 2 and Partitioning Oil was transferred into row 3. The chip was loaded on Chromium Controller to generate single cell GEMs. GEM-RT was performed in a C1000 Touch Thermal cycler (Bio-Rad) at the following conditions: 53 °C for 45 minutes, 85 °C for 5 minutes, held at 4 °C. Post GEM-RT clean-up was performed with DynaBeads MyOne Silane Beads (ThermoFisher Scientific). cDNA was amplified using C1000 Touch Thermal cycler at the following conditions: 98 °C for 3 minutes, 12 cycles of (90 °C for 15 seconds, 67 °C for 20 seconds and 72 °C for 1 minute), 72 °C for 1 minute, held at 4 °C. Amplified cDNA was cleaned with the SPRIselect Reagent Kit (Beckman Coulter) and quality

was assessed using 2100 Bioanalyser (Agilent). Libraries were constructed following the manufacturer's protocol and sequenced in pair-end mode on Hi-Seq4000 platform.

### 7.10.2 Alignment and quantification of sequencing data

Cell Ranger v2.10 was used in order to de-multiplex raw base call (BCL) files generated by Illumina sequencers into FASTQ files, perform the alignment, barcode counting and UMI counting. Ensembl BioMart version 91 was used to generate the reference genome.

### 7.10.3 Quality control of sequencing data

Data was filtered based on the Median Absolute Deviation (MAD) of the distribution of the number of detected genes. In addition, the percentage of mitochondrial content was set to less than 20%. Following quality control, 3,877 single cells from EHT D3, 2,152 from EHT D5, 1,742 from PD0332991 (CDK4/6i), 2,140 from Roscovitine (CDK2/1i) and 1,741 from RO3306 (CDK1i) were used in downstream analyses.

### 7.10.4 Seurat Alignment Strategy

In order to perform a direct comparison of clusters that belonged to the same cell type across different conditions, we adopted the Seurat Alignment workflow (Butler et al., 2018). We calculated Highly Variable Genes (HVGs) for each of the different conditions. HVGs were detected based on their average expression against their dispersion, by means of the "FindVariableGenes" Seurat command with the following parameters: mean.function equal to ExpMean, dispersion.function equal to LogVMR, x.low.cutoff equal to 0.0125, x.high.cutoff equal to 3, and y.cutoff equal to 0.5. For the analysis of EHT D3 and EHT D5 we selected 1,289 common HVGs that were expressed in both datasets. For the analysis of EHT D5, CDK4/6i,

CDK2/1i and CDK1i we selected 2,306 HVGs that were expressed in at least 2 out of 4 samples. Canonical Correlation Analysis (CCA) was then performed in order to identify shared correlation structures across the different conditions using the "RunMultiCCA" command. Twenty significant CCA components were selected by means of the shared correlation strength, using the "MetageneBicorPlot" command in both cases. Aligned CCA space was then generated with the "AlignSubspace" Seurat command. Significant CCA aligned components were then used to create the 3D tSNE space using the "RunTSNE" Seurat command.

### 7.10.5 Downstream analysis of sequencing data

For the clustering in the 3D tSNE space we used the "FindClusters" command in Seurat that performs the Shared Nearest Neighbor (SNN) modularity optimization based clustering algorithm in Correlation Component Analysis (CCA) aligned space. In total, thirteen clusters were identified using SNN modularity optimisation based clustering algorithm on the 35 significant CCA aligned components at 0.6 resolution. Positive marker genes that were expressed in at least half of the cells within the thirteen identified clusters were calculated with "FindAllMarkers" Seurat command, using Wilcoxon rank sum test with the threshold set to 0.25. We over-clustered the cells and then calculated for each cluster the average expression level of the top 20 marker genes. After calculating the correlation across the clusters we merged those with a correlation higher than 0.9. By merging the most highly correlated, we ended up with 4 clusters, and by calculating marker genes we were able to assign cell identity to the resulting clusters.

### 7.10.6 Pseudotime ordering

The set of common HVGs was used to order cells along a pseudotime trajectory using the Monocle2 R package v1.99.0. The 'tobit' expression family and 'DDRTree' reduction method were used with the default parameters. Finally, we identified genes that change as a

function of pseudotime across each of the three branches by setting the 'fullModelFormulaStr' parameter equal to '~sm.ns(Pseudotime)'. For the subclustering, we retrieved only the endothelial cluster (cluster 1) and performed separate clustering on the Monocle space using the strategy previously described for the tSNE space, ending up with 4 subclusters.

### 7.10.7 Cell cycle analysis

In order to infer cell cycle states in our single cell data we adopted the Satija's single cell scoring strategy (Butler et al., 2018). In more details, we assigned to each cell a score based on its expression of G2/M and S phase markers. These marker sets should be anticorrelated in their expression levels, and cells not expressing either are likely not cycling or in G1 phase. We assigned scores in the 'CellCycleScoring' function, which stores S and G2/M scores in 'object@meta.data', along with the predicted classification of each cell in either G2M, S or G1 phase.

# REFERENCES

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell *121*, 295–306.

Amos, T.A., and Gordon, M.Y. (1995). Sources of human hematopoietic stem cells for transplantation--a review. Cell Transplant. *4*, 547–569.

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell *121*, 295–306.

Amos, T.A., and Gordon, M.Y. (1995). Sources of human hematopoietic stem cells for transplantation--a review. Cell Transplant. *4*, 547–569.

Batsivari, A., Rybtsov, S., Souilhol, C., Binagui-Casas, A., Hills, D., Zhao, S., Travers, P., and Medvinsky, A. (2017). Understanding Hematopoietic Stem Cell Development through Functional Correlation of Their Proliferative Status with the Intra-aortic Cluster Architecture. Stem Cell Reports *8*, 1549–1562.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144–155.

Bertero, A., Madrigal, P., Galli, A., Hubner, N.C., Moreno, I., Burks, D., Brown, S., Pedersen, R.A., Gaffney, D., Mendjan, S., et al. (2015). Activin/Nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. Genes Dev. *29*, 702–717.

Bertero, A., Brown, S., Madrigal, P., Osnato, A., Ortmann, D., Yiangou, L., Kadiwala, J., Hubner, N.C., de los Mozos, I.R., Sadée, C., et al. (2018). The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. Nature *555*, 256–259.

Berthet, C., Aleem, E., Coppola, V., Tessarollo, L., and Kaldis, P. (2003). Cdk2 knockout mice are viable. Curr. Biol. *13*, 1775–1785.

Blajeski, A.L., Phan, V.A., Kottke, T.J., and Kaufmann, S.H. (2002). G(1) and G(2) cell-cycle arrest following microtubule depolymerization in human breast cancer cells. J. Clin. Invest. *110*, 91–99.

Böiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J.C.A., Azzoni, E., Woll, P.S., Mead, A.J., Hultquist, A., Swiers, G., et al. (2013). Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. Cell Stem Cell *13*, 535–548.

Boward, B., Wu, T., and Dalton, S. (2016). Concise Review: Control of Cell Fate Through Cell Cycle and Pluripotency Networks. Stem Cells *34*, 1427–1436.

Bowie, M.B., McKnight, K.D., Kent, D.G., McCaffrey, L., Hoodless, P.A., and Eaves, C.J. (2006). Hematopoietic stem cells proliferate until after birth and show a reversible phase-specific engraftment defect. J. Clin. Invest. *116*, 2808–2816.

Brazil, D.P., Church, R.H., Surae, S., Godson, C., and Martin, F. (2015). BMP signalling: agony and antagony in the family. Trends Cell Biol. *25*, 249–264.

Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C.H.-H., Lim, B., Vardy, L., Dunn, N.R., Trotter, M., Pedersen, R., et al. (2011). Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors. Stem Cells *29*, 1176–1185.

Burns, C.E., Traver, D., Mayhall, E., Shepard, J.L., and Zon, L.I. (2005). Hematopoietic stem cell fate is established by the Notch-Runx pathway. Genes Dev. *19*, 2331–2342.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell

transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Cahan, P., and Daley, G.Q. (2013). Origins and implications of pluripotent stem cell variability and heterogeneity. Nat. Rev. Mol. Cell Biol. *14*, 357–368.

Cai, Z., de Bruijn, M., Ma, X., Dortland, B., Luteijn, T., Downing, J.R., and Dzierzak, E. (2000). Haploinsufficiency of AML1 Affects the Temporal and Spatial Generation of Hematopoietic Stem Cells in the Mouse Embryo. Immunity *13*, 423–431.

Campbell, J.M., Nottle, M.B., Vassiliev, I., Mitchell, M., and Lane, M. (2012). Insulin Increases Epiblast Cell Number of In Vitro Cultured Mouse Embryos via the PI3K/GSK3/p53 Pathway. Stem Cells Dev. *21*, 2430–2441.

Camus, A., Perea-Gomez, A., Moreau, A., and Collignon, J. (2006). Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. Dev. Biol. *295*, 743–755.

Cao, H., Heazlewood, S.Y., Williams, B., Cardozo, D., Nigro, J., Oteiza, A., and Nilsson, S.K. (2016). The role of CD44 in fetal and adult hematopoietic stem cell regulation. Haematologica *101*, 26–37.

Carrelha, J., Meng, Y., Kettyle, L.M., Luis, T.C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., et al. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. Nature *554*, 106–111.

Chanda, B., Ditadi, A., Iscove, N.N., and Keller, G. (2013). Retinoic Acid Signaling Is Essential for Embryonic Hematopoietic Stem Cell Development. Cell *155*, 215–227.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011a). Chemically defined conditions for human iPSC derivation and culture. Nat. Methods *8*, 424–429.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011b). Chemically defined conditions for

human iPSC derivation and culture. Nat. Methods *8*, 424–429.

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E., and Speck, N.A. (2009). Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature *457*, 887–891.

Choi, K.D., Vodyanik, M.A., Togarrati, P.P., Suknuntha, K., Kumar, A., Samarjeet, F., Probasco, M.D., Tian, S., Stewart, R., Thomson, J.A., et al. (2012). Identification of the Hemogenic Endothelial Progenitor and Its Direct Precursor in Human Pluripotent Stem Cell Differentiation Cultures. Cell Rep. *2*, 553–567.

Clevers, H., and Nusse, R. (2012). Wnt/β-Catenin Signaling and Disease. Cell *149*, 1192–1205.

Crisan, M., Kartalaei, P.S., Vink, C.S., Yamada-Inagawa, T., Bollerot, K., van IJcken, W., van der Linden, R., de Sousa Lopes, S.M.C., Monteiro, R., Mummery, C., et al. (2015). BMP signalling differentially regulates distinct haematopoietic stem cell types. Nat. Commun. *6*, 8040.

D'Souza, A., Lee, S., Zhu, X., and Pasquini, M. (2017). Current Use and Trends in Hematopoietic Cell Transplantation in the United States. Biol. Blood Marrow Transplant. *23*, 1417–1421.

Dailey, L., Ambrosetti, D., Mansukhani, A., and Basilico, C. (2005). Mechanisms underlying differential responses to FGF signaling. Cytokine Growth Factor Rev. *16*, 233–247.

Datto, M.B., Li, Y., Panus, J.F., Howe, D.J., Xiong, Y., and Wang, X.F. (1995). Transforming growth factor beta induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. Proc. Natl. Acad. Sci. U. S. A. *92*, 5545–5549.

Diril, M.K., Ratnacaram, C.K., Padmakumar, V.C., Du, T., Wasser, M., Coppola, V., Tessarollo, L., and Kaldis, P. (2012). Cyclin-dependent kinase 1 (Cdk1) is essential for cell division and suppression of DNA re-replication but not for liver regeneration. Proc. Natl. Acad. Sci. U. S. A. *109*, 3826–3831.

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., et al. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. Nat. Cell Biol. *17*, 580–591.

Dou, D.R., Calvanese, V., Sierra, M.I., Nguyen, A.T., Minasian, A., Saarikoski, P., Sasidharan, R., Ramirez, C.M., Zack, J.A., Crooks, G.M., et al. (2016). Medial HOXA genes demarcate haematopoietic stem cell fate during human development. Nat. Cell Biol. *18*, 595–606.

Doulatov, S., Notta, F., Eppert, K., Nguyen, L.T., Ohashi, P.S., and Dick, J.E. (2010). Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. Nat. Immunol. *11*, 585–593.

Durand, C., Robin, C., Bollerot, K., Baron, M.H., Ottersbach, K., and Dzierzak, E. (2007). Embryonic stromal clones reveal developmental regulators of definitive hematopoietic stem cells. Proc. Natl. Acad. Sci. *104*, 20838–20843.

Dzierzak, E., and Bigas, A. (2018). Blood Development: Hematopoietic Stem Cell Dependence and Independence. Cell Stem Cell *22*, 639–651.

Easterbrook, J., Fidanza, A., and Forrester, L.M. (2016). Concise review: programming human pluripotent stem cells into blood. Br. J. Haematol. *173*, 671–679.

Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.-H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., et al. (2014). Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nat. Commun. *5*, 4372.

Eliasson, P., and Jönsson, J.-I. (2010). The hematopoietic stem cell niche: low in oxygen but a nice place to be. J. Cell. Physiol. *222*, 17–22.

Espín-Palazón, R., Stachura, D.L., Campbell, C.A., García-Moreno, D., Del Cid, N., Kim, A.D., Candel, S., Meseguer, J., Mulero, V., and Traver, D. (2014). Proinflammatory signaling regulates hematopoietic stem cell emergence. Cell *159*, 1070–1085.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Faial, T., Bernardo, A.S., Mendjan, S., Diamanti, E., Ortmann, D., Gentsch, G.E., Mascetti, V.L., Trotter, M.W.B., Smith, J.C., and Pedersen, R.A. (2015). Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. Development *142*, 2121–2135.

Fantl, V., Stamp, G., Andrews, A., Rosewell, I., and Dickson, C. (1995). Mice lacking cyclin D1 are small and show defects in eye and mammary gland development. Genes Dev. *9*, 2364–2372.

Ferreira, M.S.V., and Mousavi, S.H. (2018). Nanofiber technology in the ex vivo expansion of cord blood-derived hematopoietic stem cells. Nanomedicine Nanotechnology, Biol. Med. *14*, 1707–1718.

Finn, R.S., Dering, J., Conklin, D., Kalous, O., Cohen, D.J., Desai, A.J., Ginther, C., Atefi, M., Chen, I., Fowst, C., et al. (2009). PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. Breast Cancer Res. *11*, R77.

Frassoni, F. (2006). The laws covering in vitro fertilization and embryo research in Italy. Bone Marrow Transplant. *38*, 5–6.

Fujimoto, T., Anderson, K., Jacobsen, S.E.W., Nishikawa, S., and Nerlov, C. (2007). Cdk6 blocks myeloid differentiation by interfering with Runx1 DNA binding and Runx1-C/EBPα interaction. EMBO J. *26*, 2361–2370.

Fujiwara, Y., Chang, A.N., Williams, A.M., Orkin, S.H., Yoh, K., Ohta, J., Nishimura, S., Lim, K.C., Engel, J.D., and Yamamoto, M. (2004). Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. Blood *103*, 583–585.

Fuschi, P., Carrara, M., Voellenkle, C., Garcia-Manteiga, J.M., Righini, P., Maimone, B., Sangalli, E., Villa, F., Specchia, C., Picozza, M., et al. (2017). Central role of the p53 pathway in the noncoding-RNA response to oxidative stress. Aging (Albany. NY). *9*, 2559–2586.

Gama-Norton, L., Ferrando, E., Ruiz-Herguido, C., Liu, Z., Guiu, J., Islam, A.B.M.M.K., Lee, S.-U., Yan, M., Guidos, C.J., López-Bigas, N., et al. (2015). Notch signal strength controls cell fate

in the haemogenic endothelium. Nat. Commun. *6*, 8510.

Gering, M., and Patient, R. (2005). Hedgehog Signaling Is Required for Adult Blood Stem Cell Formation in Zebrafish Embryos. Dev. Cell *8*, 389–400.

Ginhoux, F., and Guilliams, M. (2016). Tissue-Resident Macrophage Ontogeny and Homeostasis. Immunity *44*, 439–449.

Glimm, H., Oh, I.H., and Eaves, C.J. (2000). Human hematopoietic stem cells stimulated to proliferate in vitro lose engraftment potential during their S/G(2)/M transit and do not reenter G(0). Blood *96*, 4185–4193.

Göke, J., Chan, Y.-S., Yan, J., Vingron, M., and Ng, H.-H. (2013). Genome-wide Kinase-Chromatin Interactions Reveal the Regulatory Network of ERK Signaling in Human Embryonic Stem Cells. Mol. Cell *50*, 844–855.

Goldie, L.C., Lucitti, J.L., Dickinson, M.E., Hirschi, K.K., Dieterlen-Lièvre, F., and Péault, B. (2008). Cell signaling directing the formation and function of hemogenic endothelium during murine embryogenesis. Blood *112*, 3194–3204.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Fibroblast Growth Factor 2 Modulates Transforming Growth Factor β Signaling in Mouse Embryonic Fibroblasts and Human ESCs (hESCs) to Support hESC Self-Renewal. Stem Cells *25*, 455–464.

Guo, G., Luc, S., Marco, E., Lin, T.-W., Peng, C., Kerenyi, M.A., Beyaz, S., Kim, W., Xu, J., Das, P.P., et al. (2013). Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. Cell Stem Cell *13*, 492–505.

Hao, Q.L., Zhu, J., Price, M.A., Payne, K.J., Barsky, L.W., and Crooks, G.M. (2001). Identification of a novel, human multilymphoid progenitor in cord blood. Blood *97*, 3683–3690.

Hellström, M., Phng, L.-K., Hofmann, J.J., Wallgard, E., Coultas, L., Lindblom, P., Alva, J., Nilsson, A.-K., Karlsson, L., Gaiano, N., et al. (2007). Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. Nature *445*, 776–780.

Hocevar, B.A., and Howe, P.H. (1998). Mechanisms of TGF-beta-induced cell cycle arrest. Miner. Electrolyte Metab. *24*, 131–135.

Hochedlinger, K., and Jaenisch, R. (2015). Induced Pluripotency and Epigenetic Reprogramming. Cold Spring Harb. Perspect. Biol. *7*, a019448.

Hoeffel, G. (2018). Fetal monocytes and the origins of tissue-resident macrophages. Cell. Immunol. *330*, 5–15.

Holden, C. (2009). Obama executive order. For Congress and NIH, headaches ahead on stem cells. Science *323*, 1552–1553.

Holyoake, T.L., Horrocks, C., Thomas, T., Eaves, C.J., and Eaves, A.C. (2000). Cell separation improves the sensitivity of detecting rare human normal and leukemic hematopoietic cells in vivo in NOD/SCID mice. Cytotherapy *2*, 411–421.

Hur, J., Choi, J.-I., Lee, H., Nham, P., Kim, T.-W., Chae, C.-W., Yun, J.-Y., Kang, J.-A., Kang, J., Lee, S.E., et al. (2016). CD82/KAI1 Maintains the Dormancy of Long-Term Hematopoietic Stem Cells through Interaction with DARC-Expressing Macrophages. Cell Stem Cell *18*, 508–521.

Hydbring, P., Bahram, F., Su, Y., Tronnersjö, S., Högstrand, K., von der Lehr, N., Sharifi, H.R., Lilischkis, R., Hein, N., Wu, S., et al. (2010). Phosphorylation by Cdk2 is required for Myc to repress Ras-induced senescence in cotransformation. Proc. Natl. Acad. Sci. U. S. A. *107*, 58–63.

Hydbring, P., Malumbres, M., and Sicinski, P. (2016). Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. Nat. Rev. Mol. Cell Biol. *17*, 280–292.

Ichikawa, M., Yoshimi, A., Nakagawa, M., Nishimoto, N., Watanabe-Okochi, N., and Kurokawa, M. (2013). A role for RUNX1 in hematopoiesis and myeloid leukemia. Int. J. Hematol. *97*, 726–734.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., and Benvenisty, N. (2000). Differentiation of human embryonic stem cells into embryoid bodies

compromising the three embryonic germ layers. Mol. Med. *6*, 88–95.

Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2011). Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-gonad-mesonephros region. J. Exp. Med. *208*, 2417–2427.

Ivanovs, A., Rybtsov, S., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2014a). Identification of the niche and phenotype of the first human hematopoietic stem cells. Stem Cell Reports *2*, 449–456.

Ivanovs, A., Rybtsov, S., Anderson, R.A., and Medvinsky, A. (2014b). CD43 but Not CD41 Marks the First Hematopoietic Stem Cells in the Human Embryo. Blood *124*.

Jirawatnotai, S., Hu, Y., Michowski, W., Elias, J.E., Becks, L., Bienvenu, F., Zagozdzon, A., Goswami, T., Wang, Y.E., Clark, A.B., et al. (2011). A function for cyclin D1 in DNA repair uncovered by protein interactome analyses in human cancers. Nature *474*, 230–234.

Kaufman, D.S. (2009). Toward clinical therapies using hematopoietic cells derived from human pluripotent stem cells. Blood *114*, 3513–3523.

Kekre, N., and Antin, J.H. (2014). Hematopoietic stem cell transplantation donor sources in the 21st century: choosing the ideal donor when a perfect match does not exist. Blood J.

Kennedy, M., D'Souza, S.L., Lynch-Kattman, M., Schwantz, S., and Keller, G. (2007). Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. Blood *109*, 2679–2687.

Kennedy, M., Awong, G., Sturgeon, C.M., Ditadi, A., LaMotte-Mohs, Ross Zúñiga-Pflücker JC, J.C., and Keller, G. (2012). T Lymphocyte Potential Marks the Emergence of Definitive Hematopoietic Progenitors in Human Pluripotent Stem Cell Differentiation Cultures. Cell Rep. *2*, 1722–1735.

Knockaert, M., Gray, N., Damiens, E., Chang, Y.-T., Grellier, P., Grant, K., Fergusson, D., Mottram, J., Soete, M., Dubremetz, J.-F., et al. (2000). Intracellular targets of cyclin-dependent

kinase inhibitors: identification by affinity chromatography using immobilised inhibitors. Chem. Biol. *7*, 411–422.

Kollmann, K., Heller, G., Schneckenleithner, C., Warsch, W., Scheicher, R., Ott, R.G., Schäfer, M., Fajmann, S., Schlederer, M., Schiefer, A.-I., et al. (2013). A Kinase-Independent Function of CDK6 Links the Cell Cycle to Tumor Angiogenesis. Cancer Cell *24*, 167–181.

Komiya, Y., and Habas, R. (2008). Wnt signal transduction pathways. Organogenesis *4*, 68–75.

Kopan, R., and Ilagan, M.X.G. (2009). The canonical Notch signaling pathway: unfolding the activation mechanism. Cell *137*, 216–233.

Kozar, K., Ciemerych, M.A., Rebel, V.I., Shigematsu, H., Zagozdzon, A., Sicinska, E., Geng, Y., Yu, Q., Bhattacharya, S., Bronson, R.T., et al. (2004). Mouse Development and Cell Proliferation in the Absence of D-Cyclins. Cell *118*, 477–491.

Krentz, N.A.J., van Hoof, D., Li, Z., Watanabe, A., Tang, M., Nian, C., German, M.S., and Lynn, F.C. (2017). Phosphorylation of NEUROG3 Links Endocrine Differentiation to the Cell Cycle in Pancreatic Progenitors. Dev. Cell *41*, 129–142.e6.

Kuleshov, M. V, Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90-7.

Kumano, K., Chiba, S., Kunisato, A., Sata, M., Saito, T., Nakagami-Yamaguchi, E., Yamaguchi, T., Masuda, S., Shimizu, K., Takahashi, T., et al. (2003). Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. Immunity *18*, 699–711.

Kumar, S., and Geiger, H. (2017). HSC Niche Biology and HSC Expansion Ex Vivo. Trends Mol. Med. *23*, 799–819.

Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., Zuyev, S., Ansell, J., and 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*, 4891–4899.

Lacombe, J., Herblot, S., Rojas-Sutterlin, S., Haman, A., Barakat, S., Iscove, N.N., Sauvageau, G., and Hoang, T. (2010). Scl regulates the quiescence and the long-term competence of hematopoietic stem cells. Blood *115*, 792–803.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N.K., Möröy, T., Bonifer, C., et al. (2012). GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. Blood *120*, 314–322.

Lanner, F., and Rossant, J. (2010). The role of FGF/Erk signaling in pluripotent cells. Development *137*, 3351–3360.

Laurenti, E., and Göttgens, B. (2018). From haematopoietic stem cells to complex differentiation landscapes. Nat. 2018 5537689 *553*, 418.

Laurenti, E., Frelin, C., Xie, S., Ferrari, R., Dunant, C.F., Zandi, S., Neumann, A., Plumb, I., Doulatov, S., Chen, J., et al. (2015). CDK6 levels regulate quiescence exit in human hematopoietic stem cells. Cell Stem Cell *16*, 302–313.

Leslie, J.D., Ariza-McNaughton, L., Bermange, A.L., McAdow, R., Johnson, S.L., and Lewis, J. (2007). Endothelial signalling by the Notch ligand Delta-like 4 restricts angiogenesis. Development *134*, 839–844.

Leung, A., Ciau-Uitz, A., Pinheiro, P., Monteiro, R., Zuo, J., Vyas, P., Patient, R., and Porcher, C. (2013). Uncoupling VEGFA functions in arteriogenesis and hematopoietic stem cell specification. Dev. Cell *24*, 144–158.

Li, X., Rong, Y., Zhang, M., Wang, X.L., LeMaire, S.A., Coselli, J.S., Zhang, Y., and Shen, Y.H. (2009). Up-regulation of thioredoxin interacting protein (Txnip) by p38 MAPK and FOXO1 contributes to the impaired thioredoxin activity and increased ROS in glucose-treated endothelial cells. Biochem. Biophys. Res. Commun. *381*, 660–665.

Li, Z., Lan, Y., He, W., Chen, D., Wang, J., Zhou, F., Wang, Y., Sun, H., Chen, X., Xu, C., et al.

(2012). Mouse embryonic head as a site for hematopoietic stem cell development. Cell Stem Cell *11*, 663–675.

Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S.A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M.S., Wu, M., et al. (2012). RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. EMBO J. *31*, 4318–4333.

Lis, R., Karrasch, C.C., Poulos, M.G., Kunar, B., Redmond, D., Duran, J.G.B., Badwe, C.R., Schachterle, W., Ginsberg, M., Xiang, J., et al. (2017). Conversion of adult endothelium to immunocompetent haematopoietic stem cells. Nature *545*, 439–445.

Liu, J.-Y., Yao, J., Li, X.-M., Song, Y.-C., Wang, X.-Q., Li, Y.-J., Yan, B., and Jiang, Q. (2014). Pathogenic role of lncRNA-MALAT1 in endothelial cell dysfunction in diabetes mellitus. Cell Death Dis. *5*, e1506.

Lizama, C.O., Hawkins, J.S., Schmitt, C.E., Bos, F.L., Zape, J.P., Cautivo, K.M., Borges Pinto, H., Rhyner, A.M., Yu, H., Donohoe, M.E., et al. (2015). Repression of arterial genes in hemogenic endothelium is sufficient for haematopoietic fate acquisition. Nat. Commun. *6*, 7739.

Luesink, M., Hollink, I.H.I.M., van der Velden, V.H.J., Knops, R.H.J.N., Boezeman, J.B.M., de Haas, V., Trka, J., Baruchel, A., Reinhardt, D., van der Reijden, B.A., et al. (2012). High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. Blood *120*, 2064–2075.

Malumbres, M., Sotillo, R., Santamaría, D., Galán, J., Cerezo, A., Ortega, S., Dubus, P., and Barbacid, M. (2004). Mammalian Cells Cycle without the D-Type Cyclin-Dependent Kinases Cdk4 and Cdk6. Cell *118*, 493–504.

Marks-Bluth, J., Khanna, A., Chandrakanthan, V., Thoms, J., Bee, T., Eich, C., Kang, Y.C., Knezevic, K., Qiao, Q., Fitch, S., et al. (2015). SMAD1 and SMAD5 Expression Is Coordinately Regulated by FLI1 and GATA2 during Endothelial Development. Mol. Cell. Biol. *35*, 2165–2172.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in

medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. *78*, 7634–7638.

McGarvey, A.C., Rybtsov, S., Souilhol, C., Tamagno, S., Rice, R., Hills, D., Godwin, D., Rice, D., Tomlinson, S.R., and Medvinsky, A. (2017). A molecular roadmap of the AGM region reveals BMPER as a novel regulator of HSC maturation. J. Exp. Med. *214*, 3731–3751.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep. *11*, 1892–1904.

McInnes, C. (2008). Progress in the evaluation of CDK inhibitors as anti-tumor agents. Drug Discov. Today *13*, 875–881.

McLaren, A. (2001). Ethical and social considerations of stem cell research. Nature *414*, 129–131.

McMahon, C., Suthiphongchai, T., DiRenzo, J., and Ewen, M.E. (1999). P/CAF associates with cyclin D1 and potentiates its activation of the estrogen receptor. Proc. Natl. Acad. Sci. U. S. A. *96*, 5382–5387.

Medvinsky, A., and Dzierzak, E. (1996). Definitive hematopoiesis is autonomously initiated by the AGM region. Cell *86*, 897–906.

Meijer, L., Borgne, A., Mulner, O., Chong, J.P., Blow, J.J., Inagaki, N., Inagaki, M., Delcros, J.G., and Moulinoux, J.P. (1997). Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. Eur. J. Biochem. *243*, 527–536.

Mendes, S.C., Robin, C., and Dzierzak, E. (2005). Mesenchymal progenitor cells localize within hematopoietic sites throughout ontogeny. Development *132*, 1127–1136.

Mendjan, S., Mascetti, V.L., Ortmann, D., Ortiz, M., Karjosukarso, D.W., Ng, Y., Moreau, T.,

and Pedersen, R.A. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. Cell Stem Cell *15*, 310–325.

Metcalf, R.A., Zhao, S., Anderson, M.W., Lu, Z.S., Galperin, I., Marinelli, R.J., Cherry, A.M., Lossos, I.S., and Natkunam, Y. (2010). Characterization of D-cyclin proteins in hematolymphoid neoplasms: lack of specificity of cyclin-D2 and D3 expression in lymphoma subtypes. Mod. Pathol. *23*, 420–433.

Moreau, T., Evans, A.L., Vasquez, L., Tijssen, M.R., Yan, Y., Trotter, M.W., Howard, D., Colzani, M., Arumugam, M., Wu, W.H., et al. (2016). Large-scale production of megakaryocytes from human pluripotent stem cells by chemically defined forward programming. Nat. Commun. *7*, 11208.

Murry, C.E., and Keller, G. (2008). Differentiation of Embryonic Stem Cells to??Clinically Relevant Populations: Lessons from Embryonic Development. Cell *132*, 661–680.

Nestorowa, S., Hamey, F.K., Pijuan Sala, B., Diamanti, E., Shepherd, M., Laurenti, E., Wilson, N.K., Kent, D.G., and Gottgens, B. (2016). A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. Blood *128*, e20–e31.

Ng, E.S., Azzola, L., Bruveris, F.F., Calvanese, V., Phipson, B., Vlahos, K., Hirst, C., Jokubaitis, V.J., Yu, Q.C., Maksimovic, J., et al. (2016). Differentiation of human embryonic stem cells to HOXA+ hemogenic vasculature that resembles the aorta-gonad-mesonephros. Nat. Biotechnol. *34*, 1168–1179.

Niwa, A., Heike, T., Umeda, K., Oshima, K., Kato, I., Sakai, H., Suemori, H., Nakahata, T., and Saito, M.K. (2011). A novel Serum-Free monolayer culture for orderly hematopoietic differentiation of human pluripotent cells via mesodermal progenitors. PLoS One *6*.

North, T., Gu, T.L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marín-Padilla, M., and Speck, N.A. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. Development *126*, 2563–2575.

Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., Mcleod,

J., Laurenti, E., Dunant, C.F., et al. (2015). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science *351*, 1–16.

Oberlin, E., Tavian, M., Blazsek, I., and Péault, B. (2002). Blood-forming potential of vascular endothelium in the human embryo. Development *129*, 4147–4157.

Ohata, S., Nawa, M., Kasama, T., Yamasaki, T., Sawanobori, K., Hata, S., Nakamura, T., Asaoka, Y., Watanabe, T., Okamoto, H., et al. (2009). Hematopoiesis-dependent expression of CD44 in murine hepatic progenitor cells. Biochem. Biophys. Res. Commun. *379*, 817–823.

Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the Target of Multiple Chromosomal Translocations in Human Leukemia, Is Essential for Normal Fetal Liver Hematopoiesis. Cell *84*, 321–330.

Oren, T., Torregroza, I., and Evans, T. (2005). An Oct-1 binding site mediates activation of the gata2 promoter by BMP signaling. Nucleic Acids Res. *33*, 4357–4367.

Org, T., Duan, D., Ferrari, R., Montel-Hagen, A., Van Handel, B., Kerényi, M.A., Sasidharan, R., Rubbi, L., Fujiwara, Y., Pellegrini, M., et al. (2015). Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. EMBO J. *34*, 759–777.

Ostergaard, P., Simpson, M.A., Connell, F.C., Steward, C.G., Brice, G., Woollard, W.J., Dafou, D., Kilo, T., Smithson, S., Lunt, P., et al. (2011). Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). Nat. Genet. *43*, 929–931.

Ottersbach, K., and Dzierzak, E. (2005). The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. Dev. Cell *8*, 377–387.

Palazzani, L. (2011). Embryo research in Italy: the bioethical and biojuridical debate. Hum. Reprod. Genet. Ethics *17*, 28–39.

Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. Front. Physiol. *5*, 3.

Passweg, J.R., Baldomero, H., Bader, P., Bonini, C., Duarte, R.F., Dufour, C., Gennery, A.,

Kröger, N., Kuball, J., Lanza, F., et al. (2017). Use of haploidentical stem cell transplantation continues to increase: the 2015 European Society for Blood and Marrow Transplant activity survey report. Bone Marrow Transplant. *52*, 811–817.

de Pater, E., Kaimakis, P., Vink, C.S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P.S., Camper, S.A., Speck, N., and Dzierzak, E. (2013). Gata2 is required for HSC generation and survival. J. Exp. Med. *210*, 2843–2850.

Pauklin, S., and Vallier, L. (2013). The cell cycle state of stem cells determines cell fate propensity. Cell *155*, 135–147.

Pauklin, S., and Vallier, L. (2015). Activin/Nodal signalling in stem cells. Development *142*, 607–619.

Pauklin, S., Pedersen, R.A., and Vallier, L. (2011). Mouse pluripotent stem cells at a glance. J. Cell Sci. *124*, 3727–3732.

Pauklin, S., Madrigal, P., Bertero, A., and Vallier, L. (2016). Initiation of stem cell differentiation involves cell cycle-dependent regulation of developmental genes by Cyclin D. Genes Dev. *30*, 421–433.

Peeters, M., Ottersbach, K., Bollerot, K., Orelio, C., de Bruijn, M., Wijgerde, M., and Dzierzak, E. (2009). Ventral embryonic tissues and Hedgehog proteins induce early AGM hematopoietic stem cell development. Development *136*, 2613–2621.

Peterson, L.F., Boyapati, A., Ranganathan, V., Iwama, A., Tenen, D.G., Tsai, S., and Zhang, D.-E. (2005). The Hematopoietic Transcription Factor AML1 (RUNX1) Is Negatively Regulated by the Cell Cycle Protein Cyclin D3. Mol. Cell. Biol. *25*, 10205–10219.

Pietras, E.M., Warr, M.R., and Passegué, E. (2011). Cell cycle regulation in hematopoietic stem cells. J. Cell Biol. *195*, 709–720.

Pietras, E.M., Reynaud, D., Kang, Y.-A., Carlin, D., Calero-Nieto, F.J., Leavitt, A.D., Stuart, J.M., Göttgens, B., and Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased

Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. Cell Stem Cell *17*, 35–46.

Poon, R.Y.C. (2016). Cell Cycle Control: A System of Interlinking Oscillators. (Humana Press, New York, NY), pp. 3–19.

Pouget, C., Peterkin, T., Simões, F.C., Lee, Y., Traver, D., and Patient, R. (2014). FGF signalling restricts haematopoietic stem cell specification via modulation of the BMP pathway. Nat. Commun. *5*, 5588.

Prevo, R., Pirovano, G., Puliyadi, R., Herbert, K.J., Rodriguez-Berriguete, G., O'Docherty, A., Greaves, W., McKenna, W.G., and Higgins, G.S. (2018). CDK1 inhibition sensitizes normal cells to DNA damage in a cell cycle dependent manner. Cell Cycle 1–11.

Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.-A., and Trapnell, C. (2017a). Single-cell mRNA quantification and differential analysis with Census. Nat. Methods *14*, 309–315.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017b). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982.

Ramos-Mej??a, V., Navarro-Montero, O., Ayll??n, V., Bueno, C., Romero, T., Real, P.J., and Menendez, P. (2014). HOXA9 promotes hematopoietic commitment of human embryonic stem cells. Blood *124*, 3065–3075.

Robert-Moreno, A., Espinosa, L., de la Pompa, J.L., Bigas, A., and Simpson, P. (2005). RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. Development *132*, 1117–1126.

Robin, C., Bollerot, K., Mendes, S., Haak, E., Crisan, M., Cerisoli, F., Lauw, I., Kaimakis, P., Jorna, R., Vermeulen, M., et al. (2009). Human Placenta Is a Potent Hematopoietic Niche Containing Hematopoietic Stem and Progenitor Cells throughout Development. Cell Stem Cell *5*, 385–395.

Rocca, A., Farolfi, A., Bravaccini, S., Schirone, A., and Amadori, D. (2014). Palbociclib (PD

0332991): targeting the cell cycle machinery in breast cancer. Expert Opin. Pharmacother. *15*, 407–420.

Ruiz-Herguido, C., Guiu, J., D'Altri, T., Inglés-Esteve, J., Dzierzak, E., Espinosa, L., and Bigas, A. (2012). Hematopoietic stem cell development requires transient Wnt/β-catenin activity. J. Exp. Med. *209*, 1457–1468.

Rybtsov, S., Sobiesiak, M., Taoudi, S., Souilhol, C., Senserrich, J., Liakhovitskaia, A., Ivanovs, A., Frampton, J., Zhao, S., and Medvinsky, A. (2011). Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. J. Exp. Med. *208*, 1305–1315.

Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature *502*, 232–236.

Sankaran, V.G., and Orkin, S.H. (2013). The switch from fetal to adult hemoglobin. Cold Spring Harb. Perspect. Med. *3*, a011643.

Santamaría, D., Barrière, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Cáceres, J.F., Dubus, P., Malumbres, M., and Barbacid, M. (2007a). Cdk1 is sufficient to drive the mammalian cell cycle. Nature *448*, 811–815.

Santamaría, D., Barrière, C., Cerqueira, A., Hunt, S., Tardy, C., Newton, K., Cáceres, J.F., Dubus, P., Malumbres, M., and Barbacid, M. (2007b). Cdk1 is sufficient to drive the mammalian cell cycle. Nature *448*, 811–815.

Sawamiphak, S., Kontarakis, Z., and Stainier, D.Y.R. (2014). Interferon gamma signaling positively regulates hematopoietic stem cell emergence. Dev. Cell *31*, 640–653.

Schafer, K.A. (1998). The cell cycle: a review. Vet. Pathol. *35*, 461–478.

Shivdasani, R.A., Mayer, E.L., and Orkin, S.H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. Nature *373*, 432–434.

Sicinska, E., Aifantis, I., Le Cam, L., Swat, W., Borowski, C., Yu, Q., Ferrando, A.A., Levin, S.D., Geng, Y., von Boehmer, H., et al. (2003). Requirement for cyclin D3 in lymphocyte development and T cell leukemias. Cancer Cell *4*, 451–461.

Siekmann, A.F., and Lawson, N.D. (2007). Notch signalling and the regulation of angiogenesis. Cell Adh. Migr. *1*, 104–106.

Singh, A.M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A.L., Sun, Y., Menendez, L., Kulik, M., and Dalton, S. (2012). Signaling Network Crosstalk in Human Pluripotent Cells: A Smad2/3-Regulated Switch that Controls the Balance between Self-Renewal and Differentiation. Cell Stem Cell *10*, 312–326.

Slukvin, I.I. (2013). Hematopoietic specification from human pluripotent stem cells : current advances and challenges toward de novo generation of hematopoietic stem cells. Blood *122*, 4035–4046.

Souilhol, C., Gonneau, C., Lendinez, J.G., Batsivari, A., Rybtsov, S., Wilson, H., Morgado-Palacin, L., Hills, D., Taoudi, S., Antonchuk, J., et al. (2016). Inductive interactions mediated by interplay of asymmetric signalling underlie development of adult haematopoietic stem cells. Nat. Commun. *7*, 10784.

Sturgeon, C.M., Ditadi, A., Clarke, R.L., and Keller, G. (2013). Defining the path to hematopoietic stem cells. Nat. Biotechnol. *31*, 416–418.

Sturgeon, C.M., Ditadi, A., Awong, G., Kennedy, M., and Keller, G. (2014). Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. Nat. Biotechnol. *32*, 554–561.

Su, S.C., and Tsai, L.-H. (2011). Cyclin-Dependent Kinases in Brain Development and Disease. Annu. Rev. Cell Dev. Biol. *27*, 465–491.

Sugimura, R., Jha, D.K., Han, A., Soria-Valles, C., da Rocha, E.L., Lu, Y.-F., Goettel, J.A., Serrao, E., Rowe, R.G., Malleshaiah, M., et al. (2017). Haematopoietic stem and progenitor cells from human pluripotent stem cells. Nature *545*, 432–438.

Sumi, T., Oki, S., Kitajima, K., and Meno, C. (2013). Epiblast ground state is controlled by canonical Wnt/β-catenin signaling in the postimplantation mouse embryo and epiblast stem cells. PLoS One *8*, e63378.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872.

Tam, P.P.L., and Behringer, R.R. (1997). Mouse gastrulation:the formation of the mammalian body plan. Mech. Dev. *68*, 3–25.

Taoudi, S. (2005). Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. Development *132*, 4179–4191.

Tavian, M., Coulombel, L., Luton, D., Clemente, H.S., Dieterlen-Lièvre, F., and Péault, B. (1996). Aorta-associated CD34+ hematopoietic cells in the early human embryo. Blood *87*, 67–72.

Tavian, M., Hallais, M.F., and Péault, B. (1999). Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo. Development *126*, 793–803.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tian, D., Dong, J., Jin, S., Teng, X., and Wu, Y. (2017). Endogenous hydrogen sulfide-mediated MAPK inhibition preserves endothelial function through TXNIP signaling. Free Radic. Biol. Med. *110*, 291–299.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak,

K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386.

Vagapova, E.R., Spirin, P. V, Lebedev, T.D., and Prassolov, V.S. (2018). The Role of TAL1 in Hematopoiesis and Leukemogenesis. Acta Naturae *10*, 15–23.

Vallier, L., Reynolds, D., and Pedersen, R.A. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. Dev. Biol. *275*, 403–421.

Vallier, L., Alexander, M., and Pedersen, R. a (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell Sci. *118*, 4495–4509.

Vallier, L., Touboul, T., Brown, S., Cho, C., Bilican, B., Alexander, M., Cedervall, J., Chandran, S., Ahrlund-Richter, L., Weber, A., et al. (2009). Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. Stem Cells *27*, 2655–2666.

Vargel, Ö., Zhang, Y., Kosim, K., Ganter, K., Foehr, S., Mardenborough, Y., Shvartsman, M., Enright, A.J., Krijgsveld, J., and Lancrin, C. (2016). Activation of the TGFβ pathway impairs endothelial to haematopoietic transition. Sci. Rep. *6*, 21518.

Vassilev, L.T. (2006). Cell Cycle Synchronization at the G2/M Phase Border by Reversible Inhibition of CDK1. Cell Cycle *5*, 2555–2556.

Vella, S., Tavanti, E., Hattinger, C.M., Fanelli, M., Versteeg, R., Koster, J., Picci, P., and Serra, M. (2016). Targeting CDKs with Roscovitine Increases Sensitivity to DNA Damaging Drugs of Human Osteosarcoma Cells. PLoS One *11*, e0166233.

Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. *19*, 271–281.

Venditti, A., Battaglia, A., Del Poeta, G., Buccisano, F., Maurillo, L., Tamburini, A., Del Moro, B., Epiceno, A., Martiradonna, M., Caravita, T., et al. (1999). Enumeration of CD34+ hematopoietic progenitor cells for clinical transplantation: comparison of three different

methods. Bone Marrow Transplant. *24*, 1019–1027.

Vicente, C., Vazquez, I., Conchillo, A., García-Sánchez, M.A., Marcotegui, N., Fuster, O., González, M., Calasanz, M.J., Lahortiga, I., and Odero, M.D. (2012). Overexpression of GATA2 predicts an adverse prognosis for patients with acute myeloid leukemia and it is associated with distinct molecular abnormalities. Leukemia *26*, 550–554.

Vo, L.T., and Daley, G.Q. (2015). De novo generation of HSCs from somatic and pluripotent stem cell sources. Blood *125*, 2641–2648.

Vodyanik, M.A., Thomson, J.A., and Slukvin, I.I. (2006). Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. Blood *108*, 2095–2105.

Wang, R.N., Green, J., Wang, Z., Deng, Y., Qiao, M., Peabody, M., Zhang, Q., Ye, J., Yan, Z., Denduluri, S., et al. (2014a). Bone Morphogenetic Protein (BMP) signaling in development and human diseases. Genes Dis. *1*, 87–105.

Wang, X.-Y., Lan, Y., He, W.-Y., Zhang, L., Yao, H.-Y., Hou, C.-M., Tong, Y., Liu, Y.-L., Yang, G., Liu, X.-D., et al. (2008). Identification of mesenchymal stem cells in aorta-gonad-mesonephros and yolk sac of human embryos. Blood *111*, 2436–2443.

Wang, Z., Fan, M., Candas, D., Zhang, T.-Q., Qin, L., Eldridge, A., Wachsmann-Hogiu, S., Ahmed, K.M., Chromy, B.A., Nantajit, D., et al. (2014b). Cyclin B1/Cdk1 Coordinates Mitochondrial Respiration for Cell-Cycle G2/M Progression. Dev. Cell *29*, 217–232.

Warren, L., Bryder, D., Weissman, I.L., and Quake, S.R. (2006). Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. Proc. Natl. Acad. Sci. U. S. A. *103*, 17807–17812.

Watt, S.M., Butler, L.H., Tavian, M., Bühring, H.J., Rappold, I., Simmons, P.J., Zannettino, a C., Buck, D., Fuchs, a, Doyonnas, R., et al. (2000). Functionally defined CD164 epitopes are expressed on CD34(+) cells throughout ontogeny but display distinct distribution patterns in adult hematopoietic and nonhematopoietic tissues. Blood *95*, 3113–3124.

Wermuth, P.J., Li, Z., Mendoza, F.A., and Jimenez, S.A. (2016). Stimulation of Transforming Growth Factor-β1-Induced Endothelial-To-Mesenchymal Transition and Tissue Fibrosis by Endothelin-1 (ET-1): A Novel Profibrotic Effect of ET-1. PLoS One *11*, e0161988.

Wiles, M. V, and Johansson, B.M. (1999). Embryonic stem cell development in a chemically defined medium. Exp. Cell Res. *247*, 241–248.

Wilkinson, R.N., Pouget, C., Gering, M., Russell, A.J., Davies, S.G., Kimelman, D., and Patient, R. (2009). Hedgehog and Bmp Polarize Hematopoietic Stem Cell Emergence in the Zebrafish Dorsal Aorta. Dev. Cell *16*, 909–916.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial Transcriptional Control In Blood Stem/Progenitor Cells: Genome-wide Analysis of Ten Major Transcriptional Regulators. Cell Stem Cell *7*, 532–544.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013a). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell *154*, 1112–1126.

Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013b). Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells. Cell *154*, 1112–1126.

Yokomizo, T., and Dzierzak, E. (2010). Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. Development *137*, 3651–3661.

Yoo, Y.D., Choi, J.Y., Lee, S.J., Kim, J.S., Min, B.R., Lee, Y.I., and Kang, Y.K. (1999). TGF-beta-induced cell-cycle arrest through the p21(WAF1/CIP1)-G1 cyclin/Cdks-p130 pathway in gastric-carcinoma cells. Int. J. Cancer *83*, 512–517.

Yusa, K., Rashid, S.T., Strick-Marchand, H., Varela, I., Liu, P.-Q., Paschon, D.E., Miranda, E., Ordóñez, A., Hannan, N.R.F., Rouhani, F.J., et al. (2011). Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. Nature *478*, 391–394.

Zhang, J., Li, H., Yabut, O., Fitzpatrick, H., D'Arcangelo, G., and Herrup, K. (2010). Cdk5 suppresses the neuronal cell cycle by disrupting the E2F1-DP1 complex. J. Neurosci. *30*, 5219–5228.