DOCTOR OF PHILOSOPHY

Using single cell transcriptomics to characterise endothelial populations in embryonic haematopoiesis

Oatley, Morgan Ann Elizabeth

Award date:
2019

Using single cell transcriptomics to characterise endothelial populations in embryonic haematopoiesis

Morgan Ann Elizabeth Oatley BSc (Hons)
EMBL Rome

March 2019

Supervisor: Dr. Christophe Lancrin

This thesis was submitted in fulfilment of the requirement for the degree of Doctor of Philosophy

University of Dundee
School of Life Sciences

EMBL

University of Dundee

Contents

List of Tables

List of Figures

Acknowledgments

Firstly, I would like to acknowledge and thank my supervisor Dr. Christophe Lancrin, who invited me into his lab and shared his science with me. Without his support and mentorship this project would not have been possible. I would like to also thank the past and present members of the Lancrin lab who provided me with support both scientifically and personally. In particular, I would like to thank Irina Pinheiro, Isabelle Bergiers and Yasmin Serina Secanechia whose friendship was essential to the making of this PhD and who so willingly gave me assistance, time and ideas whenever I asked. I am also grateful to the support I received from the facilities at EMBL Rome. In particular I would like to thank Cora Chadick who was there for countless FACS experiments.

Secondly, I would like to thank my former supervisors Professor Moira O'Bryan and Dr. Duangporn Jamsai who introduced me to the excitement of biological research, gave me confidence in my abilities and encouraged me to dream big.

Thirdly, I would like to thank my family and friends back home in Australia who encouraged and supported me from afar. And to Marcos who always looked on the brighter side and was a very good listener.

Declaration

I declare that this dissertation is an original composition and based on the results of my own research. Work that is not my own has been clearly indicated in the text with reference to the researchers or their publications. This dissertation has not previously been submitted for a higher degree.

Morgan Ann Elizabeth Oatley

I certify that Morgan Ann Elizabeth Oatley has spent the equivalent of at least 3.5 years in research work at the European Molecular Biology Laboratories, Rome, Italy and that she has fulfilled all the conditions of the Ordinance and Regulations relevant for submission. Morgan Ann Elizabeth Oatley is therefore qualified to submit and present this thesis in application for the degree of Doctor of Philosophy at the School of Life Sciences, University of Dundee (EMBL-Dundee Joint Program).

Christophe Lancrin

Summary

Haematopoietic stem cells (HSCs) possess the unique ability to reconstitute an animal's entire haematopoietic system. As such, it is inherently interesting to the fields of regenerative medicine and developmental biology to understand how these cells first form in the embryo. Haematopoietic stem and progenitor cells (HSPCs) are understood to derive from a specialised endothelial precursor termed haemogenic endothelium during a narrow window of development. This project was undertaken to better characterise endothelial populations in embryonic haematopoietic tissues in an effort to improve our understanding of the endothelial to haematopoietic transition (EHT). To this end we employed single cell technology, immunofluorescence, flow cytometry and sorting, *ex vivo* co-culturing assays and a mouse embryonic stem cell (mESC) differentiation system. This work identified two hyaluronan receptors, *Cd44* and *Stabilin-2* (*Stab2*) on the surface of endothelial cells derived from the aorta gonad mesonephros (AGM) and yolk sac, respectively. We showed that *Cd44* could be used to mark all stages of haematopoietic development arising from the AGM. This enabled us to provide transcriptional data on haemogenic endothelium, identifying potential new regulators of embryonic haematopoiesis and characterising the endothelial precursors as quiescent in nature. Furthermore, *ex vivo* and *in vitro* culturing of CD44+ cells identified a functional role for CD44 in the process of EHT. The precise mechanism by which CD44 and its ligand hyaluronan assist in the emergence of HSPCs remains unclear. We further identified the cell surface receptor Stab2 as a distinguishing feature of yolk sac vascular endothelium compared with vascular endothelium of the AGM.

Our analysis suggests that EHT is more likely to derive from yolk sac endothelium that is negative for Stab2 expression and by excluding the expression of Stab2 we can enrich for AGM-like haematopoietic progenitors. Our RNA sequencing data suggests that Stab2+ endothelial cells have a transcriptional profile similar to liver sinusoidal endothelial cells, where HSPCs migrate after emergence. This poses an interesting question of whether endothelial niche cells attract haematopoietic progenitors to the yolk sac and their role there. Overall, these two hyaluronan receptors can be used to define distinct populations in the yolk sac and AGM and likely have importance for the emergence of HSPCs.

Abbreviations

7-AAD	7-aminoactinomycin D
4-MU	4-methylumbelliferone
AF	alexa fluorophore
AGM	aorta gonad mesonephros
ANOVA	analysis of variance
APC	allophycocyanin
BMP	bone morphogenic protein
BSA	bovine serum albumin
BV	brilliant violet
CSC	cancer stem cell
CD44s	small variant isoform of CD44
CD44v	variant isoform of CD44
CFU	colony forming unit
ChIP-seq	chromatin immunoprecipitation and sequencing
CSCs	cancer stem cells
DAPI	4′,6-diamidino-2-phenylindole
dHSC	definitive haematopoietic stem cell
DMEM	dulbecco's modified eagle medium
DMEM KO	dulbecco's modified eagle medium knock out
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
EB	embryoid body
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EDU	5-ethynyl-2′-deoxyuridin
EGF	epidermal growth factor
EHT	endothelial to haematopoietic transition
EMP	erythroid myeloid progenitor, erythroid myeloid progenitor
EMT	epithelial to mesenchymal transition
ERK	extracellular signal-regulated kinases
ERM	ezrin radixin moesin

ESC	embryonic stem cell
FACS	fluorescent activated cell sorting
FBS	foetal bovine serum
FGF-1	fibroblast growth factor 1
Flt-3	fms-like tyrosine kinase 3 ligand
FMO	fluorescence minus one
G-CSF	granulocyte colony stimulating factor
GFP	green fluorescent protein
HABP	hyaluronan binding protein
HGF	hepatocyte growth factor
HPF	hours post fertilisation
hPSC	human pluripotent stem cell
HSC	haematopoietic stem cell
hPSC	human pluripotent stem cells
HSPC	haematopoietic stem and progenitor cell
IGF-2	insulin-like growth factor 2
IMDM	Iscove's modified dulbecco's medium
iPSC	induced pluripotent stem cell
IL-11	interleukin 11
IL-3	interleukin 3
IL-6	interleukin 6
iPSC	induced pluripotent stem cell
LDL	lipodensity lipoprotein
LIF	leukaemia inhibitory factor
MACS	magnetic activated cell sorting
MEF	mouse embryonic fibroblast
mESC	mous embryonic stem cell
MMP	matrix metallanoproteinases
MPP	multipotent progenitor
MTG	monothioglycerol
PAM	protospacer adjacent motif
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PE	phycoerythrin

pre-HSC	precursor of haematopoietic stem cell
qPCR	quantitative polymerase chain reaction
RBP-J	recombination signal binding protein for immunoglobulin kappa J
RNA	ribonucleic acid
SCF	stem cell factor
sgRNA	single guide ribonucleic acid
TBS	tris buffered saline
TF	transcription factor
TGF	transforming growth factor
TPO	thrombopoietin
VEGF	vascular endothelial growth factor
VSM	vascular smooth muscle

Chapter 1: Embryonic haematopoiesis

1.1 The embryonic vasculature & primitive haematopoiesis

1.1.1 Endothelial and haematopoietic lineages have a shared developmental origin

The cardiovascular network is the first organ system to develop in the embryo due to its critical role in gas exchange, the supply of nutrients and the removal of waste products. Interestingly, the two integral components of this system – the blood vessels and the blood cells are closely entwined in their development (Chong et al., 2011). The close proximity and shared gene expression of endothelial and haematopoietic cells in chick and mouse embryos led to the hypothesis that these lineages also share a developmental origin (Sabin, 1917; Kallianpur, Jordan and Brandt, 1994; Young, Baumhueter and Lasky, 1995). The loss of vascular endothelial growth factor receptor 2 (*Vegfr-2*), also known as *Flk-1,* provided some of the first direct evidence for a common progenitor; with mice displaying embryonic lethality due to loss of both vascular and haematopoietic cell types (Shalaby *et al.*, 1995). This hypothesised shared progenitor was termed the haemangioblast.

Further investigation of the bi-potentiality of *Flk-1*-expressing mesoderm, derived from mouse embryonic stem cells (mESCs) supported the existence of a haemangioblast precursor. Single Flk-1+ clones were found to give rise to both endothelial and haematopoietic cells in culture (Choi *et al.*, 1998). Lineage tracing in zebrafish embryos further supported this idea (Vogeli *et al.*, 2006). The use of laser-activated cell labelling, enabled the tracking of single cells from 6 to 30 hours post fertilisation (HPF) and identified progeny of

single cells with either expression of *Flk-1* or the haematopoietic transcription factor *Gata1* (Vogeli *et al.*, 2006). Based on this research in zebrafish and mESCs, the haemangioblast was confirmed as the source of both the vascular and haematopoietic lineages.

1.1.2 Haemangioblasts derive from the posterior primitive streak

Based on traditional microscopy techniques it was thought that these haemangioblast progenitors in mice coalesce in the yolk sac to form blood islands between embryonic day 7 and 7.5 (E7 – E7.5) (Palis and Yoder, 2001). Cells on the periphery would flatten into an endothelial phenotype and cells located centrally would up-regulate haemoglobin genes and differentiate into primitive erythrocytes in the newly formed bloodstream (Palis and Yoder, 2001). However advances in both molecular biology and fluorescent tagging techniques have challenged the traditional view of mammalian yolk sac haematopoiesis (Ferkowicz and Yoder, 2005). Although *Flk-1* has been shown to mark both endothelial and haematopoietic progenitors (Yamaguchi *et al.*, 1993; Kabrun *et al.*, 1997), further work in mouse embryos using fluorescent barcoding has revealed a lack of clonality in yolk sac derived blood islands (Ueno and Weissman, 2006). Fluorescent tagging of the murine *Flk-1* gene, in conjunction with the mesodermal marker *Brachyury* found that the endothelial and haematopoietic progenitor populations derive from the posterior primitive streak and that differentiation begins prior to migration into the yolk sac (Huber *et al.*, 2004). This is supported by expression and functional studies of the early haematopoietic transcription factors *Tal1*, *Gata-1* and *Gata-2.* RNA in-situ hybridisation experiments found these transcription

factors turned on at the mouse primitive streak stage of development prior to blood island formation (Silver and Palis, 1997). However, loss of function of *Gata-1* and *Gata-2* in mice results in severe haematopoietic but not vascular defects (Pevny *et al.*, 1991; Tsai *et al.*, 1994). Similarly, loss of *Tal1* has a severe impact on the haematopoietic system, which precedes its role in remodelling of the vasculature (Visvader, Fujiwara and Orkin, 1998; Souza *et al.*, 2005). Together, work on mouse embryonic development is supportive of a common haemangioblast progenitor, however this progenitor arises in the primitive streak region earlier in development than previously thought. These cells then migrate to the yolk sac and associate into blood islands before further development along their previously specified endothelial or haematopoietic pathway.

1.1.3 The yolk sac gives rise to both primitive and definitive haematopoietic lineages

Primitive erythrocytes emerging from the yolk sac were first distinguished from foetal liver-derived and adult erythrocytes based on their large nucleated form (Kovach *et al.*, 1967). Later gene expression analysis found that εy and βH1 haemoglobins (in mice) were associated with primitive erythrocytes of the yolk sac and silenced in definitive cell types (Leder *et al.*, 1992). These primitive erythrocytes provide the oxygen requirements of the embryo up until E12.5 when foetal liver-derived definitive erythrocytes rapidly expand in number. However, primitive erythrocytes can still be detected in the blood several days after birth (Kingsley *et al.*, 2004). As early as E7 it is also possible to detect primitive megakaryocytes and macrophages deriving from the yolk sac to

support the initial stages of embryonic development (Xu M *et al.*, 2001; Bertrand *et al.*, 2005; Tober *et al.*, 2007). These early megakaryocytes are important for the development of the vascular system. In fact lack of megakaryocyte-derived platelets in *Runx1* knockout mice is thought to result in the lethal haemorrhages at mid-gestation (Okuda *et al.*, 1996). Similarly, macrophages play an integral role in early development, clearing apoptotic cells and assisting in the morphogenesis of numerous tissues, including the emergence of HSPCs (Van Ham, Kokel and Peterson, 2012; Travnickova *et al.*, 2015). In this way, the initiation of haematopoiesis in the yolk sac is integral to the developmental progression of the embryo.

It is now also established that the yolk sac contributes to definitive haematopoiesis with the production of erythroid-myeloid progenitors (EMPs) starting at E8.5 (Palis *et al.*, 1999, 2001). These cells are thought to arise through an endothelial to haematopoietic transition in the yolk sac and go on to colonise the foetal liver prior to haematopoietic stem cell emergence (Lux *et al.*, 2008; Chen *et al.*, 2011). Lineage tracing studies have uncovered the considerable contribution of the yolk sac to adult haematopoiesis with primitive macrophages and EMPs generating the tissue-resident macrophages present in the adult organism (Ginhoux *et al.*, 2010; Gomez Perdiguero *et al.*, 2015). The yolk sac therefore, not only enables early development prior to the establishment of the HSC pool but also contributes to life-long haematopoietic functions.

1.1.4 Vasculogenesis and angiogenesis establishes a primitive circulatory system

Vasculogenesis is the formation of *de novo* blood vessels from the aggregation of angioblasts (Risau and Flamme, 1995). As with the first blood progenitors the first angioblasts derive from *Flk-1*+ cells migrating from the posterior primitive streak at E7 of mouse development (Huber *et al.*, 2004). In the absence of *Flk-1* the mesodermal progenitors do not localise appropriately in the yolk sac to form blood islands (Shalaby *et al.*, 1995, 1997). Similarly, loss of even one allele of the ligand *Vegf* or a secondary *Vegf* receptor known as *Flt-1* results in defects in the formation of the yolk sac vascular plexus at E8.5 (Fong *et al.*, 1995; Carmeliet *et al.*, 1996). Once the mesodermal cells have migrated to the periphery of the blood islands and adopted an endothelial phenotype the blood islands fuse to form a capillary plexus (Risau and Flamme, 1995). Two-photon imaging of zebrafish vessel formation has shown that lumenisation of vessels occurs through the fusion of intracellular vacuoles (Kamei *et al.*, 2006). The formation of the primary yolk sac vascular plexus is also dependent upon the TGF β-signalling pathway. Loss of function of either Tgf β 1 or Tgf β receptor II results in inadequate capillary tube formation leading to wasting and lethality in the embryo (Dickson *et al.*, 1995; Oshima, Oshima and Taketo, 1996).

Around the same time, cells from the lateral plate mesoderm coalesce along the midline of the embryo to give rise to the dorsal aorta (Coffin and Poole, 1988; Fouquet *et al.*, 1997). Again, these cells are directed in their migration by the chemo-attractant *Vegf* and concomitantly negatively regulated by BMP

antagonists released by the notochord (Carmeliet *et al.*, 1996; Cleaver and Krieg, 1998; Reese, Hall and Mikawa, 2004). The capillary network of the yolk sac then connects with the aorta of the embryo proper to form a primitive circulatory system in time for the first beat of the foetal heart (McGrath *et al.*, 2003). After the first stages of vasculogenesis, growth of the vascular network occurs through either the sprouting or splitting of pre-existing vessels, in a process termed angiogenesis (Risau, 1997). Studies in the early post-natal mouse retina have shown that the “tip” cells of vessels extend filopodia and are guided in their migration by a gradient of *Vegf-A*, which interacts with its receptor *Flk-1* (Gerhardt *et al.*, 2003). This ligand also acts upon “stalk” cells further down in the vessel to regulate their proliferation (Gerhardt *et al.*, 2003). In this way the vascular network is able to expand with the growing embryo, supplying the developing tissues with oxygen and nutrients.

At E8.25 the foetal heart begins to beat although fully functional circulation does not occur until E10 (Ji *et al.*, 2003; McGrath *et al.*, 2003). As the embryo grows the vascular plexus undergoes remodelling to form a hierarchical structure of large and small vessels. This process is dependent on the haemodynamic forces produced by the heartbeat and the increased viscosity caused by entry of erythrocytes into the bloodstream (Lucitti *et al.*, 2007). Mouse models that lack a functioning heart fail to remodel their vascular networks (Wakimoto *et al.*, 2000; Huang *et al.*, 2003). This enlargement of vessels exposed to high blood flow is achieved through fusion and targeted migration of endothelial cells (Udan, Vadakkan and Dickinson, 2013). Once

again there is a dynamic interplay between the endothelial and haematopoietic lineages during embryonic development.

1.1.5 Venous-arterial specification is both genetically pre-determined and plastic

While blood flow is important for the remodelling of the vascular plexus, the identification of arterial and venous specific receptors and ligands ignited the idea that vessel identity was genetically pre-determined. Veins were shown to specifically express *EphB4* and *Neuropilin-2* (*Np-2*) while arteries could be marked by *EphrinB2* and *Neuropilin-1* (*Np-1*) (Wang, Chen and Anderson, 1998; Herzog *et al.*, 2001). Interestingly, the knockout mouse models revealed more generalised defects in vascular development rather than one specific to either arteries or veins, indicating a connectedness of the vessels beyond their specification (Gerety *et al.*, 1999; Kawasaki *et al.*, 1999). The idea of a genetically pre-determined fate was supported however, by fate tracking experiments in the zebrafish showing that arterial or venous identity is specified in the lateral plate mesoderm prior to the migration of angioblasts to the midline (Zhong *et al.*, 2001). Similarly, in the yolk sac arterial and venous specific markers turn on prior to the initiation of blood flow (Herzog, Guttmann-Raviv and Neufeld, 2005). *Sox17* has been positioned at the apex of the arterial specification hierarchy, downstream of Vegf signalling (Kim *et al.*, 2016). Loss of *Sox17* is known to perturb arterial differentiation and is thought to act up-stream of the Notch pathway (Corada *et al.*, 2013; Chiang et al., 2017). Furthermore, loss of the notch receptors *Notch1* and *Notch4* or the

ligand *Delta-like 4* (*Dll4*) causes defects in remodelling of the embryonic vasculature (Krebs *et al.*, 2000, 2004; Duarte *et al.*, 2004).

Despite this strong evidence for early genetic specification of arteries and veins research still suggests a role for exogenous factors. Studies of mouse and chick yolk sac remodelling indicate a critical role for blood flow and haemodynamic forces in the formation of a mature vasculature (le Noble, 2003; Lucitti *et al.*, 2007). Indeed, after the initiation of blood flow, endothelial cells migrate to vessels with high flow in order to rapidly increase their diameter (Udan, Vadakkan and Dickinson, 2013). It is even possible to identify venous endothelial cells within the dorsal aorta of the early embryo suggesting that arterial-venous specification is to some degree malleable (Lindskog *et al.*, 2014). This plasticity in endothelial identity is further corroborated by studies of vascular grafts. The engraftment of quail endothelial cells into chick embryos found that while entire vessels did not adopt a new fate, individual endothelial cells could integrate into vessels and adopt arterial or venous identity depending on local cues (Moyon *et al.*, 2001; Othman-Hassan *et al.*, 2001). This evidence suggests there is a complex relationship between intrinsic and extrinsic factors in the maturation of the vasculature.

1.2 Haematopoietic stem and progenitor cells (HSPCs) arise from an endothelial precursor

1.2.1 HSCs emerge from the aorta gonad mesonephros (AGM) region at mid-gestation

Early studies on the origins of the haematopoietic lineage assumed that HSCs first arose with other early blood progenitors in the extra-embryonic yolk sac (Medvinsky *et al.,* 1993; Moore and Owen, 1967; Huang and Auerbach, 1993; Yoder *et al.*, 1997; Yoder, Hiatt and Mukherjee, 1997; Samokhvalov, Samokhvalova and Nishikawa, 2007). This idea was first challenged by studies in the avian model where chicken yolk sacs were grafted onto quail embryos prior to vascularisation (Dieterlen-Lievre, 1975). The definitive blood lineages analysed eight to ten days later were exclusively quail suggesting an intra-embryonic origin of HSCs (Dieterlen-Lievre, 1975). Almost two decades later the AGM was identified as a site of long-term, re-populating haematopoietic potential in the mouse model (Müller *et al.*, 1994). Over time and with careful transplantation studies in mice it has been shown, and is now accepted within the field that HSCs predominantly derive from the AGM region of the embryo proper (Medvinsky and Dzierzak, 1996; Cumano *et al.*, 2001). This potential was shown to exist during a precise window of development beginning at E10.5 and diminishing again by E12 (Müller *et al.*, 1994; Kumaravelu *et al.*, 2002). In further support of the AGM origin of HSCs, this process has been shown to be conserved across vertebrate evolution with studies in xenopus, zebrafish, chicken, mouse and human (Ciau-Uitz, Walmsley and Patient, 2000; Tavian *et al.*, 2001; Murayama *et al.*, 2006; Ivanovs *et al.*, 2011; Yvernogeau and Robin, 2017).

The haematopoietic cells of the AGM were first visualised in chick embryos as clusters of cells adhered to the endothelial wall; later these clusters were also identified in mice (Dieterlen-Lièvre and Martin, 1981; Garcia-Porrero, Godin and Dieterlen-Lièvre, 1995). Studies in xenopus, zebrafish, chick and human have found these intra-aortic clusters of haematopoietic cells to be localised to the ventral wall of the dorsal aorta (Tavian *et al.*, 1996; Ciau-Uitz, Walmsley and Patient, 2000; Wilkinson *et al.*, 2009; Yvernogeau and Robin, 2017). Conversely, studies in mice reveal clusters deriving from both the dorsal and ventral sides of the vessel (Taoudi and Medvinsky, 2007). However, development of HSPCs is thought to be enriched on the ventral side of the mouse AGM (Souilhol et al., 2016). Research has also revealed that other large arterial vessels such as the umbilical and vitelline arteries as well as the vasculature of the placenta can also act as sites of HSC generation (de Bruijn, 2000; Ottersbach and Dzierzak, 2005; Rhodes *et al.*, 2008; Zovein *et al.*, 2010). Detailed mapping of human and mouse embryos confirmed these observations and sought to quantify the overall number of cells emerging from the endothelium of the dorsal aorta. Serial sectioning of a five-week-old human embryo identified approximately 831 CD34+ cells attached to the endothelium (Tavian *et al.*, 1996). In mice the number of haematopoietic clusters peaked at E10.5 with 578 c-Kit+ cells identified (Yokomizo and Dzierzak, 2010). The process of HSC emergence in the embryo appears highly conserved across vertebrate species. Through numerous histological analyses it is clear that HSCs arise as clusters in the large arterial vessels of the embryo. However, these studies could not definitively determine whether

these intra-aortic clusters of haematopoietic cells had an endothelial or mesenchymal origin.

1.2.2 The endothelial to haematopoietic transition (EHT)

Development of an inducible vascular endothelial (VE-Cadherin) cre-recombinase mouse line enabled HSCs to be traced from the AGM to the bone marrow and strongly suggested an endothelial origin for this population (Zovein *et al.*, 2008). However, the idea that mesenchymal cells could migrate through the endothelial wall to the lumen of the vessel persisted. Time-lapse imaging of individual haemangioblast colonies, derived from the mESC culture system, provided the first direct evidence that HSCs emerged from an endothelial precursor, designated haemogenic endothelium (Lancrin *et al.*, 2009). This endothelial to haematopoietic transition was first visualised *in vivo* using confocal imaging of a transgenic zebrafish embryo (Bertrand *et al.*, 2010). This characterised the phenomenon as a process of trans-differentiation rather than cell division (Bertrand *et al.*, 2010; Kissa and Herbomel, 2010). Furthermore, development of an *ex vivo* culturing protocol for slices of murine AGM enabled researchers to observe the endothelial to haematopoietic transition live in a mammalian model for the first time (Boisset *et al.*, 2010). These studies lay the foundation for a new wave of haematopoietic research focussing on the haemogenic endothelium and the precise characterisation of HSC emergence.

1.2.3 Common markers of the endothelial and haematopoietic lineages

Lineage tracing and *in vivo* imaging studies have relied on key endothelial and haematopoietic markers for the characterisation of HSCs and their endothelial precursors. *VE-Cadherin* (also known as *Cdh5*) and *Tek* (also known as *Tie2*) have been extensively used to mark endothelial cells of the vasculature in conjunction with *Cd31*, a more generalised endothelial marker (Newman *et al.*, 1990; Dumont *et al.*, 1995; Breier *et al.*, 1996). Interestingly, these two common endothelial markers are also expressed intermittently on HSCs during development (North *et al.,* 2002; Baumann *et al.,* 2004; Kim, Yilmaz and Morrison, 2005). In mice and zebrafish *Cd41* is used as an early marker of haematopoietic fate while *Cd45* is present on most definitive haematopoietic cells (Thomas, 1989; Trochon *et al.*, 1996; Ferkowicz *et al.*, 2003; Mikkola *et al.*, 2003; Bertrand *et al.*, 2007). More recently, *Cd43* has been identified as the earliest marker of HSC identity in the mouse model (Rybtsov *et al.*, 2014). While *Cd41* and *Cd43* are favoured in murine studies of EHT, human experiments rely on *Cd34* and *Cd45* (Tavian, Hallais and Péault, 1999). *Cd34* has been shown in both human and mice to mark cells of the vasculature and haematopoietic clusters (Young, Baumhueter and Lasky, 1995). Similarly *c-Kit* has been extensively used due to its strong association with HSCs and the intra-aortic clusters in mice (Ikuta and Weissman, 1992; Yokomizo and Dzierzak, 2010). In conjunction with these cell surface receptors, the transcription factors *Sca1*, *Runx1* and *Tal1* have been linked to the acquisition of HSC identity (Mukouyama *et al.*, 2000; Ma *et al.*, 2002; D'Souza *et al.*, 2005). Numerous research groups have used a combination of these endothelial and haematopoietic cell surface markers in an effort to

define and isolate haemogenic endothelial cells. Others have generated reporter models using the regulatory elements of transcription factors expressed early in the adoption of the haematopoietic fate. Still the identity of haemogenic endothelium remains controversial. No cell surface marker has been identified to specifically isolate this rare cell population and little is known about its transcriptional profile.

1.2.4 Defining haemogenic endothelium

In the broadest terms, haemogenic endothelium is a subset of VE-Cadherin+ endothelial cells of the vasculature that have the potential to undergo EHT. This was first demonstrated when functional endothelial cells, capable of acetylated LDL uptake, were isolated based on VE-Cadherin+/Ter119-/CD45- markers from E9.5 embryos, and found to generate B and T lymphocytes in culture (Nishikawa *et al.*, 1998). Taking it one step further, endothelial cells of the avian vascular system were inoculated with acetylated LDL and later intra-aortic clusters were found to also be labelled (Jaffredo *et al.,* 1998). However, refining this definition and identifying a specific marker for haemogenic endothelium has proved difficult. Initial studies focused on endothelial transcription factors. Using a GFP reporter, *Etv2* was shown to be expressed in both haemangioblasts and haemogenic endothelial cells (Wareing *et al.*, 2012). Over-expression of *Etv2* in hPSCs achieved a haemato-endothelial phenotype and enabled significant haematopoietic expansion in culture (Elcheva *et al.*, 2014). Likewise, *Sox7* and *Sox17* were proposed as potential markers of haemogenic endothelium (Costa *et al.*, 2012; Clarke *et al.*, 2013). Ultimately however, these candidates lacked specificity with their expression

and function affecting other elements of the vasculature or mesodermal progenitors.

More recently researchers have sought to use the regulatory elements of transcription factors that drive haematopoietic identity, namely *Runx1, Sca-1* and the downstream target of *Runx1, Gfi1* to isolate haemogenic endothelium. Early on in the study of EHT a *Sca-1* (*Ly-6a*) GFP reporter mouse was found to be a useful tool in enriching for HSCs in the mouse dorsal aorta (de Bruijn *et al.,* 2002). Interestingly, it was not possible to enrich for HSCs using endogenous staining of *Sca-1* and less than half of the GFP+ reporter derived cells expressed the well-established HSC marker, *c-Kit* (de Bruijn *et al.,* 2002). Sub-fractionation of endothelial cells, haemogenic endothelial cells and HSCs from this reporter mouse line, and subsequent RNAseq analysis identified *Gpr56* as a cell surface marker of EHT. However, as a target gene of key haematopoietic regulators, it is probably a more suitable marker of early, transitioning HSCs (Solaimani Kartalaei *et al.,* 2014).

Two groups developed GFP reporter models using haematopoietic specific enhancers of *Runx1* and showed that they could isolate endothelial cells with haematopoietic capacity (Ng *et al.*, 2010; Swiers *et al.*, 2013). However, the endothelial cells of the *Runx1*+23GFP reporter, isolated at E10.5 showed reduced tubule formation capacity. Small-scale transcriptional analysis revealed a heterogeneous population which had already very much committed to the haematopoietic pathway and down-regulated endothelial gene expression (Swiers *et al.*, 2013). The *Runx1* target *Gfi1* was proposed

as a marker of haemogenic endothelium based on transcriptional data linking *Gfi1*+ cells to endothelial gene expression (Thambyrajah *et al.*, 2015). However, further single cell analysis of haemogenic endothelium and non-haemogenic endothelium as defined by *Gfi1* identified very few transcriptional differences between the two populations (Baron *et al.*, 2018). It's possible that by targeting transcriptional regulators of haematopoiesis, the cell populations isolated are already too late in their transition and haematopoietic commitment has already been established. Indeed research has indicated that *Runx1* is required for the maturation of CD41+ cells and not for their formation (Liakhovitskaia *et al.*, 2014).

1.2.5 HSC development occurs in a step-wise fashion

Despite the number of cells contributing to intra-aortic clusters efforts to quantify the number of HSCs emerging from the AGM using transplantation into irradiated mice found surprisingly few definitive HSCs (dHSCs). Only one HSC per AGM at E11 and a peak of three HSCs by E12 was detected (Kumaravelu *et al.*, 2002). Conversely, the foetal liver dramatically expanded its HSC potential from one at E11 to approximately 50 by E12 (Ema and Nakauchi, 2000; Kumaravelu *et al.*, 2002). The rapid expansion in the HSC population was hypothesised to result from differentiation, given the average time of the mammalian cell cycle. This idea was supported by the development of an *in vitro* co-aggregation system that enabled the maturation of the precursors of HSCs (termed pre-HSCs) in culture; dramatically increasing the measurable HSC potential of the AGM (Taoudi *et al.*, 2008). From these *ex vivo* studies, utilising re-aggregation and co-culturing

techniques with haematopoietic promoting OP9 stromal cells, the pre-HSC populations were described as part of the HSC hierarchy. The expression of the early haematopoietic marker CD41 and the definitive haematopoietic marker CD45 were used to distinguish pre-HSC type I cells (VE-Cad+/CD41+/CD45-) from pre-HSC type II cells (VE-Cad+/CD45+) which displayed differing maturation times in culture (Rybtsov *et al.*, 2011). These experiments re-defined HSC development as a multi-step process (Fig. 1). HSC maturation was further delineated by the use of *Cd43* which enabled the isolation of an even more primitive HSC precursor denoted the pro-HSC population (Rybtsov et al., 2014). These VE-Cad+/CD41+/CD43- cells take longer culturing time to mature into dHSCs but still lack endothelial potential (Rybtsov *et al.*, 2014). Extensive limiting dilution analysis of this pre-HSC pool quantified approximately 50 cells per embryo by the end of embryonic day ten (39 somite pairs). The pre-HSC pool was found to reach a peak of 65 cells during E11 before a rapid drop in potential as foetal liver haematopoiesis commences (Rybtsov *et al.*, 2016). Previous methods of assessing the HSC potential of the AGM have failed to quantify the proliferative capacity of the pre-HSCs that seed the foetal liver. The development of the HSC pool in the embryo is a progressive process with several intermediary steps, the power of which ultimately lies in the dorsal aorta and other large vessels from which the intra-aortic clusters of pre-HSPCs emerge (Fig. 1). The next obvious question is once haemogenic endothelium is specified what promotes the induction of EHT?

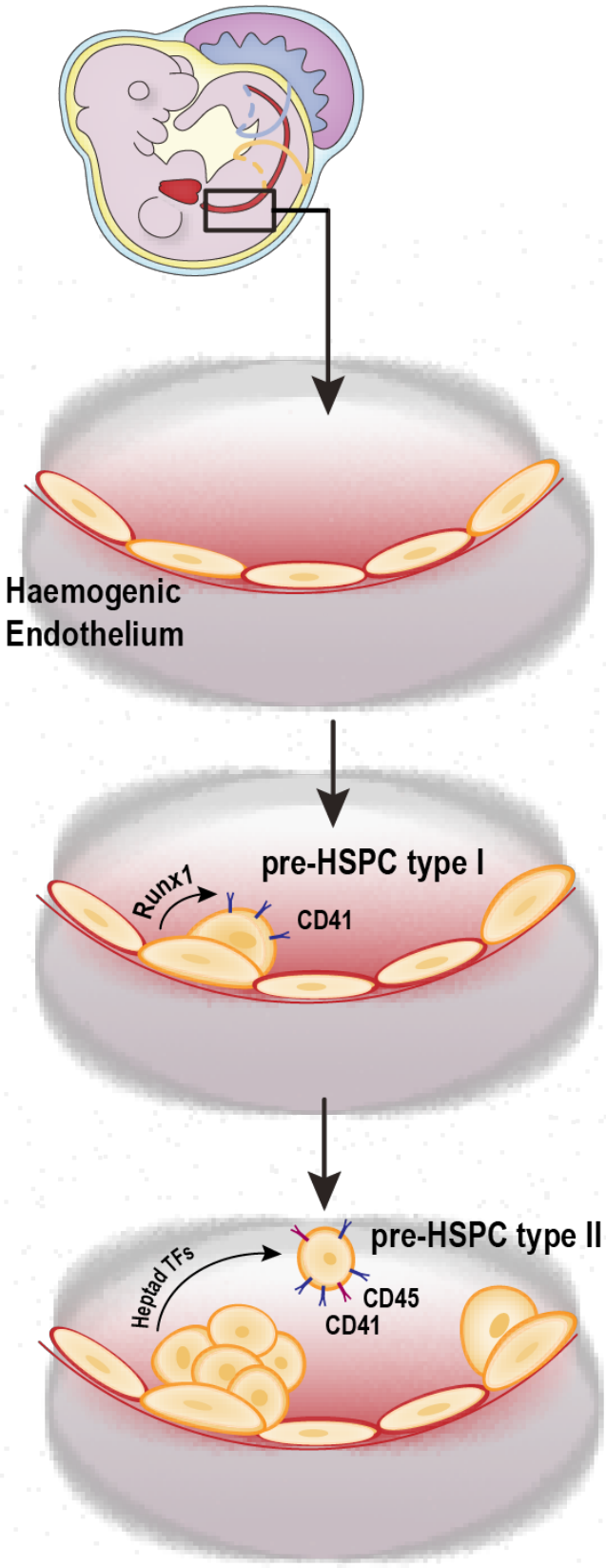


Figure 1: Model of endothelial to haematopoietic transition in the murine AGM

Endothelial to haematopoietic transition occurs in the AGM. Haemogenic endothelium differentiates into CD41 expressing pre-HSPC type I cells with the help of the transcription factor Runx1. Pre-HSPC type I cells then mature into pre-HSPC type II cells with the help of a heptad of transcription factors (TFs).

1.2.6 Transcriptional control of HSC development

Runx1 is well established as the key transcriptional driver of haematopoietic stem and progenitor cell development. *Runx1* was originally associated with haematopoiesis due to its frequent translocations in acute myeloid leukaemia (Miyoshi *et al.*, 1991). When knockout mice were generated, they observed embryonic lethality at E12.5 due to cerebral haemorrhaging and a complete loss of foetal liver haematopoiesis (Okuda *et al.*, 1996; Wang *et al.,* 1996). Further investigation found that loss of *Runx1* resulted in the complete absence of intra-aortic haematopoietic clusters in the embryo and an inability of endothelial cells from the embryo and yolk sac to form definitive blood cells *ex vivo* (Yokomizo *et al.*, 2001). Thus, *Runx1* is absolutely necessary for the formation of both HSCs in the AGM and definitive EMPs from the yolk sac (Yokomizo *et al.*, 2001). Conditional loss of *Runx1* in endothelial cells showed that the transcription factor is necessary for EHT but is not required once haematopoietic identity is established (Chen *et al.*, 2009; Tober *et al.*, 2013).

A key function of *Runx1* is the down-regulation of endothelial identity through the direct repression of *Flk-1* and the up-regulation of the transcriptional repressors *Gfi1* and *Gfi1b* (Hirai *et al.*, 2005; Lancrin *et al.*, 2012). *Runx1* is also known to target both the transcription factor *Spi1* which is involved in the development of all haematopoietic lineages and the cytokine *IL-3* which acts as a survival and proliferation factor for HSCs (Uchida, Zhang and Nimer, 1997; Robin *et al.*, 2006; Huang *et al.*, 2008). In addition to the direct targeting of down-stream effectors, *Runx1* has been implicated in the modification of the chromatin landscape. Studies have shown *Runx1* to function in the

unfolding of chromatin around target promoters, resulting in an increase in local histone acetylation and the recruitment of methyltransferases to the chromatin structure in the early stages of haematopoietic specification (Hoogenkamp *et al.*, 2009; Lichtinger *et al.*, 2012; Herglotz *et al.*, 2013).

Furthermore, discovery of an essential role for the *Runx1* trans-activation domain in the establishment of haematopoietic identity indicated a role for transcriptional complexes in HSC formation (Dowdy *et al.*, 2010). Numerous other transcription factors have been implicated in HSC development based on disease associations and haematopoietic defects observed in knockout mouse models. Both *Tal1* and *Lmo2* were factors initially linked to T-cell leukaemia and later implicated in the establishment of the haematopoietic system (Shivdasanl, Mayer and Orkin, 1995; Yamada *et al.*, 1998). Similarly, loss of *Gata1* or *Gata2* was found to severely impact on erythropoiesis (Pevny *et al.*, 1991; Tsai *et al.*, 1994). Interestingly, these haematopoietic regulators have also been identified as binding partners. *Runx1* and *Gata1* were found to functionally and physically interact during megakaryocyte development (Elagib *et al.*, 2003). Likewise, *Tal1* and *Lmo2* are known to form a complex in the haemangioblast (Patterson *et al.*, 2007).

In an effort to investigate this combinatorial control of haematopoietic differentiation a genome-wide ChIP-seq analysis was performed on ten transcription factors in a haematopoietic cell line. This study identified 927 targets bound by a combination of seven factors (Wilson *et al.*, 2010). This led to the proposal of a heptad complex of transcription factors responsible for the

adoption of haematopoietic fate. Through single cell transcriptome analysis the simultaneous expression of this heptad of genes has been linked with pre-HSC identity (Bergiers *et al.*, 2018). Further investigation showed that enforced expression of this heptad locked cells into a pre-HSC state (Bergiers *et al.*, 2018). However, analysis of the resulting gene regulatory network led to a different model of haematopoietic identity acquisition, whereby the haematopoietic factors *Runx1* and *Gata2* functioned in opposition to *Fli1* and *Erg* to specify lineage (Bergiers *et al.*, 2018). Although numerous transcription factors are known to be important in early haematopoietic development, the way in which they interact with each other and respond to cell extrinsic factors is yet to be fully understood.

1.2.7 Hedgehog and Bmp/Tgfβ signalling patterns the dorsal aorta in preparation for EHT

While gene regulatory networks are instrumental in the acquisition of HSPC identity, cell extrinsic factors are still essential to HSPC development and proliferation, as demonstrated by the precise location and timing of their emergence. Numerous signalling pathways including Hedgehog, Bmp/Tgf-β, Wnt and Notch are known to converge and interact during EHT (Fig. 2). As previously discussed, the intra-aortic clusters are often restricted to the ventral side of the dorsal aorta. This patterning of haematopoietic potential is thought to be achieved through a dynamic interplay of these signalling pathways.

The Hedgehog ligand, Sonic hedgehog (*Shh*) is released from the notochord helping to pattern the dorsal aorta (Wilkinson et al., 2009). Although increased

activation of hedgehog signalling with purmorphamine did not impact on HSPC development in zebrafish, depletion of Shh did result in loss of definitive haematopoiesis (Gering & Patient, 2005; Wilkinson et al., 2009). Addition of exogenous Shh to mouse co-aggregation cultures instead increased HSPC output (Peeters et al., 2009; Souilhol et al., 2016). Further studies indicated that over-expression of the Notch intracellular domain could be used to rescue haematopoietic defects caused by cyclopamine-induced inhibition of Hedgehog signalling, positioning the Hedgehog pathway upstream of Notch signalling (Kim et al., 2013). Haematopoietic output could also be rescued by the transcription factor *Tal1* (Kim et al., 2013).

In contrast to *Shh*, the *Bmp4* ligand shows spatial restriction to the ventral subaortic region (Marshall, Kinnon and Thrasher, 2007; Pimanda et al., 2007; Durand *et al.,* 2007). Further investigation found *Bmp4* to be essential for the specification of HSPC precursors, however must be down-regulated for the maturation of these cells (Souilhol et al., 2016). This produces a complex interplay between the expression of Bmp target genes and inhibitors. Creation of a Bmp-responsive GFP mouse line demonstrated that only Bmp activated cells from the intra-aortic clusters were capable of reconstitution of the haematopoietic system after transplantation into an irradiated recipient (Crisan et al., 2016). Furthermore, the Bmp antagonists *Bmper*, *Noggin*, *Smad6* and *Smad7* were found to be specifically expressed in the ventral region of the dorsal aorta helping to modulate Bmp signalling and enabling the maturation of HSPCs (Pimanda et al., 2007; Souilhol et al., 2015; McGarvey et al., 2017). Given the role of Bmp signalling in the bone marrow to control

niche size (Zhang et al., 2003), it is possible that *Bmp4* and its inhibitors function in the embryo to regulate the emergence of haematopoietic cells. Although HSPCs preferentially emerge from the ventral wall, maximum HSPC production is achieved with cultures containing both dorsal and ventral regions (Souihol et al., 2016). This supports a model where Hedgehog and Bmp signals combine for the establishment of the definitive haematopoietic system (Fig. 2).

1.2.8 Notch acts down-stream of other signalling pathways but up-stream of *Runx1* expression

While Notch signalling has been previously implicated in the differentiation of arteries and veins, there is now also strong evidence for the involvement of Notch in haematopoietic development (Lawson et al., 2001; Krebs et al., 2000). Loss of the *Notch1* receptor, the notch ligand, *Jagged1* and the down-stream Notch effector *RBP-j*κ all result in embryonic lethality due to severe defects in definitive haematopoiesis (Kumano et al., 2003; Robert-Moreno et al., 2004; Robert-Moreno et al., 2008). Similarly, the double knockout of the Notch targets *Hes1* and *Hes5*, which act as transcriptional repressors, results in loss of HSC activity (Guiu et al., 2013). Work in zebrafish haematopoiesis has shown that transient expression of the Notch intracellular domain, using the Gal4/UAS system, can ectopically expand the HSPC clusters to the roof of the dorsal aorta and the neighbouring cardinal vein (Burns et al., 2005). This effect could be abolished by the addition of a *Runx1* morpholino, positioning Notch signalling upstream of the major transcriptional driver of haematopoietic fate (Burns et al., 2005). In support of this need for transient Notch signalling,

enforced expression of Notch through co-culture with Delta-like 4 (*Dll4*) expressing OP9 cells inhibited the progression of immature HSCs (Souilhol et al., 2016). This dependence on Notch signalling strength can be mediated through the differing binding capacities of *Dll4* and *Jagged1* ligands. Generation of low and high sensitivity notch reporters found that the haematopoietic lineages derived from cells with a history of low Notch signal compared to arterial cells, which were previously exposed to high levels of notch (Gama-Norton et al., 2015). Furthermore, treatment of AGM explant cultures with a *Dll4* blocking antibody increased haematopoietic development (Gama-Norton et al., 2015). Once again, it is not only the signal but the strength of that signal that is important for cell fate decisions.

While the Wnt pathway is also thought to be involved in embryonic haematopoiesis, its exact role in the jigsaw is still unclear. The down-stream effector of the canonical Wnt pathway, β-*catenin* is known to be important, as the stabilisation of this protein was found to increase haematopoietic output in AGM explants (Ruiz-Herguido et al., 2012). Knockdown and rescue experiments have shown that *Wnt16a* is important in zebrafish haematopoiesis through a non-canonical pathway and is thought to act in parallel with *Shh* to up-regulate Notch genes in haemogenic endothelium (Clements et al., 2011). Much still needs to be explored however into the role of Wnt signalling during EHT. Altogether, it is the dosage and intersection of several signalling cascades that enables EHT to occur in the AGM at the precise time and spatial location for successful development of the blood system (Fig. 2).

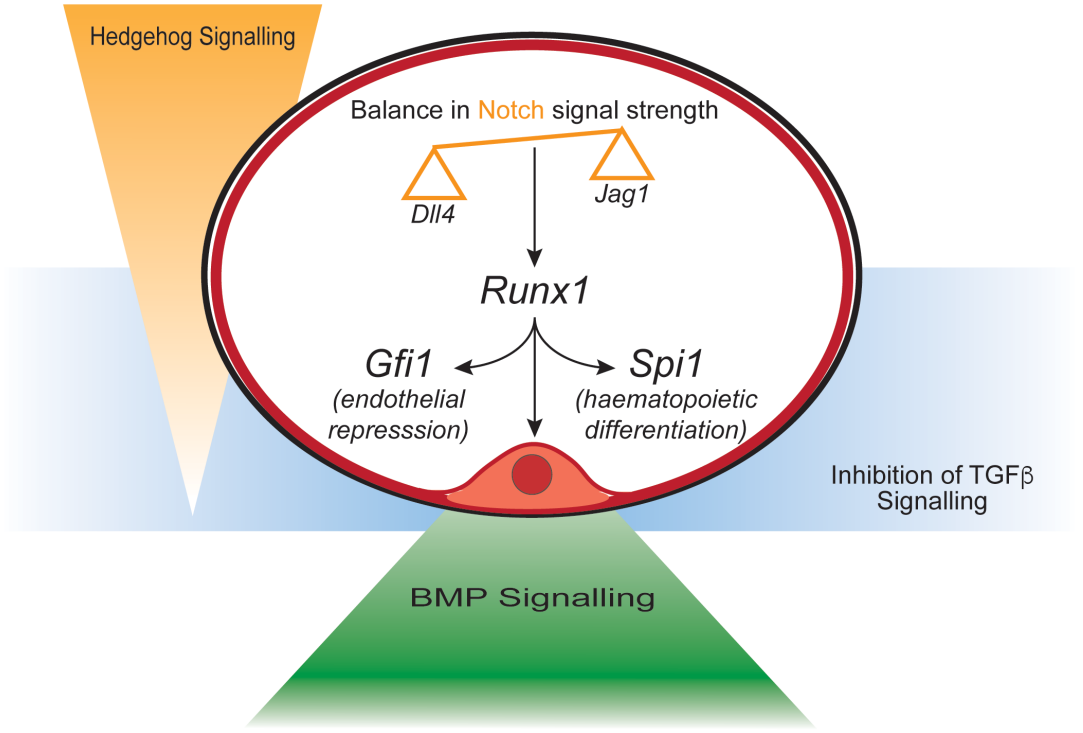


Figure 2: Convergence of signalling pathways during EHT

Scheme displaying the key signalling pathways involved in the induction of EHT in the dorsal aorta. Bmp, hedgehog and inhibition of TGFβ signalling are known to be important for the emergence of HSCs. Transient or low Notch signal is necessary for the up-regulation of the key haematopoietic transcription factor Runx1 which drives the expression of targets responsible for the down-regulation of endothelial genes and the up-regulation of haematopoietic genes.

1.2.9 Haematopoietic stem cells migrate to the foetal liver and bone marrow

Once the yolk sac and AGM regions have established the definitive progenitor populations, these cells then migrate to seed the foetal liver (Fig. 3) (Johnson and Moore, 1975). Although little research has been done on the migratory path of haematopoietic cells to the foetal liver, it is thought that adhesion receptors, selectins and integrins are involved. Single cell analysis of HSCs during foetal liver migration found high expression of several integrin genes (Ciriza *et al.*, 2012). More specifically, the loss of $\beta 1$ integrin leads to a

complete inability of haematopoietic cells to seed both the foetal liver and bone marrow (Hirsch *et al.*, 1996; Potocnik, Brakebusch and Fa, 2000). Furthermore, impairment in the haematopoietic lineage of the Rho GTPase protein *Rac1*, that acts downstream of several cell surface receptors, resulted in haematopoietic progenitors with impaired migratory capabilities (Ghiaur *et al.*, 2008). Once the pre-HSCs have seeded the foetal liver at E12 they undergo rapid expansion, with the tissue increasing its transplantation potential ten-fold by E14 (Ema and Nakauchi, 2000). The proliferation and survival of HSCs is strongly influenced by the microenvironment. Populations of hepatic progenitors and pericytes have been isolated from the foetal liver and shown to support HSC proliferation *ex vivo* (Chou and Lodish, 2010; Khan *et al.*, 2016). This function is in part due to the cytokines they express, which include SCF, IGF2 and Angiopoietin-like-2 or Angiopoietin-like-3 (Chou and Lodish, 2010; Khan *et al.*, 2016).

Beginning at E15 the embryonic HSCs migrate again to the spleen and bone marrow, although the HSCs cannot be detected in the bone marrow until E17.5 (Fig. 3) (Christensen *et al.*, 2004). This seeding occurs in a gradual way, as analysis of foetal blood between E12.5 and E17.5 found a consistent, low level of haematopoietic stem cell potential, indicating a continual release of HSCs into the bloodstream (Christensen *et al.*, 2004). The migration of haematopoietic progenitors to the bone marrow is driven by the chemo-attractant *Cxcl12* (also known as *Sdf-1*) and its receptor *Cxcr4* (Aiuti *et al.*, 1997; Ara *et al.*, 2003). This signalling axis not only regulates the initial

colonisation of the bone marrow but is also important for the maintenance of the stem cell pool (Sugiyama *et al.*, 2006).

Current estimates of the adult haematopoietic stem cell pool suggest that there are approximately 17 000 long term re-populating HSCs residing in the mouse bone marrow (Busch *et al.*, 2015). These cells possess enormous self-renewal and proliferation potential. Early transplantation studies demonstrated that a single HSC is capable of reconstituting the entire haematopoietic system of an irradiated mouse (Osawa *et al.*, 1996). Through *in vivo* barcoding of *Tie2*+ HSC progenitors in the embryo it could be estimated how much one HSC could contribute to the adult stem cell population (Pei *et al.*, 2017). This innovative technique found that a single embryonic HSC could give rise to several hundred HSC clones in later life (Pei *et al.*, 2017). This estimate was confirmed in humans with one embryonic HSC transplanted into irradiated, immune-deficient mice giving rise to at least 300 daughter HSCs (Ivanovs *et al.*, 2011). Given these intrinsic properties and the utility of both HSCs and differentiated blood cells to clinical treatments there has been a strong drive for biomedical researchers to develop methods for both expanding endogenous populations of HSCs *ex vivo* and producing *de novo* HSCs *in vitro*. This applicability of HSCs to regenerative medicine has in part also driven the research into HSC ontogeny.

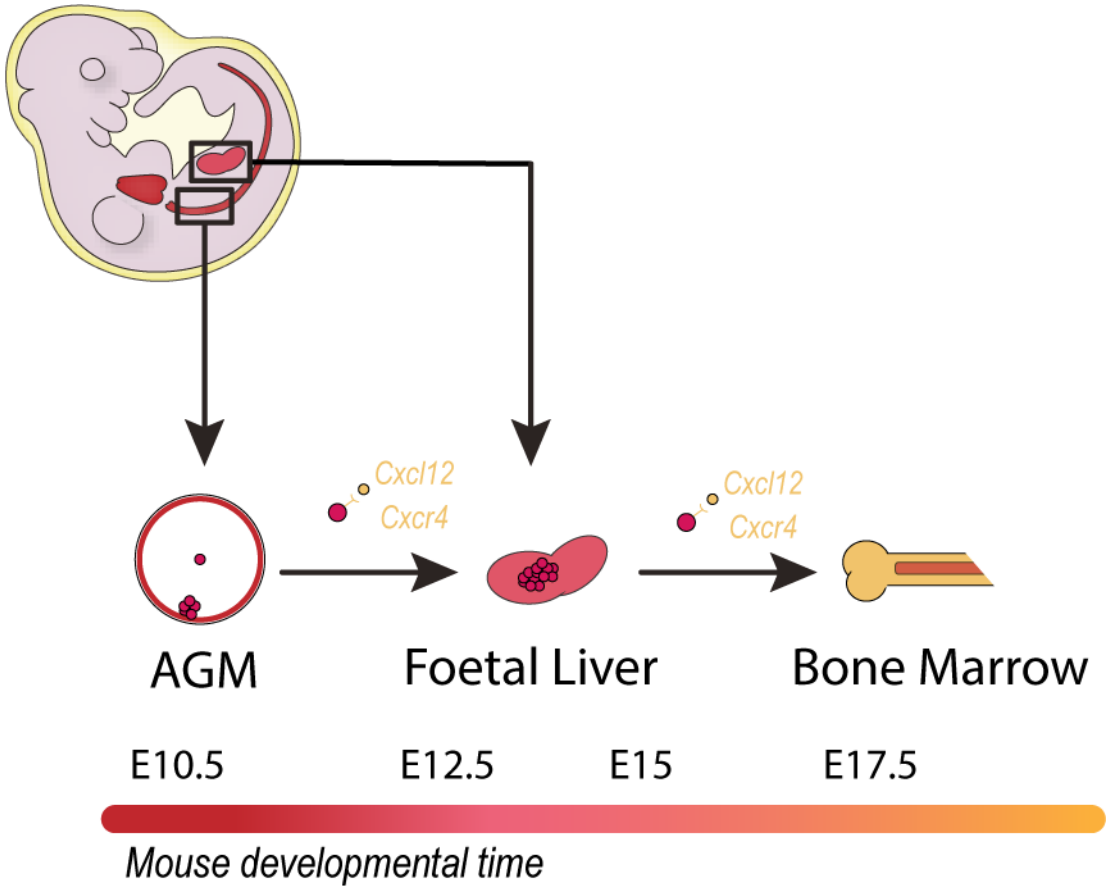


Figure 3: HSC migration in the mouse embryo

Timeline of HSC emergence and migration in the murine embryo. HSCs emerge in the AGM region and migrate first to the foetal liver and then to the bone marrow with the assistance of the cell surface receptor *Cxcr4* and its ligand *Cxcl12*.

1.3 ESC differentiation towards blood

1.3.1 The ESC model of haematopoietic differentiation

The establishment of the first mouse embryonic stem cell (mESC) lines in the early 1980s opened up enormous opportunity for the development of new model systems of mammalian development (Evans and Kaufman, 1981; Martin, 1981). Further investigation identified leukaemia inhibitory factor (LIF) to be the key cytokine maintaining embryonic stem cell (ESC) pluripotency through the activation of *Stat3* (Smith *et al.*, 1988; Williams *et al.*, 1988; Matsuda *et al.*, 1999). More recently the LIF-Stat3 axis has been shown to preserve pluripotency through the transcriptional regulation of two of the Yamanaka factors, c-*Myc* and *Klf4* (Cartwright, 2005; Hall *et al.*, 2009). Removal of feeder cells and LIF from the ESC culture system leads to the formation of cell aggregates of endoderm, mesoderm and ectoderm, known as embryoid bodies (EBs) (Doetschman *et al.*, 1985). Further culturing of the EBs was found to readily generate multiple cell types including the haematopoietic lineage (Wiles and Keller, 1991; Keller *et al.*, 1993).

As in the embryo, Flk-1 was identified in the ESC culture system as a tool to isolate mesodermal progenitors with haematopoietic capacity along with endothelial and vascular smooth muscle (VSM) cell potential (Fig. 4) (Choi *et al.*, 1998; Ema *et al.*, 2003). Continued refinement of the EB culturing conditions and cell surface marker analysis of Flk-1+ (haemangioblast) progeny established the ESC system as a close model of embryonic haematopoiesis (Keller, 1995; Kabrun *et al.*, 1997). Through further cytokine

exposure and culturing it is possible to produce most differentiated blood cell types from ESC-derived progenitors, including, erythrocytes, megakaryocytes and platelets, neutrophils, macrophages, as well as B and T lymphocytes (Nakano, Kodama and Honjo, 1994; Eto *et al.*, 2003; Fujimoto *et al.*, 2003; Pooter *et al.*, 2003; Carotta *et al.*, 2004; Lieber *et al.*, 2004). Despite this success the isolation of HSCs from this system and the transplantation of ESC-derived cells with multi-potent potential has proven difficult (Müller and Dzierzak, 1993; Hole *et al.*, 1996). Although it has not been possible to induce ESCs into functional HSCs the in *vitro* differentiation model has provided invaluable insights into the specification and development of the haematopoietic system.

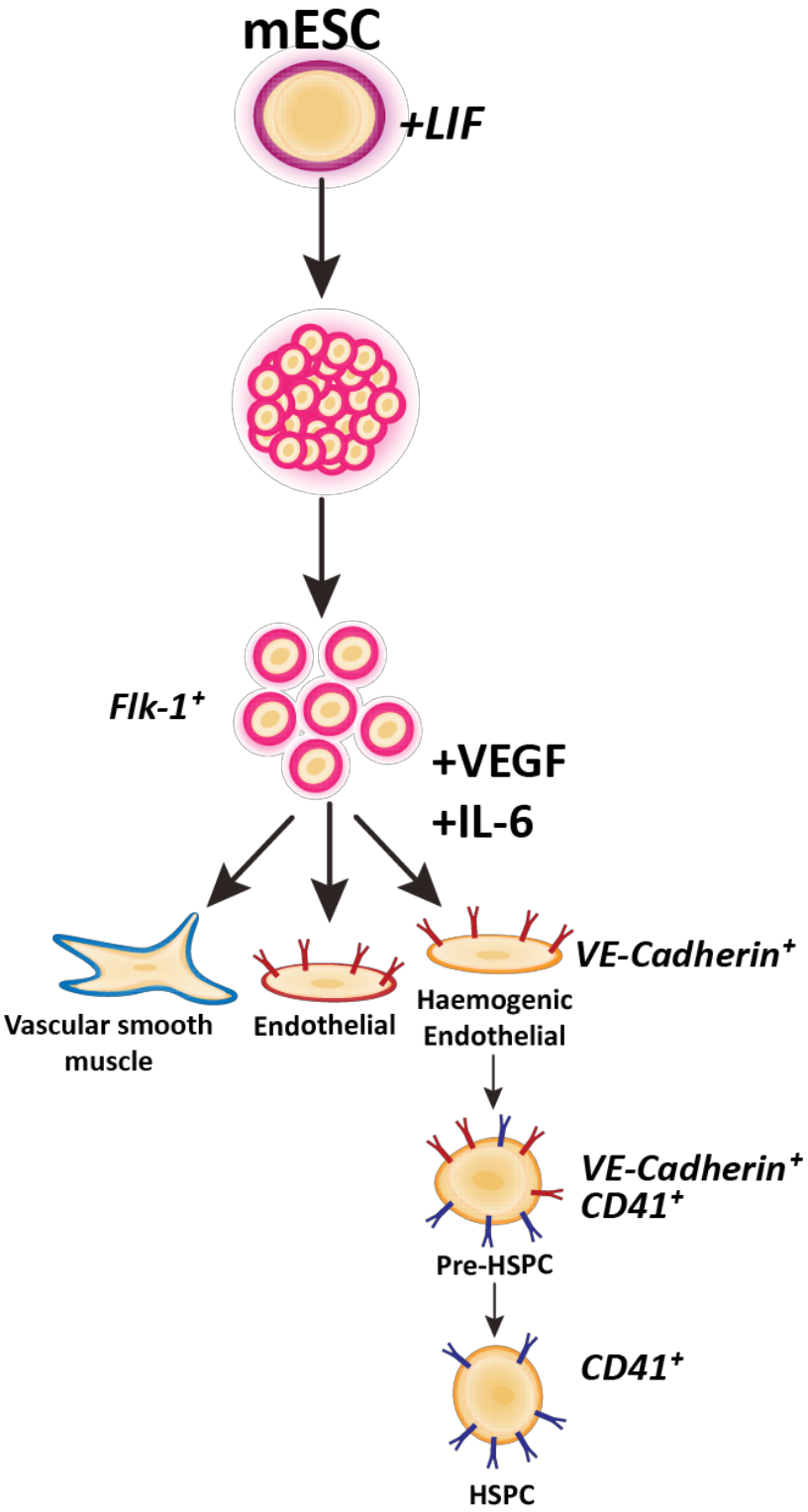


Figure 4: In vitro model of haematopoietic development

The pluripotency of mouse embryonic stem cells can be maintained with LIF. Removal of LIF results in the differentiation of ESCs and clusters of cells called embryoid bodies form. Flk1+ mesoderm can be isolated from embryoid bodies and can differentiate in the presence of VEGF and IL-6 into vascular smooth muscle cells, endothelium and haematopoietic cells.

1.3.2 Contribution of the ESC model to haematopoietic research

The scalability of the ESC differentiation system to blood and our ability to directly observe and manipulate the culture has made this a powerful model for understanding haematopoietic development. Use of the mESC system has provided access to a developmental time point not easily accessible in the embryo, resulting in the first direct observation of EHT (Lancrin *et al.*, 2009). The ability to manipulate specific aspects of the culturing conditions has enabled us to understand the importance of exogenous factors to HSPC development including shear stress and the impact of specific cytokines such as *Bmp4* and *Vegf* (Park *et al.*, 2004; Nostro *et al.*, 2008; Adamo *et al.*, 2009). It has often been the case that an improved understanding in the *in vitro* model of haematopoiesis has spurred on research *in vivo*. Isolation of the Flk-1+ haemangioblast progenitors was first discovered in vitro and later confirmed in mouse embryos (Choi *et al.*, 1998; Huber *et al.*, 2004). As such the mESC system has proved an invaluable tool enabling the field to test ideas not yet possible in the embryo and drive research of haematopoietic development forward.

1.4 Production of HSCs *in vitro*

1.4.1 Expansion of HSCs *ex vivo*

The ability to expand HSCs in culture would be hugely beneficial for both research and regenerative medicine. However, HSCs *in vitro* show a strong preference towards differentiation over self-renewal. Nevertheless researchers have been able to demonstrate self-renewal capacity *in vitro* through the transplantation of HSCs derived from single cells or short-term

cultures (Glimm and Eaves, 1999; Ema *et al.*, 2000). Although several cytokines have been identified that aid in the proliferation of HSCs including SCF, Flt-3, IL-3, IL-6, IL-11 and G-CSF, incubation with a combination of these factors produced only modest two- and three-fold increases in HSC populations (Conneally *et al.*, 1997; Miller and Eaves, 1997). Analysis of the foetal liver microenvironment, which promotes HSC expansion, helped to improve culturing conditions. By combining SCF, TPO and FGF-1 with the foetal liver expressed IGF-2, Angptl2 and Angptl3, HSC numbers could be increased up to 24-fold (Zhang and Lodish, 2005; Zhang *et al.*, 2006). Wnt signalling was also found to play an important role in HSC self-renewal as incubation of isolated HSCs with the Wnt3a ligand or transduction with constitutively active β-catenin drastically expanded the HSC population (Reya *et al.*, 2003; Willert *et al.*, 2003). However, the gains achieved through signal induction are much smaller in comparison to the self-renewal capacity of transcription factor transduced HSCs. Exogenous expression of Hoxb4 can expand the HSC population of a ten-day *ex vivo* culture by 40-fold (Antonchuk, Sauvageau and Humphries, 2002). However, clinical issues with the use of transcription factor induction continue to spur research into other methods of stimulating HSC expansion *ex vivo*. Research into HSC metabolism found that enforced glycolysis through the blocking of a mitochondrial enzyme could help to maintain HSC potency and led to a modest three-fold increase in HSCs over a five-day culture period (Liu *et al.*, 2015). Although multiple methods have been identified that can help in the expansion of HSCs *ex vivo* the gains remain small, limiting the clinical use of these cells.

1.4.2 Transcription factor reprogramming generates functional HSCs *in vitro*

With the inability to generate HSCs from ESCs and the difficulty obtaining sufficient numbers of HSCs *ex vivo* many researchers have focussed on the idea of reprogramming either induced pluripotent stem cells (iPSCs) or else adult fibroblasts and endothelial cells. The observation during iPSC generation that a portion of cells with high *Oct4* expression also express the pan-haematopoietic marker *Cd45* led researchers to induce *Oct4* expression in human fibroblasts (Szabo *et al.*, 2010). These *Oct4*-expressing fibroblasts had a similar transcriptional profile to haematopoietic progenitors and when exposed to specific cytokines were able to develop into functional myeloid, erythroid and megakaryocyte lineages (Szabo *et al.*, 2010). Soon researchers were using transcription factor cocktails in order to reshape the gene regulatory network into a haematopoietic configuration. It was found that multi-potent haematopoietic progenitors could be specified from CD34+CD45+ human iPSCs with the induction of *Hoxa9*, *Erg*, *Rora*, *Sox4* and *Myb* expression (Doulatov *et al.*, 2013). Similarly, adult mouse fibroblasts exposed to the ectopic expression of five of the heptad of haematopoietic transcriptional regulators (*Erg*, *Lmo2*, *Gata2*, *Runx1c* and *Tal1*) were also induced into multi-potent progenitors (MPPs) (Batta *et al.*, 2014). Interestingly, transcriptional analysis of this transformation found that the fibroblasts transitioned through a haemogenic endothelial stage in the process (Batta *et al.*, 2014). The recognition that definitive haematopoietic lineages arise through an endothelial intermediate spurred research into understanding the role of endothelium in HSC production. Use of fibroblasts from a GFP reporter

mouse line, under the control of *Cd34* regulatory elements enabled the screening of transcription factors that promoted both endothelial and haematopoietic identity (Pereira *et al.*, 2013). Activation of four transcription factors (*Gata2*, *cFos*, *Gfi1b* and *Etv6*) was found to specify an endothelial population with haematopoietic potential (Pereira *et al.*, 2013). Another research team focussed on developing a co-culturing system with modified primary endothelial cells that have constitutively active Akt signalling and support HSC self-renewal (Kobayashi *et al.*, 2010). This vascular induction combined with transcription factor reprogramming (using *Foxb*, *Runx1*, *Gfi1* and *Spi1*) enabled the transformation of human umbilical vein endothelial cells into functional MPPs (Sandler *et al.*, 2014). By following on from these approaches and taking into account the critical role endothelial cells play in HSC emergence researchers were able for the first time through transcriptional reprogramming to produce fully functional HSCs *in vitro* (Lis *et al.*, 2017; Sugimura *et al.*, 2017). HSC emergence from iPSCs was achieved by first isolating CD34+ haemogenic endothelial cells before over-expressing seven transcription factors (*Erg*, *Hoxa5*, *Hoxa9*, *Hoxa10*, *Lcor*, *Runx1* and *Spi1*) which enabled the primary and secondary reconstitution of mouse haematopoietic systems (Sugimura *et al.*, 2017). Similarly, researchers were able to produce lymphoid-competent HSCs through the conversion of adult mouse endothelial cells using vascular induction and transient expression of *Foxb*, *Runx1*, *Gfi1* and *Spi1* (Lis *et al.*, 2017). This ground-breaking research not only advances the hope of using HSCs in a clinical context but signifies the importance of the endothelial origin of HSCs and the vascular microenvironment in which they are born.

Chapter 2: Hyaluronan and hyaluronan binding proteins (HABPs)

2.1 Hyaluronan is an inductive signal and provides a matrix for cell mobility

2.1.1 Hyaluronan is an integral part of the extra-cellular space

Hyaluronan is a large unbranched polymer composed of 2,000 to 25,000 disaccharide units that exists commonly on the cell surface and within the extracellular matrix (ECM) (Weissmann and Meyer, 1954; Toole, 2004). Hyaluronan is synthesised by one of three trans-membrane enzymes that directly extrude hyaluronan into the extra-cellular space (Weigel, Hascall and Tammi, 1997). This high molecular weight glycosaminoglycan forms mesh-like structures between cells, using its hydrophilic nature to expand the extra-cellular space and is frequently associated with migrating cell types and tissue morphogenesis (Toole, 2004). For this reason it is interesting to consider the role of hyaluronan during EHT. Loss of function of Hyaluronan synthase-2 (Has2) leads to embryonic lethality at E9.5 due to cardiac defects (Camenisch *et al.*, 2000). The lack of hyaluronan present in *Has2* knockout embryos suggests that this enzyme is primarily responsible for hyaluronan synthesis during early development (Camenisch *et al.*, 2000).

The broad spectrum of cell processes related to hyaluronan can be explained by both its size range and the number and diversity of HABPs (Day and Prestwich, 2002; Toole, 2004). In mice and humans there are approximately twenty-five HABPs described and six hyaluronan-degrading enzymes (Csoka, Frost and Stern, 2001; Day and Prestwich, 2002). Exogenous expression of

HAS1, HAS2 and HAS3 resulted in the production of predominantly high molecular weight hyaluronan, suggesting that low molecular weight hyaluronan is probably the result of enzymatic activity (Itano *et al.*, 1999). High molecular weight hyaluronan is generally associated with quiescence and has been shown to suppress the immune response, angiogenesis and cellular differentiation (Feinberg and Beebe, 1983; Delmage *et al.*, 1986; Su *et al.*, 2017; Wong *et al.*, 2017). Conversely, low molecular weight hyaluronan has been shown to be stimulatory in nature, promoting angiogenesis and inflammation (West et al., 1985; Campo *et al.*, 2010). These opposing responses generated based on the size of the macromolecule greatly increases the complexity of the role hyaluronan plays in cellular processes.

2.1.2 The formation of pericellular matrices assists in migration and tissue morphogenesis

A key role for hyaluronan in the extracellular space is the formation of pericellular matrices around motile, load bearing and transitioning cell types. The hyaluronan-enriched matrix has been visualised by exclusion assays and electron microscopy on vascular smooth muscle cells, chondrocytes and epithelial cells (Knudson *et al.*, 1996; Evanko, Angello and Wight, 1999; Cohen *et al.*, 2003). This layer of hyaluronan associates with aggrecan and versican as well as other binding proteins to form a complex, highly hydrated buffer, which can affect the adhesion, signalling and motility of a cell (Evanko *et al.*, 2007). Pericellular matrix formation is mediated often by the attraction of hyaluronan to the cell surface by its principal binding protein CD44 (Knudson *et al.*, 1996). The exogenous expression of CD44 or the addition of

high molecular weight hyaluronan to CD44 expressing cells has been shown to be sufficient to precipitate the formation of the pericellular matrix, indicating the key role HABPs play in this process (Knudson and Knudson, 1991; Knudson, Bartnik and Knudson, 1993).

A key function of the pericellular matrix is to aid in cell migration, enabling cells the space to move. Time-lapse microscopy observed that the rapid formation of the pericellular matrix coordinated with cell detachment during smooth muscle cell migration (Evanko, Angello and Wight, 1999). Similarly, formation of the matrix in cancer cell lines increased cell motility, while over-expression of hyaluronan synthases resulted in greater freedom of movement through the repression of contact mediated inhibition (Itano *et al.*, 2002; Ricciardelli *et al.*, 2007). Furthermore, a pericellular matrix is known to form around pre-ovulatory oocytes in preparation for their journey (Camaioni *et al.*, 1996).. In addition to promoting migration, the loose, hydrated nature of the pericellular matrix allows cells to undergo shape changes and transitions. The pericellular matrix has been implicated in both specific cell transformations such as the process of condensation that precipitates limb bud formation and more generalised changes in cell shape that occur during mitosis (Maleski and Knudson, 1996; Evanko, Angello and Wight, 1999). As such, it seems hyaluronan is often involved in the movement and morphogenesis of multiple different cell types.

2.1.3 Hyaluronan deposition is associated with cancer metastasis

Not only is hyaluronan important to normal cell movement, its disturbance is also associated with disease states. This has drawn further attention since the

discovery that the resistance to cancer displayed by the naked mole rat is attributable to the production of extremely high molecular weight hyaluronan (Tian *et al.*, 2013). This unusual property could be reversed by over-expressing a hyaluronan degrading enzyme in naked mole rat cells, inducing tumourigenesis (Tian *et al.*, 2013). A high level of hyaluronan in the stroma around tumours has been used as a general marker of poor cancer prognosis. Increased hyaluronan deposition is associated with a high risk of metastasis in colorectal, gastric, ovarian, prostate and breast cancers (Ropponen *et al.*, 1998; Setälä *et al.*, 1999; Anttila *et al.*, 2000; Auvinen *et al.*, 2000; Lipponen *et al.*, 2001). In addition to these correlative studies, it has been shown that knockdown of Has2 in breast cancer cell lines reduces their aggressive cell growth and detachment (Li *et al.*, 2007). Thus, even in perturbed cases hyaluronan is critically involved in cell movement and transformation.

2.2 CD44 is the principal receptor of hyaluronan and is associated with migrating cell types

2.2.1 CD44 has a complicated relationship with its principal ligand hyaluronan

CD44 was first identified by monoclonal antibodies as a glycoprotein on the cell surface of myeloid and lymphoid cells (Trowbridge *et al.*, 1982; Hughes, Colombatti and August, 1983). This led to its characterisation as a lymphoid homing receptor, mediating the interaction of lymphoid cells with vascular endothelium (Jalkanen *et al.*, 1986, 1987). Elucidation of the genomic structure revealed *Cd44* to be a highly dynamic receptor with multiple splice isoforms (Screaton *et al.*, 1992). In fact, the *Cd44* genomic sequence consists

of twenty exons, ten of which undergo alternative splicing (Tölg *et al.*, 1993). All *Cd44* isoforms share with a high degree of homology the hyaluronan binding domain, stem region, trans-membrane domain and cytoplasmic domains (Thorne, 2003). The smallest and most abundant form, *Cd44s* consists of only ten constitutive exons which encode these regions while variant isoforms (Cd44v) include a combination of exons six through fifteen which add extra elements to the stem region of the receptor (Tölg *et al.*, 1993; Ponta, Sherman and Herrlich, 2003).

This heterogeneity could indicate a high degree of functional diversity. Binding affinity assays confirmed this, revealing that at least *in vitro* CD44 was capable of interacting with numerous elements of the ECM including collagen, laminin, fibronectin, osteopontin and hyaluronan (Aruffo *et al.*, 1990; Jalkanen and Jalkanent, 1992; Weber *et al.*, 1996). Additionally, CD44 is also known to recruit matrix-metalloproteineases (MMPs) to the cell surface (specifically MMP-9 and MMP-7) where they can interact with growth factors to affect cell signalling or function in the degradation of the ECM (Yu and Stamenkovic, 1999, 2000; Yu *et al.*, 2002). Despite these varied roles, CD44 is frequently referred to and most well characterised as the principal cell surface receptor for hyaluronan.

CD44 interacts with hyaluronan through a conserved ~100 amino acid motif known as the link domain, which is shared with many other HABPs (Neame, Christner and Baker, 1986; Yang *et al.*, 1994). CD44 differs however from other link-domain proteins as binding of hyaluronan is influenced by flanking

regions, glycosylation and the clustering of multiple CD44 receptors (Sleeman *et al.*, 1996; English, Lesley and Hyman, 1998; Teriete *et al.*, 2004). Although it is thought that a main function of CD44 binding to hyaluronan is simply to anchor the glycosaminoglycan to the cell membrane, structural studies have uncovered ligand induced conformational changes indicating a more complex interaction (Takeda *et al.*, 2006; Banerji *et al.*, 2007). The situation is further complicated by the diversity in size of hyaluronan polymers. CD44 was found to have higher affinity for high molecular weight hyaluronan due to multivalent binding (Wolny *et al.*, 2010; Mizrahy *et al.*, 2011). Overall, it appears that the relationship between CD44 and hyaluronan and the outcomes of their interaction are highly context dependent.

2.2.2 Down-stream targets and effectors of the CD44 receptor

Given the heterogeneity of CD44 proteins generated through alternative splicing and post-translational modifications, it's possible this cell surface receptor performs multiple functions. Firstly, CD44 acts as a docking site for ECM components, MMPs and growth factors, as previously discussed. In addition to these ligands, CD44 is known to act as a co-receptor for other cell surface proteins, thereby modulating signalling events in the cell. Furthermore, CD44 is known to directly interact with the cytoskeleton effecting change in cell shape. Much of the interest in CD44 is due to its association with cancers; as such many of the signalling pathways described are in the context of a disease state.

The tyrosine kinase receptor *c-Met* is critical to embryonic development, associated with cancer metastasis and a co-receptor of the variant isoform CD44v6 (Orian-rousseau *et al.*, 2002). Specific blocking antibodies against CD44v6 are capable of completely abolishing c-MET activation by its ligand Hepatocyte growth factor (HGF) (Orian-rousseau *et al.*, 2002). Targeted mutation studies have shown that cooperation between these two receptors is dependent upon a functional variable exon six and the intracellular domain of CD44, that assists in the activation of the c-MET target ERK (Orian-rousseau *et al.*, 2002; Orian-Rousseau *et al.*, 2007). Activation of *c-Met* by HGF results in the co-internalisation of both c-MET and CD44v6 (Hasenauer *et al.*, 2013). Confirmation of a CD44-cMET-HGF signalling complex was provided by the haplo-insufficiency of *c-Met* on a *Cd44* knockout background, with animals dying at birth from lung and nerve defects (Matzke *et al.*, 2007). Indeed, in the absence of CD44, c-MET recruits ICAM-1 to act as a co-receptor, providing evidence of a compensation mechanism which could explain the lack of phenotype in *Cd44* null mice (Olaku *et al.*, 2011). It is postulated that this relationship between CD44v6 and c-MET is the main driver of the metastatic phenotype associated with CD44v6 up-regulation. It is worth noting that this interaction has been observed in both epithelial and endothelial cells (Orian-rousseau *et al.*, 2002; Tremmel *et al.*, 2009). Furthermore, preliminary data suggest that CD44v6 also complexes with FLK-1 to mediate Vegf-A signalling (Tremmel *et al.*, 2009). This novel function of CD44 as a co-receptor of FLK-1 has not been further explored, although could explain the vascular phenotype observed in *Cd44* null mice where Matrigel tube formation was impaired and the endothelium appeared retracted and thin (Cao *et al.*, 2006).

CD44 has also been shown to form a co-receptor complex with the ErbB family of tyrosine kinases which are activated by Epidermal growth factor (EGF) and Heregulin (HER2) (Ghatak, Misra and Toole, 2005). CD44 was previously implicated in the EGF signalling pathways as *Cd44* knockout keratinocytes fail to proliferate in response to EGF stimulus (Kaya *et al.*, 1997). In cancer cell lines CD44, ERBB1 and ERBB2 were found to cluster at the cell membrane and undergo co-internalisation and proteolysis in response to EGF, enhancing migration capacity (Pályi-Krekk *et al.*, 2008; Hernández *et al.*, 2011). *In vivo* the CD44 variant CD44v3, which contains a binding site for heparin sulphate, was found to recruit both a heparin-binding EGF precursor and the processing enzyme MMP-7 to its co-receptor ERBB4 enabling efficient activation of the signalling pathway (Yu *et al.*, 2002). As a result *Cd44* null mice showed reduced phosphorylation of the ERBB4 receptor and increased apoptosis of epithelial and smooth muscle cells where this complex is present (Yu *et al.*, 2002). This evidence shows that CD44 can play an important role in localising ligands to their receptors to facilitate signal transduction.

2.2.3 CD44 is expressed on both HSPCs and lymphocytes

The association of CD44 with lymphocyte homing was first established due to the effect of blocking antibodies. This interference in the binding of lymphocytes to endothelial cells was found to prevent their return to the lymphatic organs (Jalkanen *et al.*, 1986). The link was supported by studies in *Cd44* knockout mice where loss of *Cd44*, although resulting in no dramatic

phenotype, impaired lymphocyte homing to the thymus and lymph nodes (Protin *et al.*, 1999). Further research has shown that the function of CD44 on lymphocytes is related to their activation (Galandrini et al., 1994). Stimulation of T cells induces the binding of CD44 to hyaluronan and enables the rolling of lymphocytes across endothelial cells of the vasculature for homing and extravasation to sites of inflammation (Degrendele *et al.*, 1997; Bonder *et al.*, 2006).

Although *Cd44* null mice were viable and displayed few obvious defects, *Cd44* is thought to play a role in HSPC migration. Not only is Cd44 expressed on HSPCs, hyaluronan plays an integral role in their lodgement in the bone marrow, and with the loss of *Cd44,* myeloid progenitors showed difficulties in exiting the bone marrow for the blood stream (Schmits *et al.*, 1997; Dimitroff *et al.*, 2001; Nilsson *et al.*, 2003). Primitive CD34+ HSPCs have been shown to utilise CD44 in their adherence to hyaluronan and human cord blood cells could be inhibited in their migration to the bone marrow by pre-treatment with a CD44 blocking antibody (Legras *et al.*, 1997; Avigdor *et al.*, 2004). Re-examination of *Cd44* null mice found delayed migration of HSCs from the foetal liver to the bone marrow and reduced transplantation efficiency (Cao *et al.*, 2015). Together, these findings suggest that CD44 is most important during the establishment of the haematopoietic system. Interestingly, in the adult it is during disease states such as leukaemia that we can observe a more significant role for *Cd44* in relation to HSPCs. CD44 has been shown to be indispensable to leukaemic stem cells, which rely more extensively on the CD44-hyaluronan interaction for engraftment and maintenance in the stem

cell niche (Jin *et al.*, 2006; Krause *et al.*, 2006). Preliminary studies have also shown that by targeting CD44 it is possible to specifically eradicate the leukaemic stem cell population (Jin *et al.*, 2006).

2.2.4 Up-regulation of CD44 is associated with cancer stem cells and metastasis

Since the association of CD44 expression *in vitro* with metastasising rat pancreatic cancer cell lines, much research on CD44 has been focussed on its association with cancer (Günthert *et al.*, 1991). Of particular interest is the fact that CD44 and its splice variants have been linked to populations of cancer stem cells (CSCs) (Zöller, 2011). CSCs represent a small subset of tumours, which display self-renewing and differentiation capacity through xenograft transplants and are thought to be responsible for the recurrence of disease (Wicha, Liu and Dontu, 2006). In addition to leukaemic stem cells, (Jin *et al.*, 2006; Krause *et al.*, 2006), CD44 has been associated with cancer stem cells in gastric, colorectal, head and neck, breast, prostate and pancreatic cancers (Al-Hajj *et al.*, 2003; Patrawala *et al.*, 2006; Li *et al.*, 2007; Prince *et al.*, 2007; Du *et al.*, 2008; Takaishi *et al.*, 2009). Despite its extensive use as a maker of CSCs, the benefit CD44 confers on tumours is still not fully understood.

One mechanism by which CD44 is thought to be useful to CSCs is through the promotion of survival pathways and inhibition of apoptosis. The PI3K/Akt survival pathway is used by cancer cells to override cell cycle checkpoints and evade apoptosis (Sabbatini and McCormick, 1999; Liang and Slingerland,

2003). CD44 has been shown to activate PI3K signalling resulting in the downstream phosphorylation of Akt through its interaction with its co-receptor ERBB2 or through the binding of its ligand osteopontin (Lin and Yang-Yen, 2001; Misra *et al.*, 2008). Furthermore, hyaluronan oligomers were found to interfere with CD44 signalling and inhibit the phosphorylation of Akt in a murine mammary cancer cell line (Ghatak, Misra and Toole, 2002). Additionally, CD44 is thought to confer an advantage on CSCs by promoting drug resistance through the up-regulation of the drug efflux pump MDR1 and the promotion of a glycolytic metabolic state through direct association with PKM2 (Bourguignon *et al.*, 2008; Tamada *et al.*, 2012).

CD44 is not only a marker of CSC populations, but has been implicated in cancer progression and metastasis with cell surface expression linked to migratory capacity in several cancer cell lines (Günthert *et al.*, 1991; Okamoto *et al.*, 1999; Patrawala *et al.*, 2006). However some studies have suggested that in fact CD44 acts as a suppressor of metastasis (Gao *et al.*, 1997; Lopez *et al.*, 2005). Nevertheless, it has been shown that the p53 regulated microRNA, miR-34a can specifically down-regulate CD44 expression inhibiting prostate cancer metastasis in immune suppressed mice (Liu *et al.*, 2011). This discrepancy in results could be attributable to the differing effects and binding capacities of the CD44 splice variants. Indeed, it has been shown that different splice variants of CD44 can confer different oncogenic potential (Hofmann *et al.*, 1991). Switching between the standard and variant CD44 isoforms has also been shown to be instrumental to the epithelial to mesenchymal transition (EMT) (Brown *et al.*, 2011). Overall CD44

participates in a diverse array of cell biological processes and its function is modulated by the splice isoforms, ligands and binding partners present. As such, unravelling the contribution of CD44 to a biological system is a complex task.

2.3 Stabilin-2 acts as the major clearance receptor of hyaluronan

2.3.1 A new hyaluronan receptor

In comparison to *Cd44*, far less research has been undertaken on the cell surface receptor *Stabilin-2 (Stab2)*. *Stab2* was first identified as an hyaluronan binding protein expressed on liver sinusoidal endothelial cells (McCourt *et al.*, 1999; Zhou *et al.*, 2000). Further investigation has since expanded upon its role and expression within the body. In addition to the liver endothelium, Stab2 expression is found in lymph nodes, spleen, heart valves and bone marrow as well as the epithelium of the eye and kidney (Falkowski *et al.*, 2003). Its role as a clearance receptor involves not only the engulfment of metabolic waste products such as hyaluronan and heparin but also the phagocytosis of necrotic tissue, apoptotic cells and aging erythrocytes (Harris and Weigel, 2008; Harris, Weigel and Weigel, 2008; Park *et al.*, 2008; Kim *et al.*, 2010; Lee *et al.*, 2011; D'Souza, Park and Kim, 2013). Loss of function of *Stab2* in mice results in little obvious defects, although an increase in serum hyaluronan levels was observed (Hirose *et al.*, 2012). Interestingly, a lack of *Stab2* resulted in a dramatic reduction in the ability of cancer cells to metastasise within the mouse which runs counter to the idea that hyaluronan deposition aids cancer progression (Hirose *et al.*, 2012).

In addition to functioning as a scavenger receptor Stab2 has recently been attributed several novel functions relating to cell adhesion and aggregation. Over-expression of Stab2 in mouse fibroblasts found that Stab2 mediated cell aggregation through homophilic interactions using fasciclin-1 domains (Park, Jung and Kim, 2009). *In vivo*, Stab2 has been shown to enable the adhesion of lymphocytes to endothelial cells where it is expressed and aids in the fusion of myoblasts during muscle growth and regeneration (Jung, Park and Kim, 2007; Kim, Park and Kim, 2016). Research into Stab2 has been hampered in part by its large coding region (~8kb) and lack of suitable antibodies. With more tools now available and the possibility to specify Stab2+ endothelial cells from ESCs using a TGF β inhibitor it has become easier to better elucidate the function of Stab2 and its association with hyaluronan (Nonaka *et al.*, 2008).

Chapter 3: Materials and Methods

3.1 Chemicals

Table 1: Chemicals

Chemical	Manufacturer	Reference Number
2-Propanol	Sigma-Aldrich	33539
Agarose	Sigma-Aldrich	A9539
Bovine serum albumin (BSA)	Sigma-Aldrich	A9418
Collagenase	Sigma-Aldrich	C9722
Dimethyl sulphoxide (DMSO)	Sigma-Aldrich	D5879
Ethanol (absolute)	Sigma-Aldrich	32205
Ethidium bromide	Sigma-Aldrich	E1510
Glycine	Sigma-Aldrich	G8898
Paraformaldehyde	Santa Cruz	SC281692
Phosphate Buffered Saline (PBS) (without Ca2+ or Mg2+)	Gibco	14190-094
Sucrose	Sigma-Aldrich	S5016
Triton-X 100	Sigma-Aldrich	T8787
Tween-20	Sigma-Aldrich	P9416

3.2 Mouse embryo protocols

3.2.1 Embryo dissection

All mouse experiments were performed on either C57BL/6N wild-type mice or *Runx1* null mice. Timed matings were set up overnight and embryos collected from pregnant females in PBS supplemented with 10% FBS (PAA Clone, A15-02) between embryonic day 9.5 and 11.5. Forceps were used to remove embryos from the uterine tissue under a stereomicroscope. In the case of embryos derived from *Runx1*$^{+/-}$ mice the yolk sac or head was used for genotyping. In order to determine the developmental age, somite pairs were

counted and embryos grouped according to *table 2*. The yolk sac and AGM region were isolated as shown in *figure 3*. Once the embryo was removed from the uterine tissue forceps were used to make an incision in the yolk sac and the tissue was peeled away from the embryo. Cuts were made in the vitelline and umbilical vessels to detach the yolk sac from the embryo. The head and tail were then removed above and below the limb buds using 0.45mm needles. Next, a cut was made ventral to the AGM region to remove the limb buds and organs. Another cut was made dorsal to the AGM region to remove the somite tissue above. To create a single cell suspension AGMs and yolk sacs were pooled and incubated in 1.25mg/mL of collagenase for 30 minutes at 37°C. Cells were then pipetted up and down to create a homogenous cell suspension.

Table 2: Embryonic stages of mouse development

Number of somite pairs	Embryonic days post coitum
21-29	E9.5
30-34	E10
35-39	E10.5
40-44	E11
45-48	E11.5

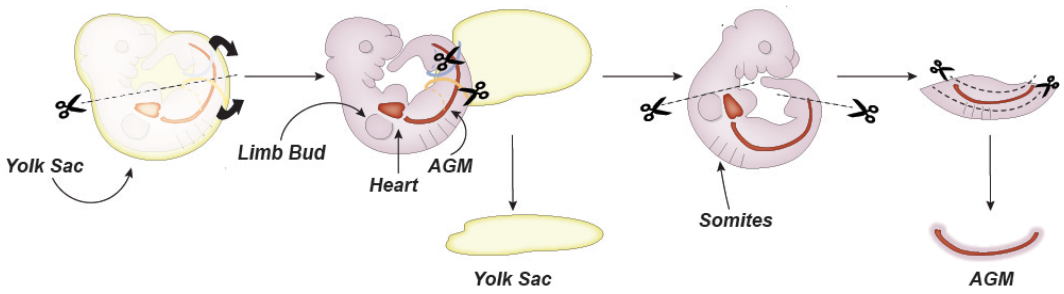


Figure 5: Dissection of yolk sac and AGM from mouse embryo

The yolk Sac and AGM region can be isolated from mid-gestation mouse embryos using forceps and 0.45mm needles. Cuts were made as indicated by the dotted lines.

3.2.2 Genotyping

Genotyping of mouse embryos was performed using the Kappa hot start mouse genotyping kit (KAPA Biosystems; KK7352) according to the manufacturers instructions. Briefly, the head or yolk sac of an embryo was incubated at 75°C for ten minutes in extract buffer with two units of Kappa express extract enzyme. Samples were then vortexed and centrifuged for one minute. The PCR reaction was then performed using 1µL of the supernatant in a 25µL reaction with the KAPA 2G mix containing DNA polymerase and DNTPs as well as 0.5µM final concentration of each of the *Runx1* primers listed below.

Runx1 Forward Primer 1: 5’ CCAATGAGAAACAGTAGTAGC 3’

Runx1 Forward Primer 2: 5’ TGCTTAATGGTGACACTTTCT 3’

Runx1 Reverse Primer: 5’GCCTTCAGAATCAGTAGAAC 3’

The reaction was then run on the thermocycler for 35 cycles with a fifteen second denaturation step at 95°C, annealing for fifteen seconds at 60°C and elongation for thirty seconds at 72°C. The products were then visualised with electrophoresis using ethidium bromide on a 2% agarose gel.

3.3 In vitro cell culturing methods

3.3.1 Mouse embryonic fibroblast (MEF) preparation

Mouse embryonic fibroblasts were isolated from E13.5 embryos and the primary cells stored in liquid nitrogen in 40% DMEM (Dulbeco's modified eagle medium) (Gibco; 11965), 50% foetal bovine serum (FBS) (PAA Clone, A15-02) and 10% DMSO. To prepare MEFs for ESC culturing a primary vial was thawed in a 15cm^2 dish with DMEM media (supplemented with 1% Penicilliin/Streptomycin (Gibco; 15140122), 1% L-glutamine (Gibco; 25030-024) and 1% non-essential amino acids (Gibco; 11140-035)). Fifteen percent FBS and 0.12mM β-mercaptoethanol (Gibco; 31350-010) was then added. The MEFs were cultured for five days in order to expand as much as possible. Cells were split when they reached 80-90% confluency. To inhibit further cell growth MEFs were treated with 10µg/mL of mitomycin C (Sigma: M4287) and incubated for two hours at 37°C. The cells were then harvested with TryplE-Express (Gibco; 12605-010), counted and frozen in liquid nitrogen for later use.

3.3.2 Mouse embryonic stem cell (mESC) culture

The A2lox-Empty embryonic stem cell line (kindly provided by Dr. Michael Kyba, University of Minnesota) was maintained in DMEM-ES culture medium

containing DMEM KO (Invitrogen; 10829018) supplemented with 1% Penicilliin/Streptomycin, 1% L-glutamine and 1% non-essential amino acids. In addition, the media contained 15% FBS (PAA Clone, A15-02), 0.024µg/mL of LIF (produced by the protein expression facility at EMBL, Heidelberg) and 0.12mM β-mercaptoethanol. All media was sterile filtered using a Millipore stericup (SCGPU 01RE) with a 0.22µm filter. Cells were maintained on MEFs and incubated at 37°C with 5% CO_2 and 95% relative humidity. TryplE-Express was used to detach cells for passaging, collection or to create a single cell suspension. Cells were incubated in TryplE-Express for three to five minutes at 37°C.

3.3.3 Growth factors

The growth factors listed in *table 3* were added to the culture media for *in vitro* and *ex vivo* experiments with mESC derived cells and cells from primary embryonic mouse tissue.

Table 3: Growth Factors

Name	Manufacturer	Reference Number
Basic Fibroblast Growth Factor (bFGF)	R&D Systems	233-FB-025
Fms-like tyrosine kinase 3 ligand (Flt-3)	Preprotech	250-31L
Interleukin-3 (Il-3)	Preprotech	213-13
Interleukin-6 (Il-6)	Preprotech	216-16
Interleukin-7 (Il-7	Preprotech	217-17
Interleukin-11 (Il-11)	Preprotech	220-11
Leukocyte Inhibitory Factor (LIF)	EMBL Heidelberg	
Oncostatin-M	R&D Systems	495-MO-025

Stem Cell Factor (SCF)	Preprotech	250-03
Vascular endothelial growth factor (VEGF)	Preprotech	500-P131

3.3.4 Embryoid body (EB) differentiation

To generate embryoid bodies and isolate the haemangioblast progenitors of blood development we used an embryonic stem cell (ESC) differentiation system based on a previously described protocol (Sroczynska *et al.*, 2009). To begin differentiation the ESCs were passaged twice on 0.1% gelatin to remove feeder cells first in DMEM-ES medium and then using IMDM-ES medium containing IMDM (Iscove's modified Dulbecco medium) (Lonza; BE12-726F) supplemented with 1% Penicilliin/Streptomycin and 1% L-glutamine. In addition, 15% FBS (PAA Clone, A15-02), 0.024µg/mL LIF and 0.12mM β-mercaptoethanol was added. Once feeders were removed cells were harvested and cultured in untreated $10cm^2$ petri dishes at a density of $0.3x10^6$ cells per dish with EB medium containing IMDM (supplemented with 1% Penicilliin/Streptomycin and 1% L-glutamine), 10% FBS, 0.6% Transferrin (Roche; 10652), 0.03% monothioglycerol (MTG) (Sigma; M6145) and 50µg/mL ascorbic acid (Sigma; A4544). After three days in culture EBs were harvested and the Flk1+ haemangioblast progenitors were isolated. The EB clusters were collected and dissociated using TryplE-Express for five minutes. The TryplE-Express was neutralised using IMDM with 20% FBS before cells were filtered through a BD cup filcon filter (BD Biosciences; 340629) to create a single cell suspension. The cells were centrifuged and resuspended in 5mL of IMDM with 20% FBS. The same volume of lymphocyte separation medium (Lonza; 17-829E) was carefully added to remove dead cells by density

gradient separation. Cells were then centrifuged for twenty minutes at 800G with no brake to maintain the phases. The central white phase was then transferred to a new tube with 10mL of IMDM with 20% FBS. Cells were counted using a haemocytometer and resuspended in 1x PBS supplemented with 0.5% FBS and 0.5mM EDTA (MACS buffer) to create a concentration of $1x10^8$ cells per mL. A magnetic activated cell sorting (MACS) system was used to isolate *Flk1*+ progenitors. Cells were stained with anti-Flk1, APC conjugated antibody (eBiosciences; 17-5821-81) at a concentration of 3µL per $1x10^7$ cells and incubated at 4°C for five minutes. Cells were then washed with 10mL of MACS buffer, centrifuged and resuspended in a one in five dilution of anti-APC microbeads (Miltenyi Biotech; 130-090-855) diluted in MACS buffer at a concentration of 1mL for every $1x10^8$ cells. Cells were then incubated at 4°C for fifteen minutes. Cells were washed again with 10mL of MACS buffer, centrifuged and resuspended in 500µL volume of MACS buffer. The LS MACS column (Miltenyi Biotech, 130-042-401) was attached to a magnet to perform the magnetic separation and rinsed twice with 5mL of MACS buffer before loading the cells. The column is then washed three times with 3mL of MACS buffer. Finally, the column is removed from the magnet and the *Flk1*+ cells eluted with 5mL of MACS buffer. The positive cell fraction was then centrifuged, resuspended and counted before freezing cells for later use in haemangioblast assays. The purity of *Flk-1*+ cells was evaluated using FACS analysis on the FACS Canto cytometer (Becton Dickson) using FACS Diva software. This method routinely results in *Flk-1*+ cell fractions with greater than 95% purity.

3.3.5 Haemangioblast culture assay

To further differentiate the haemangioblast progenitors into vascular smooth muscle (VSM), endothelial cells and blood cells the *Flk-1*+ cells were first thawed on 0.1% gelatin in haemangioblast medium containing IMDM (supplemented with 1% Penicillin/Streptomycin and 1% L-glutamine), 10% FBS, 0.6% transferrin, 0.3% MTG, 50µg/mL ascorbic acid, 0.05% VEGF (10µg/mL) and 0.1% IL-6 (10µg/mL). In addition, 15% D4T supernatant was added. D4T supernatant is produced by culturing D4T endothelial cells for 48 hours in IMDM media with 10% FBS and 30mg of endothelial growth supplement (BD, 354006). The cultured media is then collected for use in the haemangioblast culture. Cells were grown for between 24 and 72 hours and used for downstream assays such as flow cytometry, quantitative RT-PCR and time-lapse imaging analysis.

3.3.6 Blocking CD44-hyaluronan interaction in vitro

For *in vitro* experiments to inhibit CD44-hyaluronan interaction, *Flk-1*+ haemangioblast cells were plated as per the haemangioblast assay (see 3.3.5) with either no intervention, 10µg/mL of anti-CD44 blocking antibody [KM201] (Abcam; ab25340), 300µg/mL of hyaluronidase enzyme (Sigma; H4272) or 50µM of 4-methylumbelliferone (4-MU) (Sigma; M1381-25G). Cell populations were then analysed using flow cytometry (see 3.5.1) to understand the relative proportion of VSM, endothelial and haematopoietic populations using CD41 PE, VE-Cadherin AF660 and Kit BV421 antibodies (see *table 4*).

3.4 Generation of a CD44 ESC knockout line with CrispR-Cas9

3.4.1 Generation of plasmid for CD44 knockout allele

The CD44 knockout mESC line was generated based on a previously published protocol (Wettstein et al., 2016). Firstly, the pX458 plasmid containing a Cas9 nuclease construct and a GFP reporter (pSpCas9-2A-GFP, Addgene ID: 48138) (kindly provided by Dr. James Hackett, EMBL Rome) (*fig. 6*) was amplified; linearised using the BpiI restriction enzyme and gel purified using the QIAquick gel extraction kit (Qiagen; 27804)(Fig, 6). A single guide RNA was designed to target the constitutive exon 2 of the *Cd44* transcript using the Broad Institute GPP web portal (5' CATGGAATACACCTGCGTAGCGG 3'). Complementary oligo DNA sequences were ordered with the PAM motif removed and the addition of "CACCG" 5' to the sense strand and "CAAA" on the 5' of the antisense strand as well as the addition of a "C" nucleotide 3' of the antisense strand. These overhangs are complementary to the sticky ends produced by BpiI digestion of the pX458 plasmid. To generate double stranded sgRNAs the oligos were placed on the heat block for four minutes at 94°C, ten minutes at 70°C and twenty minutes at 37°C. A ligation reaction was then performed for 1 hour at room temperature using 20μM of the double stranded sgRNAs, 50ng of pX458 plasmid and 400U of T4 ligase (NEB: M0202S). The plasmid was transformed into competent DH10b E.Coli using a heat shock method and the bacteria grown overnight on agar plates with ampicillin. DNA from the subsequent clones was purified using Qiagen miniprep columns and sequenced to ensure sgRNA insertion.

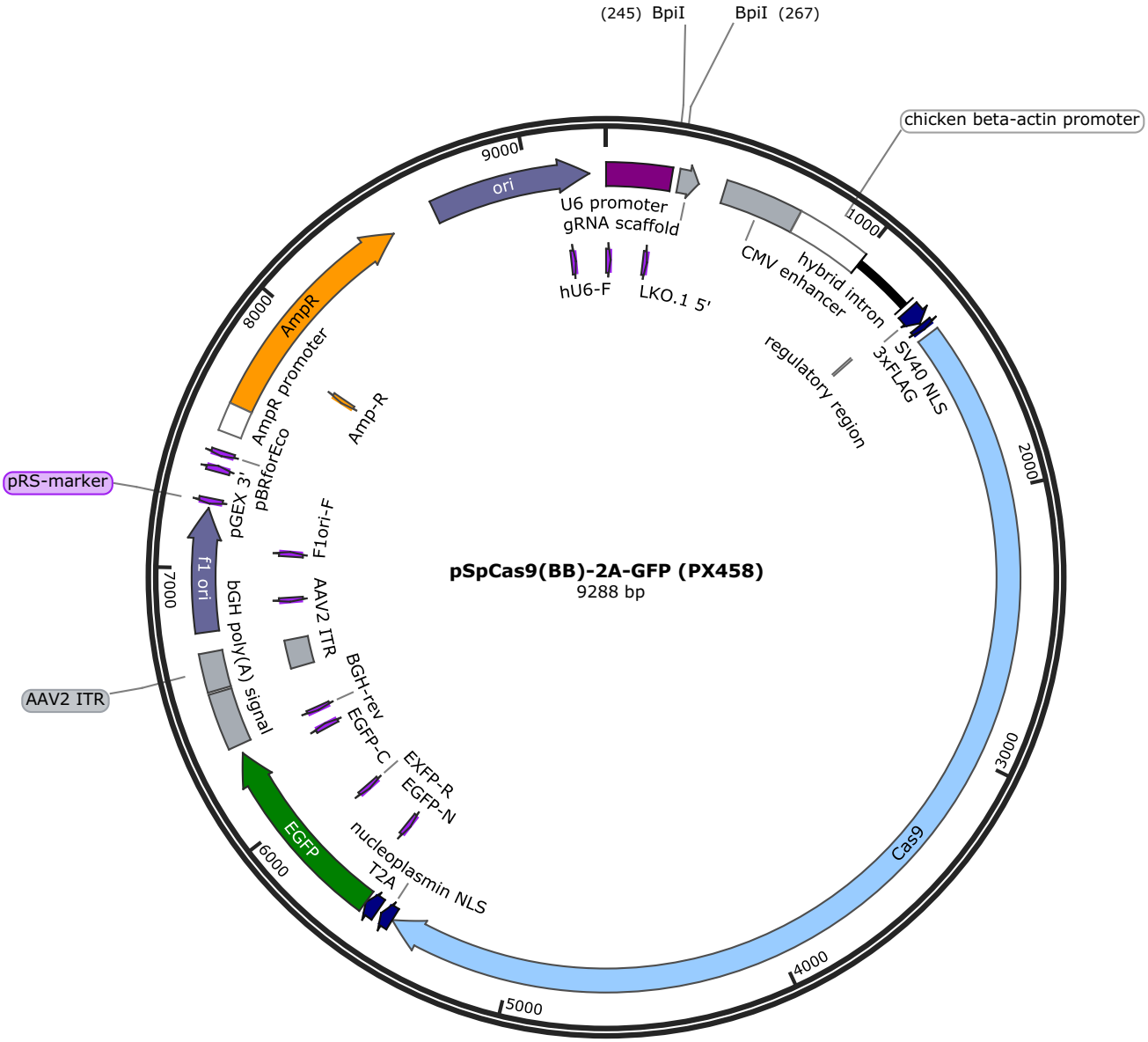


Figure 6: PX458 plasmid map detailing key features

PX458 plasmid map created with SnapGene Viewer details the key features including Cas9 nuclease construct (light blue), EGFP reporter (green), ampicillin-resistance gene (orange) and sgRNA scaffold with BpiI restriction sites upstream.

3.4.2 Transfection of mESCs

To transfect A2lox empty mESCs we used the Nanojuice transfection kit (Merck Millipore; 719020-3). Cells were treated with 2.5µL of Nanojuice transfection reagent, 1.25µL of Nanojuice booster and 1.25µg of plasmid DNA for 48 hours. The ESCs were then harvested and single GFP+ clones were FACS sorted onto MEFs in a 96 well plate. Single clones were expanded and validated for loss of CD44 through FACS analysis with PE conjugated anti-CD44 antibody.

3.5 Cell staining techniques

3.5.1 Flow cytometry and cell sorting

For flow cytometry analysis and sorting, single cell suspensions were obtained from either ESC culture or embryo dissections and kept in 1x PBS supplemented with 10% FBS (PAA clone, A15-02). Cells were counted and stained for 10 minutes at room temperature with fluorescence-conjugated antibodies (*table 4*). Cells were washed, filtered and analysed using a FACS Aria III (Becton Dickinson) and FACS Diva software. Fluorescence minus one (FMO) controls were used to determine background fluorescence and establish gating of populations. Data were analysed using FlowJo v10.1r5 (Tree Star Inc.). Cells were sorted using an 85µm nozzle for downstream molecular analysis or a 100µm nozzle for live cell culturing experiments.

Table 4: Antibodies used for flow cytometry and sorting

Antibody	Clone	Fluorochrome	Reference	Concentration
CD309	Avas12a1	APC	eBioscience; 17-5821-81	6µg/mL
CD144	BV13	eFluor 660	eBioscience; 50-1441-82	1µg/mL
CD41	MWReg30	PE	eBioscience; 12-0411-82	0.5 µg/mL
CD41	MWReg30	FITC	BD Biosciences; 561849	5µg/mL
CD117	2B8	BV421	BD Biosciences; 562609	1µg/mL
CD44	IM7	PE	BD Biosciences; 553134	0.08 µg/mL
CD45	30-F11	FITC	BD Biosciences; 553079	2µg/mL
CD45	30-F11	BV605	BD Biosciences; 563053	2µg/mL
CD43	S7	PerCP-Cy5.5	BD Biosciences; 562865	4µg/mL
CD19	MB19.1	PE	eBioscience; 12-0191-81	2µg/mL
CD11b	M1/70	APC	eBioscience; 17-0112-81	1µg/mL
CD4	RM4-5	PE-Cy7	BD Biosciences; 561099	0.25µg/mL
CD8a	53-6.7	FITC	BD Biosciences; 557668	2µg/mL
CD61	2C9.G2	BB700	BD Biosciences; 742110	1µg/mL
Stabilin-2	34-2	AF488	MBL; D317-A48	5µg/mL

3.5.2 Cell cycle analysis

To analyse the cell cycle status of AGM derived cells we used the Click-iT plus EDU flow cytometry kit (Life technologies; C10633). After dissection, AGMs were incubated in 10μM of EDU in PBS with 10% FBS for 1 hour at 37°C. The tissue was then washed and incubated in 1.25mg/mL of collagenase for 30 minutes. The sample was washed again and stained for cell surface markers (CD44, VE-Cadherin and Kit). Approximately 1000 cells were sorted based on cell surface expression into their different populations. Cells were then washed in PBS with 1% BSA then fixed in paraformaldehyde. The cells were permeabilised with a saponin-based wash solution before the addition of a copper protectant and AF488 picolyl azide for the detection of EDU. Finally the cell populations were stained with Hoechst (Invitrogen; R37605) before re-analysis on the FACS Aria III to determine the proportion of cells from each population in G0/G1, S and G2 phase.

3.5.3 Immunofluorescence on mouse cryo-sections

Mid-gestation mouse embryos were dissected and fixed in 4% paraformaldehyde for 15 minutes at room temperature then incubated in 15% sucrose solution for 2 hours before snap freezing in OCT (Tissue-Tek; 4583). 10μm transverse cryo-sections of the AGM region were then placed on superfrost plus slides (Thermoscientific; J1800AMNZ). Sections were washed in PBS, incubated in 1M glycine solution and permeabilised with 0.3% Triton X-100. Blocking solution consisting of 5% donkey serum, 5% chicken serum and 0.1% Tween-20 in 1x TBS buffer was applied to sections for 2 hours at room temperature. Sections were incubated in primary antibodies (*table 5*)

overnight at 4°C and then washed. Secondary antibodies were applied for 1 hour at room temperature and washed before DAPI nuclear stain was applied for 15 minutes (*table 6*). Slides were washed before being mounted with Prolong gold (Life Technologies; P36970) and imaged on a Leica SP5 confocal microscope.

Table 5: Primary antibodies for immunofluorescence

Antibody	Clone	Reference	Concentration
CD44	Rabbit polyclonal	Abcam; 157107	2µg/mL
VE-Cadherin	Rat monoclonal eBioBV13	eBioscience; 14-1441-81	2.5µg/mL

Table 6: Secondary antibodies for immunofluorescence

Antibody	Reference	Concentration
Alexa Fluor 488 donkey anti-rabbit	Life Technologies; A21206	4µg/mL
Alexa Fluor 568 goat anti-rat	Life Technologies; A11077	4µg/mL
DAPI nuclear stain	Invitrogen; 101635	5µg/mL

3.6 *Ex vivo* cell culturing methods

3.6.1 OP9 co-culturing assay for haematopoietic potential

OP9 cells were maintained in MEM alpha medium (Gibco; 22561-021) with 20% FBS (ATCC 30-2020). One day before sorting 3000 OP9 cells per well were seeded onto a 96 well plate in OP9 basic medium. On the day of sorting the medium was changed to a haemogenic endothelium rich medium containing IMDM (Lonza; BE12726F) treated with 1% penicillin-streptomycin (Gibco; 15140-122) and supplemented with 10% FBS (PAA Clone, A15-02), L-glutamine, transferrin, MTG, abscorbic acid, LIF, 50ng ml^{-1} SCF, 25ng ml^{-1} IL3, 5ng ml^{-1} IL11, 10ng ml^{-1} IL6, 10ng ml^{-1} Oncostatin M and 1ng ml^{-1} bFGF. For limiting dilution co-cultures cells were sorted using a 100µm nozzle directly onto the confluent OP9 stromal layer and incubated for three days. Round cell colonies were quantified manually using a light microscope.

3.6.2 Haematopoietic colony forming assay

One hundred cells were initially sorted onto a confluent OP9 stromal layer as per OP9 co-culturing assay. After three days in culture cells were harvested with TrypLE express (Gibco) and colony-forming unit-culture (CFU-C) assays were initiated using Methocult complete medium (Stem Cell Technologies; M3434). Cells were grown in 35mm culture dishes for a further 7 days. Colonies were quantified manually using a light microscope and scored as either CFU-Erythroid, CFU-Macrophage or CFU-Mixed.

3.6.3 Lymphocyte progenitor assay

Fifty cells were sorted onto confluent OP9 or OP9-DL1 stromal layers as per the OP9 co-culturing assay. Instead of using the rich haemogenic endothelium medium cells were incubated with factors conducive to lymphocyte development - MEM-alpha medium supplemented with 20% FBS (PAA Laboratories), 50ng ml^{-1} SCF, 5ng ml^{-1} Flt-3L and 1ng ml^{-1} IL7. Medium was changed every 4-5 days and cells were split as necessary. Cells were cultured for 21 days before harvesting with TrypLE express for flow cytometry analysis with anti-CD45, anti-CD19, anti-CD11b, anti-CD4 and anti-CD8a antibodies (see *table 1*).

3.6.4 Blocking CD44-hyaluronan interaction *ex vivo*

For *ex vivo* CD44-hyaluronan blocking experiments 20 VE-Cadherin+/CD44High cells were sorted, as per OP9 co-culturing assay (see 3.6.1), into a medium containing no antibody, 5µg/mL or 10µg/mL of anti-CD44 antibody [KM20] (Abcam; ab25340). Colonies were quantified manually using a light microscope after three days in culture.

3.7 Molecular biology techniques

3.7.1 RNA extraction

RNA extraction was performed on tissue culture derived cells and mouse embryo derived FACS sorted populations using the RNeasy micro-kit (Quiagen; 74004). In both cases cells were pelleted and resuspended in lysis buffer, vortexed and mixed with 70% ethanol before binding to the column. The column was washed, treated with DNase I and washed again. Next, 80%

ethanol was added to the column before drying the membrane and eluting the RNA in RNase-free water.

3.7.2 cDNA conversion

Fifty to two hundred nanograms of RNA was converted to single-stranded cDNA using the Revertaid H minus first strand cDNA synthesis kit (Thermo Scientific; K1632). RNA samples were incubated for five minutes at 65°C with random hexamer primers for annealing. RNase inhibitor, dNTPs and 20 Units of RT enzyme were added and reverse transcription carried out for one hour at 42°C. Denaturation was performed at 70°C for five minutes. PCR reaction was carried out immediately to avoid degradation of splice variants.

3.7.3 PCR amplification of CD44 splice variants

Five microlitres of cDNA was amplified using 2.5 Units of DreamTaq DNA polymerase (Thermo Scientific; EP0702). Primers were designed against the constitutive exons 5 and 16 to capture variant exons present in the CD44 transcript (see below). The reaction was run for 35 cycles with an annealing temperature of 58°C and an elongation time of 90 seconds.

Forward Primer: 5’ AGCACCCCAGAAAGCTACAT 3’

Reverse Primer: 5’ TTCTTGCATCTTTAGCGCCG 3’

Splice variants were then visualised on a 1% agarose gel using ethidium bromide. Bands were cut and DNA purified using a gel extraction kit (Qiagen; 28704) before sequencing to identify variant exons expressed.

3.7.4 Single-cell quantitative PCR

For single cell qPCR analysis cells were sorted using an 85µm nozzle directly into lysis buffer (Cells Direct qRT-PCR Kit, Invitrogen) and snap frozen. Samples were then reverse transcribed with superscript III reverse transcriptase from the Cells Direct one-step qRT-PCR Kit for 15 minutes at 50°C. The cDNA was then pre-amplified for twenty cycles with 25nM final concentration of each outer-primer for the genes of interest. The cDNA was then diluted with loading reagent (Fluidigm) and SoFastTM EvaGreen supermix (Biorad) and loaded onto a chip with 50µM of inner primer mix. Amplification of the 96 target genes (Supplementar*y table 1*) was measured with the Fluidigm Biomark HD system with the Biomark Data Collection software and the GE96 x 96 + Meltv2.pcl program.

3.7.5 Single-cell quantitative PCR analysis

Analysis of single cell qPCR data was performed as previously described (Bergiers *et al.*, 2018). Briefly, initial analysis was performed using the Fluidigm Real Time PCR analysis software. Hierarchical clustering and principal component analysis were performed using the SINGuLAR analysis toolkit (Fluidigm version 3.5) in R software (version 3.2.1). This analysis was done with the assistance of Christophe Lancrin and EMBL Rome Bioinformatician Andreas Buness.

3.8 RNA sequencing

3.8.1 Smart-seq2 25 bulk RNA sequencing

Bulk RNA sequencing was performed as described by the SmartSeq2 protocol (Picelli *et al.*, 2014). Briefly, cells were FACS sorted directly into lysis buffer containing 0.2% Triton X-100, oligo-dT primers and dNTP mix, and then snap frozen. Reverse transcription was then performed, followed by pre-amplification for 14 cycles. Nextera libraries were then prepared and sequenced on the Illumina Next Seq sequencer. The preparation of Smart-seq2 libraries was performed within our lab by Kerstin Ganter and sequencing undertaken by the EMBL GeneCore Facility in Heidelberg.

3.8.2 Bulk RNA sequencing analysis

Sequencing data was analysed with the aid of the EMBL Galaxy tools (galaxy.embl.de) (Afgan *et al.*, 2016) - specifically, FASTX for adaptor clipping, RNA STAR for mapping and htseq-count for obtaining raw gene expression counts. Maya Shvartsman and Polina Pavlovich performed pre-processing of the bulk RNA sequencing data within the lab. The R software (version 3.5.1, http://www.R-project.org.) along with R studio (version 0.99.879) was then used to perform differential expression analysis using the DESeq2 package. t-SNE plots were created using the Rtsne and plotTsne packages and differentially expressed genes were visualised using ggplot and pHeatmap packages (appendix I).

3.8.3 Takara single cell isolation and RNA sequencing

AGM derived cells were FACS sorted into VE-Cadherin+ and VE-Cadherin- populations before staining with the viability dye propidium iodide. Cells were

counted and diluted to 1 cell per 50 nL for dispensing onto the ICELL8 Chip (Takara) with the Multi-sample Nano-dispenser (Takara). All nano-wells were then imaged with a fluorescence microscope (Olympus) and the images were analysed with CellSelect software (Takara) to determine viability and select wells containing a single cell. Two hundred wells were selected and dispensed with a reverse transcription mix containing 5x RT buffer, dNTPs, RT e5 oligo and Maxima H minus RT. The RNA was then reverse transcribed at 42° for 90 minutes followed by a denaturation step at 85° for five minutes. The cDNA was concentrated using the DNA Clean and Concentrator-5 kit (Zymo Research). The cDNA was then treated with Exonuclease I, incubating at 37° for 30 minutes and 80° for 20 minutes before amplification with the Advantage 2 PCR kit (Clontech Takara) for 18 cycles. The cDNA was then purified with Ampure XP Beads (Beckmann Coulter) and the size distribution measured on a DNA Bioanalyser. Nextera libraries were then prepared and samples sequenced on an illumina NextSeq sequencer.

3.8.4 Single cell RNA sequencing analysis

The single cell sequencing data again underwent pre-processing using the EMBL Galaxy tools (galaxy.embl.de) (Afgan *et al.*, 2016). Quality control was performed using fastqc software before reads were attributed to single cells using the well barcodes and saved into a single file. Adaptors and poly-A sequences were then removed using Cutadapt and reads were aligned to the mouse genome GRCm38.86 using STAR. Reads were counted if they overlapped with exactly one gene and where the UMI was found in greater than two reads. This generated a count matrix of 200 cells with 11,775 genes detected in more than 3 cells. These pre-processing steps were performed by

the bioinformatician at EMBL Rome, Andreas Buness with Polina Pavlovich in our lab. Down-stream analysis was then performed in R (version 3.5.1, http://www.R-project.org) and R studio (version 0.99.879) using the Seurat package for single cell analysis (Macosko et al., 2015). After quality control to remove potential doublets and cells with high mitochondrial gene content we were left with 178 single cells for analysis. Marker genes were then visualised using TSNEPlot(), FeaturePlot() and pHeatmap() functions.

Chapter 4: Identifying a new marker and regulator of the endothelial to haematopoietic transition in the AGM

4.1 CD44 marks all stages of early blood development

4.1.1 Expression of CD44 correlates with haematopoietic development

It is well established that the first haematopoietic stem and progenitor cells (HSPCs) in the embryo have an endothelial origin. Time-lapse imaging *in vitro* (Lancrin *et al.*, 2009), live imaging of zebrafish embryos (Bertrand *et al.*, 2010; Kissa & Herbomel, 2010) and even *ex vivo* imaging of mouse endothelium (Boisset *et al.*, 2010) have been able to capture the precise moments of EHT. Despite the wealth of visual data much remains to be learned in regards to the molecular characterisation and regulation of this process. In particular, it has proved difficult to isolate the endothelial precursor of blood development known as haemogenic endothelium. With the advent of single cell technology isolating rare populations and characterising developmental transitions has become significantly easier. To better understand the endothelial origin of HSPCs a single cell RNA sequencing (RNAseq) experiment was performed in the lab by Özge Vargel (Supplementary Fig. 1a). Subsequent bioinformatics analysis by our collaborator Valentine Svensson at EMBL-EBI identified the cell surface receptor CD44 as a robust marker of a population with dual endothelial and haematopoietic gene expression (Supplementary Fig. 1b).

In order to validate the hypothesis that CD44 could be used to mark cells undergoing EHT I performed FACS analysis on both our *in vitro* model of haematopoietic differentiation and on embryonic mouse tissue. CD44 was

highly expressed in the haemangioblast (blast) culture of mouse embryonic stem cells (mESCs) that mimics early blood development. The proportion of cells expressing both the endothelial marker VE-Cadherin and CD44 peaked after 48 hours in culture when the process of EHT is most active (Fig. 7a). To investigate *in vivo* I dissected the AGM region of E9.5 and E10.5 mouse embryos. We gated on cells with high VE-Cadherin expression to ensure we were selecting cells connected to the vessel (Fig. 7b). Within the VE-Cadherin+ fraction we could identify a proportion of cells that also expressed CD44 on their cell surface (Fig. 7c). The frequency of this population of VE-Cadherin+/CD44+ cells increased significantly from E9.5 to E10.5 as haematopoietic clusters develop in the AGM (Fig. 7d). In this way we were able to confirm our RNAseq data; identifying at the protein level a population of cells with both VE-Cadherin and CD44 cell surface expression *in vitro* and *in vivo*.

To understand the localisation of this cell population I performed immunofluorescence on transverse sections of the AGM region of mouse embryos, using a monoclonal VE-Cadherin and a polyclonal CD44 antibody in conjunction with DAPI nuclear staining (Fig. 8a). We found co-expression of VE-Cadherin and CD44 on a number of cells in the AGM region; interestingly expression of these cell surface markers did not appear to overlap. We identified double positive cells belonging to the haematopoietic clusters (Fig 8b) and cells attached to the endothelial wall of the AGM (Fig. 8c). This supported the hypothesis that CD44 could be used to mark cells undergoing EHT.

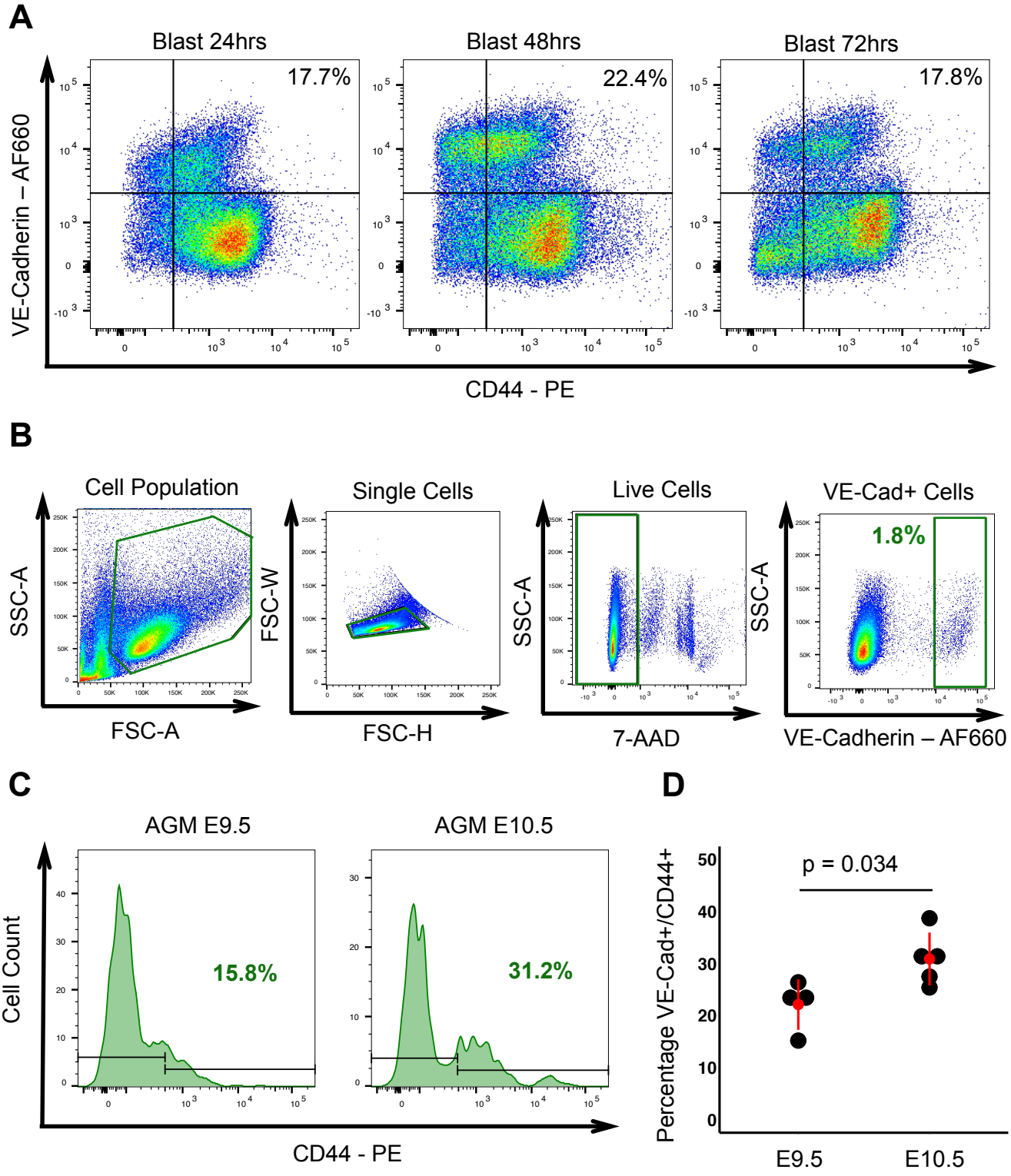


Figure 7: CD44 marks a subset of VE-Cadherin+ cells *in vitro* and *in vivo*

FACS analysis of CD44 and VE-Cadherin expression in haemangioblast culture (A) and AGM tissue (B-D). Gating strategy for E10.5 *in vivo* analysis removing debris, doublets, dead cells and selecting cells with high VE-Cadherin expression (B). The VE-Cadherin+/CD44+ population increases between E9.5 and E10.5 (C). Comparison of population percentages between E9.5 and E10.5 (D), n = 4 independent pools of embryos. Red dot indicates mean value and error lines represent the standard deviation. Significance was determined by an unpaired Student's T test, p = 0.034.

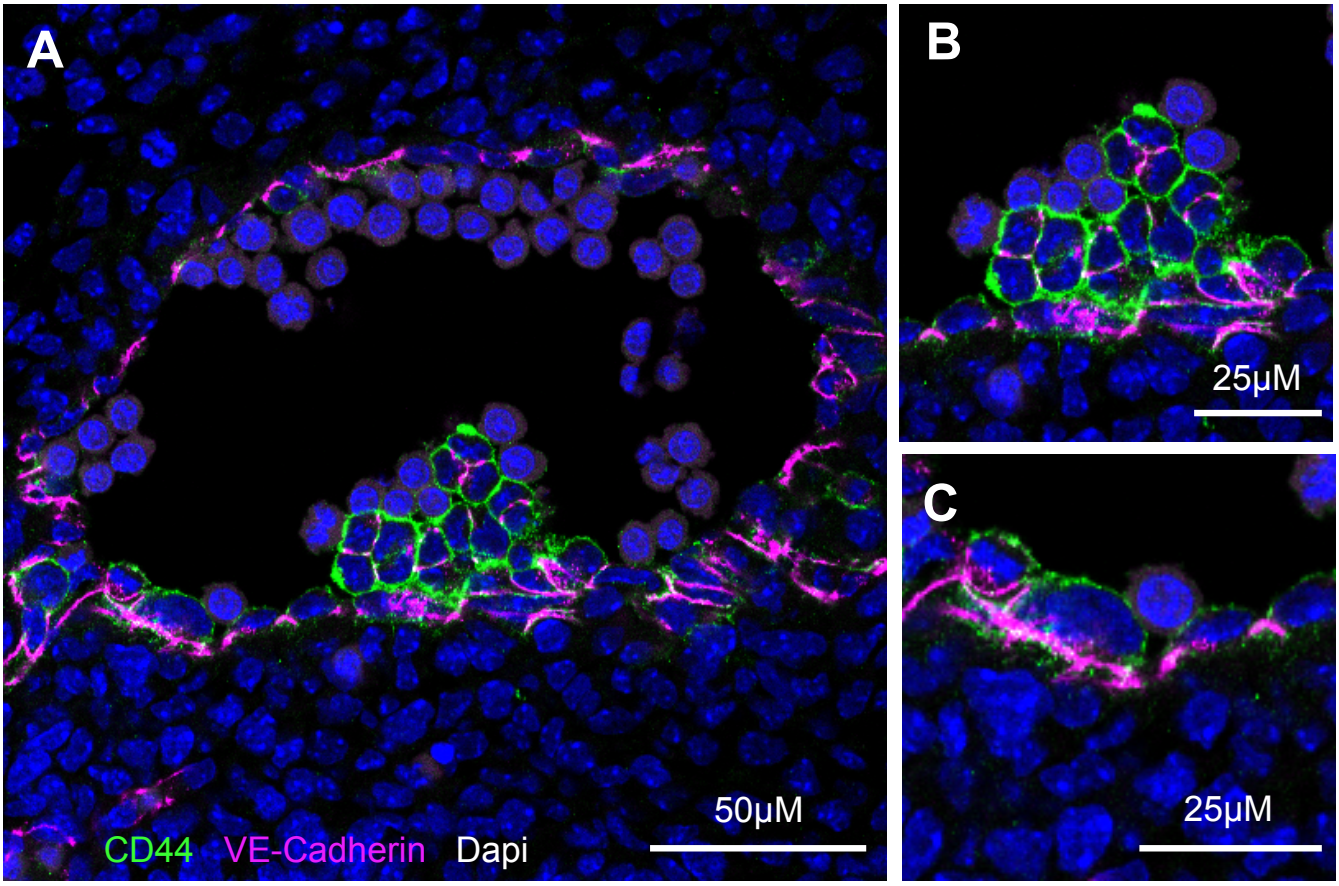


Figure 8: CD44 marks cells both haematopoietic clusters and endothelium

Immunofluorescence staining of 10μM cryosections of the AGM region of an E10 mouse embryo with CD44, VE-Cadherin and DAPI nuclear stain (A). Enlarged image of CD44 and VE-Cadherin staining of cells of a haematopoietic cluster (B). Enlarged image of a VE-Cadherin+/CD44+ cell attached to the endothelial wall (C).

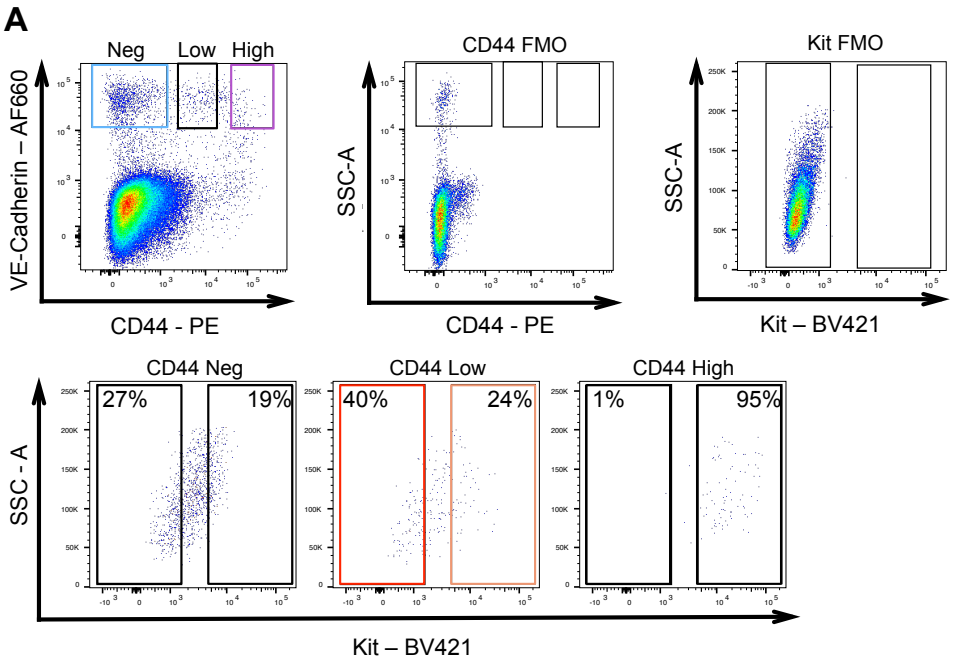
4.1.2 A subset of cells expressing CD44 lack expression of the haematopoietic marker c-Kit and are quiescent in nature

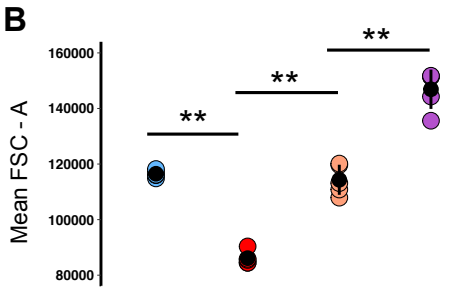
A marker of haematopoietic clusters in the mouse embryo is the stem cell factor receptor, c-Kit (CD117) (Yokomizo and Dzierzak, 2010). To understand if our VE-Cadherin+/CD44+ population overlapped with c-Kit expression we performed further FACS analysis on mid-gestation mouse embryos. We found that cells expressing high levels of CD44 also expressed high levels of c-Kit at the protein level (Fig. 9a). However, for cells with lower levels of CD44 expression only a subset expressed c-Kit (Fig. 9a). This could mean there is a portion of the VE-Cadherin+/CD44+ population that do not belong to a haematopoietic cluster.

Mean forward scatter is a flow cytometric property often used to evaluate cell size. By comparing the forward scatter to the expression of CD44 and c-Kit within the vascular endothelium (VE-Cadherin+) (Fig. 9b), we identified a significant difference of cell size between the CD44 populations based on a one-way ANOVA ($F(3, 16) = 142.01$, $p < 0.00001$) (Fig. 9c). Tukey HSD post-hoc tests found that cells with low expression of CD44 and no c-Kit were significantly smaller in size compared to both cells expressing no CD44 ($p < 0.01$) and cells with low levels of CD44 and c-Kit ($p < 0.01$). Additionally, cells with low CD44 expression and c-Kit were significantly smaller than cells expressing high levels of CD44 ($p < 0.01$). This difference in cell size indicated that these cell populations could be part of a transition process and may have many more molecular differences in addition to these physical dissimilarities. Therefore, for subsequent analysis we divided the AGM-

derived cells into four populations based on CD44 and c-Kit staining, namely, CD44Negative, CD44Low/Kit-, CD44Low/Kit+ and CD44High (Fig. 9a).

We went on to perform cell cycle analysis on the sorted populations using EDU incorporation and Hoechst DNA staining. We found that on average approximately 16% of CD44Negative cells and 25% of CD44High cells had entered S phase during the one hour of incubation with EDU compared to less than 5% of both CD44Low/Kit- and CD44Low/Kit+ populations (Fig. 10a). Overall, the vast majority of CD44Low/Kit- cells were found to be in G0/G1 phase indicating a quiescent phenotype. One-way ANOVAs were performed to compare the cell cycle status of the different populations and determine statistical significance. CD44Low/Kit- cells were found to have statistically reduced cycling compared to both CD44Negative and CD44High populations (Fig. 10b). Although not statistically significant a larger proportion of CD44Low/Kit+ cells were observed to be in G2 phase, which could indicate that these cells are primed and ready for proliferation. There are clear differences in the surface markers, size and cycling rates of these populations that led us to further investigate their cellular identities.





C

ANOVA

Cell size

	SS	df	MS	F
Between Groups	9.3 x 10^9	3	3.1 x10^9	142.01
Within Groups	3.4 x 10^8	16	2.2 x 10^7	
Total	9.6 x 10^9	19		

$p < 0.00001$

Figure 9: Not all CD44+ cells express c-Kit and vary in cell size

Flow cytometry analysis at E10 mouse AGMs reveals that a subset of CD44+ cells does not express the haematopoietic marker c-Kit. Here we show the expression of Kit within the populations of CD44Neg, CD44Low and CD44High cells based on CD44 and Kit fluorescence minus one (FMO) controls (A). Forward scatter area was compared across populations of cells expressing CD44 and c-Kit (B), n = 4 independent pools of E10 embryos. A one-way ANOVA was performed on the four independent experiments, followed by Tukey's HSD post-hoc tests to determine significance (C). Significant differences were identified between CD44Negative and CD44Low/Kit-, CD44Negative and CD44Low/Kit+, CD44Low/Kit- and CD44Low/Kit+, CD44Low/Kit- and CD44High, CD44Low/Kit+ and CD44High, with p-values < 0.01 **.

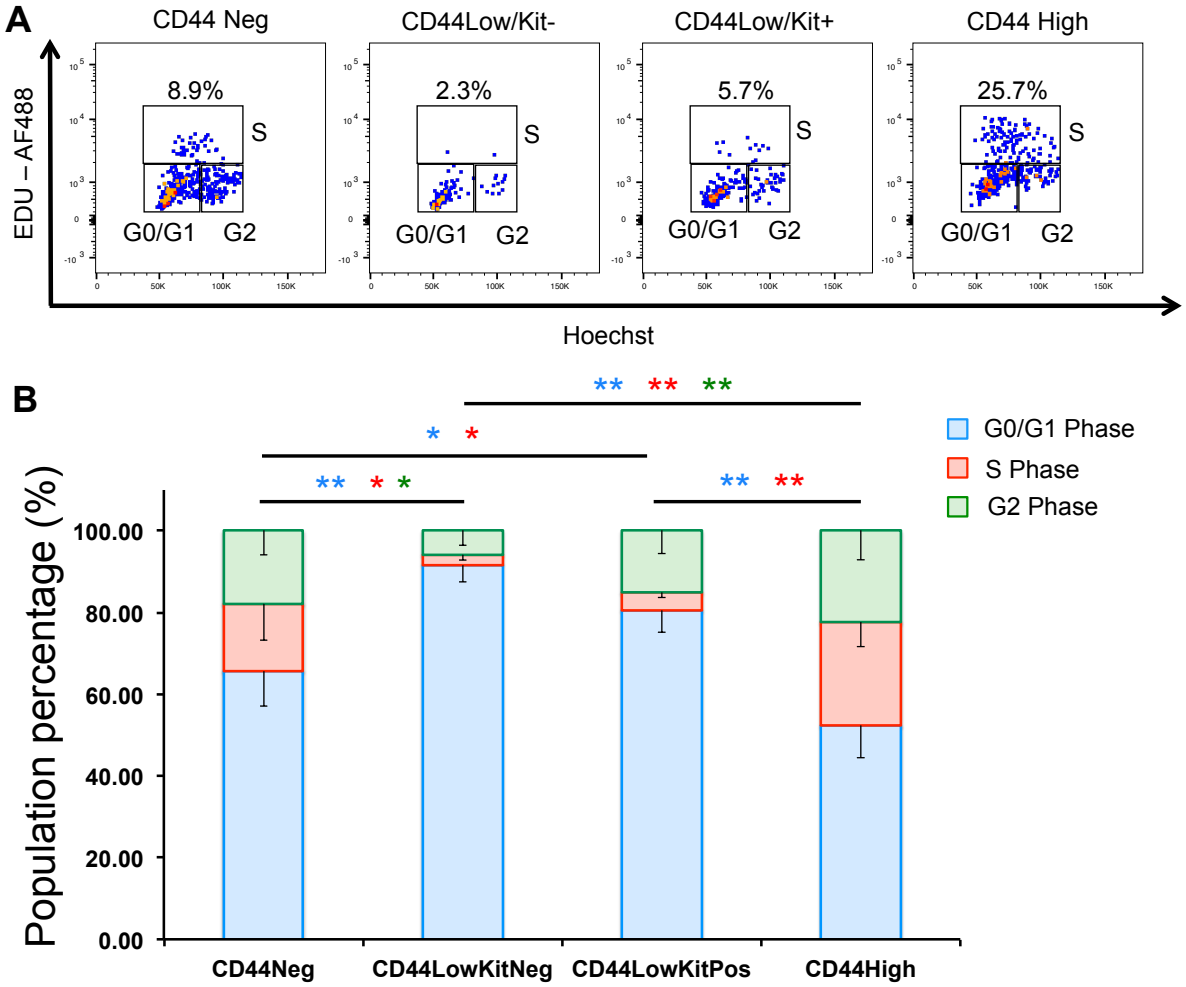


Figure 10: CD44+ cells cycle at different rates

Representative FACS plots of EDU and Hoechst staining on CD44Negative, CD44Low/Kit-, CD44Low/Kit+ and CD44High populations. Approximately 1000 cells from each population were sorted from pooled E10 mouse AGMs into 1.5mL micro-centrifuge tubes, fixed and re-analysed for EDU incorporation (A). One-way ANOVAs and Tukey's HSD post-hoc tests were used to evaluate the significance of differences in cell cycle status between CD44 populations (B). Bar plots represent the average of 4 independent experiments and the error bars are equal to the standard deviation, * $p < 0.05$ and ** $p < 0.01$.

4.1.3 Transcriptional analysis of the CD44 populations reveals a developmental progression from endothelial to haematopoietic identity

To investigate the molecular identities of the CD44 populations previously defined in the embryonic mouse vasculature we decided to perform single-cell quantitative PCR (sc-qPCR) using the Biomark HD platform. Primer sets were designed against 95 genes related to endothelial and haematopoietic development (Supplementary table 1). This technique minimises the background noise inherent to single cell analyses and enables the detection of low abundant genes such as transcription factors that drive changes in cell fate. In total we assessed the transcriptional profiles of 213 single cells across the four different populations previously defined (Fig. 11a). Unsupervised clustering identified five homogenous clusters with distinct transcriptional profiles. Four major groups were identified that closely resembled the populations defined by our sorting strategy and a smaller cluster of 7 cells was found that showed intermediate expression between the two CD44Low populations (Fig. 11a). Cells were FACS sorted from pooled AGMs derived from both E10 and E11 staged embryos, however no difference in transcription was observed between the time-points (Fig. 11a).

Of the four main populations identified, two appeared largely endothelial in nature (CD44 Negative and CD44Low/Kit-) with high expression of Tek (Tie2), Kdr (Flk-1) and Pecam1 (Cd31). Interestingly, one of the endothelial populations specifically expressed high levels of the TGF β inhibitors Smad6 and Smad7, which had previously been shown to promote the endothelial to haematopoietic transition *in vitro* (Vargel *et al.*, 2016). The CD44Low/Kit+

population showed dual expression of endothelial and haematopoietic genes with the up-regulation of Gfi1, Spi1, Bcl11a and Ikzf2. This profile is consistent with previous single cell data of pre-HSPCs (Zhou *et al.*, 2016). Similarly, the CD44Low/Kit+ cells showed simultaneous expression of the haematopoietic heptad of transcription factors (Gata2, Runx1, Lyl1, Erg, Fli1, Lmo2, Tal1), which is also associated with pre-HSPC identity (Bergiers *et al.*, 2018). Importantly, CD44Low/Kit+ cells also expressed high levels of Smad6 and Smad7 suggesting a link to the more endothelial CD44Low/Kit- population (Fig. 11a & 12b). Finally, CD44 High cells showed a significant down-regulation of endothelial gene expression while retaining high levels of haematopoietic transcripts. This is in line with a previous report that suggested that endothelial identity is rapidly lost upon haematopoietic commitment (Swiers *et al.*, 2013). Interestingly, the common haematopoietic marker CD41 (Itga2b) showed weak and stochastic expression in the pre-HSPC population while it's partner receptor CD61 (Itgb3) was expressed more highly and specifically in the CD44Low/Kit+ fraction (Fig. 11a & 12b).

Dimensionality reduction analysis using the t-SNE algorithm showed that although two endothelial populations were identified, the CD44 Low/Kit- population clustered closer to the pre-HSPCs (CD44 Low/Kit+) than to the other endothelial cell cluster (CD44 Negative) (Fig. 12a). Indeed, we identified distinct commonalities in gene expression between these two populations. By averaging the expression profiles of all single cells belonging to a particular cluster we could more easily visualise and compare profiles (Fig. 12b). The two CD44Low populations clearly shared expression of Smad6, Smad7,

Pde3a, Fbn1, CD44 and Emb indicating a developmental link. This suggested that by using CD44 expression we could distinguish endothelial cells primed for transition into pre-HSPCs from other vascular endothelium. By comparing the transcriptional data across these populations, we could identify increasing haematopoietic gene expression and decreasing endothelial gene expression as cells increase their surface levels of CD44 and c-Kit (Fig. 12b). Furthermore, the idea that the CD44Low/Kit- population represents haemogenic endothelium was further supported by the isolation of the small intermediate cell cluster that formed a transcriptional bridge between the two CD44Low groups. Importantly, it is in this small subset where expression of the key haematopoietic transcription factor, Runx1, is first detected (Fig. 12b).

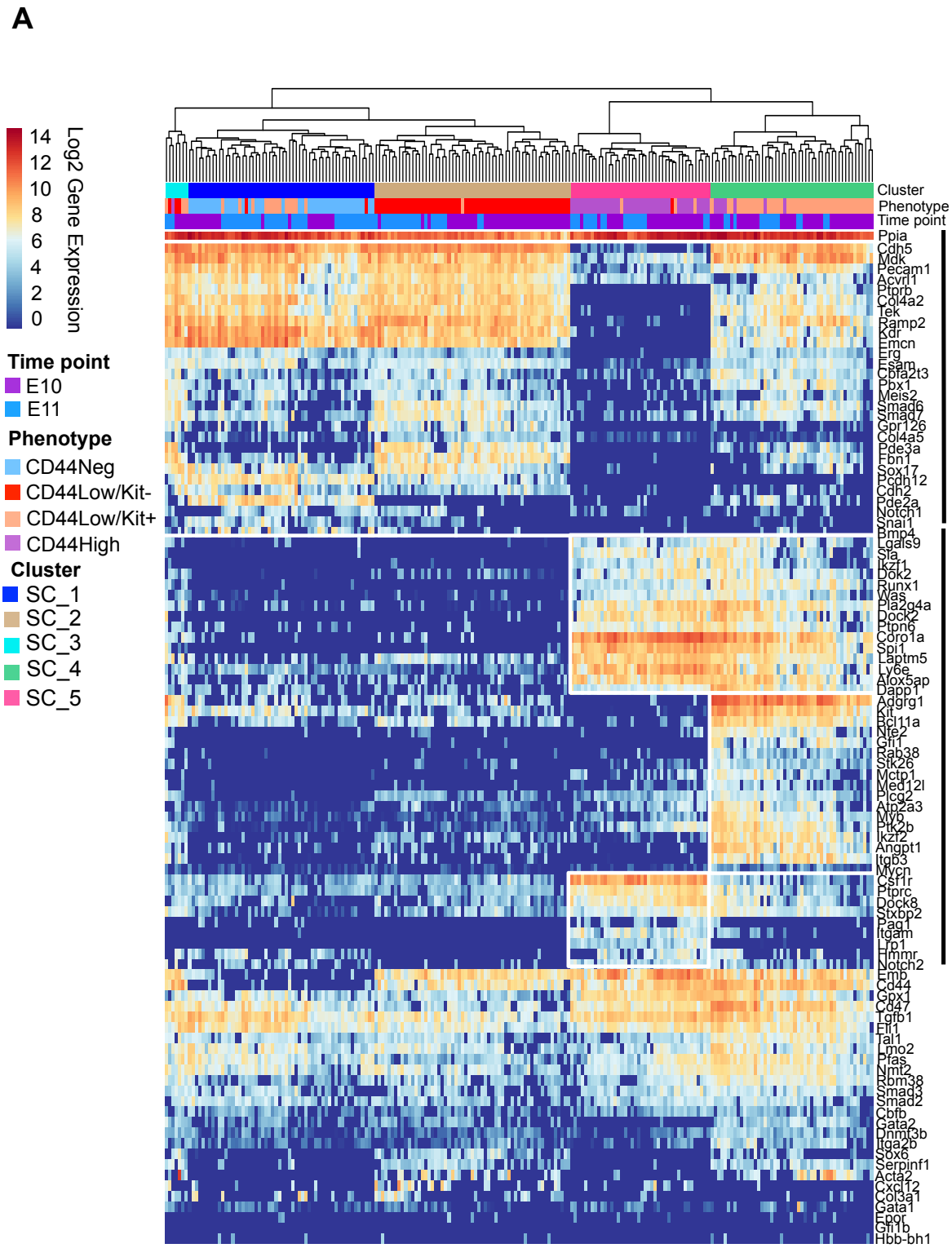


Figure 11: Single cell qPCR identifies five clusters with varying degrees of endothelial and haematopoietic gene expression

213 single cells from E10 and E11 mouse AGMs (time-point is indicated in figure by purple and blue markers) were FACS sorted directly into 96-well Bio-Rad PCR plates containing lysis buffer. Cells were sorted according to the phenotype detailed in the figure (either CD44Neg, CD44Low/Kit-, CD44Low/Kit+ or CD44High) and reverse transcribed and profiled based on 95 genes related to endothelial and haematopoietic development (see supplementary table 1). Unsupervised clustering identified five clusters, which closely correlated to their cell surface phenotype (A). Genes related to endothelial and haematopoietic processes are marked on the right-hand side.

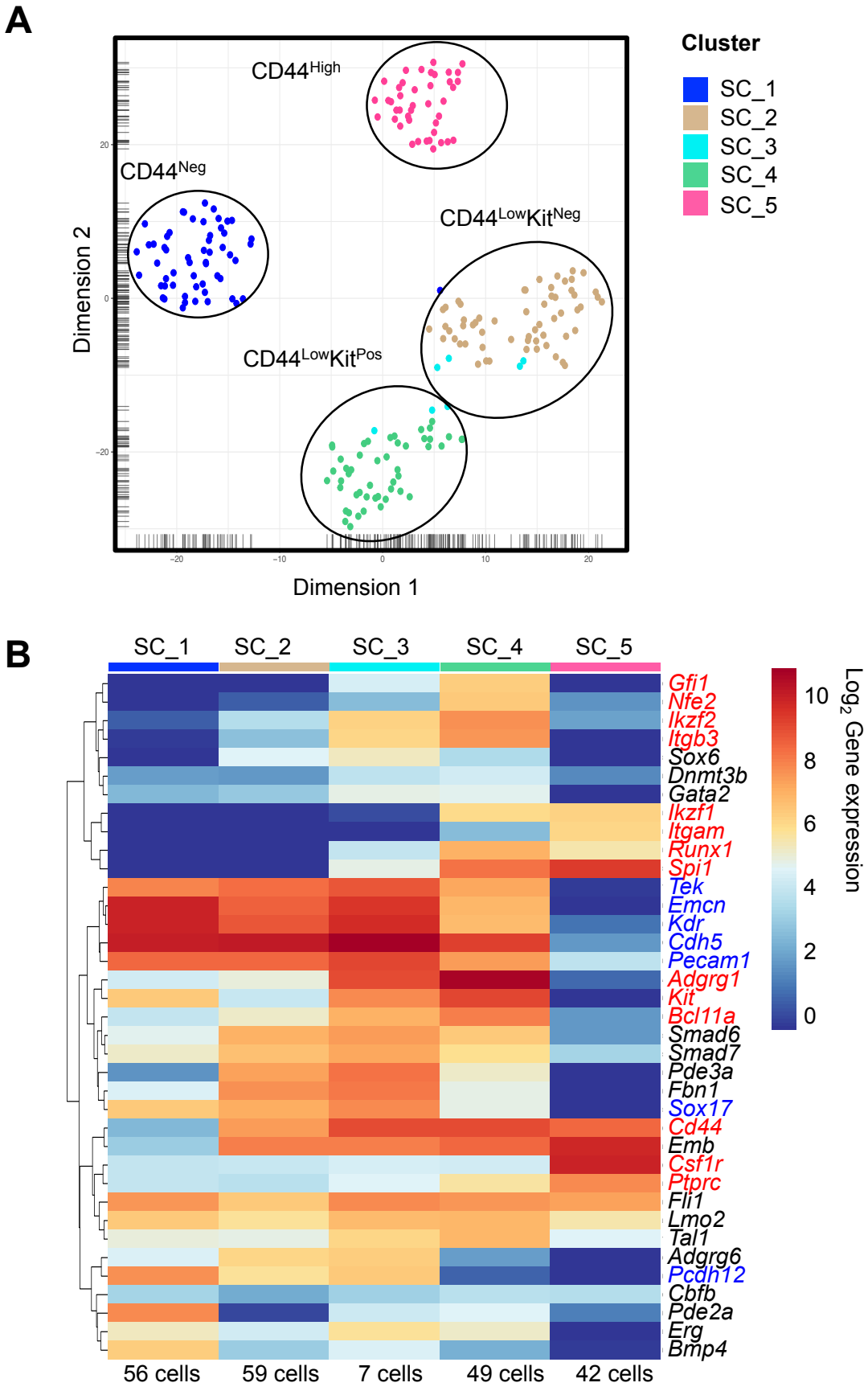

A
CD44High
CD44Neg
CD44LowKitNeg
CD44LowKitPos
Dimension 2
Dimension 1
Cluster
SC_1
SC_2
SC_3
SC_4
SC_5
B
SC_1
SC_2
SC_3
SC_4
SC_5
Gfi1
Nfe2
Ikzf2
Itgb3
Sox6
Dnmt3b
Gata2
Ikzf1
Itgam
Runx1
Spi1
Tek
Emcn
Kdr
Cdh5
Pecam1
Adgrg1
Kit
Bcl11a
Smad6
Smad7
Pde3a
Fbn1
Sox17
Cd44
Emb
Csf1r
Ptprc
Fli1
Lmo2
Tal1
Adgrg6
Pcdh12
Cbfb
Pde2a
Erg
Bmp4
56 cells
59 cells
7 cells
49 cells
42 cells
Log2 Gene expression
10
8
6
4
2
0

Figure 12: CD44 and KIT can be used to distinguish homogenous populations with increasing haematopoietic identity

The 213 cells sorted and profiled in figure 11 were used in a t-SNE dimension reduction analysis in (A). This analysis shows that cell surface expression of CD44 and c-Kit can be used to distinguish homogenous populations of cells. Average expression of the cells sorted and profiled in figure 11 across a subset of endothelial and haematopoietic genes are displayed in a simplified heatmap. Groups were identified based on unsupervised clustering (B). Cell numbers are indicated below each group. Key endothelial genes are marked in blue and key haematopoietic genes are marked in red.

4.1.4 CD44 can be used to isolate haemogenic endothelium

Previous work on EHT characterised several intermediary populations between haemogenic endothelium and definitive HSCs (pro-HSC, pre-HSC type I and pre-HSC type II) based on the cell surface markers VE-Cadherin, CD41, CD43 and CD45 (Rybtsov *et al.*, 2011, 2014). These populations required differing maturation times in culture to form definitive HSCs with transplantation potential. In order to place the CD44+ populations in context we compared them to the transcriptional profiles of cells sorted from these previously defined populations using our sc-qPCR protocol (Fig. 13a).

Interestingly we identified two separate pro-HSC clusters, which displayed different gene expression profiles, based on the 95 genes that we assessed (Supplementary table 1) (Fig. 13c). This heterogeneity brings into doubt whether there is a clear distinction between pro-HSCs and pre-HSCs type I cells. As expected, the pro-HSCs displayed lower levels of haematopoietic gene expression and higher levels of endothelial gene expression (Fig. 13c). We then combined analysis of the pro-HSCs, pre-HSCs type I and type II with the populations defined by CD44 and c-Kit. t-SNE clustering revealed strong similarities between the CD44High and pre-HSC type II cells as well as the

CD44Low/Kit+ and pre-HSC type I population (Fig. 13d). Interestingly the CD44Low/Kit- population clustered with the more endothelial pro-HSC group lending further support to the idea that this population in fact represents haemogenic endothelium (Fig. 13d). When combining all markers together with multi-colour flow cytometry, we found that on average 98% of pre-HSCs type II, 93% of pre-HSCs type I and 80% of pro-HSCs expressed CD44 on their cell surface (Fig. 13b). From our results it appears that CD44 cell surface expression can be used to isolate more homogenous populations in blood cell development and to distinguish the earliest stage of EHT - haemogenic endothelium.

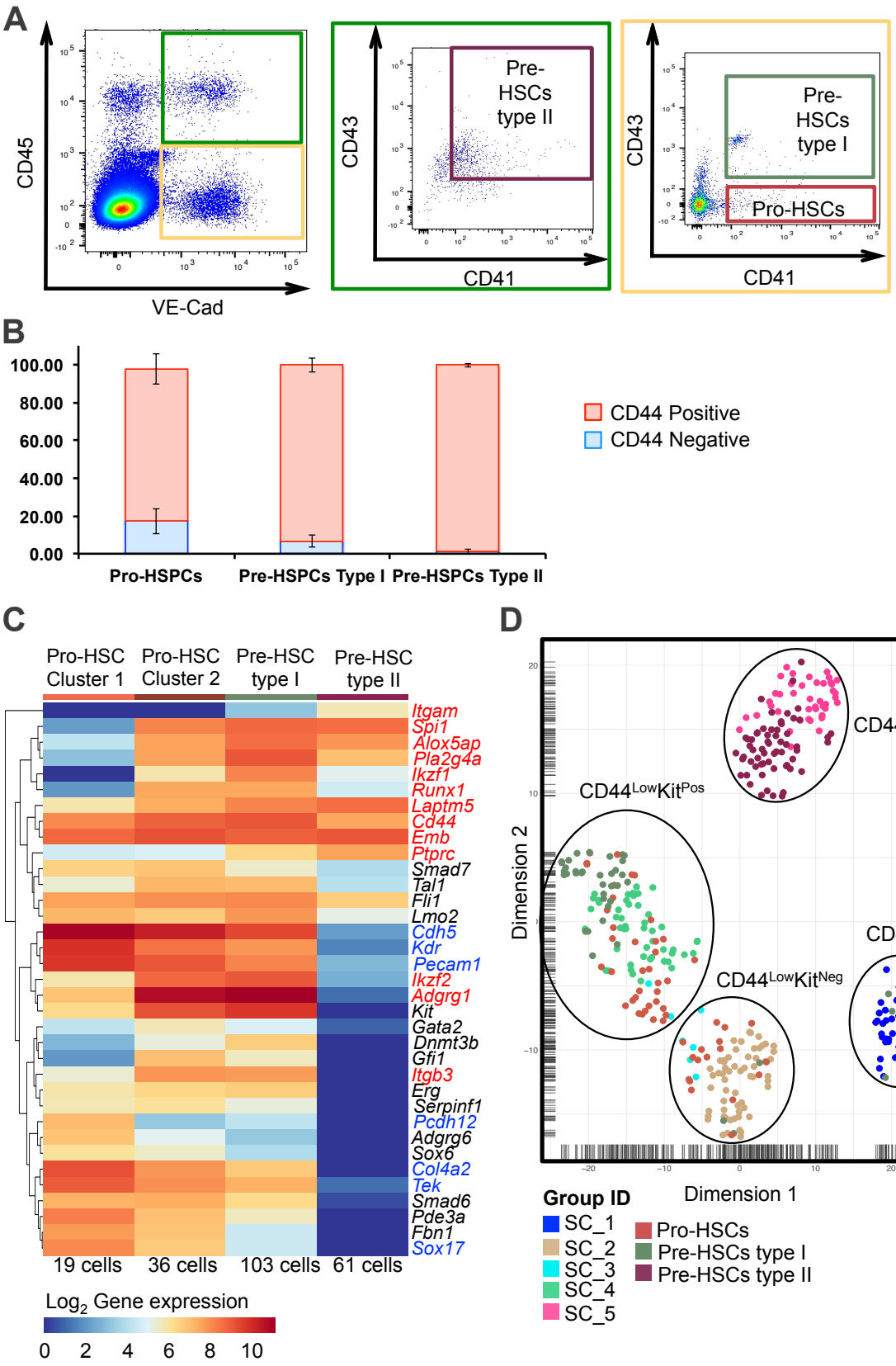

A
CD45
VE-Cad
CD43
CD41
Pre-
HSCs
type II
Pre-
HSCs
type I
Pro-HSCs
B
100.00
80.00
60.00
40.00
20.00
0.00
Pro-HSPCs
Pre-HSPCs Type I
Pre-HSPCs Type II
CD44 Positive
CD44 Negative
C
Pro-HSC
Cluster 1
Pro-HSC
Cluster 2
Pre-HSC
type I
Pre-HSC
type II
Itgam
Spi1
Alox5ap
Pla2g4a
Ikzf1
Runx1
Laptm5
Cd44
Emb
Ptprc
Smad7
Tal1
Fli1
Lmo2
Cdh5
Kdr
Pecam1
Ikzf2
Adgrg1
Kit
Gata2
Dnmt3b
Gfi1
Itgb3
Erg
Serpinf1
Pcdh12
Adgrg6
Sox6
Col4a2
Tek
Smad6
Pde3a
Fbn1
Sox17
19 cells
36 cells
103 cells
61 cells
Log2 Gene expression
0 2 4 6 8 10
D
CD44High
CD44LowKitPos
CD44Neg
CD44LowKitNeg
Dimension 2
Dimension 1
Group ID
SC_1
SC_2
SC_3
SC_4
SC_5
Pro-HSCs
Pre-HSCs type I
Pre-HSCs type II

Figure 13: Comparison of CD44 expression with previous markers of EHT

FACS plots showing the sorting gates used to isolate pro-HSCs (VE-Cad+/CD45-/CD41+/CD43-), pre-HSCs type I (VE-Cad+/CD45-/CD41+/CD43+) and type II (VE-Cad+/CD45+/CD41+/CD43+) populations from AGM tissue (A). Bar graph showing the percentage of the pro-HSCs, pre-HSCs type I and type II that expressed CD44 on their cell surface in the AGM of E10 mouse embryos (B). Graph indicates the average of n = 3 independent experiments and error bars show standard deviation. Cells were sorted based on the gating in (A) from E9.5 and E11 mouse AGM into Bio-Rad 96-well plates for single cell qPCR profiling. Average expression of a subset of endothelial and haematopoietic genes is displayed in the heatmap based on the groups identified by unsupervised clustering (C). Cell numbers are indicated below each group. t-SNE dimension reduction analysis combines the clustering of pro-HSCs, pre-HSCs type I and type II with the CD44-defined populations from figure 12 based on single cell qPCR expression profiles (D).

4.1.5 RNA sequencing of the CD44 populations reveals the earliest changes in EHT and identifies potential new regulators of haematopoietic development

So far, we have established that CD44 cell surface expression when combined with c-Kit can be used to track the progression of EHT. To better characterise the different stages of this transition we performed a further RNA sequencing experiment; this was done with the help of our lab manager Kerstin Ganter and the sequencing facility at EMBL Heidelberg. To maximise the purity of our cell populations and the detection of low abundant genes we opted to analyse 25 cell samples comparing CD44 Negative, CD44Low/Kit- and CD44Low/Kit+ and CD44 High populations using the SmartSeq2 protocol (Picelli *et al.*, 2014). Overall, our results built upon what we had previously observed using single cell qPCR, showing that haematopoietic identity increases with CD44 and c-Kit cell surface expression.

Differential expression analysis comparing the CD44Negative and CD44 Low/Kit- populations using the DESeq2 package in R identified distinct transcriptional differences between these two endothelial populations

(Appendix I). In total 452 genes were found to be differentially expressed with an adjusted p-value < 0.01 and a log fold change greater than two (Fig. 14a & Supplementary table 2). Amongst these significantly different genes were 48 transcription factors 34 of which were up-regulated in the CD44Low/Kit- population (Fig. 14b & Supplementary table 2). This is interesting as it indicates that this small, quiescent group of cells could be in the midst of significant changes to its expression profile, and supports the idea that these cells are primed for transition. While within this group of regulators we see a number of factors already associated with haematopoietic development such as *Spi1*, *Ikzf1*, *Ikzf2* and *Gfi1*, there are also many transcription factors that have not previously been implicated in embryonic haematopoiesis (Fig. 14b). Given the utility of transcription factors in reprogramming studies, it would be interesting to further investigate the role of these factors in EHT.

Exploration of signalling pathway dynamics revealed a slight up-regulation in the Hedgehog ligand Desert hedgehog (*Dhh*) but down-regulation in the Hedgehog target genes *Myc*, Ccnd1 and *Snai1* in the CD44Low/Kit- population (Fig. 15a). Both Hedgehog and Bmp signalling are known to pattern the dorsal aorta prior to EHT (Wilkinson et al., 2009). We observed a down-regulation of *Bmp4* and an up-regulation of several inhibitors of the Bmp/Tgfβ signalling pathway, namely, *Smad6*, *Smad7*, *Bmper*, *Id1* and *Id2* (Fig. 15a). Previous research has shown that although Bmp signalling is necessary for HSC development in the AGM the pathway must be down-regulated for the emergence of intra-aortic clusters (McGarvey et al., 2017). As such, the strong inhibition of Bmp signalling we observe in the

CD44Low/Kit- population could indicate that this group of cells is preparing for EHT. In addition, this population shows a strong up-regulation of genes related to Notch signalling including *Notch1*, *Dll4*, *Jag1*, *Hey1*, and *Hey2*. This is also consistent with previous reports of the critical role Notch signalling plays in EHT (Burns et al., 2005). In particular, the high levels of the notch ligand *Jag1* displayed in CD44low/Kit- cells compared to CD44Negative is indicative of a population that is diverging from the *Dll4* responsive arterial fate, towards the haematopoietic lineage (Gamma-Norton et al., 2015). The close association of our sequencing results to what is known about EHT in the literature provides strong support for our hypothesis that we can use CD44 surface expression to isolate both haemogenic endothelial cells and pre-HSPCs.

In conjunction with Kiran Raosaheb Patil and his PhD student Katharina Zirngibl we used the RNA sequencing data to compare the metabolic status of our CD44Negative and CD44Low/Kit- populations. For this analysis 1605 genes were identified using the Wald test to be differentially expressed between the CD44Negative and CD44Low/Kit- populations (p-value < 0.01). Amongst these genes, they found a 1.32 fold enrichment (p-value < 0.05, Fisher's Exact Test) of metabolic genes based on reporter metabolite analysis. Overall, they observed a down-regulation of glycolysis and TCA cycle related genes and an up-regulation of autophagy related genes in the CD44Low/Kit- population (Fig. 16a-c). Both glycolysis and autophagy are known to be used by adult HSCs to maintain quiescence (Takubo et al., 2013; Ho et al., 2017). Furthermore, these results are consistent with our previous

observations that cells from this population are both smaller in size and proliferate less.

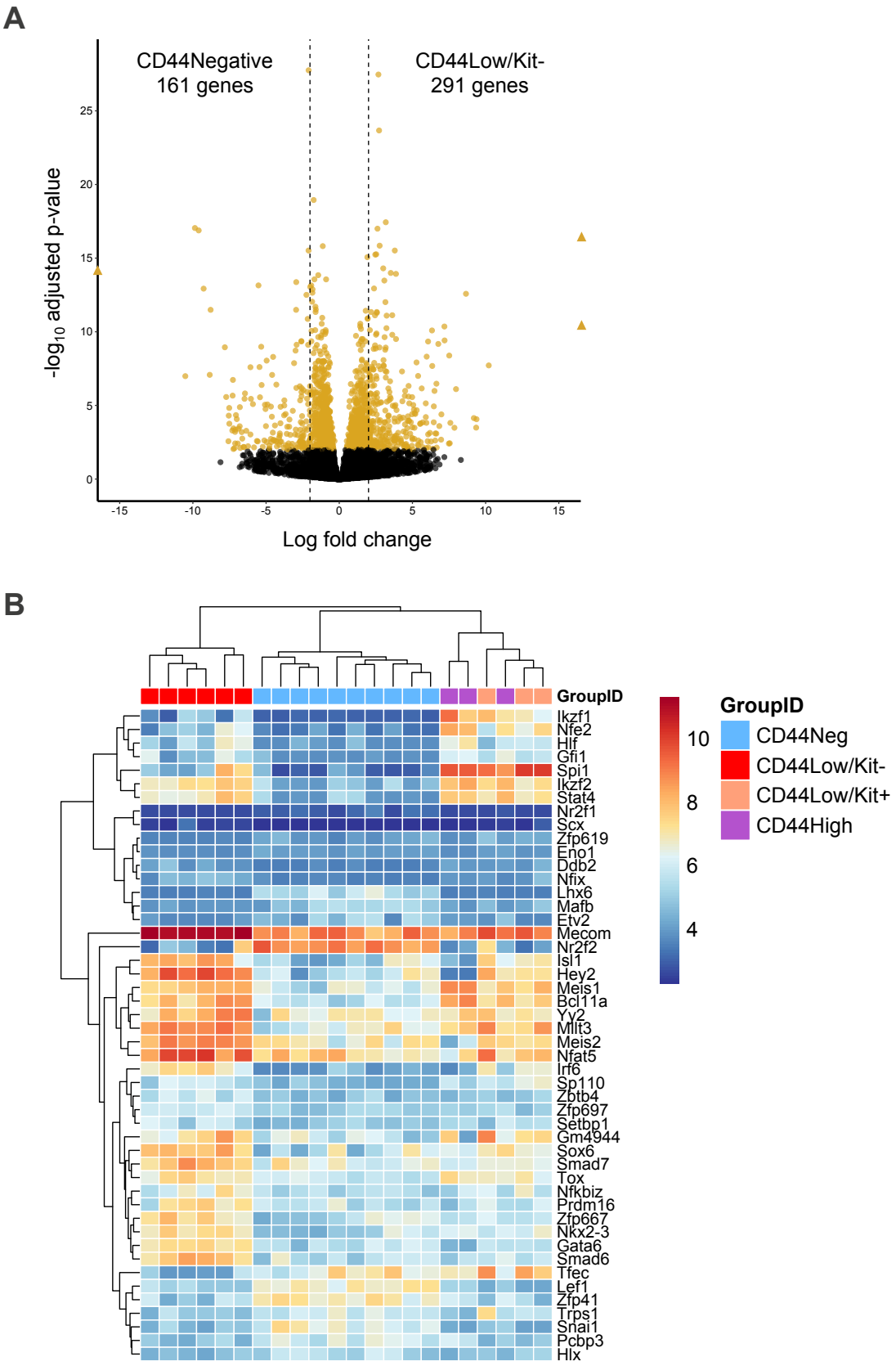


Figure 14: RNA sequencing reveals distinct differences between CD44Neg and CD44Low/Kit- endothelial populations

Cells were bulk sorted into PCR strip tubes based on their expression of VE-Cadherin, CD44 and c-Kit from AGM tissue derived from either E9.5, E10 or E11 mouse embryos. 25 cells were sorted per sample resulting in ten replicates of CD44Neg, six replicates of CD44Low/Kit-, three replicates of CD44Low/Kit+ and three replicates of CD44High, based on gating in Figure 9a. The volcano plot represents the differential gene expression between CD44Negative and CD44Low/Kit- populations (A). Significant genes with an adjusted p-value < 0.01 are shown in yellow and dots located above or below the dotted lines represent genes with a greater than 2 log fold change. Heatmap showing the expression of 48 transcription factors differentially expressed between CD44Negative and CD44Low/Kit- populations (B).

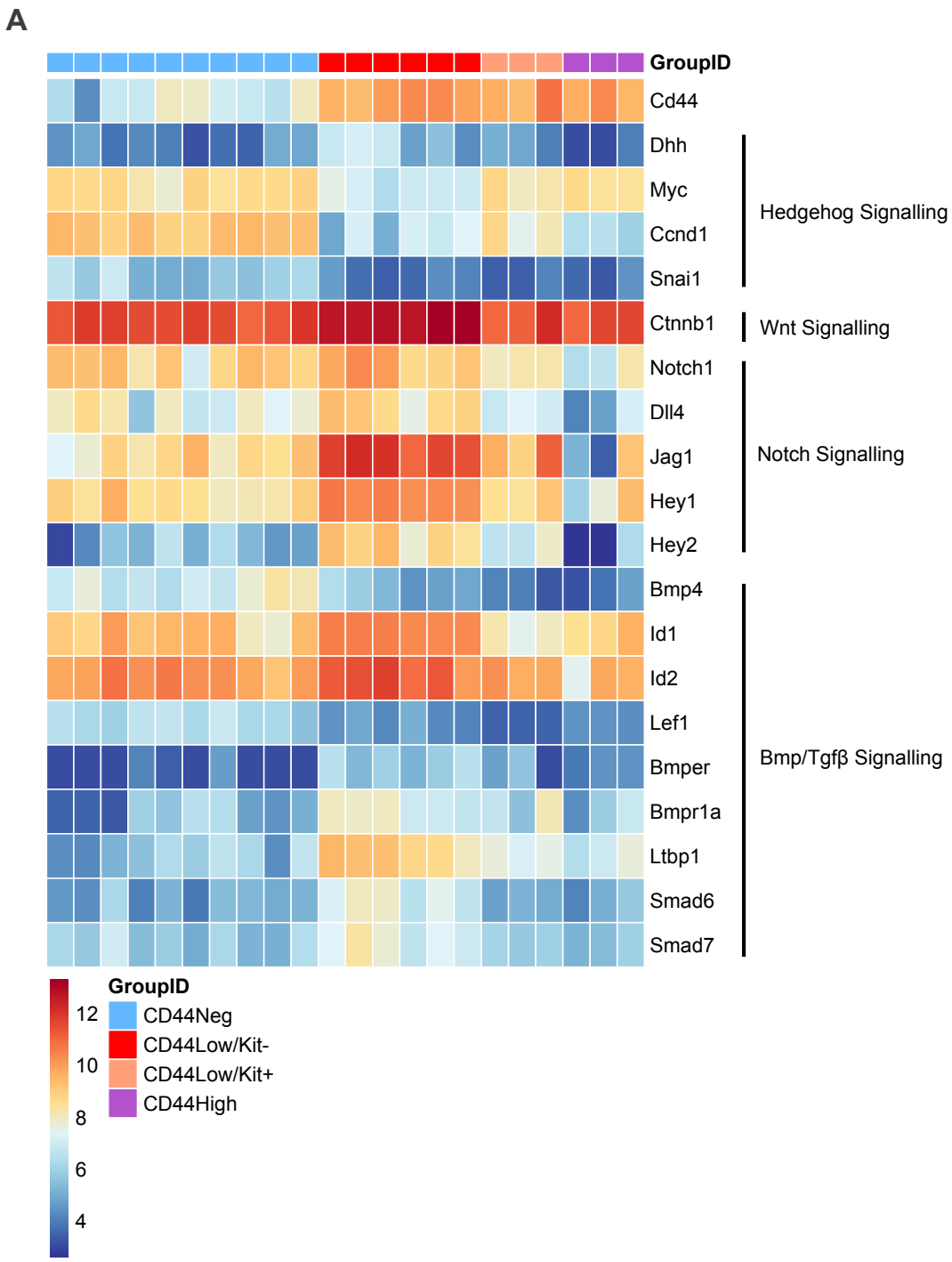


Figure 15: RNA sequencing reveals convergence of signalling pathways in CD44Low/Kit- cells

Further analysis of 25-cell bulk RNA sequencing from Figure 14. Heatmap showing changes in expression of key genes related to Hedgehog, Wnt, Notch and Bmp/Tgfβ signalling across the four different CD44 populations (A).

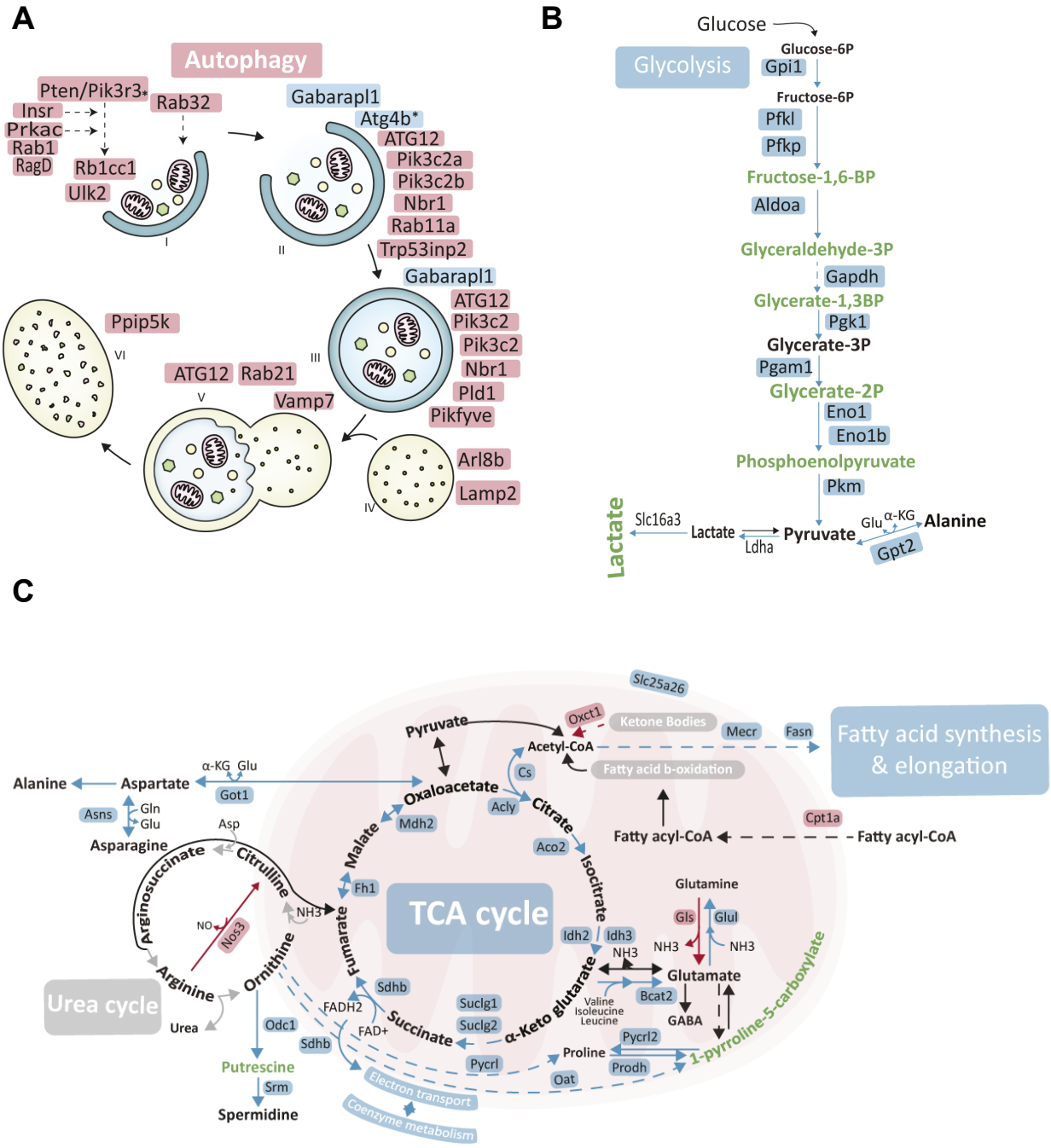

A
Autophagy
Pten/Pik3r3*
Insr
Prkac
Rab1
RagD
Rab32
Rb1cc1
Ulk2
I
Gabarapl1
Atg4b*
ATG12
Pik3c2a
Pik3c2b
Nbr1
Rab11a
Trp53inp2
II
Gabarapl1
ATG12
Pik3c2
Pik3c2
Nbr1
Pld1
Pikfyve
III
Ppip5k
VI
ATG12
Rab21
V
Vamp7
Arl8b
Lamp2
IV
B
Glucose
Glycolysis
Glucose-6P
Gpi1
Fructose-6P
Pfkl
Pfkp
Fructose-1,6-BP
Aldoa
Glyceraldehyde-3P
Gapdh
Glycerate-1,3BP
Pgk1
Glycerate-3P
Pgam1
Glycerate-2P
Eno1
Eno1b
Phosphoenolpyruvate
Pkm
Lactate
Slc16a3
Lactate
Pyruvate
Ldha
Glu
α-KG
Alanine
Gpt2
C
Slc25a26
Pyruvate
Oxct1
Ketone Bodies
Mecr
Fasn
Fatty acid synthesis & elongation
Acetyl-CoA
Fatty acid b-oxidation
Cs
Acly
Citrate
Oxaloacetate
Malate
Mdh2
Aco2
Isocitrate
Fatty acyl-CoA
Cpt1a
Fatty acyl-CoA
Alanine
Aspartate
α-KG
Glu
Got1
Asns
Gln
Glu
Asparagine
Asp
Argininosuccinate
Citrulline
NO
Nos3
NH3
Fh1
Fumarate
TCA cycle
Glutamine
Gls
Glul
Idh2
Idh3
NH3
NH3
NH3
Glutamate
Arginine
Ornithine
Urea cycle
Urea
Odc1
Sdhb
FADH2
FAD+
Sdhb
Succinate
Suclg1
Suclg2
α-Keto glutarate
Valine
Isoleucine
Leucine
Bcat2
GABA
Pycrl2
Proline
Prodh
1-pyrroline-5-carboxylate
Pycrl
Oat
Putrescine
Srm
Spermidine
Electron transport
Coenzyme metabolism

Figure 16: Comparison of CD44Negative and CD44Low/Kit- reveals differences in metabolic profile

Metabolic gene enrichment analysis was performed on 25-cell bulk RNA sequencing data from Figure 14, comparing the CD44Neg and CD44Low/Kit- samples. This figure summarises the key metabolic nodes and pathways enriched in differentially expressed genes between CD44Negative and CD44Low/Kit- populations. These were selected based on reporter metabolite analysis. Up-regulated genes are marked in red and down-regulated genes in blue. Genes related to autophagy are depicted in (A) where the genes associated with autophagosome formation are shown to be up-regulated and remain so as the autophagosome fuses to the lysosome to create an autolysosome and the cell products within are enzymatically digested. The glycolysis pathway is depicted in (B), where glucose is broken down into pyruvate to release energy for the cell. The enzymes involved in this pathway are down-regulated between these two populations. Finally, the TCA cycle is depicted in (C) where energy is released from the cell through the oxidation of acetyl-CoA. Analysis for this figure was done in Kiran Raosaheb Patil's lab by his PhD student Katharina Zirngibl and the figure was made in collaboration with scientific illustrator Gisela Luz.

4.2 CD44 is a regulator of EHT

4.2.1 Cells expressing CD44 and Kit can robustly generate haematopoietic colonies *ex vivo* and differentiate into all the blood cell lineages

In order to verify the haematopoietic potential of our CD44+ populations we performed OP9 co-culturing assays. Firstly, single cells were sorted directly onto a 96 well plate of confluent OP9 stromal cells in a media conducive to haematopoietic development and grown for three days before scoring for round cell colony formation (Fig. 17a). At single cell frequency we observed an average of approximately 11% growth from the CD44 High population and an average of 42% growth from the CD44Low/Kit+ population (Fig. 17b). This difference was found to be significant ($p = 0.025$) and is in line with the highly proliferative nature of pre-HSPCs. It also reveals a significant difference in the

proliferative potential between early type I pre-HSPCs and more mature type II pre-HSPCs. No round cell colonies were generated at the single cell level from the CD44 Negative and CD44Low/Kit- populations (Fig. 17b). We further tested the haematopoietic potential of these two endothelial-like populations at a higher cell density, plating 300 cells per well of a 96 well plate (Fig. 17c). Although we observed some haematopoietic cell growth from the CD44Low/Kit- population, while we never detected any from the CD44Negative one, this observation was not statistically different based on a Chi-square test for independence ($p= 0.052$) (Fig. 17d). More repetitions are needed to prove that the CD44Low/Kit- population has a higher haematopoietic capacity than the CD44Negative endothelial cells.

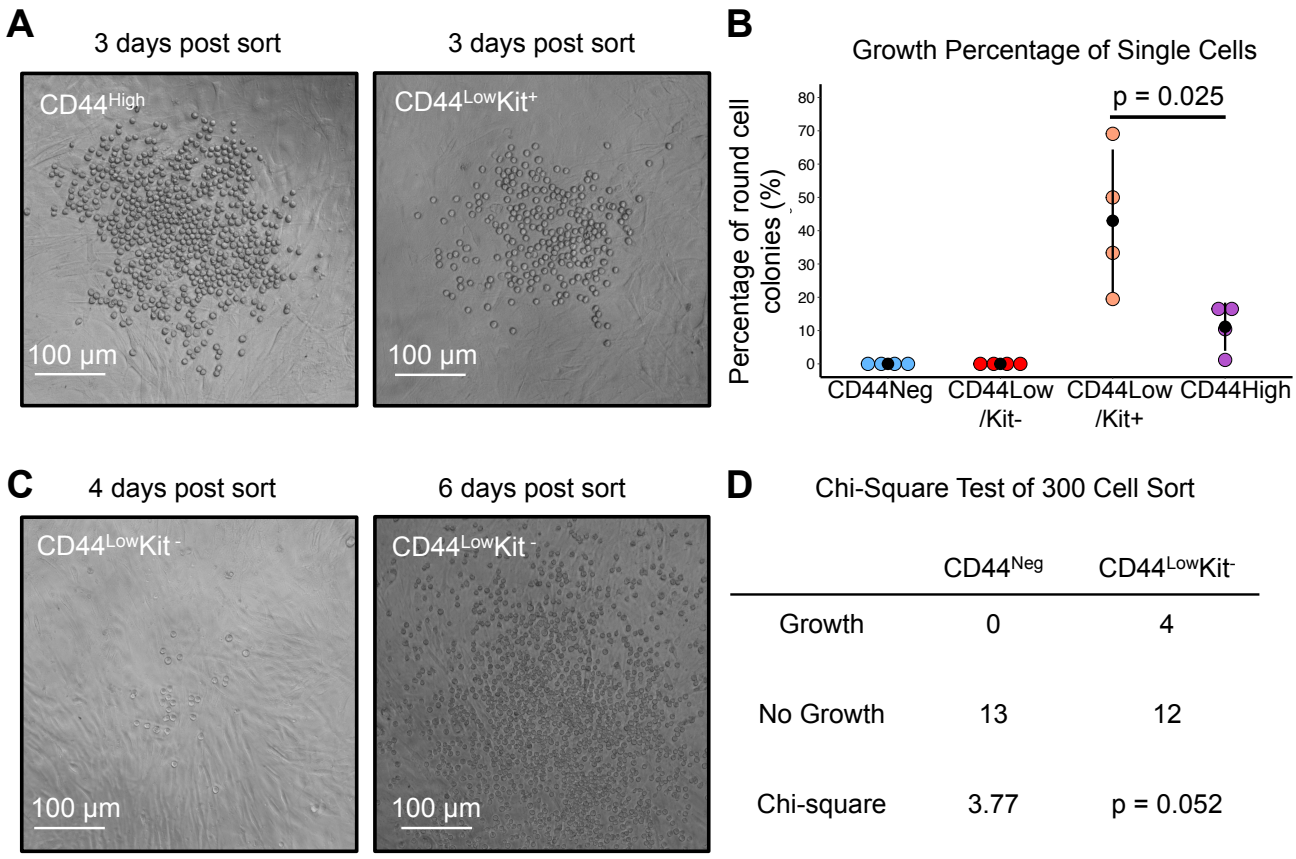


Figure 17: CD44+ cells show haematopoietic potential *ex vivo*

Single cells were FACS sorted from CD44Neg, CD44Low/Kit-, CD44Low/Kit+ or CD44High populations based on gating in Figure 9a. Cells were sorted directly onto OP9 stromal layers in 96-well culture plates from the AGM tissue of E10 mouse embryos. Round cell colonies were visualised and counted after three days incubation (A). Percentage of positive wells was quantified for four independent experiments generated from pooled E10 mouse AGMs (B). Significance was determined by a two-tailed, paired Student's T test. Three hundred CD44Negative and CD44Low/Kit- cells were sorted onto an OP9 stromal layer and quantified after six days in culture (C). A chi-square test for independence was performed between the two conditions (D).

We further used methocult colony forming unit (CFU) assays to assess the ability of CD44+ populations to contribute to erythroid and myeloid blood cell lineages. Both CD44High and CD44Low/Kit+ populations readily generated both erythroid, myeloid or mixed colonies *ex vivo* (Fig. 18a). One hundred cells were first co-cultured with OP9 cells for three days. As with the single cell assay more colonies were generated from the CD44Low/Kit+ population than the CD44High (Fig. 18b). These cells were then transferred to a methylcellulose-based culture and grown for a further seven days. The CD44Low/Kit+ population produced more erythroid and mixed colonies overall (Fig. 18c). However, the differentiation potential of CD44High cells appeared to be higher than CD44Low/Kit+, if the number of round cell colonies initially generated on the OP9 stromal layer was taken into account (Fig. 18d). This could be linked to the more mature state of the CD44High population. Further co-culturing experiments were performed using OP9 and OP9-DL1 stromal cells in conjunction with lymphocyte inductive cytokines (IL-7 and Flt-3) to assess the potential of CD44High cells to contribute to the lymphocyte lineage. CD44High cells were cultured for 21 days prior to flow cytometric analysis of cell surface marker expression. CD44High cells were found to be capable of generating CD19-expressing B cells and CD4+/CD8a+ T cells *ex vivo*, as well as a small proportion of CD11b expressing macrophages (Fig. 18e). These results indicate that cells derived from the AGM and marked with the CD44 cell surface receptor possess multi-potent haematopoietic potential.

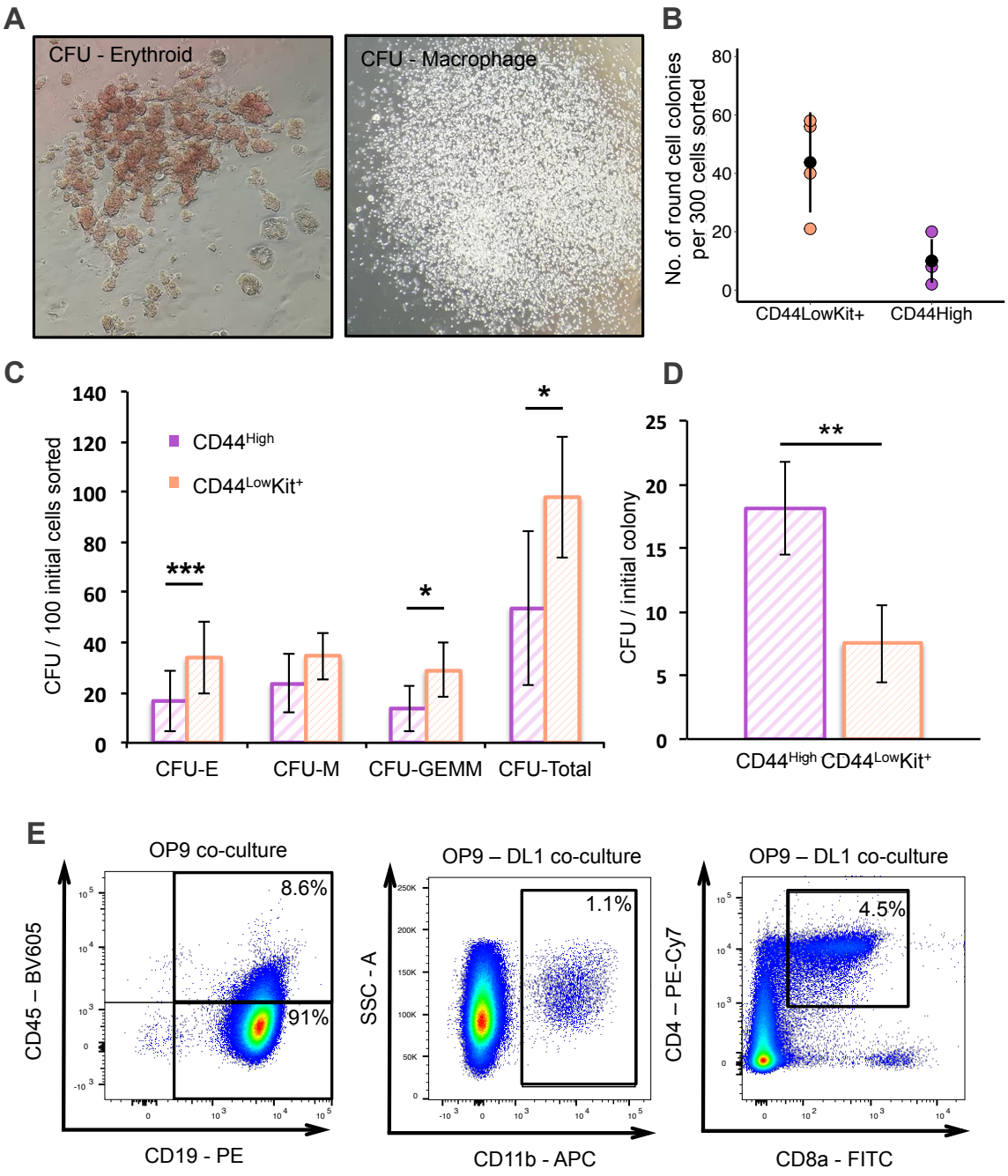

A
CFU - Erythroid
CFU - Macrophage
B
No. of round cell colonies per 300 cells sorted
80
60
40
20
0
CD44LowKit+
CD44High
C
CFU / 100 initial cells sorted
140
120
100
80
60
40
20
0
CD44High
CD44LowKit+
***
*
*
CFU-E
CFU-M
CFU-GEMM
CFU-Total
D
CFU / initial colony
25
20
15
10
5
0
**
CD44High
CD44LowKit+
E
OP9 co-culture
8.6%
91%
CD45 – BV605
CD19 - PE
OP9 – DL1 co-culture
1.1%
SSC - A
CD11b - APC
OP9 – DL1 co-culture
4.5%
CD4 – PE-Cy7
CD8a - FITC

Figure 18: CD44+ cells possess multi-lineage potential

Cells were first FACS sorted from the AGMs of E10 mouse embryos directly onto OP9 stromal layers on 96-well culture plates before being transferred and cultured further in methylcellulose. CFU assays were visualised and counted after three days of OP9 co-culture and seven days in methylcellulose-based medium (A). Three wells with 100 cells each of CD44High and CD44Low/Kit+ populations were plated on OP9 stromal cells and the number of colonies quantified after three days before transferring cells into methylcellulose (B). Erythroid, Myeloid and mixed colonies were quantified after a total of 10 days in culture (C). The number of CFUs generated relative to the number of colonies originally generated on OP9 was also quantified (D). These Bar graphs are the result of four independent experiments and the error bars represent standard deviation. Comparisons between CD44High and CD44Low/Kit+ populations were made using a two-tailed paired Student's T-test where * indicates p-value < 0.05, ** indicates p-value < 0.01 and *** indicates p-value < 0.001. CD44High and CD44Low/Kit+ cells were also cultured on OP9 and OP9-Dl1 stromal cells to generate lymphocyte colonies. Representative FACS plots of cells after 21 days in culture showing expression of B-cell marker CD19, macrophage marker CD11b and T cell markers CD4 and CD8a (E). For this experiment n = 2.

4.2.2 Runx1 is not required for the formation of haemogenic endothelium but for the progression to the pre-HSPC type I stage

Runx1 is well characterised as a critical driver of the definitive haematopoietic program, acting to down-regulate endothelial identity through the activation of its targets Gfi1 and Gfi1b (Lancrin *et al.*, 2012). In the absence of this transcription factor no haematopoietic clusters are formed in the embryonic vasculature (Yokomizo et al., 2001). To understand precisely when Runx1 is needed within our CD44+ populations we performed flow cytometric analysis and transcriptional profiling in a Runx1 knockout mouse model (Fig. 19). Staining of Runx1 knockout embryos with VE-Cadherin, CD44 and c-Kit revealed a loss of the pre-HSPC type I (CD44Low/Kit+) and pre-HSPC type II (CD44High) populations (Fig. 19a). This was confirmed by single cell qPCR where cells from the knockout embryos only cluster with the CD44Negative and CD44Low/Kit- populations (Fig. 19b & c). Interestingly, we found no transcriptional differences in the CD44 Negative and CD44Low/kit- cells derived from wild-type or knockout littermates. Unsupervised clustering found

these cells to segregate indiscriminately (Fig. 19d). As such Runx1 is not necessary for the formation of this early EHT population.

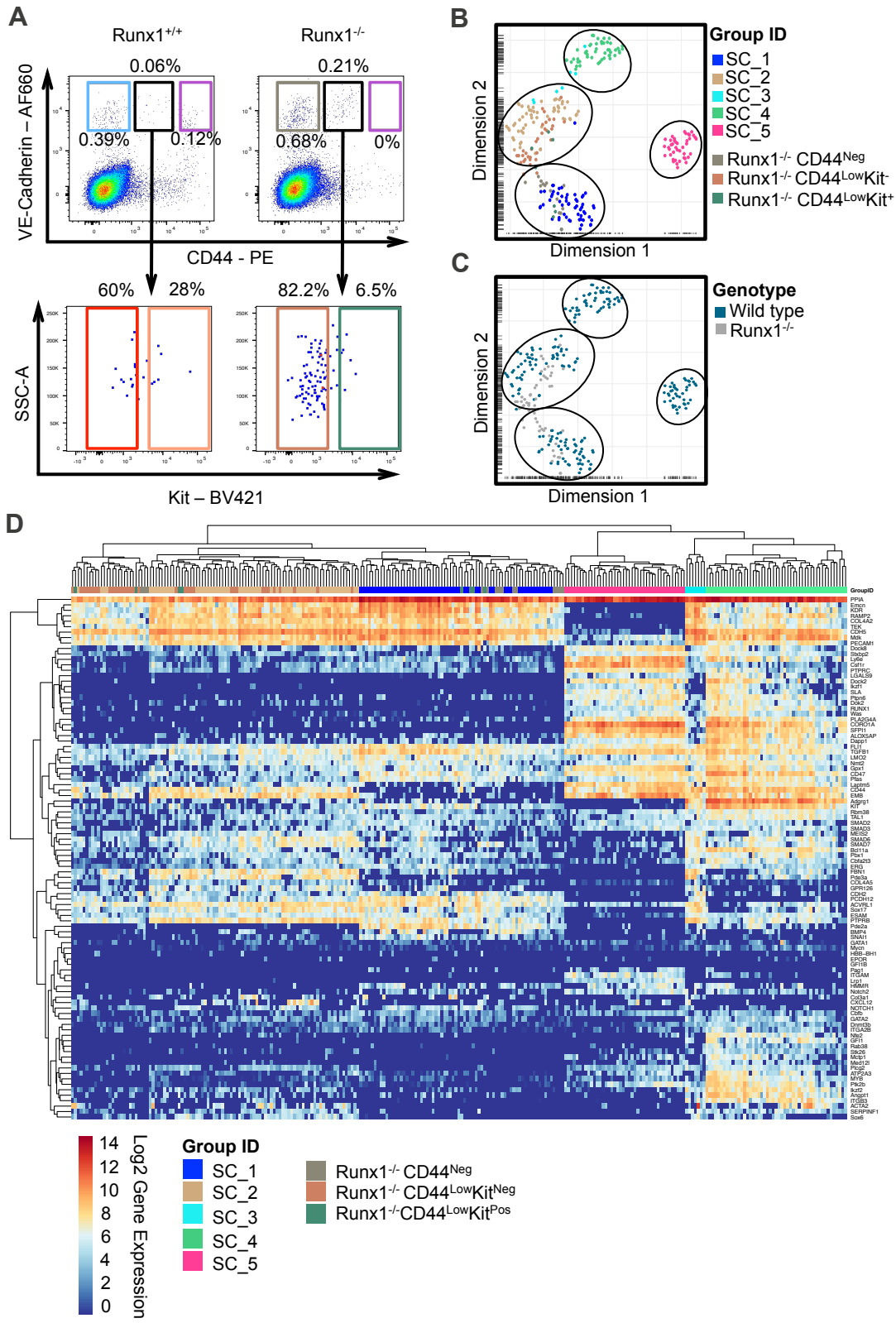


Figure 19: CD44+ pre-HSPCs type I and type II are absent in Runx1 knockout embryos

FACS analysis comparing wild-type and Runx1 knockout embryos (A). Single-cells were FACS sorted from E10 mouse AGMs directly into 96-well plates for qPCR analysis. t-SNE clustering combining single-cell qPCR analysis of wild-type and knockout embryos, showing both CD44 populations (B) and their genotypes (C). Heatmap of single cell qPCR transcriptional data and unsupervised clustering where Runx1 knockout samples are interspersed amongst CD44Low/Kit- and CD44Negative populations (D).

4.2.3 Blocking CD44 inhibits haematopoietic development both *ex vivo* and *in vitro*

To understand whether CD44 expression played an active role in the development of HSPCs we used a CD44 blocking antibody to inhibit its function. By performing the *ex vivo* OP9 co-culturing assays on CD44+ cells in the presence of the KM201 blocking antibody we were able to inhibit round cell colony formation in a dose dependent manner (Fig. 20b). We also noticed that in the presence of the blocking antibody the round cell colonies that formed were smaller in size (Fig. 20a). Quantification of cell number found a significant decrease in the size of round cell colonies in the presence of the inhibitory antibody (Fig. 20c).

To further investigate the role of CD44 in EHT we applied the CD44 blocking antibody to our *in vitro* system of early blood development. Exposure of the haemangioblast (blast) assay to CD44 inhibition for 48 hours resulted in a significant decrease in CD41+/VE-Cadherin- haematopoietic progenitors. We could also observe a compensatory increase in the proportion of endothelial (CD41-/VE-Cadherin+) and vascular smooth muscle (VSM) cells (CD41-/VE-Cadherin-) in the culture (Fig. 21a & b). Overall, this showed for the first time a functional role for CD44 in the emergence of HSPCs.

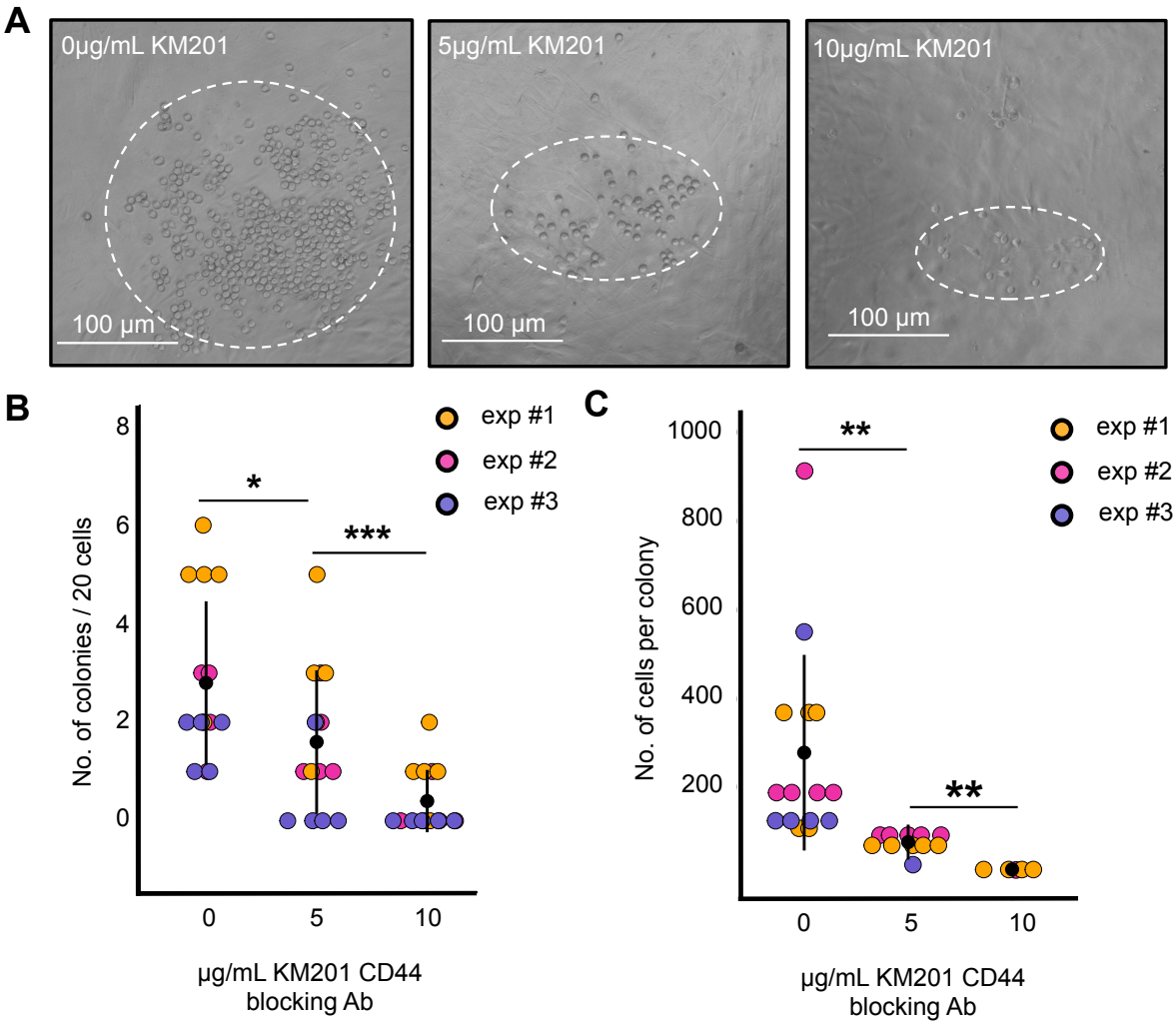
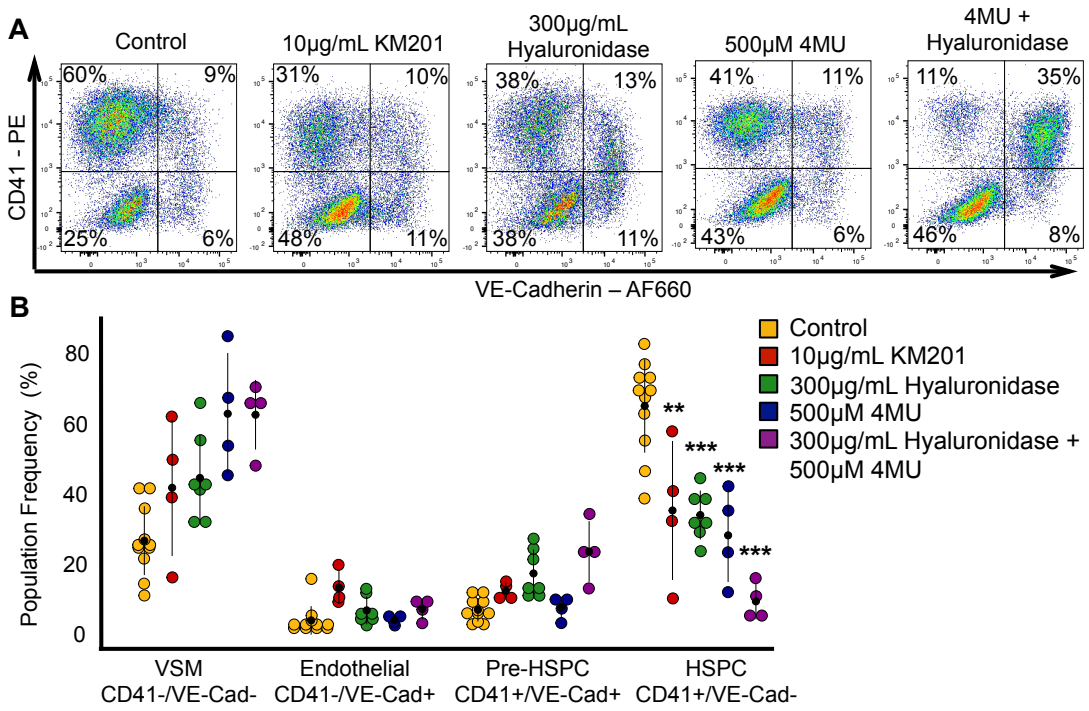


Figure 20: *Ex vivo* culturing with CD44 blocking antibody inhibits round cell colony formation

For these experiments 20 CD44High cells were FACS sorted from E10 mouse AGMs directly onto OP9 stromal layers in 96 well culture plates. Images of round cell colonies after three days in culture, generated from CD44High cells sorted onto OP9 stromal layers with increasing concentrations of KM201 CD44 blocking antibody (A). Quantification of the number of round cell colonies generated from 20 CD44High cells with 5 and 10µg/mL of KM201 (B) and quantification of the size of colonies generated (C). Results are based on three independent experiments and significance determined by two-tailed independent Student's T-tests where * indicates p-value < 0.05, ** indicates p-value < 0.01 and *** indicates p-value < 0.001.

4.2.4 Alterations in hyaluronan content *in vitro* inhibits HSPC formation

The KM201 blocking antibody binds close to the hyaluronan-binding domain of CD44 and most likely inhibits its function through prohibiting interaction with its major ligand, hyaluronan. To explore this hypothesis we further utilised the *in vitro* system to model EHT and treated the cells with an enzyme that catalyses hyaluronan degradation, hyaluronidase and an inhibitor of hyaluronan synthesis, 4-methylumbelliferone (4-MU). Treatment with either of these two compounds resulted in a block in EHT, significantly reducing the proportion of haematopoietic progenitors forming and again showing an increase in the proportion of pre-HSPCS or VSM cells in culture (Fig. 21a & b). By using these two compounds together we found a combined effect with an even greater reduction in HSPC output (Fig. 21a & b). As such, the interaction between CD44 and hyaluronan appears to play a regulatory role in EHT.



C

ANOVA

HSPC output

	SS	df	MS	F
Between Groups	10,404	4	2,601	18.66
Within Groups	3,485	25	139	
Total	13,888	29		

$p < 0.000001$

Figure 21: Differentiation of blast culture towards blood can be blocked by inhibitors of CD44 and its ligand hyaluronan

Representative FACS plots of analysis of blast differentiation culture when untreated or treated with a CD44 blocking antibody – KM201, a hyaluronan degrading enzyme – hyaluronidase, a hyaluronan synthase inhibitor 4-methylumbelliferone (4-MU) or a combination of hyaluronidase and 4-MU (A). Dot plot summarising the changes in populations generated from the blast culture under different treatment conditions (B). Data is based on at least four independent experiments for each condition. Significance was determined on HSPC output of the differentiation culture based on a one-way ANOVA, finding a p-value of < 0.0000001 (C). A Dunnett's post-hoc test was then used to determine the significance of each condition compared to the control whereby ** indicates a p-value of < 0.01 and *** indicates a p-value of < 0.001.

4.2.5 Knockout of CD44 *in vitro* does not affect EHT

To further explore the function of CD44 in HSPC development we created a CD44 knockout mESC cell line using Crispr-Cas9 technology. A single guide RNA was designed to target exon 2 of the CD44 transcript and transfected into mESCs along with a GFP-tagged Cas9 nuclease using the pX458 plasmid. Clones were isolated using single cell FACS sorting. Loss of CD44 was validated by loss of the CD44 protein in flow cytometric analysis (Fig. 22a). Haemangioblast culture of the CD44 knockout cell line revealed no obvious defects in the differentiation of Flk-1+ mesoderm into blood (Fig. 22b & c). This could be due to compensation mechanism whereby another hyaluronan binding protein is up-regulated in response to the loss of CD44 on the cell surface.

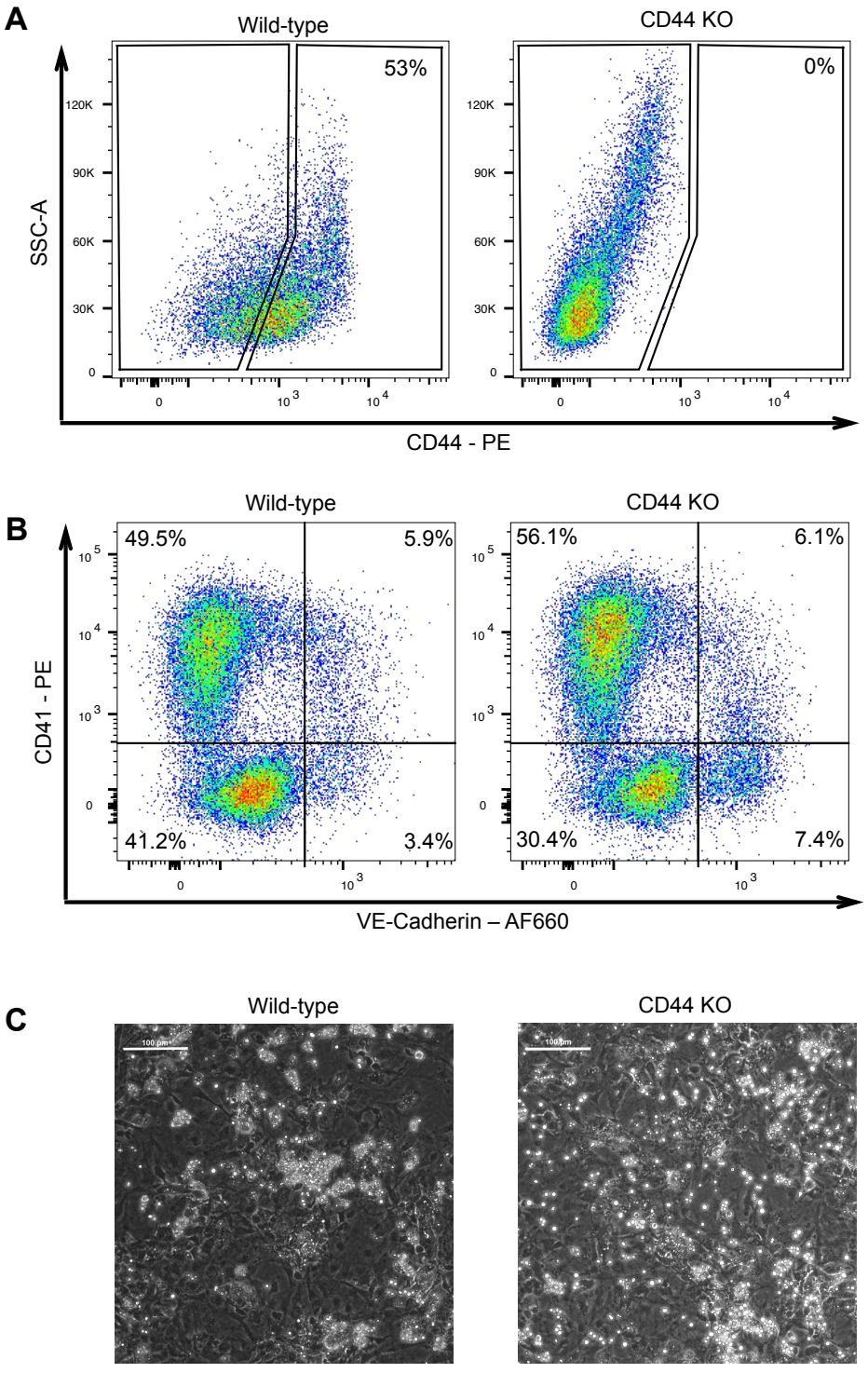


Figure 22: Knockout of CD44 *in vitro* does not impact on EHT

FACS plot demonstrating loss of CD44 cell surface expression upon transfection with CD44 targeting single guide RNA and Cas9 nuclease (A). Representative FACS plots of day two of blast culture comparing wild-type and CD44 knockout mESCs, (n = 2) (B). Representative images of day two of blast culture showing formation of round cell colonies both in wild-type and CD44 knockout cultures, scale bar is equal to 100μM (n = 2).

Chapter 5: Haematopoietic potential appears to be restricted to the Stab2 negative population in the yolk sac

5.1 Stab2 expression distinguishes the endothelium of the yolk sac from the AGM

5.1.1 Single cell RNA sequencing identifies Stab2 as a marker of yolk sac endothelium

To better understand the properties of endothelial cells that transition into HSPCs it is interesting to consider the similarities and differences that exist between the AGM and yolk sac. While both tissues undergo EHT in a Runx1 dependent manner it is not known whether the yolk sac can give rise to HSCs. Furthermore, it is useful to the field of developmental haematopoiesis to characterise genes that are able to distinguish from which tissue haematopoietic cells are derived. To this end, we performed a single cell RNA sequencing experiment to compare VE-Cadherin+ and VE-Cadherin- cells from the AGM and yolk sac using the Takara ICELL8 platform. Two hundred cells were sequenced and the data analysed using the Seurat package in R. After quality control measures were performed we analysed the expression of 11,767 genes across 178 single cells.

Dimension reduction analysis using the t-SNE algorithm and unsupervised clustering found that the populations roughly clustered according to the tissue from which they derived and their expression of VE-Cadherin (Fig. 23a & 24a). Based on this clustering we could identify 47 differentially expressed genes across the four groups with an adjusted p-value of < 0.05 and a log2

fold change > 1.5, displayed in the heatmap in figure 23 (Supplementary table 3). Amongst these marker genes a cluster of seven genes appeared to be more highly expressed in yolk sac endothelium and two genes up-regulated in the endothelium derived from the AGM (Fig. 23a). Interestingly, of the seven genes associated with the yolk sac endothelium, four are known to be expressed in liver sinusoidal endothelial cells – *Maf*, *Mrc1*, *Lyve1* and *Stab2* (Fig. 24c) (Nonaka et al., 2007; Géraud et al., 2010; Sørensen et al., 2015). Given the liver is also a site of embryonic haematopoiesis it is interesting that these two tissues share this endothelial signature. Furthermore, the genes *Colec12*, *Mrc1* and *Stab2* are characterised as scavenger receptors (Jang et al., 2009; Schledzewski et al., 2011; Sørensen et al., 2015). *Stab2* in particular is a scavenger for hyaluronan (Hirose et al., 2012). The two markers we found to be highly characteristics of AGM endothelial cells, we identified previously in our single cell qPCR and 25-cell RNA sequencing – *Mecom* and *Emcn* (Fig 24c). Given the surprising coincidence that another hyaluronan binding protein is highly expressed and specific to the yolk sac vascular endothelium, we decided to investigate the role of Stab2 in combination with CD44.

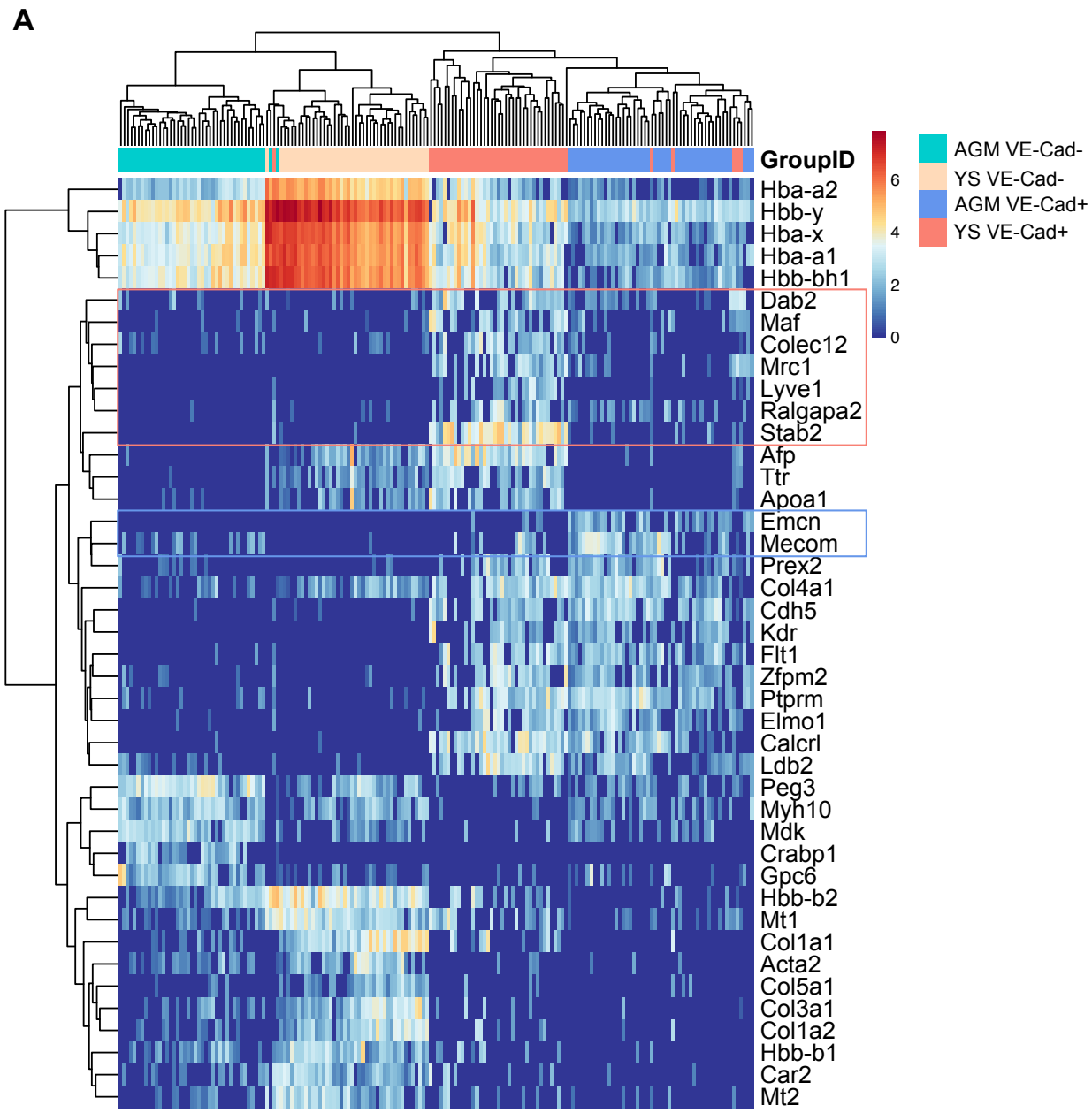


Figure 23: Single cell RNA sequencing analysis identifies markers that distinguish vascular endothelium from the yolk sac and AGM

Cells were FACS sorted from the AGM and yolk sac of E10 mouse embryos into VE-Cad+ and VE-Cad- fractions before being distributed by a nano-dispenser onto a 5,184 nano-well chip at very low concentration to capture single-cells. After quality control and normalisation steps the transcriptional profiles of 178 single cells were analysed using the Seurat package on R (version 3.5.1). The FindMarkers function was used to identify 42 differentially expressed genes with an adjusted p-value of < 0.05 and a log 2-fold change of > 1.5 and are displayed in the heatmap. Above. Of these 42 genes 7 were identified as being specifically up-regulated in yolk sac vascular endothelium (marked in orange) and 2 genes were identified that are specific to AGM vascular endothelium (marked in blue).

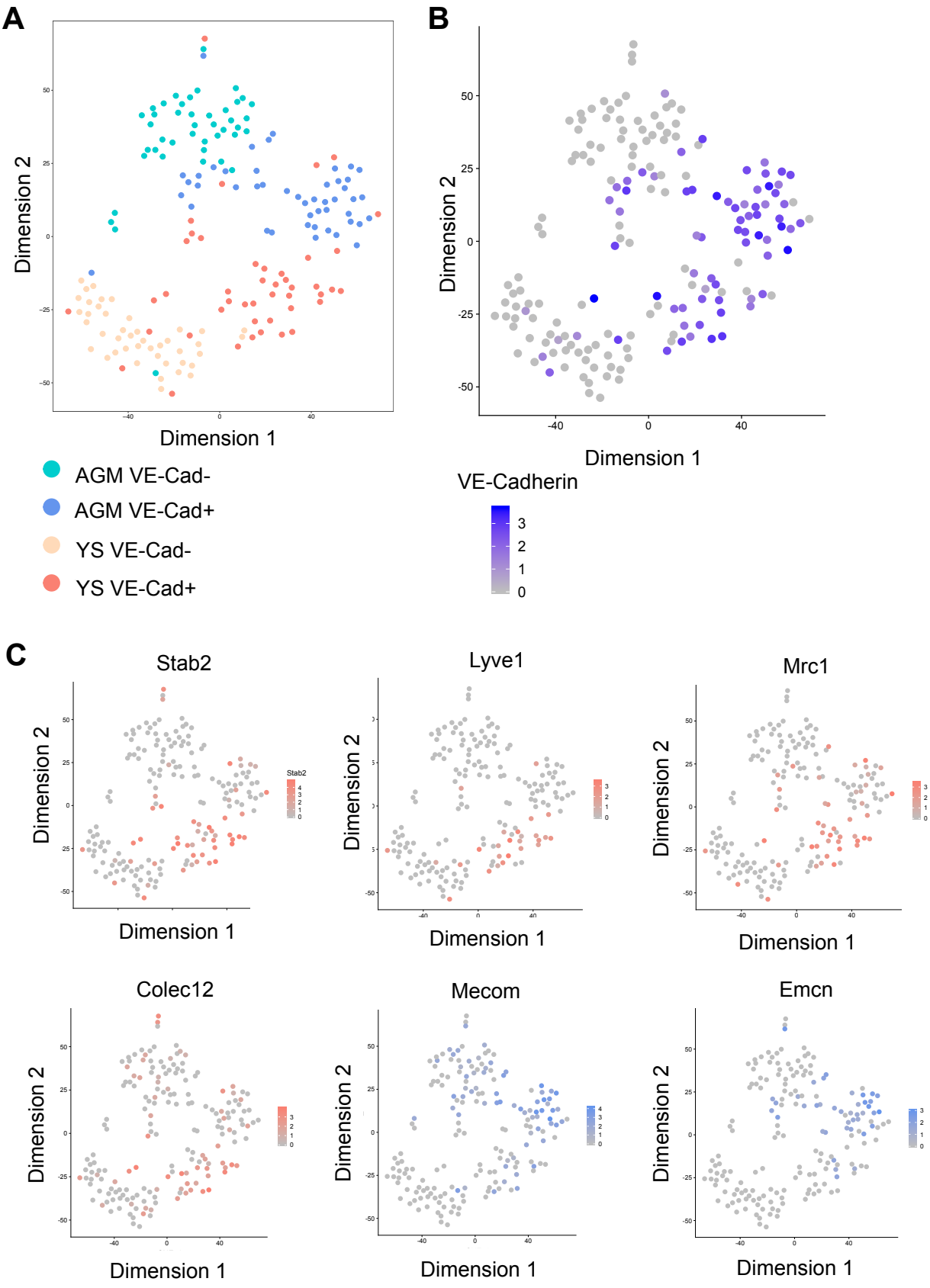

A
Dimension 2
Dimension 1
AGM VE-Cad-
AGM VE-Cad+
YS VE-Cad-
YS VE-Cad+
B
Dimension 2
Dimension 1
VE-Cadherin
3
2
1
0
C
Stab2
Lyve1
Mrc1
Colec12
Mecom
Emcn
Dimension 2
Dimension 1

Figure 24: Expression distribution of yolk sac and AGM specific markers of vascular endothelium

Dimension reduction analysis was run on the 178 single cells analysed in Figure 23 that passed quality control measures. The t-SNE plot shows that cells tend to cluster depending on whether they expressed VE-Cadherin and from which embryonic tissue they derived (A). VE-Cadherin was expressed, as expected in the endothelium of the yolk sac and AGM (B). Stab2, Lyve1, Mrc1 and Colec12 expression closely maps to cells that derive from the yolk sac endothelium (C). Conversely, Mecom and Emcn are mostly expressed in VE-Cadherin+ cells of the AGM (C).

5.1.2 The majority of yolk sac vascular endothelium expresses the Stab2 receptor on the cell surface

To validate the idea that Stab2 could be used as a marker to distinguish yolk sac derived vascular endothelium from the AGM we performed FACS analysis using a monoclonal Stab2 antibody. We confirmed that Stab2 cell surface expression was restricted to VE-cadherin+ yolk sac endothelial cells, with little to no marker expression detected in the AGM (Fig. 25a-c). In fact, the majority of VE-cadherin+ cells were found to express Stab2 in the yolk sac at E9.5, E10 and E11 (Fig. 25a & b). Interestingly, in the haemangioblast culture system, which was designed for robust production of haematopoietic cells, Stab2 is not expressed (Fig. 25d). This finding is of note as the haemangioblast culture is often considered to more closely resemble yolk sac haematopoiesis due to the lack of HSCs produced in the system. Given this difference in endothelial cell output, perhaps this is not the case.

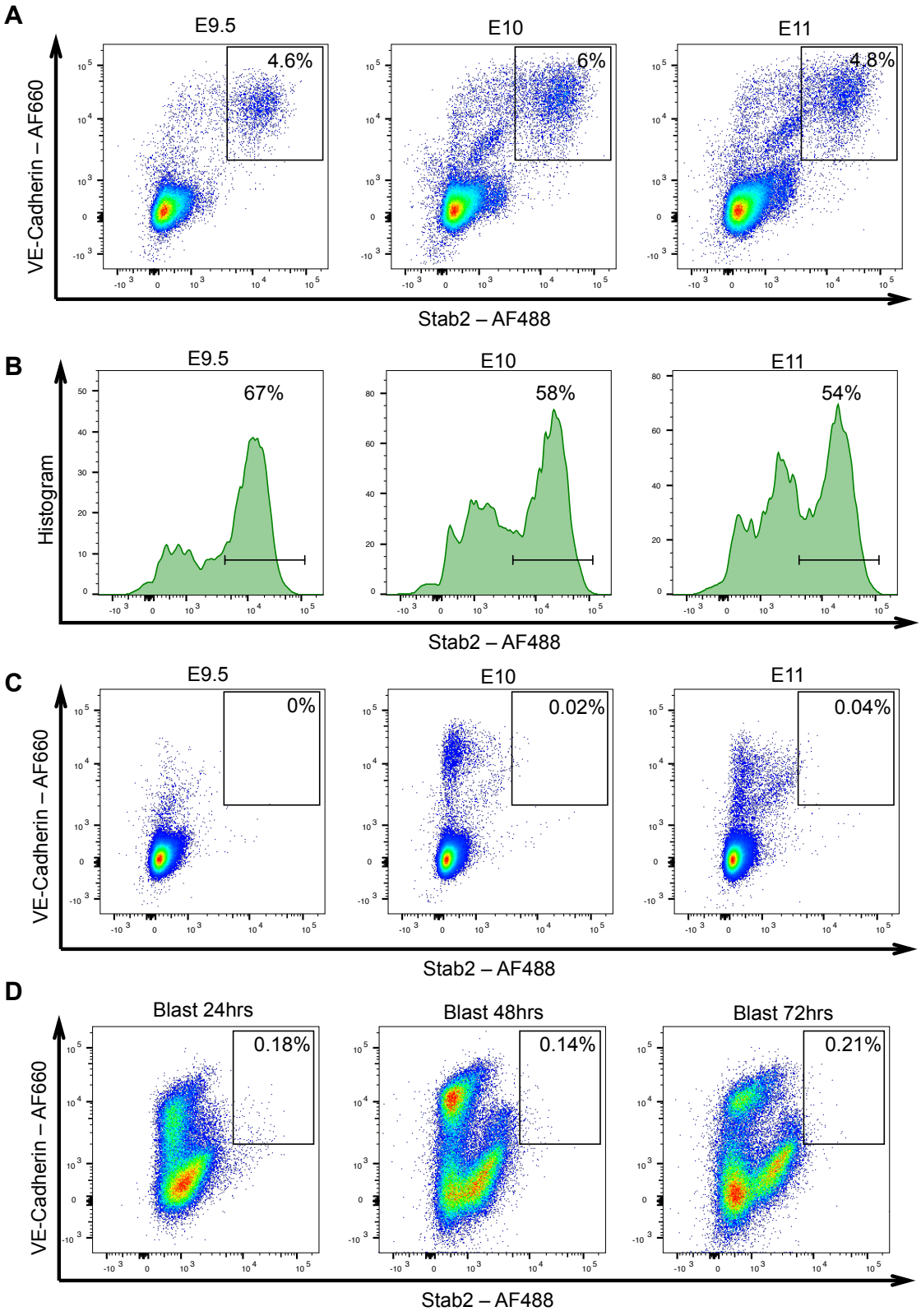


Figure 25: Flow cytometry analysis of Stab2 expression in yolk sac, AGM and haemangioblast culture

Flow cytometry analysis showed that Stab2 is expressed at the protein level in the yolk sac of E9.5, E10 and E11 mouse embryos (A) and that within the VE-Cadherin+ fraction is present on greater than 50% of endothelial cells across this time-course (B). Conversely, Stab2 protein expression was absent in the AGM at E9.5, E10 and E11 (C) and haemangioblast culture (D) after 24, 48 and 72 hours of incubation.

5.1.3 Stab2 appears to be down-regulated as cells become more haematopoietic

Like CD44, Stab2 is also a HABP, however, unlike CD44, which is often used to tether cells to hyaluronan, Stab2 is characterised for its role as a scavenger receptor, acting to remove hyaluronan from the blood stream (Hirose et al., 2012). Given the utility of CD44 in the AGM to track EHT and the recently uncovered role for hyaluronan in blood cell emergence we decided to combine staining of CD44 expression with Stab2 in the yolk sac. We found that cells expressing high levels of CD44 had no Stab2 expression on their cell surface while the majority of cells with lower expression of CD44 also expressed Stab2 (Fig. 26a & b). If, as we see in the AGM, CD44Low cells transition into CD44High cells, then the CD44Low fraction appears to either arise from Stab2- vascular endothelium or Stab2 is down-regulated as CD44 is up-regulated. To further explore this hypothesis we looked at the pre-HSPC specific marker CD61 (Itgb3) that we identified in our previous transcriptional analysis of EHT in the AGM (Fig. 12b). We found that by excluding Stab2+ cells from the CD44Low/Kit+ fraction of the vascular endothelium of the yolk sac we could dramatically enrich for CD61+ cells (Fig. 26b). This further supported the idea that the haematopoietic transition was occurring in the Stab2 negative portion of the yolk sac endothelium.

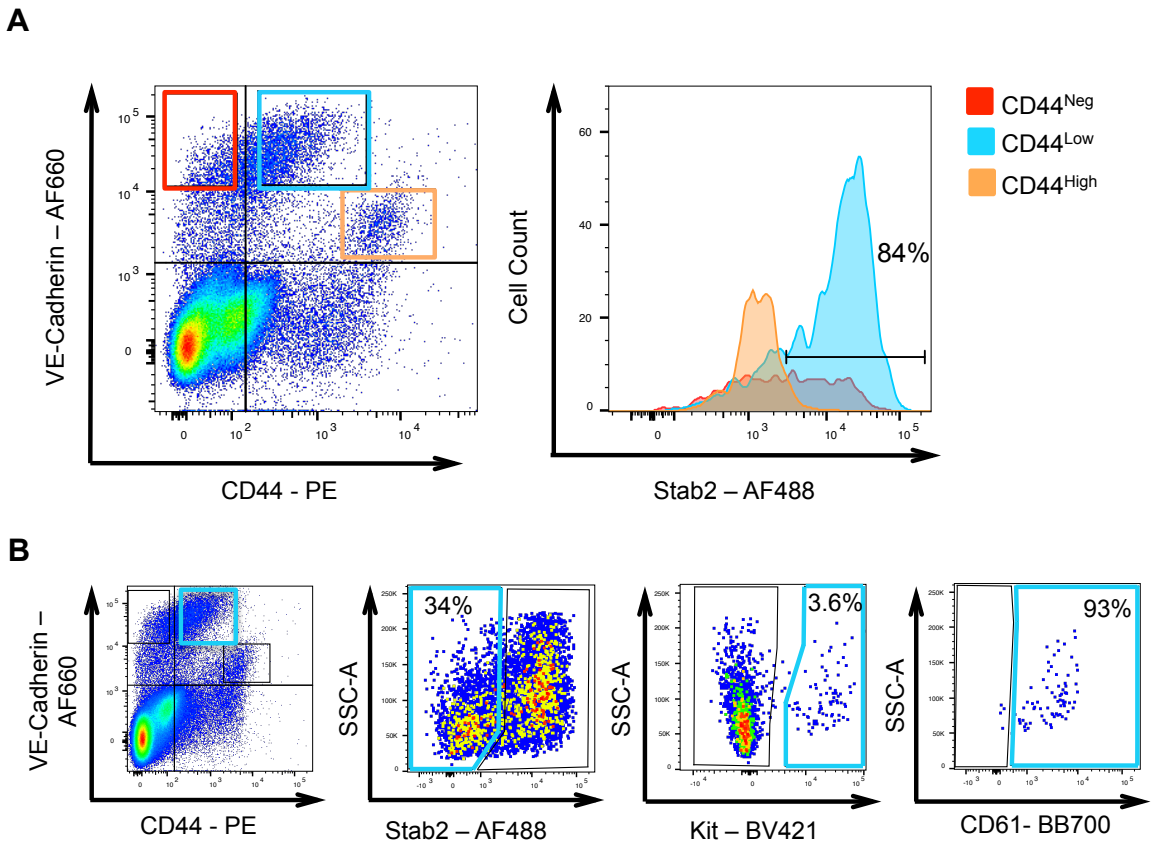


Figure 26: Stab2 appears to be down-regulated in more haematopoietic populations

Flow cytometry analysis of mouse yolk sacs at E10 found that CD44High cells lack Stab2 expression while the majority of CD44Low cells are Stab2+ (A). VE-Cadherin+/CD44+/Stab2-/Kit+ cells are enriched for the pre-HSPC marker CD61 (B).

5.2 Stab2 can be used to isolate a haematopoietic progenitor population in the yolk sac

5.2.1 Single cell qPCR isolates AGM-like haematopoietic progenitors from the yolk sac by excluding Stab2 expression

To understand whether we could isolate AGM-like haematopoietic progenitors from the yolk sac by excluding Stab2 surface expression we performed single-cell qPCR on sorted populations from E10 and E11 staged embryos. By analysing the 95 haematopoietic and endothelial related genes (Supplementary table 1), across 212 cells expressing various combinations of VE-Cadherin, CD44, c-Kit and Stab2 we could isolate four populations with varying degrees of haematopoietic and endothelial gene expression (Fig. 27a). We found that the CD44High population of the yolk sac closely resembled that of the AGM with expression of haematopoietic genes such as *Runx1*, *Ikzf1*, *Csf1r* and the pan-haematopoietic marker *Cd45* (*Ptprc*) as well as down-regulation of endothelial genes such as *Emcn*, *Pecam1*, *Tek*, *Kdr* and *Sox17* (Fig. 27b & Supplementary Fig. 4). Furthermore, the CD44Low/Kit+/Stab2- cells also resembled our previously characterised pre-HSPC population with dual expression of endothelial and haematopoietic genes alongside specific markers like *Adgrg1* and *Itgb3* (Fig. 27b & Supplementary Fig. 4). We also identified a small endothelial population mostly composed of CD44Low/Kit-/Stab2- cells which shared some similarity with the haemogenic endothelial population we found in the AGM (Fig. 27b & Supplementary Fig. 4). This population showed some small changes in gene expression that could indicate it is the precursor of the pre-HSPC-like population with slight decreases in *Emcn* and *Pcdh12* as well as increases in

Emb, *Pde3a* and *Sox6* (Fig. 27b). With this small set of genes we were not able to further disentangle any differences between the CD44Neg and CD44Low/Stab2+ populations although they appeared largely endothelial in nature. Further transcriptional and functional investigation of these populations is needed to understand their contribution, if any, to embryonic haematopoiesis.

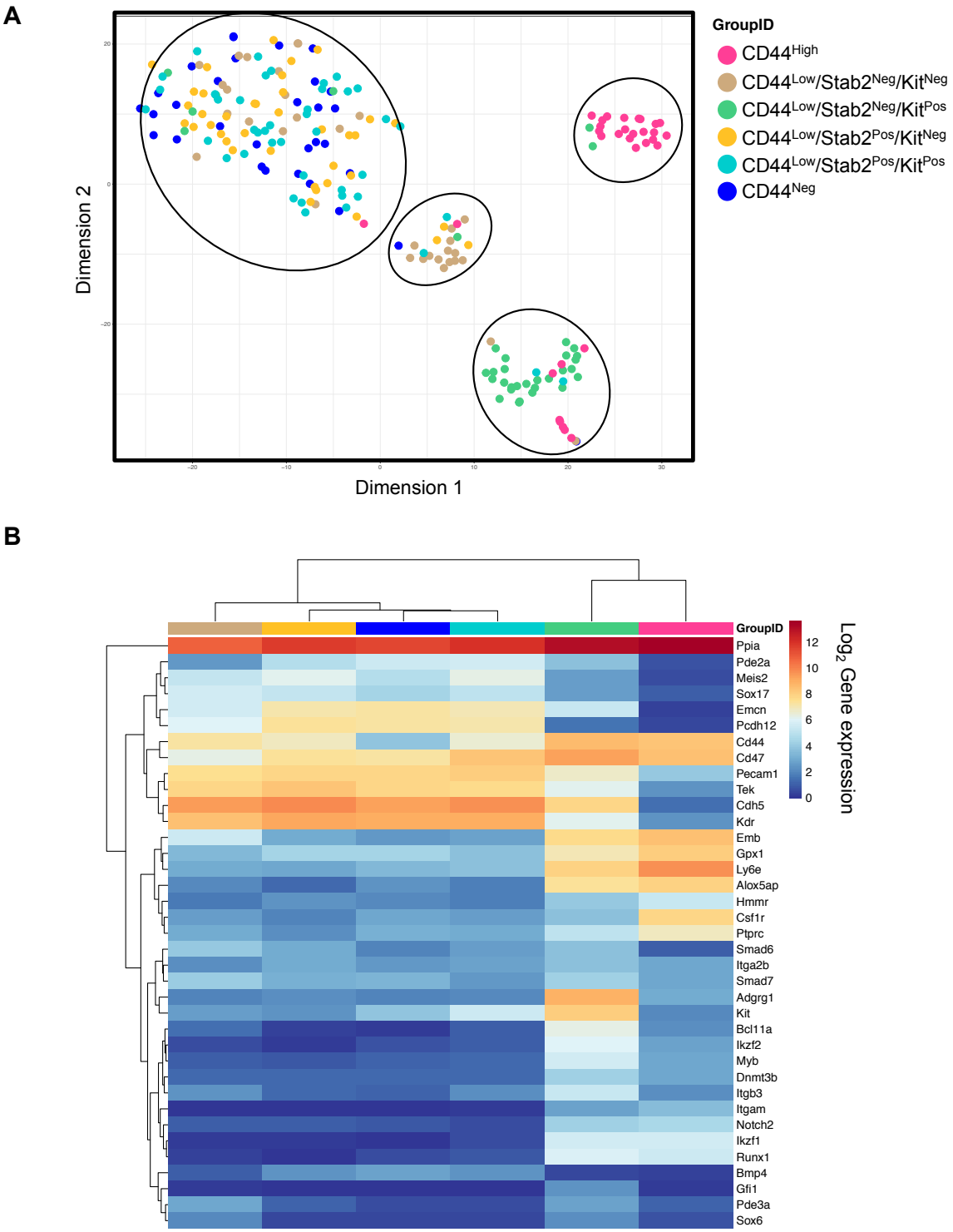


Figure 27: Single-cell qPCR identifies haematopoietic progenitors in the Stab2- cell populations

Single cells were FACS sorted into 96-well plates based on their expression of VE-Cadherin, CD44, Stab2 and c-Kit from the yolk sacs of E10 mouse embryos. t-SNE dimension reduction analysis identified four clusters of cells corresponding mainly to cells with CD44High, CD44Low/Stab2-/Kit+, CD44Low/Stab2-/Kit- and a combined group of endothelial-like cells (A). Average log2 expression of each group based on single-cell qPCR showed that CD44High cells from the yolk sac resembled those isolated from the AGM and that CD44Low/Stab2-/Kit+ cells expressed both endothelial and haematopoietic marker genes (B).

Chapter 6: Discussion

6.1 Summary of results

Using single-cell transcriptomic analysis we sought to characterise the endothelial populations present in embryonic haematopoietic tissues of the mouse embryo. We identified the HABP, CD44 as a cell surface receptor that could be used in conjunction with Kit and VE-Cadherin to track the progression of EHT. With this combination of markers, we were able to robustly isolate different stages of HSPC formation. This enabled us to perform in depth transcriptional analysis and characterise the physical properties of vascular endothelium, haemogenic endothelium, pre-HSPC type I and type II cells. Our transcriptional results were in line with previous reports on the importance of the heptad of haematopoietic transcription factors and the dynamic role of Hedgehog, Notch, Wnt and Bmp/Tgf β signalling pathways. We uncovered distinct differences in cell size, cell cycle and metabolic status as endothelial cells transition towards HSPCs, and with RNA sequencing were able to identify new potential regulators of EHT for future investigation. In addition, we found a functional role for CD44 and its ligand hyaluronan in EHT through *ex vivo* co-culturing assays and manipulation of the mouse *in vitro* model of blood development. This emphasises the importance of the microenvironment in cellular transitions and the need for further research into the role the extracellular matrix components in this process.

By comparing endothelial populations from the yolk sac and AGM we were able to identify a second HABP, Stabilin-2, as a cell surface receptor that is

present on endothelial cells of the yolk sac and absent in the AGM. We found that Stab2 was not expressed on cells advanced in the transition towards haematopoietic progenitors and that by excluding Stab2+ cells we could easily isolate AGM-like haematopoietic progenitors from the yolk sac. Further work is needed to characterise these progenitors and to understand the role of the Stab2+ endothelium in the yolk sac microenvironment. Our discovery of two hyaluronan receptors as markers for endothelial populations in embryonic haematopoiesis poses intriguing questions on the overall effect of hyaluronan on EHT.

6.2 CD44 is a new marker of EHT

6.2.1 CD44 has broader expression than Kit but is more specific than CD41

Previous work in developmental haematopoiesis has relied heavily on the cell surface marker CD41 for emerging HSPC populations, with CD45 serving to mark more mature haematopoietic cell types (Ferkowicz *et al.*, 2003; Rybtsov *et al.*, 2011, 2014). However, there are a few issues with the use of CD41 as a marker in this system. CD41 is expressed at the focal point between cells of the haematopoietic cluster, making it more difficult to identify the positive cells (Boisset *et al.*, 2010). Furthermore, the receptor displays a variable expression pattern, whereby it is up-regulated during pre-HSPC development before being down-regulated again as the HSPCs migrate to the foetal liver and bone marrow (Robin, Ottersbach and Boisset, 2011; Boisset *et al.*, 2013). Transplantation studies performed using cells marked by CD41 and its partner

receptor CD61 found that stem cell potential resided only within the double positive population (Boisset *et al.*, 2013). Furthermore, the first single cell RNA sequencing study published on EHT dynamics using CD41, CD43 and CD45 as markers failed to capture the earliest steps of the transition, namely haemogenic endothelium (Zhou *et al.*, 2016). Indeed, our own single cell transcriptome data reveals that CD41 has weak and variable expression, however, CD61 appears to strongly and specifically mark pre-HSPC type I cells (Fig. 11a). Cartography studies of the mid-gestation mouse embryonic vasculature opted to use Kit for the quantification of haematopoietic clusters, most likely due to its greater specificity (Yokomizo and Dzierzak, 2010). However, the Kit receptor is up-regulated later in EHT than CD41 (Rybtsov *et al.*, 2014). Conversely, CD44 is expressed on the outer surface of cells in the haematopoietic clusters (Fig. 8a) and is up-regulated prior to CD41 (Fig. 11a), maintaining its expression as cells migrate to the foetal liver and bone marrow (Avigdor *et al.*, 2004). As such CD44 can be used to mark the earliest stage of EHT and presents itself as a more reliable marker of HSPC emergence as expression of CD44 increases with haematopoietic identity.

6.2.2 Utility of CD44 could be improved through analysis of tissue-specific isoforms

A significant drawback of using CD44 as a marker is its widespread expression in many non-haematopoietic tissues. The smallest variant of CD44, CD44s is expressed on a wide variety of cell types and is the target of most commercial anti-bodies. Alternative splicing, however, of the ten variant exons present in the CD44 transcript are thought to provide significant tissue

specificity (Hirano *et al.*, 1993). To understand if different splice variants of CD44 are expressed on haematopoietic precursors we extracted RNA from CD44+ and CD44- cells from the AGM, yolk sac and blast culture. Primers were designed in the constitutive exons either side of the variable region (Supplementary Fig. 2a). Amplification of the resulting cDNA revealed 5 bands in the VE-Cadherin+/CD44+ cell fraction (Supplementary Fig. 2b). The smallest band corresponded to the expected size for the CD44s isoform. All tissues displayed the same banding pattern (data not shown). Sequencing of AGM derived cDNA revealed that the smallest band (E) indeed corresponded to the common CD44s isoform and that the larger bands (A, B and C) aligned to regions of the variable exons 8, 9 and 10 (Supplementary Fig. 2c). Interestingly, despite the larger size on the gel of band A compared to band B sequencing revealed a larger region of homology between bands B and the CD44 coding sequence. This could be explained by the presence of an alternative splice site in the intervening intron. Bands D and F did not align to the CD44 coding sequence. Thus, it appears that several CD44 variant isoforms are present on haematopoietic progenitors of the AGM. This preliminary data, although intriguing, requires further investigation to identify and validate the significance of these isoforms to the process of EHT.

Further investigation of the expression of CD44 isoform(s) in EHT could also lead to a more complete understanding of its function. As previously discussed, addition of alternative exons to the stem region of the CD44 receptor can modify its binding capacity. Much of the research centres on the effect CD44 variant isoforms have in various cancer models. The presence of

CD44v8-10 has been linked to increased metastatic potential in some cancers (Lau *et al.*, 2014). Interestingly, this potential is thought to occur by CD44v8-10 promoting endothelial junction disassembly through the phosphorylation of VE-Cadherin (Zhang *et al.*, 2014). Given that VE-Cadherin is also important during EHT, this would be an interesting mechanism of action to explore in the haematopoietic system.

6.3 CD44 has enabled in-depth transcriptional analysis of EHT dynamics

6.3.1 CD44Negative cells express both venous and arterial markers

The identification of CD44 as a marker of pre-HSPCs led to the isolation of two endothelial populations from the VE-Cadherin+ fraction of the AGM. Both CD44Negative and CD44Low/Kit- populations displayed broadly endothelial profiles, however, the shared expression of Smad6, Smad7, Fbn1 and Pde3a between CD44Low/Kit- and CD44Low/Kit+ cells led us to hypothesise that these cells represented the earliest stage of EHT – haemogenic endothelium (Fig. 11 & 12). This transcriptional association was further investigated and confirmed using 25-cell RNA sequencing (Fig. 14 & 15).

Some reports have suggested that CD44 is simply an arterial cell marker, leading us to question whether our second endothelial population derives from the nearby cardinal vein (Robert-Moreno et al., 2008). Previous single cell analysis of EHT, which isolated CD31+/VE-Cadherin+/CD41-/CD43-/CD45-endothelial cells from the AGM, identified only one population (Zhou et al., 2016). This population clearly corresponded with our CD44Negative cells. It seems unlikely that two independent research groups would have such high

contamination of endothelial cells from a smaller nearby vessel, especially given that CD44Negative cells are the most numerous VE-Cadherin+ population represented in our FACS data of the AGM region (Fig. 9a).

Furthermore Zhou and colleagues (2016) published an interesting finding identifying both venous and arterial markers in their CD31+/VE-Cadherin+/CD41-/CD43-/CD45- endothelial cells and a down-regulation of venous markers in their CD31+/CD45-/CD41^{Low}/c-Kit+/CD201^{High} type I pre-HSCs. We observed a similar phenomenon whereby our CD44Negative endothelial population expressed both venous (Nr2f2, Nrp2, Ephb4) and arterial makers (Efnb2, Sox17) while our haemogenic endothelial cells showed a loss of Nrp2 and Nr2f2 expression (Supplementary Fig. 3). Interestingly, expression of the traditional venous marker Ephb4 was maintained alongside the arterial marker Efnb2 in the haemogenic endothelial cells, and only lost upon acquisition of a haematopoietic profile.

This down-regulation of venous related markers reflects the view in the field that the haematopoietic lineage is more closely related to the arterial endothelium. Although it is the case that HSC development is restricted to the major arteries, this could be the result of the arterial microenvironment rather than an inherent property of arterial endothelial cells. Indeed, over-expression of the Notch intracellular domain was found to expand HSPC production to the venous vessels, indicating that venous endothelial cells could be competent for EHT given the right conditions (Burns *et al.*, 2005). Furthermore, these findings suggest that the expression of venous and

arterial markers may not be as definite as previously thought in embryonic development.

6.3.2 Transcriptional regulation of pre-HSCs

It is important to understand the transcriptional dynamics of this process as haemogenic endothelium represents the starting point of HSC development. Recently, promising results were achieved in HSC reprogramming using transcription factor cocktails to induce fate change. Interestingly, this success was achieved through the use of either endothelial cells as a starting point or transitional state, highlighting the importance of understanding the endothelial origin of HSCs (Lis *et al.*, 2017; Sugimura *et al.*, 2017).

When comparing our transcriptional profiles to the previously published single-cell dataset by Zhou and colleagues (2016), we find many of the same haematopoietic transcription factors up-regulated at the pre-HSPC type I stage (*Nfe2*, *Runx1*, *Gfi1*, *Spi1* and *Ikzf2*) (Fig. 11a & 14b). However, several of the genes they highlight as part of the pre-HSPC type I signature we observe up-regulated at the haemogenic endothelial stage which was not captured in the previous analysis, including, *Nkx2-3*, *Hlf*, *Sox6*, *Bcl11a* and *Stat4* (Fig. 14b). Furthermore, we can identify several novel transcription factors not previously associated with embryonic haematopoiesis, including, *Irf6*, *Nfat5*, *Mllt3*, *Yy2*, *Isl1* and *Mecom* (Fig. 14b). We also found two receptors *Itgb3* (*Cd61*) and *Adgrg1* (*Gpr56*) that act as specific markers of early pre-HSPCs. Although Zhou and colleagues also identified these

receptors, they reported no difference in the expression of these markers between pre-HSPCs type I and type II populations (Fig. 11a).

Overall, the advantage of our analysis is that we can utilise a simple and robust sorting strategy that can isolate all stages of HSPC development and importantly distinguish vascular and haemogenic endothelium. From our RNA sequencing results we have been able to uncover new potential regulators of haematopoietic development, finding 32 transcription factors up-regulated at the haemogenic endothelial stage (Fig. 14b). This could have important implications for the regenerative medicine field given the utility of transcription factor reprogramming to HSC production.

6.3.3 The CD44Low/Kit- haemogenic endothelial population is at the intersection of key signal transduction pathways

Transcriptional analysis of the EHT process also provided us with insight into the dynamic signalling interactions that occur during HSC development. Previous research has shown that Hedgehog and Bmp signalling are needed to pattern the AGM region in preparation for EHT (Wilkinson *et al.*, 2009). In the CD44Low/Kit- haemogenic endothelial population we observed a down-regulation of target genes of the Hedgehog signalling pathway, including, Myc, Ccnnd1 and Snai1 (Fig. 15). Reduction in Snai1 expression is surprising given its link to a similar cellular transformation, the epithelial to mesenchymal transition (EMT) that occurs in cancer cells (Carver *et al.*, 2001).

In regards to the Bmp/ Tgfβ signalling pathway, we observed more distinct changes. *Bmp4* and its target gene *Lef1* were both down-regulated at the haemogenic endothelial stage. This is consistent with previous work that showed down-regulation of the Bmp pathway was necessary for pre-HSPC emergence (Marshall, Kinnon and Thrasher, 2007; Pimanda *et al.*, 2007). We also observed an up-reguation of the Bmp receptor *Bmpr1a*, which was previously found to play a role in maintaining the size of the HSC niche in the bone marrow (Zhang *et al.*, 2003). Furthermore we saw specific up-regulation of the Bmp target genes *Id1* and *Id2* that are known to act as transcriptional repressors often used to prevent cellular differentiation (Benezra *et al.*, 1990; Norton *et al.*, 1998). As such, their high expression in haemogenic endothelium and subsequent down-regulation could indicate that these cells are primed and waiting to undergo EHT.

The related Tgfβ pathway, is also known to play an essential role in haematopoietic development with knockout of Tgfβ receptor II and Tgfβ1 ligand resulting in embryonic lethality from haematopoietic defects (Dickson *et al.*, 1995; Oshima, Oshima and Taketo, 1996). *In vitro* treatment of immature HSPCs with Tgfβ inhibited their proliferation while exposure of the mouse ESC differentiation system to a Tgfβ inhibitor greatly increased haematopoietic output (Keller *et al.*, 1990; Vargel *et al.*, 2016). Our study further supports the idea that Bmp/Tgfβ signalling must be inhibited for EHT to occur as we see a specific up-regulation of the Bmp antagonist Bmper and the Tgfβ inhibitors Smad6 and Smad7 at the haemogenic endothelial stage

(Fig. 15). Previous studies have shown that these genes are up-regulated on the ventral side of the dorsal aorta, providing further support for our isolation of haemogenic endothelium using CD44 cell surface expression (Pimanda *et al.*, 2007; McGarvey *et al.*, 2017).

Down-stream of Hedgehog, Wnt and Bmp/Tgfβ signalling pathways is the Notch pathway (Clements *et al.,* 2011; Kim *et al.,* 2013). Notch signalling is also strongly associated with early HSPC development as the knockout mouse models of *Notch1*, *Jagged1* and *RBP-J*κ show significant defects in definitive haematopoiesis (Kumano *et al.*, 2003; Robert-Moreno *et al.*, 2005, 2008). In accordance with this we saw a strong and specific up-regulation in both *Notch1* and *Jagged1* in the haemogenic endothelial population along with their target genes *Hey1* and *Hey2* (Fig. 15). It is possible that CD44 is an early down-stream target of Notch signalling as knockout of the Notch transcription factors *Hey1* and *Hey2* results in loss of CD44 expression (Fischer *et al.*, 2004). It is this transient Notch expression that also leads to the up-regulation of the master transcriptional regulator of haematopoietic development, *Runx1* (Burns *et al.*, 2005; Nakagawa *et al.*, 2006).

The idea that Bmp/Tgfβ inhibited and Notch responsive endothelial cells up-regulate Runx1 in order to initiate EHT is not a new model of HSC development. What is new is our isolation of a potential early Notch target, *Cd44* which is a cell surface receptor up-regulated prior to Runx1 expression that enables the easy isolation of this population for in depth investigation. Although haemogenic endothelium displays differences in other receptor

proteins such as *Notch1*, *Jagged1* and *Bmpr1a* these receptors show less dynamic range in their expression when compared to the neighbouring endothelial cells. This is due to the dynamic nature of Notch signalling, which regulates both arterial and haematopoietic fate, the outcome of which is also dependent on the strength and timing of that signal (Gama-Norton *et al.*, 2015; Souilhol *et al.*, 2016).

This series of events is supported by our experiments with the Runx1 knockout mouse model. We showed that loss of Runx1 resulted in the absence of both the CD44Low/Kit+ and CD44High pre-HSPC populations but did not affect the formation of CD44Low/Kit- cells (Fig. 19). Indeed, single cell qPCR identified no transcriptional changes between cells derived from wild-type versus Runx1-/- embryos (Fig. 19). Runx1 is not necessary for the specification of haemogenic endothelium but for its initiation of EHT. This idea is supported by studies where expansion of Runx1 and Notch intracellular domain expression into non-arterial endothelial cells resulted in the exogenous formation of HSPCs (Burns *et al.*, 2005; Yzaguirre *et al.*, 2018).

6.3.4 The quiescence and metabolic changes in haemogenic endothelium is reminiscent of adult HSCs

HSCs are known to rely on glycolysis for their energy needs. This is in part due to their lodgement in the hypoxic bone marrow niche but is also a mechanism used to protect their DNA from reactive oxygen species across their long life-span (Chen *et al.*, 2008; Simsek *et al.*, 2010). This carefully regulated metabolic program also allows HSCs to maintain their quiescent

nature (Takubo *et al.*, 2013). Maintaining this low metabolic state and hence the longevity of the HSC compartment is facilitated by autophagy which degrades active mitochondria in the cytoplasm (Ho *et al.*, 2017). Similarly, endothelial cells also function mostly through the glycolytic pathway, however the reasons are slightly different. Vascular endothelial cells rely on glycolysis in order to rapidly switch from quiescence to proliferation during angiogenesis and function effectively in avascular areas (Mertens *et al.*, 1990; De Bock *et al.*, 2013). Endothelial cells are also thought to utilise glycolysis as a pathway to synthesise important macromolecules, to reduce their exposure to reactive oxygen species and to preserve oxygen molecules for transfer (De Bock, Georgiadou and Carmeliet, 2013; Wilhelm *et al.*, 2016).

Comparison of the transcriptomes of our vascular and haemogenic endothelial populations found a difference in the expression of numerous metabolic genes. Overall, we observed a global down-regulation of genes related to the glycolysis pathway and the TCA (tricarboxylic acid) cycle as well as an up-regulation of genes related to autophagy (Fig. 16). This finding was consistent with our previous observation that haemogenic endothelial cells were significantly smaller in size compared to both vascular endothelial cells and pre-HSPCs (Fig. 9b). Furthermore, cell cycle analysis revealed that this cell population was quiescent in nature with on average approximately 92 per cent of cells in the G0/G1 phase (Fig. 10). This high level of quiescence could also explain the lack of growth we observed in our OP9 co-culture assay. It is possible that this cell population is more preoccupied with transcriptional changes, which could also explain the significant increase in expression of

many transcription factors we found in comparison to the vascular endothelial population (Fig. 14b).

The EHT process is known to occur through trans-differentiation rather than cell division (Kissa and Herbomel, 2010). As such, it is also possible that these future HSCs are tightly managing their exposure to reactive oxygen species in an effort to protect themselves from DNA damage, which could limit their lifespan in the haematopoietic niche. Changes in metabolic state has also been linked to epigenetic reprogramming which could also explain the differences observed in endothelial cells primed for haematopoietic transition (Ryall *et al.*, 2015). Interestingly, we found that a methyl-transferase, Dnmt3b, was up-regulated in both haemogenic endothelial cells and pre-HSPCs type I (Fig. 11a & 12b). Dnmt3b is responsible for *de novo* methylation during embryogenesis and is known to combine with Dnmt3a to enable the differentiation of adult HSCs (Challen *et al.*, 2014). This observation is further supportive of the idea that these cells could be undergoing a change of fate.

6.4 Hyaluronan and HABPs are implicated in the progression of EHT

6.4.1 The binding of CD44 to hyaluronan could enable cell shape change and migration

Much of what we know about CD44 is due to its role in the migration of cancer cells. Both increased deposition of hyaluronan around cancer cells and the presence of CD44 on the cell surface are strongly associated with EMT, a transition that has been speculated to share commonalities with EHT (Günthert *et al.*, 1991; Okamoto *et al.*, 1999; Auvinen *et al.*, 2000; Lipponen *et*

al., 2001; Patrawala *et al.*, 2006). Similarly, deletion of the hyaluronan synthase enzyme Has2 impairs migration of cells during gastrulation (Bakkers, 2004). It is now known that CD44 is up-regulated at the very beginning of HSPC emergence from the dorsal aorta (Fig. 8). And from previous research, we know that CD44 expression is maintained as HSPCs migrate to the bone marrow (Cao *et al.*, 2015). Thus, a likely function of CD44 on the surface of HSPCs is to aid in migration to the different sites of haematopoiesis. Given the strong interaction between CD44 and hyaluronan and its role in the formation of pericellular matrices, it is possible that CD44 also functions to enable the shape change required during EHT or to protect the travelling HSPCs in their journey to the foetal liver (Knudson, Bartnik and Knudson, 1993).

6.4.2 Loss of CD44 could be compensated for by related HABPs

Although we could perturb EHT and the haematopoietic output by applying a CD44 blocking antibody to *ex vivo* cells or our *in vitro* culture system, a similar phenomenon does not occur when we knock out the *Cd44* gene entirely from our ES cells (Fig. 20-22). Loss of CD44 in mice also results in only mild perturbations in lymphocyte homing and HSC migration (Schmits *et al.*, 1997; Cao *et al.*, 2015). This could be a result of redundancy in the system whereby another HABP or integrin receptor can compensate for the loss of CD44. This has been demonstrated in a collagen-induced arthritis mouse model, whereby the loss of CD44 increases inflammation as hyaluronan instead binds to the Hyaluronan mediated motility receptor (*Hmmr*) (Nedvetzki *et al.*, 2004). The arthritis is only alleviated through the knockdown of both CD44 and *Hmmr*

(Nedvetzki *et al.*, 2004). Similarly, a compensation mechanism has also been demonstrated in primary liver cultures, where use of CD44 blocking antibodies inhibited regeneration but no phenotype was observed in CD44 knockout animals. Instead the heterologous receptor Intercellular adhesion molecule-1 (ICAM-1) compensated for the loss of CD44 (Olaku et al., 2011). Thus, it seems plausible that within the haematopoietic system the function of CD44 could also be performed by other cell surface receptors. This is especially likely given the critical nature of HSPC development and the evolutionary conservation we observe in EHT.

6.4.3 Preliminary data suggest that haematopoietic capacity is restricted to Stab2- cell populations

Based on flow cytometry analysis and transcriptional data we found an anti-correlation between Stab2 expression and haematopoietic markers (Fig. 26). Thus, by excluding Stab2+ cells we could enrich for haematopoietic progenitors amongst VE-Cadherin+ cells of the yolk sac (Fig. 27). While the transcriptional profile of these progenitors appears very similar to the type I pre-HSPCs derived from the AGM, it would be interesting to investigate their functional capacity and perform lineage tracing to understand their origin and eventual fate. Interestingly, Stab2 expression is also absent from the *in vitro* haemangioblast culture (Fig. 25d). These results indicate that Stab2 can be a useful tool to enrich for haematopoietic progenitors in the yolk sac. It remains to be established whether Stab2+ endothelium down-regulates this receptor on the path to haematopoietic development or whether two independent vascular endothelial populations co-exist in the blood vessels of the yolk sac.

The question still remains, what role do Stab2+ endothelial cells play in the yolk sac and are they involved in embryonic haematopoiesis? Stab2 is characterised as a hyaluronan binding protein and scavenger receptor (Schledzewski *et al.*, 2011; Hirose *et al.*, 2012). Interestingly, within the top markers identified for VE-Cadherin+ yolk sac cells from our single-cell RNA sequencing was another hyaluronan binding protein, Lyve1 and two other scavenger receptors Mrc1 and Colec12 (Fig. 23 & 24). Furthermore, both Lyve1 and Mrc1 in addition to Stab2 are known markers of liver sinusoidal endothelial cells (Nonaka *et al.*, 2008; Sørensen *et al.*, 2015). Stab2 has also been reported as a cell surface receptor on bone marrow sinusoidal endothelial cells, indicating a role for Stab2 in HSPC homing and adhesion which is analogous to a report that Stab2 is involved in lymphocyte adhesion in the liver (Jung, Park and Kim, 2007; Qian *et al.*, 2009). Given the association of HSPCs with sinusoidal endothelial cells in the bone marrow and the expression of Stab2 on endothelial cells in all three haematopoietic tissues (bone marrow, liver and yolk sac) we hypothesise that Stab2 could play a supportive role for HSPCs in the yolk sac as part of an endothelial niche (Kiel *et al.*, 2005; Kunisaki *et al.*, 2013). Although this is highly speculative, the shared expression of Stab2 across several haematopoietic sites is very interesting and worth further exploration.

6.5 Conclusions and future plans

Overall, we have identified two hyaluronan receptors that act as interesting markers for distinct endothelial and haematopoietic populations in the yolk sac

and AGM. CD44 is particularly intriguing, as it appears to be a cell surface marker up-regulated just prior to Runx1 expression. Use of CD44 opens up the possibility to better study the earliest stages of EHT and we hope that this marker will prove useful to the field of embryonic haematopoiesis. In the future, it would be interesting to increase the specificity of CD44 as a marker by investigating the expression of variant exons 8, 9 10 with quantitative PCR, full length transcriptome sequencing or isoform specific antibodies. It would be further useful to explore the idea that CD44 is a down-stream target of Notch signalling through FACS analysis of the *Jag1* or *Rbp-j*κ knockout mouse models. This could also provide insight into the receptors that could compensate for CD44 function leading to the development of double knockout mice. One likely candidate would be the hyaluronan binding protein *Hmmr*. In this way we could more fully understand the role of CD44 in EHT. Understanding the mechanism of CD44 action could also be explored through the use of fluorescently tagged hyaluronan molecules enabling us to see how pre-HSPCs associate with hyaluronan as they emerge and migrate in the embryo.

In addition to understanding how CD44 is functionally involved in embryonic haematopoiesis, this work has also opened up several new avenues of research to explore in regards to EHT. We uncovered several new potential regulators of HSPC emergence. These candidates could be further investigated through over-expression studies in our *in vitro* system of EHT or through the use of conditional knockout mouse models. Furthermore, it would be interesting to look at the epigenetic regulator *Dnmt3b* to understand

whether it plays a role in re-shaping the chromatin during this transition. Being able to isolate cells so early in this transition process by using CD44 also means that we can explore the very first changes in the chromatin using more global techniques such as ATAC-seq. Finally, it would be worthwhile to further investigate the metabolic changes we observed during EHT. As the differences we observe in glycolysis, the TCA cycle and autophagy are based on transcriptional data it would be useful to validate these results with direct measurements of metabolic by-products such as lactate or with fluorescent staining of autophagosomes or active mitochondrial potential.

The work we report here on Stab2 expression in the yolk sac is in the preliminary stages of research. Our main finding is that by excluding Stab2 expression we can greatly enrich for haematopoietic progenitors in this tissue. However, this project poses interesting questions and possibilities due to the similarity of these yolk sac endothelial cells with the liver and bone marrow. The next steps will be to perform imaging analysis to see if these endothelial cells interact with haematopoietic progenitors and to over-express Stab2 in our mESC culture system to see whether EHT can be inhibited. Ultimately it would be interesting to use Stab2 in a lineage tracing experiment to understand whether Stab2+ endothelium contributes to the haematopoietic lineage.

References

Adamo, L., Naveiras, O., Wenzel, P. L., McKinney-Freeman, S., Mack, P. J., Gracia-Sancho, J., Suchy-Dicey, A., Yoshimoto, M., Lensch, M. W., Yoder, M. C., Garcia-Cardeña, G. and Daley, G. Q. (2009) 'Biomechanical forces promote embryonic haematopoiesis', *Nature*, 459(7250), pp. 1131–1135.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A. and Goecks, J. (2016) 'The Galaxy platform for accessible, reproducible and collaborative biomedical analyses', *Nucleic acids research*, 44(W1), pp. W3–W10.

Aiuti, B. a, Webb, I. J., Bleul, C., Springer, T. and Gutierrez-Ramos, J. C. (1997) 'The Chemokine SDF-1 Is a Chemoattractant for Human CD34+ Hematopoietic Progenitor Cells and Provides a New Mechanism to Explain the Mobilization of CD34+ Progenitors to Peripheral Blood', *The Journal of Experimental Medicine*, 185(1), pp. 111–20.

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. and Clarke, M. F. (2003) 'Prospective identification of tumorigenic breast cancer cells', *Proceedings of the National Academy of Sciences*, 100(7), pp. 3983–3988.

Antonchuk, J., Sauvageau, G. and Humphries, R. (2002) 'HOXB4-induced expansion of adult hematopoietic stem cells ex vivo', *Cell*, 109, pp. 39–45.

Anttila, M. A., Tammi, R. H., Tammi, M. I., Syrja, K. J. and Saarikoski, S. V (2000) 'High Levels of Stromal Hyaluronan Predict Poor Disease Outcome in Epithelial Ovarian Cancer High Levels of Stromal Hyaluronan Predict Poor Disease Outcome in Epithelial Ovarian Cancer 1', *Cancer research*, pp. 150–155.

Ara, T., Tokoyoda, K., Sugiyama, T., Egawa, T., Kawabata, K. and Nagasawa, T. (2003) 'Long-Term Hematopoietic Stem Cells Require Stromal Cell-Derived Factor-1 for Colonizing Bone Marrow during Ontogeny.', *Immunity,* 19(2), pp. 257-267.

Aruffo, a, Stamenkovic, I., Melnick, M., Underhill, C. B. and Seed, B. (1990) 'CD44 is the principal cell surface receptor for hyaluronate.', *Cell*, 61(7), pp. 1303–1313.

Auvinen, P., Tammi, R., Parkkinen, J., Tammi, M., Ågren, U., Johansson, R., Hirvikoski, P., Eskelinen, M. and Kosma, V. M. (2000) 'Hyaluronan in peritumoral stroma and malignant cells associates with breast cancer spreading and predicts survival', *American Journal of Pathology*, 156(2), pp. 529–536.

Avigdor, A., Goichberg, P., Shivtiel, S., Dar, A., Peled, A., Samira, S., Kollet, O., Hershkoviz, R., Alon, R., Hardan, I., Ben-hur, H., Naor, D., Nagler, A. and

Lapidot, T. (2004) ‘CD44 and hyaluronic acid cooperate with SDF-1 in the trafficking of human CD34+ stem/progenitor cells to bone marrow’, *Blood*, 103(8), pp. 2981–2989.

Bakkers, J. (2004) ‘Has2 is required upstream of Rac1 to govern dorsal migration of lateral cells during zebrafish gastrulation’, *Development*, 131(3), pp. 525–537.

Banerji, S., Wright, A. J., Noble, M., Mahoney, D. J., Campbell, I. D., Day, A. J. and Jackson, D. G. (2007) ‘Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction’, *Nature Structural and Molecular Biology*, 14(3), pp. 234–239.

Baron, C. S., Kester, L., Klaus, A., Boisset, J. C., Thambyrajah, R., Yvernogeau, L., Kouskoff, V., Lacaud, G., Van Oudenaarden, A. and Robin, C. (2018) ‘Single-cell transcriptomics reveal the dynamic of haematopoietic stem cell production in the aorta’, *Nature Communications*, 9(1).

Batta, K., Florkowska, M., Kouskoff, V. and Lacaud, G. (2014) ‘Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells’, *Cell Reports*, 9(5), pp. 1871–1884.

Baumann, C. I., Bailey, A. S., Li, W., Ferkowicz M.J., Yoder, M. C. and Fleming W. H. (2005) 'CD144 (VE-Cadherin) is transiently expressed by fetal liver hematopoietic stem cells', *Blood,* 104(4), pp. 1010-1016.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. and Weintraub, H. (1990) ‘The protein Id: A negative regulator of helix-loop-helix DNA binding proteins’, *Cell*, 61(1), pp. 49–59.

Bergiers, I., Andrews, T., Vargel Bölükbaşı, Ö., Buness, A., Janosz, E., Lopez-Anguita, N., Ganter, K., Kosim, K., Celen, C., Itır Perçin, G., Collier, P., Baying, B., Benes, V., Hemberg, M. and Lancrin, C. (2018) ‘Single-cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis’, *eLife*, 7, pp. 1–38.

Bertrand, J. Y., Chi, N. C., Santoso, B., Teng, S., Stainier, D. Y. R. and Traver, D. (2010) ‘Haematopoietic stem cells derive directly from aortic endothelium during development.’, *Nature*. Nature Publishing Group, 464(7285), pp. 108–111.

Bertrand, J. Y., Jalil, A., Klaine, M., Jung, S., Cumano, A., Godin, I. and Dc, W. (2005) ‘Three pathways to mature macrophages in the early mouse yolk sac’, *Blood*, 106(9), pp. 3004–3011.

Bertrand, J. Y., Kim, A. D., Violette, E. P., Stachura, D. L., Cisson, J. L. and Traver, D. (2007) ‘Definitive hematopoiesis initiates through a committed erythromyeloid progenitor in the zebrafish embryo’, *Development*, 134(23), pp. 4147–4156.

De Bock, K., Georgiadou, M. and Carmeliet, P. (2013) 'Role of endothelial cell metabolism in vessel sprouting', *Cell Metabolism*. Elsevier Inc., 18(5), pp. 634–647.

De Bock, K., Georgiadou, M., Schoors, S., Kuchnio, A., Wong, B. W., Cantelmo, A. R., Quaegebeur, A., Ghesquière, B., Cauwenberghs, S., Eelen, G., Phng, L. K., Betz, I., Tembuyser, B., Brepoels, K., Welti, J., Geudens, I., Segura, I., Cruys, B., Bifari, F., Decimo, I., Blanco, R., Wyns, S., Vangindertael, J., Rocha, S., Collins, R. T., Munck, S., Daelemans, D., Imamura, H., Devlieger, R., Rider, M., Van Veldhoven, P. P., Schuit, F., Bartrons, R., Hofkens, J., Fraisl, P., Telang, S., Deberardinis, R. J., Schoonjans, L., Vinckier, S., Chesney, J., Gerhardt, H., Dewerchin, M. and Carmeliet, P. (2013) 'Role of PFKFB3-driven glycolysis in vessel sprouting', *Cell*, 154(3), pp. 651–663.

Boisset, J.-C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E. and Robin, C. (2010) 'In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium.', *Nature*. Nature Publishing Group, 464(7285), pp. 116–120.

Boisset, J.-C., Clapes, T., Van Der Linden, R., Dzierzak, E. and Robin, C. (2013) 'Integrin IIb (CD41) plays a role in the maintenance of hematopoietic stem cell activity in the mouse embryonic aorta', *Biology Open*, 2(5), pp. 525–532.

Bonder, C. S., Clark, S. R., Norman, M. U., Johnson, P. and Kubes, P. (2006) 'Use of CD44 by CD4+Th1 and Th2 lymphocytes to roll and adhere', *Blood*, 107(12), pp. 4798–4806.

Bourguignon, L. Y. W., Peyrollier, K., Xia, W. and Gilad, E. (2008) 'Hyaluronan-CD44 interaction activates stem cell marker Nanog, Stat-3-mediated MDR1 gene expression, and ankyrin-regulated multidrug efflux in breast and ovarian tumor cells', *Journal of Biological Chemistry*, 283(25), pp. 17635–17651.

Breier, G., Breviario, F., Caveda, L., Berthier, R., Schnurch, H., Gotsch, U., Vestweber, D., Risau, W. and Dejana, E. (1996) 'Molecular cloning and expression of murine vascular endothelial- cadherin in early stage development of cardiovascular system', *Blood*, 87(2), pp. 630–641.

Brown, R. L., Reinke, L. M., Damerow, M. S., Perez, D., Chodosh, L. A., Yang, J. and Cheng, C. (2011) 'CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression', *Journal of Clinical Investigation*, 121(3), pp. 1064–1074.

de Bruijn, M. F. T. R. (2000) 'Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo', *The EMBO Journal*, 19(11), pp. 2465–2474.

de Bruijn, M. F. T. R., Ma, X., Robin, C., Ottersbach, K., Sanchez, M. J.,

Dzierzak, E. (2002). 'Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta', *Immunity,* 16, pp. 673-683.

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 and Development*, 19(19), pp. 2331–2342.

Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S. M., Reth, M., Höfer, T. and Rodewald, H.-R. (2015) 'Fundamental properties of unperturbed haematopoiesis from stem cells in vivo', *Nature*. 518, pp. 542-546.

Camaioni, A., Salustri, A., Yanagishita, M. and Hascall, V. C. (1996) 'Proteoglycans and proteins in the extracellular matrix of mouse cumulus cell-oocyte complexes', *Archives of Biochemistry and Biophysics*, 325(2), pp. 190–198.

Camenisch, T. D., Spicer, A. P., Brehm-Gibson, T., Biesterfeldt, J., Augustine, M. Lou, Calabro, A., Kubalak, S., Klewer, S. E. and McDonald, J. A. (2000) 'Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme', *Journal of Clinical Investigation*, 106(3), pp. 349–360.

Campo, G. M., Avenoso, A., Campo, S., D'Ascola, A., Nastasi, G. and Calatroni, A. (2010) 'Small hyaluronan oligosaccharides induce inflammation by engaging both toll-like-4 and CD44 receptors in human chondrocytes', *Biochemical Pharmacology*, 80(4), pp. 480–490.

Cao, G., Savani, R. C., Fehrenbach, M., Lyons, C., Zhang, L., Coukos, G. and DeLisser, H. M. (2006) 'Involvement of endothelial CD44 during in vivo angiogenesis', *American Journal of Pathology*, 169(1), pp. 325–336.

Cao, H., Heazlewood, S. Y., Williams, B., Cardozo, D., Nigro, J., Oteiza, A. and Nilsson, S. K. (2015) 'The role of CD44 in fetal and adult hematopoietic stem cell regulation'. *Haematologica*, 101(1), pp. 26-37

Carmeliet, P., Ferreira, V., Breier, G., Pollefeyt, S., Kieckens, L., Gertsenstein, M., Fahrig, M., Vandenhoeck, A., Harpal, K., Eberhardt, C., Declercq, C., Pawling, J., Moons, L., Collen, D., Risaut, W. and Nagy, A. (1996) 'Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele', *Nature*, 380, pp. 435–439.

Carotta, S., Pilat, S., Mairhofer, A., Schmidt, U., Dolznig, H., Steinlein, P., Dc, W. and Beug, H. (2004) 'Directed differentiation and mass cultivation of pure erythroid progenitors from mouse embryonic stem cells Directed differentiation and mass cultivation of pure erythroid progenitors from mouse embryonic stem cells', *Blood*, 104(6), pp. 1873–1880.

Cartwright, P. (2005) 'LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism', *Development*, 132(5), pp.

885–896.

Carver, E. A., Jiang, R., Lan, Y., Oram, K. F. and Gridley, T. (2001) ‘The mouse snail gene encode a key regulator of teh epithelial-mesenchymal transition’, *Molecular and cellular biology*, 21(23), pp. 8184–8188.

Challen, G. A., Sun, D., Mayle, A., Jeong, M., Luo, M., Rodriguez, B., Mallaney, C., Celik, H., Yang, L., Xia, Z., Cullen, S., Berg, J., Zheng, Y., Darlington, G. J., Li, W. and Goodell, M. A. (2014) ‘Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells’, *Cell Stem Cell*, 15(3), pp. 350–364.

Chen, C., Liu, Y., Liu, R., Ikenoue, T., Guan, K.-L., Liu, Y. and Zheng, P. (2008) ‘TSC–mTOR maintains quiescence and function of hematopoietic stem cells by repressing mitochondrial biogenesis and reactive oxygen species’, *The Journal of Experimental Medicine*, 205(10), pp. 2397–2408.

Chen, M. J., Li, Y., De Obaldia, M. E., Yang, Q., Yzaguirre, A. D., Yamada-Inagawa, T., Vink, C. S., Bhandoola, A., Dzierzak, E. and Speck, N. A. (2011) ‘Erythroid/Myeloid Progenitors and Hematopoietic Stem Cells Originate from Distinct Populations of Endothelial Cells’, *Cell Stem Cell*, 9(6), pp. 541–552.

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*. Nature, 457(7231), pp. 887–891.

Chiang, I. K.-N., Fritzsche, M., Pichol-Thievend, C., Neal, A., Holmes, K., Lagendijk, A., Overman, J., D’Angelo, D., Omini, A., Hermkens, D., Lesieur, E., Fossat, N., Radziewic, T., Liu, K., Ratnayaka, I., Corada, M., Bou-Gharios, G., Tam, P. P. L., Carroll, J., Dejana, E., Schulte-Merker, S., Hogan, B. M., Beltrame, M., De Val, S. and Francois, M. (2017) ‘SoxF factors induce Notch1 expression via direct transcriptional regulation during early arterial development.’, *Development*, 144(20), pp. 2629–2639.

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C. and Keller, G. (1998) ‘A common precursor for hematopoietic and endothelial cells.’, *Development*, 125(4), pp. 725–32.

Chong, D. C., Koo, Y., Xu, K., Fu, S. and Cleaver, O. (2011) ‘Stepwise arteriovenous fate acquisition during mammalian vasculogenesis’, *Developmental Dynamics*, 240(9), pp. 2153–2165.

Chou, S. and Lodish, H. F. (2010) ‘Fetal liver hepatic progenitors are supportive stromal cells for hematopoietic stem cells’, *Proceedings of the National Academy of Sciences*, 107(17), pp. 7799–7804.

Christensen, J. L., Wright, D. E., Wagers, A. J. and Weissman, I. L. (2004) ‘Circulation and chemotaxis of fetal hematopoietic stem cells’, *PLoS Biology*, 2(3), pp. 368–377.

Ciau-Uitz, A., Walmsley, M. and Patient, R. (2000) 'Distinct origins of adult and embryonic blood in Xenopus', *Cell*, 102(6), pp. 787–796.

Ciriza, J., Hall, D., Lu, A., de Sena, J. R., Al-Kuhlani, M. and García-Ojeda, M. E. (2012) 'Single-cell analysis of murine long-term hematopoietic stem cells reveals distinct patterns of gene expression during fetal migration', *PLoS ONE*, 7(1), pp. 1–13.

Clarke, R. L., Yzaguirre, A. D., Yashiro-Ohtani, Y., Bondue, A., Blanpain, C., Pear, W. S., Speck, N. a and Keller, G. (2013) 'The expression of Sox17 identifies and regulates haemogenic endothelium.', *Nature cell biology*, 15(5), pp. 502–10.

Cleaver, O. and Krieg, P. a (1998) 'VEGF mediates angioblast migration during development of the dorsal aorta in Xenopus.', *Development*, 125(19), pp. 3905–3914.

Clements, W. K., Kim, A. D., Ong, K. G., Moore, J. C., Lawson, N. D. and Traver, D. (2011) 'A somitic Wnt16/Notch pathway specifies haematopoietic stem cells.', *Nature*, 474(7350), pp. 220–224.

Coffin, J. D. and Poole, T. J. (1988) 'Embryonic vascular development: immunohistochemical identification of the origin and subsequent morphogenesis of the major vessel primordia in quail embryos.', *Development*, 102, pp. 735–748.

Cohen, M., Klein, E., Geiger, B. and Addadi, L. (2003) 'Organization and adhesive properties of the hyaluronan pericellular coat of chondrocytes and epithelial cells', *Biophysical Journal*, 85(3), pp. 1996–2005.

Conneally, E., Cashman, J., Petzer, A. and Eaves, C. (1997) 'Expansion in vitro of Transplantable Human Cord Blood Stem Cells Demonstrated Using a Quantitative Assay of their Lympho-Myeloid Repopulating Activity in Nonobese Diabetic- Scid / Scid Mice', *Proceedings of the National Academy of Sciences*, 94, pp. 9836–9841.

Corada, M., Orsenigo, F., Morini, M. F., Pitulescu, M. E., Bhat, G., Nyqvist, D., Breviario, F., Conti, V., Briot, A., Iruela-Arispe, M. L., Adams, R. H. and Dejana, E. (2013) 'Sox17 is indispensable for acquisition and maintenance of arterial identity', *Nature Communications*, 4, pp. 1–14.

Costa, G., Mazan, A., Gandillet, A., Pearson, S., Lacaud, G. and Kouskoff, V. (2012) 'SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development', *Development*, 139(9), pp. 1587–1598.

Crisan, M., Solaimani Kartalaei, P., Neagu, A., Karkanpouna, S., Yamada-Inagawa, T., Purini, C., Vink, C. S., van der Linden, R., van Ijcken, W., Chuva de Sousa Lopes, S. M., Monteiro, R., Mummery, C. and Dzierzak, E. (2016) 'BMP and Hedgehog Regulate Distinct AGM Hematopoietic Stem Cells

Ex Vivo', *Stem Cell Reports*, 6, pp. 1–13.

Csoka, A. B., Frost, G. I. and Stern, R. (2001) 'The six hyaluronidase-like genes in the human and mouse genomes', *Matrix Biology*, 20(8), pp. 499–508.

Cumano, A., Ferraz, J. C., Klaine, M., Di Santo, J. P. and Godin, I. (2001) 'Intraembryonic, but not yolk sac hematopoietic precursors, isolated before circulation, provide long-term multilineage reconstitution', *Immunity*, 15(3), pp. 477–485.

D'Souza, S. L., Elefanty, A. G. and Keller, G. (2005) 'SCL / Tal-1 is essential for hematopoietic commitment of the hemangioblast but not for its development', *Blood*, 105(10), pp. 3862–3870.

D'Souza, S., Park, S. Y. and Kim, I. S. (2013) 'Stabilin-2 acts as an engulfment receptor for the phosphatidylserine-dependent clearance of primary necrotic cells', *Biochemical and Biophysical Research Communications*, 432(3), pp. 412–417.

Day, A. J. and Prestwich, G. D. (2002) 'Hyaluronan-binding proteins: Tying up the giant', *Journal of Biological Chemistry*, 277(7), pp. 4585–4588.

DeGrendele, H. C., Estess, P. and Siegelman, M. H. (1997) 'Requirement for CD44 in activated T cell extravasation into an inflammatory site.', *Science*, 278(5338), pp. 672–675.

Delmage, J. M., Powars, D. R., Jaynes, P. K. and Allerton, S. E. (1986) 'The selective suppression of immunogenicity by hyaluronic acid.', *Annals of clinical laboratory science*, 16(4), pp. 303–310.

Dickson, M. C., Martin, J. S., Cousins, F. M., Kulkarni, A. B., Karlsson, S. and Akhurst, R. J. (1995) 'Defective haematopoiesis and vasculogenesis in transforming growth factor-beta 1 knock out mice.', *Development*, 121(6), pp. 1845–54.

Dieterlen-Lievre, F. (1975) 'On the origin of haemopoietic stem cells in the avian embryo: an experimental approach.', *Journal of embryology and experimental morphology*, 33(3), pp. 607–619.

Dieterlen-Lièvre, F. and Martin, C. (1981) 'Diffuse intraembryonic hemopoiesis in normal and chimeric avian development', *Developmental Biology*, 88(1), pp. 180–191.

Dimitroff, C. J., Lee, J. Y., Rafii, S., Fuhlbrigge, R. C. and Sackstein, R. (2001) 'CD44 is a major E-selectin ligand on human hematopoietic progenitor cells', *J.Cell Biol.*, 153, pp. 1277–1286.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. and Kemler, R. (1985) 'The in vitro development of blastocyst-derived embryonic stem cell lines:

formation of visceral yolk sac, blood islands and myocardium.', *Journal of embryology and experimental morphology*, 87, pp. 27–45.

Doulatov, S., Vo, L. T., Chou, S. S., Kim, P. G., Arora, N., Li, H., Hadland, B. K., Bernstein, I. D., Collins, J. J., Zon, L. I. and Daley, G. Q. (2013) 'Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors', *Cell Stem Cell*, 13(4), pp. 459–470.

Dowdy, C. R., Xie, R., Frederick, D., Hussain, S., Zaidi, S. K., Vradii, D., Javed, A., Li, X., Jones, S. N., Lian, J. B., van Wijnen, A. J., Stein, J. L. and Stein, G. S. (2010) 'Definitive hematopoiesis requires Runx1 C-terminal-mediated subnuclear targeting and transactivation', *Human Molecular Genetics*, 19(6), pp. 1048–1057.

Du, L., Wang, H., He, L., Zhang, J., Ni, B., Wang, X., Jin, H., Cahuzac, N., Mehrpour, M., Lu, Y. and Chen, Q. (2008) 'CD44 is of functional importance for colorectal cancer stem cells', *Clinical Cancer Research*, 14(21), pp. 6751–6760.

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', *PNAS,* 104(52), pp. 20838-20843.

Duarte, A., Hirashima, M., Benedito, R., Trindade, A., Diniz, P., Bekman, E., Costa, L., Henrique, D. and Rossant, J. (2004) 'Dosage-sensitive requirement for mouse Dll4 in artery development', *Genes and Development*, 18(20), pp. 2474–2478.

Dumont, D. J., Fong, G. ‐H, Puri, M. C., Gradwohl, G., Alitalo, K. and Breitman, M. L. (1995) 'Vascularization of the mouse embryo: A study of flk ‐1, tek, tie, and vascular endothelial growth factor expression during development', *Developmental Dynamics*, 203(1), pp. 80–92.

Elagib, K. E., Racke, F. K., Mogass, M., Khetawat, R., Delehanty, L. L. and Goldfarb, A. N. (2003) 'RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation', *Blood*, 101(11), pp. 4333–4341.

Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J. H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., Yakubov, E., Cooke, J., Thomson, J. A. and Slukvin, I. (2014) 'Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators', *Nature Communications*, 5, pp. 1–11.

Ema, H. and Nakauchi, H. (2000) 'Expansion of hematopoietic stem cells in the developing liver of a mouse embryo', 95(7), pp. 2284–2289.

Ema, M., Faloon, P., Zhang, W. J., Hirashima, M., Reid, T., Stanford, W. L., Orkin, S., Choi, K. and Rossant, J. (2003) 'Combinatorial effects of Flk1 and Tal1 on vascular and hematopoietic development in the mouse', *Genes &*

Development, 17, pp. 380–393.

English, N. M., Lesley, J. F. and Hyman, R. (1998) 'Site-specific de-N-glycosylation of CD44 can activate hyaluronan binding, and CD44 activation states show distinct threshold densities for hyaluronan binding', *Cancer Research*, 58(16), pp. 3736–3742.

Eto, K., Leavitt, A. L., Nakano, T. and Shattil, S. J. (2003) 'Development and Analysis of Megakaryocytes from Murine Embryonic Stem Cells', *Methods in Enzymology*, 365, pp. 142–158.

Evanko, S. P., Angello, J. C. and Wight, T. N. (1999) 'Formation of Hyaluronan- and Versican-Rich Pericellular Matrix Is Required for Proliferation and Migration of Vascular Smooth Muscle Cells', *Arteriosclerosis, Thrombosis, and Vascular Biology*, 19, pp. 1004–1013.

Evanko, S. P., Tammi, M. I., Tammi, R. H. and Wight, T. N. (2007) 'Hyaluronan-dependent pericellular matrix', *Advanced Drug Delivery Reviews*, 59(13), pp. 1351–1365.

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

Falkowski, M., Schledzewski, K., Hansen, B. and Goerdt, S. (2003) 'Expression of stabilin-2, a novel fasciclin-like hyaluronan receptor protein, in murine sinusoidal endothelia, avascular tissues, and at solid/ liquid interfaces', *Histochemistry and Cell Biology*, 120(5), pp. 361–369.

Feinberg, R. N. and Beebe, D. C. (1983) 'Hyaluronate in vasculogenesis', *Science*, 220(4602), pp. 1177–1179.

Ferkowicz, M. J., Starr, M., Xie, X., Li, W., Johnson, S. a, Shelley, W. C., Morrison, P. R. and Yoder, M. C. (2003) 'CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo.', *Development*, 130(18), pp. 4393–4403.

Ferkowicz, M. J. and Yoder, M. C. (2005) 'Blood island formation: Longstanding observations and modern interpretations', *Experimental Hematology*, 33(9), pp. 1041–1047.

Fischer, A., Schumacher, N., Maier, M., Sendtner, M. and Gessler, M. (2004) 'The Notch target gene Hey1 and Hey2 are required for embryonic vascular development', *Genes & Development*, 18, pp. 901–911.

Fong, G. H., Rossant, J., Gertsenstein, M. and Breitman, M. L. (1995) 'Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium', *Nature*, 376, pp. 66–70.

Fouquet, B., Weinstein, B. M., Serluca, F. C. and Fishman, M. C. (1997) 'Vessel patterning in the embryo of the zebrafish: Guidance by notochord',

Developmental Biology, 183(1), pp. 37–48.

Fujimoto, T., Kohata, S., Suzuki, H., Miyazaki, H., Fujimura, K. and Dc, W. (2003) 'Production of functional platelets by differentiated embryonic stem (ES) cells in vitro', *Blood*, 102(12), pp. 4044–4051.

Galandrini, R., Galluzzo, E., Albi, N., Grossi, C. E. and Velardi, A. (1994) 'Hyaluronate is costimulatory for human T cell effector functions and binds to CD44 on activated T cells.', *J.Immunol.*, 153, pp. 21–31.

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., Maeda, T., Espinosa, L., Kopan, R. and Bigas, A. (2015) 'Notch signal strength controls cell fate in the haemogenic endothelium', *Nature Communications*, 6, pp. 8510

Gao, A. C., Lou, W., Dong, J. and Isaacs, J. T. (1997) 'CD44 is a metastasis suppressor gene for prostatic cancer located on human chromosome 11p13'', *Cancer Research*, 57, pp. 846–849.

Garcia-Porrero, J. A., Godin, I. E. and Dieterlen-Lièvre, F. (1995) 'Potential intraembryonic hemogenic sites at pre-liver stages in the mouse', *Anatomy and Embryology*, 192(5), pp. 425–435.

Géraud, C., Schledzewski, K., Demory, A., Klein, D., Kaus, M., Peyre, F., Sticht, C., Evdokimov, K., Lu, S., Schmieder, A. and Goerdt, S. (2010) 'Liver sinusoidal endothelium: A microenvironment-dependent differentiation program in rat including the novel junctional protein liver endothelial differentiation-associated protein-1', *Hepatology*, 52(1), pp. 313–326.

Gerety, S. S., Wang, H. U., Chen, Z. F. and Anderson, D. J. (1999) 'Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development', *Molecular Cell*, 4(3), pp. 403–414.

Gerhardt, H., Golding, M., Fruttiger, M., Ruhrberg, C., Lundkvist, A., Abramsson, A., Jeltsch, M., Mitchell, C., Alitalo, K., Shima, D. and Betsholtz, C. (2003) 'VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia', *Journal of Cell Biology*, 161(6), pp. 1163–1177.

Gering, M. and Patient, R. (2005) 'Hedgehog signaling is required for adult blood stem cell formation in zebrafish embryos', *Developmental Cell*, 8(3), pp. 389–400.

Ghatak, S., Misra, S. and Toole, B. P. (2002) 'Hyaluronan oligosaccharides inhibit anchorage-independent growth of tumor cells by suppressing the phosphoinositide 3-kinase/Akt cell survival pathway', *Journal of Biological Chemistry*, 277(41), pp. 38013–38020.

Ghatak, S., Misra, S. and Toole, B. P. (2005) 'Hyaluronan constitutively

regulates ErbB2 phosphorylation and signaling complex formation in carcinoma cells', *Journal of Biological Chemistry*, 280(10), pp. 8875–8883.

Ghiaur, G., Ferkowicz, M. J., Milsom, M. D., Bailey, J., Witte, D., Cancelas, J. A., Yoder, M. C. and Williams, D. A. (2008) 'Rac1 is essential for intraembryonic hematopoiesis and for the initial seeding of fetal liver with definitive hematopoietic progenitor cells', *Blood*, 111(7), pp. 3322–3330.

Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M. F., Conway, S. J., Ng, L. G., Stanley, E. R., Samokhvalov, I. M. and Merad, M. (2010) 'Fate mapping analysis reveals that adult microglia derive from primitive macrophages', *Science*, 330(6005), pp. 841–845.

Glimm, H. and Eaves, C. J. (1999) 'Direct evidence for multiple self-renewal divisions of human in vivo repopulating hematopoietic cells in short-term culture', *The Journal of the American Society of Hematology*, 94(7), pp. 2161–2168.

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., De Bruijn, M. F., Geissmann, F. and Rodewald, H. R. (2015) 'Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors', *Nature*, 518(7540), pp. 547–551.

Guiu, J., Shimizu, R., D'Altri, T., Fraser, S. T., Hatakeyama, J., Bresnick, E. H., Kageyama, R., Dzierzak, E., Yamamoto, M., Espinosa, L. and Bigas, A. (2013) 'Hes repressors are essential regulators of hematopoietic stem cell development downstream of Notch signaling', *The Journal of Experimental Medicine*, 210(1), pp. 71–84.

Günthert, U., Hofmann, M., Rudy, W., Reber, S., Zöller, M., Hauβmann, I., Matzku, S., Wenzel, A., Ponta, H. and Herrlich, P. (1991) 'A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells', *Cell*, 65(1), pp. 13–24.

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., Tomlinson, S. and Smith, A. (2009) 'Oct4 and LIF/Stat3 Additively Induce Krüppel Factors to Sustain Embryonic Stem Cell Self-Renewal', *Cell Stem Cell*, 5(6), pp. 597–609.

Van Ham, T. J., Kokel, D. and Peterson, R. T. (2012) 'Apoptotic cells are cleared by directional migration and elmo1-dependent macrophage engulfment', *Current Biology*, 22(9), pp. 830–836.

Harris, E. N., Weigel, J. A. and Weigel, P. H. (2008) 'The human hyaluronan receptor for endocytosis (HARE/stabilin-2) is a systemic clearance receptor for heparin', *Journal of Biological Chemistry*, 283(25), pp. 17341–17350.

Harris, E. N. and Weigel, P. H. (2008) 'The ligand-binding profile of HARE: Hyaluronan and chondroitin sulfates A, C, and D bind to overlapping sites distinct from the sites for heparin, acetylated low-density lipoprotein, dermatan

sulfate, and CS-E', *Glycobiology*, 18(8), pp. 638–648.

Hasenauer, S., Malinger, D., Koschut, D., Pace, G., Matzke, A., von Au, A. and Orian-Rousseau, V. (2013) 'Internalization of Met Requires the Co-Receptor CD44v6 and Its Link to ERM Proteins', *PLoS ONE*, 8(4), p. e62357.

Herglotz, J., Kuvardina, O. N., Kolodziej, S., Kumar, A., Hussong, H., Grez, M. and Lausen, J. (2013) 'Histone arginine methylation keeps RUNX1 target genes in an intermediate state', *Oncogene*, 32(20), pp. 2565–2575.

Hernández, D., Miquel-Serra, L., Docampo, M. J., Marco-Ramell, A., Cabrera, J., Fabra, A. and Bassols, A. (2011) 'V3 versican isoform alters the behavior of human melanoma cells by interfering with CD44/ErbB-dependent signaling', *Journal of Biological Chemistry*, 286(2), pp. 1475–1485.

Herzog, Y., Guttmann-Raviv, N. and Neufeld, G. (2005) 'Segregation of arterial and venous markers in subpopulations of blood islands before vessel formation', *Developmental Dynamics*, 232(4), pp. 1047–1055.

Herzog, Y., Kalcheim, C., Kahane, N., Reshef, R. and Neufeld, G. (2001) 'Differential expression of neuropilin-1 and neuropilin-2 in arteries and veins', *Mechanisms of Development*, 109(1), pp. 115–119.

Hirai, H., Samokhvalov, I. M., Fujimoto, T., Nishikawa, S., Imanishi, J. and Nishikawa, S. I. (2005) 'Involvement of Runx1 in the down-regulation of fetal liver kinase-1 expression during transition of endothelial cells to hematopoietic cells', *Blood*, 106(6), pp. 1948–1955.

Hirano, H., Screaton, G. R., Bell, M. V, Jackson, D. G., Bell, J. I. and Hodes, R. J. (1993) 'CD44 isoform expression mediated by alternative splicing: tissue-specific regulation in mice', *International Immunology*, 6(1), pp. 49–59.

Hirose, Y., Saijou, E., Sugano, Y., Takeshita, F., Nishimura, S., Nonaka, H., Chen, Y.-R., Sekine, K., Kido, T., Nakamura, T., Kato, S., Kanke, T., Nakamura, K., Nagai, R., Ochiya, T. and Miyajima, A. (2012) 'Inhibition of Stabilin-2 elevates circulating hyaluronic acid levels and prevents tumor metastasis', *Proceedings of the National Academy of Sciences*, 109(11), pp. 4263–4268.

Hirsch, E., Iglesias, a, Potocnik, a J., Hartmann, U. and Fässler, R. (1996) 'Impaired migration but not differentiation of haematopoietic stem cells in the absence of beta1 integrins.', *Nature*, 380, pp. 171–175.

Ho, T. T., Warr, M. R., Adelman, E. R., Lansinger, O. M., Flach, J., Verovskaya, E. V, Figueroa, M. E. and Passegué, E. (2017) 'Autophagy maintains the metabolism and function of young and old stem cells', *Nature*, 543(7644), pp. 205–210.

Hofmann, M., Rudy, W., Zöller, M., Zã, M., Tã, C., Ponã-a, H., Herrlich, P. and Giinthert, U. (1991) 'CD44 Splice Variants Confer Metastatic Behavior in

Rats□: Homologous Sequences Are Expressed in Human Tumor Cell Lines', *Cancer Research*, 51, pp. 5292–5297.

Hole, N., Graham, G. J., Menzel, U. and Ansell, J. D. (1996) 'A limited temporal window for the derivation of multilineage repopulating hematopoietic progenitors during embryonal stem cell differentiation in vitro.', *Blood*, 88(4), pp. 1266–76.

Hoogenkamp, M., Lichtinger, M., Krysinska, H., Lancrin, C., Clarke, D., Williamson, A., Mazzarella, L., Ingram, R., Jorgensen, H., Fisher, A., Tenen, D. G., Kouskoff, V., Lacaud, G. and Bonifer, C. (2009) 'Early chromatin unfolding by RUNX1: A molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program', *Blood*, 114(2), pp. 299–309.

Huang, C., Sheikh, F., Hollander, M., Cai, C., Becker, D., Chu, P.-H., Evans, S. and Chen, J. (2003) 'Embryonic atrial function is essential for mouse embryogenesis, cardiac morphogenesis and angiogenesis', *Development*, 130(24), p. 6111 LP-6119.

Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J. D., Shivdasani, R. A., Gilliland, D. G., Speck, N. A., Nimer, S. D. and Tenen, D. G. (2008) 'PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis', *Nature Genetics*, 40(1), pp. 51–60.

Huang, H. and Auerbach, R. (1993) 'Identification and characterization of hematopoietic stem cells from the yolk sac of the early mouse embryo.', *Proceedings of the National Academy of Sciences*, 90(21), pp. 10110–10114.

Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. and Keller, G. (2004) 'Haemangioblast commitment is initiated in the primitive streak of the mouse embryo', *Nature*, 432, pp. 625–630.

Hughes, E. N., Colombatti, A. and August, J. T. (1983) 'Murine cell surface glycoproteins. Purification of the polymorphic Pgp-1 antigen and analysis of its expression on macrophages and other myeloid cells', *Journal of Biological Chemistry*, 258(2), pp. 1014–1021.

Ikuta, K. and Weissman, I. L. (1992) 'Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation', *Proceedings of the National Academy of Sciences*, 89(4), pp. 1502–1506.

Itano, N., Atsumi, F., Sawai, T., Yamada, Y., Miyaishi, O., Senga, T., Hamaguchi, M. and Kimata, K. (2002) 'Abnormal accumulation of hyaluronan matrix diminishes contact inhibition of cell growth and promotes cell migration.', *Proceedings of the National Academy of Sciences*, 99(6), pp. 3609–14.

Itano, N., Sawai, T., Lenas, P., Yamada, Y., Imagawa, M., Shinomura, T.,

Hamaguchi, M., Yoshida, Y., Miyauchi, S., Spicer, A. P., Mcdonald, J. a, Kimata, K. and Ohnuki, Y. (1999) 'Three Isoforms of Mammalian Hyaluronan Synthases Have Distinct Enzymatic Properties', *Journal of Biological Chemistry*, 274(35), pp. 25085–25092.

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', *The Journal of Experimental Medicine*, 208(12), pp. 2417–2427.

Jaffredo, T., Gautier, R., Eichmann, A., Dieterlen-Lièvre, F. (1998) 'Intra-aortic hemopoietic cells are derived from endothelial cells during ontogeny.', *Development,* 125, pp. 4575-4583.

Jalkanen, S., Bargatze, R. E., de los Toyos, J. and Butcher, E. C. (1987) 'Lymphocyte Recognition of High Endothelium: Antibodies to Distinct Epitopes of an 85-95-kD Glycoprotein Antigen Differentially Inhibit Lymphocyte Binding to Lymph Node, Mucosal, or Synovial Endothelial Cells', *The Journal of Cell Biology,* 105, pp. 983–990.

Jalkanen, S., Bargatze, R. F., Herron, L. R. and Butcher, E. C. (1986) 'A lymphoid cell surface glycoprotein involved in endothelial cell recognition and lymphocyte homing in man.', *European Journal of Immunology*, 16(10), pp. 1195–202.

Jalkanen, S. and Jalkanent, M. (1992) 'Lymphocyte CD44 Binds the COOH-terminal Heparin-binding Domain of Fibronectin'. *The Journal of Cell Biology,* 116, 817-825.

Jang, S. J., Ohtani, K., Fukuoh, A., Yoshizaki, T., Fukuda, M., Motomura, W., Mori, K., Fukuzawa, J., Kitamoto, N., Yoshida, I., Suzuki, Y. and Wakamiya, N. (2009) 'Scavenger receptor collectin placenta 1 (CL-P1) predominantly mediates zymosan phagocytosis by human vascular endothelial cells', *Journal of Biological Chemistry*, 284(6), pp. 3956–3965.

Ji, R. P., Phoon, C. K. L., Aristizábal, O., McGrath, K. E., Palis, J. and Turnbull, D. H. (2003) 'Onset of cardiac function during early mouse embryogenesis coincides with entry of primitive erythroblasts into the embryo proper', *Circulation Research*, 92(2), pp. 133–135.

Jin, L., Hope, K. J., Zhai, Q., Smadja-Joffe, F. and Dick, J. E. (2006) 'Targeting of CD44 eradicates human acute myeloid leukemic stem cells', *Nat.Med.*, 12, pp. 1167–1174.

Johnson, G. R. and Moore, M. A. S. (1975) 'Role of stem cell migration in initiation of mouse foetal liver haemopoiesis', *Nature*, 258(5537), pp. 726–728.

Jung, M.-Y., Park, S.-Y. and Kim, I.-S. (2007) 'Stabilin-2 is involved in lymphocyte adhesion to the hepatic sinusoidal endothelium via the interaction

with M 2 integrin', *Journal of Leukocyte Biology*, 82(5), pp. 1156–1165.

Kabrun, N., Bühring, H. J., Choi, K., Ullrich, a, Risau, W. and Keller, G. (1997) 'Flk-1 expression defines a population of early embryonic hematopoietic precursors.', *Development*, 124(10), pp. 2039–2048.

Kallianpur, B. A. R., Jordan, J. E. and Brandt, S. J. (1994) 'The SCL/TAL-I Gene Is Expressed in Progenitors of Both the Hematopoietic and Vascular Systems During Embryogenesis', 83(5), pp. 1200–1208.

Kamei, M., Brian Saunders, W., Bayless, K. J., Dye, L., Davis, G. E. and Weinstein, B. M. (2006) 'Endothelial tubes assemble from intracellular vacuoles in vivo', *Nature*, 442(7101), pp. 453–456.

Kawasaki, T., Kitsukawa, T., Bekku, Y., Matsuda, Y., Sanbo, M., Yagi, T. and Fujisawa, H. (1999) 'A requirement for neuropilin-1 in embryonic vessel formation.', *Development*, 126(21), pp. 4895–4902.

Kaya, G., Rodriguez, I., Jorcano, J. L., Vassalli, P. and Stamenkovic, I. (1997) 'Selective suppression of CD44 in keratinocytes of mice bearing an antisense CD44 transgene driven by a tissue-specific promoter disrupts hyaluronate metabolism in the skin and impairs keratinocyte proliferation', *Genes and Development*, 11(8), pp. 996–1007.

Keller, G., Kennedy, M., Papayannopoulou, T. and Wiles, M. V (1993) 'Hematopoietic commitment during embryonic stem cell differentiation in culture.', *Molecular and cellular biology*, 13(1), pp. 473–86.

Keller, G. M. (1995) 'In vitro differentiation of embryonic stem cells', *Current Opinion in Cell Biology*, 7(6), pp. 862–869.

Keller, J. R., McNiece, I. K., Sill, K. T., Ellingsworth, L. R., Quesenberry, P. J., Sing, G. K. and Ruscetti, F. W. (1990) 'Transforming growth factor beta directly regulates primitive murine hematopoietic cell proliferation', *Blood*, 75(3), pp. 596–602.

Khan, J. A., Mendelson, A., Kunisaki, Y., Birbrair, A., Kou, Y., Arnal-Estape, A., Pinho, S., Ciero, P., Nakahara, F., Ma'ayan, A., Bergman, A., Merad, M. and Frenette, P. S. (2016) 'Fetal liver hematopoietic stem cell niches associate with portal vessels', *Science*, 351(6269), pp. 176–180.

Kiel, M. J., Yilmaz, Ö. H., Iwashita, T., Yilmaz, O. H., Terhorst, C. and Morrison, S. J. (2005) 'SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells', *Cell*, 121(7), pp. 1109–1121.

Kim, G. W., Park, S. Y. and Kim, I. S. (2016) 'Novel function of stabilin-2 in myoblast fusion: the recognition of extracellular phosphatidylserine as a "fuse-me" signal', *BMB reports*, 49(6), pp. 303–304.

Kim, I., Yilmaz, O. H. and Morrison, S. J. (2005) 'CD144 (VE-Cadherin) is transiently expressed by fetal liver hematopoietic stem cells', *Blood,* 106(3), pp. 903-905.

Kim, K., Kim, I. K., Yang, J. M., Lee, E., Koh, B. I., Song, S., Park, J., Lee, S., Choi, C., Kim, J. W., Kubota, Y., Koh, G. Y. and Kim, I. (2016) 'SoxF Transcription Factors Are Positive Feedback Regulators of VEGF Signaling', *Circulation Research*, 119(7), pp. 839–852.

Kim, P. G., Albacker, C. E., Lu, Y. -f., Jang, I. -h., Lim, Y., Heffner, G. C., Arora, N., Bowman, T. V., Lin, M. I., Lensch, M. W., De Los Angeles, A., Zon, L. I., Loewer, S. and Daley, G. Q. (2013) 'Signaling axis involving Hedgehog, Notch, and Scl promotes the embryonic endothelial-to-hematopoietic transition', *Proceedings of the National Academy of Sciences*, 110(2), pp. E141–E150.

Kim, S., Bae, D. J., Hong, M., Park, S. Y. and Kim, I. S. (2010) 'The conserved histidine in epidermal growth factor-like domains of stabilin-2 modulates pH-dependent recognition of phosphatidylserine in apoptotic cells', *International Journal of Biochemistry and Cell Biology*, 42(7), pp. 1154–1163.

Kingsley, P. D., Malik, J., Fantauzzo, K. A. and Palis, J. (2004) 'Yolk sac – derived primitive erythroblasts enucleate during mammalian embryogenesis', 104(1), pp. 19–26.

Kissa, K. and Herbomel, P. (2010) 'Blood stem cells emerge from aortic endothelium by a novel type of cell transition.', *Nature*, 464(7285), pp. 112–115.

Knudson, W., Aguiar, D. J., Hua, Q. and Knudson, C. B. (1996) 'CD44-anchored hyaluronan-rich pericellular matrices: An ultrastructural and biochemical analysis', *Experimental Cell Research*, 228(2), pp. 216–228.

Knudson, W., Bartnik, E. and Knudson, C. B. (1993) 'Assembly of pericellular matrices by COS-7 cells transfected with CD44 lymphocyte-homing receptor genes.', *Proceedings of the National Academy of Sciences*, 90(9), pp. 4003–7.

Knudson, W. and Knudson, C. B. (1991) 'Assembly of a chondrocyte-like pericellular matrix on non-chondrogenic cells. Role of the cell surface hyaluronan receptors in the assembly of a pericellular matrix.', *Journal of cell science*, 99 (Pt 2), pp. 227–235.

Kobayashi, H., Butler, J. M., O'Donnell, R., Kobayashi, M., Ding, B. Sen, Bonner, B., Chiu, V. K., Nolan, D. J., Shido, K., Benjamin, L. and Rafii, S. (2010) 'Angiocrine factors from Akt-activated endothelial cells balance self-renewal and differentiation of haematopoietic stem cells', *Nature Cell Biology*, 12(11), pp. 1046–1056.

Kovach, J. S., Marks, P. A., Russell, E. S. and Epler, H. (1967) 'Erythroid cell

development in fetal mice: ultrastructural characteristics and hemoglobin synthesis', *Journal of Molecular Biology*, 25(1), pp. 131–142.

Krause, D. S., Lazarides, K., Von Andrian, U. H. and Van Etten, R. A. (2006) 'Requirement for CD44 in homing and engraftment of BCR-ABL-expressing leukemic stem cells', *Nature Medicine*, 12(10), pp. 1175–1180.

Krebs, L. T., Shutter, J. R., Tanigaki, K., Krebs, L. T., Shutter, J. R., Tanigaki, K., Honjo, T., Stark, K. L. and Gridley, T. (2004) 'Haploinsufficient lethality and formation of arteriovenous malformations in Notch pathway mutants', *Genes & Development*, 18(20), pp. 2469–2473.

Krebs, L. T., Xue, Y., Norton, C. R., Shutter, J. R., Maguire, M., Sundberg, J. P., Gallahan, D., Closson, V., Kitajewski, J., Callahan, R., Smith, G. H., Stark, K. L. and Gridley, T. (2000) 'Notch signaling is essential for vascular morphogenesis in mice.', *Genes & development*, 14(11), pp. 1343–52.

Kumano, K., Chiba, S., Kunisato, A., Sata, M., Saito, T., Nakagami-Yamaguchi, E., Yamaguchi, T., Masuda, S., Shimizu, K., Takahashi, T., Ogawa, S., Hamada, Y. and Hirai, H. (2003) 'Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells', *Immunity*, 18(5), pp. 699–711.

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(21), pp. 4891–4899.

Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K., Mar, J. C., Bergman, A. and Frenette, P. S. (2013) 'Arteriolar niches maintain haematopoietic stem cell quiescence', *Nature*, 502(7473), pp. 637–643.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, Ö., Wilson, N. K., Möröy, T., Bonifer, C., Göttgens, B., Kouskoff, V. and Lacaud, G. (2012) 'GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment', *Blood*, 120(2), pp. 314–322.

Lancrin, C., Sroczynska, P., Stephenson, C., Allen, T., Kouskoff, V. and Lacaud, G. (2009) 'The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage.', *Nature*, 457(7231), pp. 892–895.

Lau, W. M., Teng, E., Chong, H. S., Lopez, K. A. P., Tay, A. Y. L., Salto-Tellez, M., Shabbir, A., So, J. B. Y. and Chan, S. L. (2014) 'CD44v8-10 is a cancer-specific marker for gastric cancer stem cells', *Cancer Research*, 74(9), pp. 2630–2641.

Lawson, N. D., Scheer, N., Pham, V. N., Kim, C.-H. H., Chitnis, A. B., Campos-Ortega, J. A. and Weinstein, B. M. (2001) 'Notch signaling is required for arterial-venous differentiation during embryonic vascular development.', *Development*, 128(19), pp. 3675–3683.

Leder, a, Kuo, a, Shen, M. M. and Leder, P. (1992) 'In situ hybridization reveals co-expression of embryonic and adult alpha globin genes in the earliest murine erythrocyte progenitors.', *Development*, 116(4), pp. 1041–1049.

Lee, S. J., Park, S. Y., Jung, M. Y., Bae, S. M. and Kim, I. S. (2011) 'Mechanism for phosphatidylserine-dependent erythrophagocytosis in mouse liver', *Blood*, 117(19), pp. 5215–5223.

Legras, S., Levesque, J. P., Charrad, R., Morimoto, K., Le Bousse, C., Clay, D., Jasmin, C. and Smadja-Joffe, F. (1997) 'CD44-mediated adhesiveness of human hematopoietic progenitors to hyaluronan is modulated by cytokines.', *Blood*, 89(6), pp. 1905–1914.

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., Wicha, M., Clarke, M. F. and Simeone, D. M. (2007) 'Identification of pancreatic cancer stem cells', *Cancer Research*, 67(3), pp. 1030–1037.

Li, Y., Li, L., Brown, T. J. and Heldin, P. (2007) 'Silencing of hyaluronan synthase 2 suppresses the malignant phenotype of invasive breast cancer cells', *International Journal of Cancer*, 120(12), pp. 2557–2567.

Liang, J. and Slingerland, J. M. (2003) 'Multiple Roles of the PI3K/PKB (Akt) Pathway in Cell Cycle Progression', *Cell Cycle*, 2(4), pp. 336–342.

Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S. A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M. S., Wu, M., Tenen, D. G., Westhead, D. R., Kouskoff, V., Lacaud, G., Göttgens, B. and Bonifer, C. (2012) 'RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis', *EMBO Journal*, 31(22), pp. 4318–4333.

Lieber, J. G., Webb, S., Suratt, B. T., Young, S. K., Johnson, G. L., Keller, G. M. and Worthen, G. S. (2004) 'The in vitro production and characterization of neutrophils from embryonic stem cells', *Blood*, 103(3), pp. 852–859.

Lin, Y. H. and Yang-Yen, H. F. (2001) 'The Osteopontin-CD44 Survival Signal Involves Activation of the Phosphatidylinositol 3-Kinase/Akt Signaling Pathway', *Journal of Biological Chemistry*, 276(49), pp. 46024–46030.

Lindskog, H., Kim, Y. H., Jelin, E. B., Kong, Y., Guevara-Gallardo, S., Kim, T. N. and Wang, R. A. (2014) 'Molecular identification of venous progenitors in the dorsal aorta reveals an aortic origin for the cardinal vein in mammals', *Development*, 141(5), pp.

Lipponen, P., Aaltomaa, S., Tammi, R., Tammi, M., Ågren, U. and Kosma, V.

M. (2001) ‘High stromal hyaluronan level is associated with poor differentiation and metastasis in prostate cancer’, *European Journal of Cancer*, 37(7), pp. 849–856.

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., Tabrizi, A. R., Shido, K., Rosenwaks, Z., Elemento, O., Speck, N. A., Butler, J. M., Scandura, J. M. and Rafii, S. (2017) ‘Conversion of adult endothelium to immunocompetent haematopoietic stem cells’, *Nature*, 545(7655), pp. 439–445.

Liu, C., Kelnar, K., Liu, B., Chen, X., Calhoun-davis, T., Li, H., Yan, H., Jeter, C., Honorio, S., Wiggins, J., Bader, A. G., Fagin, R., Brown, D. and Tang, D. G. (2011) ‘The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44’, *Nature Medicine*, 17(2), pp. 211–215.

Liu, X., Zheng, H., Yu, W.-M., Cooper, T. M., Bunting, K. D. and Qu, C.-K. (2015) ‘Maintenance of mouse hematopoietic stem cells ex vivo by reprogramming cellular metabolism’, *Blood*, 125(10), pp. 1562-1565.

Lopez, J. I., Camenisch, T. D., Stevens, M. V., Sands, B. J., McDonald, J. and Schroeder, J. A. (2005) ‘CD44 attenuates metastatic invasion during breast cancer progression’, *Cancer Research*, 65(15), pp. 6755–6763.

Lucitti, J. L., Jones, E. A. V., Huang, C., Chen, J., Fraser, S. E. and Dickinson, M. E. (2007) ‘Vascular remodeling of the mouse yolk sac requires hemodynamic force’, *Development*, 134(18), pp. 3317–3326.

Lux, C. T., Yoshimoto, M., Mcgrath, K., Conway, S. J., Palis, J. and Yoder, M. C. (2008) ‘All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac’, *Blood*, 111(7), pp. 3435–3438.

Ma, X., Robin, C., Ottersbach, K. and Dzierzak, E. (2002) ‘The Ly-6A (Sca-1) GFP transgene is expressed in all adult mouse hematopoietic stem cells.’, *Stem cells*, 20(6), pp. 514–21.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A. and McCarroll, S. A. (2015) ‘Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets’, *Cell*, 161(5), pp. 1202–1214.

Maleski, M. P. and Knudson, C. B. (1996) ‘Hyaluronan meditated aggregation of limb bud mesoderm and mesenchymal condensation during chondrogenesis’, *Experimental Cell Research*, 225(1), pp. 55–66.

Marshall, C. J., Kinnon, C. and Thrasher, A. J. (2007) ‘Polarized expression of bone morphogenetic protein-4 in the human aorta-gonad-mesonephros

region', *Blood*, 96(4), pp. 1591–1593.

Martin, G. R. (1981) 'Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.', *Proceedings of the National Academy of Sciences*, 78(12), pp. 7634–7638.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T. and Yokota, T. (1999) 'STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells', *Embo J.*, 18(15), pp. 4261–4269.

Matzke, A., Sargsyan, V., Holtmann, B., Aramuni, G., Asan, E., Sendtner, M., Pace, G., Howells, N., Zhang, W., Ponta, H. and Orian-Rousseau, V. (2007) 'Haploinsufficiency of c-Met in cd44-/- mice identifies a collaboration of CD44 and c-Met in vivo.', *Molecular and cellular biology*, 27(24), pp. 8797–806.

McCourt, P. A. G., Smedsrød, B. H., Melkko, J. and Johansson, S. (1999) 'Characterization of a hyaluronan receptor on rat sinusoidal liver endothelial cells and its functional relationship to scavenger receptors', *Hepatology*, 30(5), pp. 1276–1286.

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', *The Journal of Experimental Medicine*, 214(12), pp. 3731–3751.

McGrath, K. E., Koniski, A. D., Malik, J. and Palis, J. (2003) 'Circulation is established in a stepwise pattern in the mammalian embryo', *Blood*, 101(5), pp. 1669–1676.

Medvinsky, A., Samoylina, N. L., Müller, A. M. and Dzierzak, E. A. (1993) 'An early pre-liver intraembryonic source of CFU-S in the developing mouse', *Nature,* 364, pp. 64-67.

Medvinsky, A. and Dzierzak, E. (1996) 'Definitive hematopoiesis is autonomously initiated by the AGM region', *Cell*, 86(6), pp. 897–906.

Mertens, S., Noll, T., Spahr, R., Krutzfeldt, A. and Piper, H. M. (1990) 'Energetic response of coronary endothelial cells to hypoxia', *American Journal of Physiology*, 258(3), pp. H689–H694.

Mikkola, H. K. A., Fujiwara, Y., Schlaeger, T. M., Traver, D. and Orkin, S. H. (2003) 'Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo', *Blood*, 101(2), pp. 508–516.

Miller, C. L. and Eaves, C. J. (1997) 'Expansion in vitro of adult murine hematopoietic stem cells with transplantable lympho-myeloid reconstituting ability', *Proceedings of the National Academy of Sciences*, 94(25), pp. 13648–13653.

Misra, S., Obeid, L. M., Hannun, Y. A., Minamisawa, S., Berger, F. G., Markwald, R. R., Toole, B. P. and Ghatak, S. (2008) 'Hyaluronan constitutively regulates activation of COX-2-mediated cell survival activity in intestinal epithelial and colon carcinoma cells', *Journal of Biological Chemistry*, 283(21), pp. 14335–14344.

Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y. and Ohki, M. (1991) 't(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1.', *Proceedings of the National Academy of Sciences*, 88(23), pp. 10431–10434.

Mizrahy, S., Raz, S. R., Hasgaard, M., Liu, H., Soffer-Tsur, N., Cohen, K., Dvash, R., Landsman-Milo, D., Bremer, M. G. E. G., Moghimi, S. M. and Peer, D. (2011) 'Hyaluronan-coated nanoparticles: The influence of the molecular weight on CD44-hyaluronan interactions and on the immune response', *Journal of Controlled Release*, 156(2), pp. 231–238.

Moore, M. A. . and Owen, J. J. . (1967) 'Stem-Cell Migration in Developing Myeloid and Lymphoid Systems', *The Lancet*, 290(7517), pp. 658–659.

Moyon, D., Pardanaud, L., Yuan, L., Bréant, C. and Eichmann, A. (2001) 'Plasticity of endothelial cells during arterial-venous differentiation in the avian embryo', *Development (Cambridge, England)*, 128(17), pp. 3359–3370.

Mukouyama, Y. suke, Chiba, N., Hara, T., Okada, H., Ito, Y., Kanamaru, R., Miyajima, A., Satake, M. and Watanabe, T. (2000) 'The AML1 transcription factor functions to develop and maintain hematogenic precursor cells in the embryonic aorta-gonad-mesonephros region', *Developmental Biology*, 220(1), pp. 27–36.

Müller, A. M. and Dzierzak, E. a (1993) 'ES cells have only a limited lymphopoietic potential after adoptive transfer into mouse recipients.', *Development*, 118(4), pp. 1343–1351.

Müller, A. M., Medvinsky, A., Strouboulis, J., Grosveld, F. and Dzierzakt, E. (1994) 'Development of hematopoietic stem cell activity in the mouse embryo', *Immunity*, 1(4), pp. 291–301.

Murayama, E., Kissa, K., Zapata, A., Mordelet, E., Briolat, V., Lin, H. F., Handin, R. I. and Herbomel, P. (2006) 'Tracing Hematopoietic Precursor Migration to Successive Hematopoietic Organs during Zebrafish Development', *Immunity*, 25(6), pp. 963–975.

Nakagawa, M., Ichikawa, M., Kumano, K., Goyama, S., Kawazu, M., Asai, T., Ogawa, S., Kurokawa, M. and Chiba, S. (2006) 'AML1/Runx1 rescues Notch1-null mutation-induced deficiency of para-aortic splanchnopleural hematopoiesis', *Blood*, 108(10), pp. 3329–3334.

Nakano, T., Kodama, H. and Honjo, T. (1994) 'Generation of lymphohematopoietic cells from embryonic stem cells in culture.', *Science*,

265(5175), pp. 1098–1101.

Neame, P. J., Christner, J. E. and Baker, J. R. (1986) 'The primary structure of link protein from rat chondrosarcoma proteoglycan aggregate', *Journal of Biological Chemistry*, 261(8), pp. 3519–3525.

Nedvetzki, S., Gonen, E., Assayag, N., Reich, R., Williams, R. O., Thurmond, R. L., Huang, J.-F., Neudecker, B. A., Wang, F.-S., Wang, F.-S., Turley, E. A. and Naor, D. (2004) 'RHAMM, a receptor for hyaluronan-mediated motility, compensates for CD44 in inflamed CD44-knockout mice: a different interpretation of redundancy.', *Proceedings of the National Academy of Sciences*, 101(52), pp. 18081–6.

Newman, P. J., Berndt, M. C., Gorski, J., Ii, G. C. W., Lyman, S., Paddock, C. and Muller, W. a (1990) 'Molecules of the immunoglobulin gene superfamily', *Cell*, 247, pp. 1219–1222.

Ng, C. E. L., Yokomizo, T., Yamashita, N., Cirovic, B., Jin, H., Wen, Z., Ito, Y. and Osato, M. (2010) 'A Runx1 intronic enhancer marks hemogenic endothelial cells and hematopoietic stem cells', *Stem Cells*, 28(10), pp. 1869–1881.

Nilsson, S. K., Haylock, D. N., Johnston, H. M., Occhiodoro, T., Brown, T. J. and Simmons, P. J. (2003) 'Hyaluronan is synthesized by primitive hemopoietic cells, participates in their lodgment at the endosteum following transplantation, and is involved in the regulation of their proliferation and differentiation in vitro', *Blood*, 101(3), pp. 856–862.

Nishikawa, S. I., Nishikawa, S., Kawamoto, H., Yoshida, H., Kizumoto, M., Kataoka, H. and Katsura, Y. (1998) 'In vitro generation of lymphohematopoietic cells from endothelial cells purified from murine embryos', *Immunity*, 8(6), pp. 761–769.

le Noble, F. (2003) 'Flow regulates arterial-venous differentiation in the chick embryo yolk sac', *Development*, 131(2), pp. 361–375.

Nonaka, H., Tanaka, M., Suzuki, K. and Miyajima, A. (2007) 'Development of murine hepatic sinusoidal endothelial cells characterized by the expression of hyaluronan receptors', *Developmental Dynamics*, 236(8), pp. 2258–2267.

Nonaka, H., Watabe, T., Saito, S., Miyazono, K. and Miyajima, A. (2008) 'Development of stabilin2+endothelial cells from mouse embryonic stem cells by inhibition of TGF β /activin signaling', *Biochemical and Biophysical Research Communications*, 375(2), pp. 256–260.

North, T. E., de Bruijn, M. F. T. R., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E., Speck, N. A. (2002) 'Runx1 expression marks long-term re-populating hematopoietic stem cells in the midgestation mouse embryo', *Immunity,* 16(5), pp. 661-672.

Norton, J. D., Deed, R. W., Craggs, G. and Sablitzky, F. (1998) 'Id helix-loop-helix proteins in cell growth and differentiation', *Trends in Cell Biology*, 8, pp. 58–65.

Nostro, M. C., Cheng, X., Keller, G. M. and Gadue, P. (2008) 'Wnt, Activin, and BMP Signaling Regulate Distinct Stages in the Developmental Pathway from Embryonic Stem Cells to Blood', *Cell Stem Cell*, 2(1), pp. 60–71.

Okamoto, I., Kawano, Y., Tsuiki, H., Sasaki, J., Nakao, M., Matsumoto, M., Suga, M., Ando, M., Nakajima, M. and Saya, H. (1999) 'CD44 cleavage induced by a membrane-associated metalloprotease plays a critical role in tumor cell migration', *Oncogene*, 18(7), pp. 1435–1446.

Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G., Downing, J. R., (1996) 'AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis.', *Cell*, 84(2), pp. 321–30.

Olaku, V., Matzke, A., Mitchell, C., Hasenauer, S., Sakkaravarthi, A., Pace, G., Ponta, H. and Orian-Rousseau, V. (2011) 'c-Met recruits ICAM-1 as a coreceptor to compensate for the loss of CD44 in Cd44 null mice', *Molecular Biology of the Cell*, 22(15), pp. 2777–2786.

Orian-rousseau, V., Chen, L., Sleeman, J. P., Herrlich, P. and Ponta, H. (2002) 'CD44 is required for two consecutive steps in HGF / c-Met signaling CD44 is required for two consecutive steps in HGF / c-Met signaling', *Genes & Development*, pp. 3074–3086.

Orian-Rousseau, V., Morrison, H., Matzke, A., Kastilan, T., Pace, G., Herrlich, P. and Ponta, H. (2007) 'Hepatocyte Growth Factor-induced Ras Activation Requires ERM Proteins Linked to Both CD44v6 and F-Actin', *Molecular biology of the cell*, 18(December), pp. 986–994.

Osawa, M., Hanada, K., Hamada, H. and Nakauchi, H. (1996) 'Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematpoietic stem cell', *Science*, 273(1995), pp. 242–246.

Oshima, M., Oshima, H. and Taketo, M. M. (1996) 'TGF-beta receptor type II deficiency results in defects of yolk sac hematopoiesis and vasculogenesis.', *Developmental biology*, 179(0259), pp. 297–302.

Ottersbach, K. and Dzierzak, E. (2005) 'The murine placenta contains hematopoietic stem cells within the vascular labyrinth region', *Developmental Cell*, 8(3), pp. 377–387.

Palis, J., Robertson, S., Kennedy, M., Wall, C. and Keller, G. (1999) 'Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse.', *Development*, 126(22), pp. 5073–84.

Palis, J. and Yoder, M. C. (2001) 'Yolk-sac hematopoiesis: The first blood

cells of mouse and man', *Experimental Hematology*, 29(8), pp. 927–936.

Pályi-Krekk, Z., Barok, M., Kovács, T., Saya, H., Nagano, O., Szöllosi, J. and Nagy, P. (2008) 'EGFR and ErbB2 are functionally coupled to CD44 and regulate shedding, internalization and motogenic effect of CD44', *Cancer Letters*, 263(2), pp. 231–242.

Park, C., Afrikanova, I., Chung, Y. S., Zhang, W. J., Arentson, E., Fong Gh, G. H., Rosendahl, A. and Choi, K. (2004) 'A hierarchical order of factors in the generation of FLK1- and SCL-expressing hematopoietic and endothelial progenitors from embryonic stem cells.', *Development,* 131(11), pp. 2749–2762.

Park, S. Y., Jung, M. Y., Kim, H. J., Lee, S. J., Kim, S. Y., Lee, B. H., Kwon, T. H., Park, R. W. and Kim, I. S. (2008) 'Rapid cell corpse clearance by stabilin-2, a membrane phosphatidylserine receptor', *Cell Death and Differentiation*, 15(1), pp. 192–201.

Park, S. Y., Jung, M. Y. and Kim, I. S. (2009) 'Stabilin-2 mediates homophilic cell-cell interactions via its FAS1 domains', *FEBS Letters*, 583(8), pp. 1375–1380.

Patrawala, L., Calhoun, T., Schneider-Broussard, R., Li, H., Bhatia, B., Tang, S., Reilly, J. G., Chandra, D., Zhou, J., Claypool, K., Coghlan, L. and Tang, D. G. (2006) 'Highly purified CD44+ prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells', *Oncogene*, 25(12), pp. 1696–1708.

Patterson, L. J., Gering, M., Eckfeldt, C. E., Green, A. R., Verfaillie, C. M., Ekker, S. C. and Patient, R. (2007) 'The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish', *Blood*, 109(6), pp. 2389–2398.

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(15), pp. 2613–2621.

Pei, W., Feyerabend, T. B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., Chen, W., Sauer, S., Wolf, S., Höfer, T. and Rodewald, H. R. (2017) 'Polylox barcoding reveals haematopoietic stem cell fates realized in vivo', *Nature*, 548(7668), pp. 456–460.

Pereira, C. F., Chang, B., Qiu, J., Niu, X., Papatsenko, D., Hendry, C. E., Clark, N. R., Nomura-Kitabayashi, A., Kovacic, J. C., Ma'Ayan, A., Schaniel, C., Lemischka, I. R. and Moore, K. (2013) 'Induction of a hemogenic program in mouse fibroblasts', *Cell Stem Cell*, 13(2), pp. 205–218.

Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V.,

Orkin, S. H. and Costantini, F. (1991) 'Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1', *Nature*, 349(6306), pp. 257–260.

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S. and Sandberg, R. (2014) 'Full-length RNA-seq from single cells using Smart-seq2', *Nature Protocols*, 9(1), pp. 171–181.

Pimanda, J. E., Donaldson, I. J., de Bruijn, M. F. T. R., Kinston, S., Knezevic, K., Huckle, L., Piltz, S., Landry, J.-R., Green, A. R., Tannahill, D. and Göttgens, B. (2007) 'The SCL transcriptional network and BMP signaling pathway interact to regulate RUNX1 activity.', *Proceedings of the National Academy of Sciences*, 104(3), pp. 840–845.

Ponta, H., Sherman, L. and Herrlich, P. a. (2003) 'CD44: From adhesion molecules to signalling regulators', *Nature Reviews Molecular Cell Biology*, 4(1), pp. 33–45.

Pooter, F. De, Cho, S. K., Carlyle, J. R. and Zu, J. C. (2003) 'In vitro generation of T lymphocytes from embryonic stem cell – derived prehematopoietic progenitors', *Blood*, 102(5), pp. 1649–1653.

Potocnik, A. J., Brakebusch, C. and Fa, R. (2000) 'Fetal and Adult Hematopoietic Stem Cells Require β1 Integrin Function for Colonizing Fetal Liver , Spleen , and Bone Marrow', 12, pp. 653–663.

Prince, M. E., Sivanandan, R., Kaczorowski, A., Wolf, G. T., Kaplan, M. J., Dalerba, P., Weissman, I. L., Clarke, M. F. and Ailles, L. E. (2007) 'Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma', *Proceedings of the National Academy of Sciences*, 104(3), pp. 973–8.

Protin, U., Schweighoffer, T., Jochum, W. and Hilberg, F. (1999) 'CD44-deficient mice develop normally with changes in subpopulations and recirculation of lymphocyte subsets.', *Journal of immunology*, 163(9), pp. 4917–4923.

Qian, H., Johansson, S., McCourt, P., Smedsrød, B., Ekblom, M. and Johansson, S. (2009) 'Stabilins are expressed in bone marrow sinusoidal endothelial cells and mediate scavenging and cell adhesive functions', *Biochemical and Biophysical Research Communications*, 390(3), pp. 883–886.

Reese, D. E., Hall, C. E. and Mikawa, T. (2004) 'Negative regulation of midline vascular development by the notochord', *Developmental Cell*, 6(5), pp. 699–708.

Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R. and Weissman, I. L. (2003) 'A role for Wnt signalling in self-renewal of haematopoietic stem cells', *Nature*, 423(6938), pp. 409–414.

Rhodes, K. E., Gekas, C., Wang, Y., Lux, C. T., Francis, C. S., Chan, D. N., Conway, S., Orkin, S. H., Yoder, M. C. and Mikkola, H. K. A. (2008) 'The Emergence of Hematopoietic Stem Cells Is Initiated in the Placental Vasculature in the Absence of Circulation', *Cell Stem Cell*, 2(3), pp. 252–263.

Ricciardelli, C., Russell, D. L., Ween, M. P., Mayne, K., Suwiwat, S., Byers, S., Marshall, V. R., Tilley, W. D. and Horsfall, D. J. (2007) 'Formation of hyaluronan- and versican-rich pericellular matrix by prostate cancer cells promotes cell motility', *Journal of Biological Chemistry*, 282(14), pp. 10814–10825.

Risau, W. (1997) 'Mechanisms of angiogenesis', *Nature*, pp. 671–674.

Risau, W. and Flamme, I. (1995) 'Vasculogenesis', *Annu. Rev. Cell Dev. Biol.*, 11, pp. 73–91.

Robert-Moreno, A., Espinosa, L., de la Pompa, J. L. and Bigas, A. (2005) 'RBPj -dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells', *Development*, 132(5), pp. 1117–1126.

Robert-Moreno, À., Guiu, J., Ruiz-Herguido, C., López, M. E., Inglés-Esteve, J., Riera, L., Tipping, A., Enver, T., Dzierzak, E., Gridley, T., Espinosa, L. and Bigas, A. (2008) 'Impaired embryonic haematopoiesis yet normal arterial development in the absence of the Notch ligand Jagged1', *EMBO Journal*, 27(13), pp. 1886–1895.

Robin, C., Ottersbach, K. and Boisset, J. (2011) 'CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells', *Blood*, 117(19), pp. 5088–5092.

Robin, C., Ottersbach, K., Durand, C., Peeters, M., Vanes, L., Tybulewicz, V. and Dzierzak, E. (2006) 'An Unexpected Role for IL-3 in the Embryonic Development of Hematopoietic Stem Cells', *Developmental Cell*, 11(2), pp. 171–180.

Ropponen, K., Tammi, M., Parkkinen, J., Eskelinen, M., Tammi, R., Lipponen, P., Agren, U. and Aihava, E. (1998) 'Tumor Cell-associated Hyaluronan as an Unfavorable Prognostic Factor in Colorectal Cancer', *Cancer Research*, 58, pp. 342–347.

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', *The Journal of Experimental Medicine*, 209(8), pp. 1457-1468.

Ryall, J. G., Cliff, T., Dalton, S. and Sartorelli, V. (2015) 'Metabolic Reprogramming of Stem Cell Epigenetics', *Cell Stem Cell*, 17(6), pp. 651–662.

Rybtsov, S., Batsivari, A., Bilotkach, K., Paruzina, D., Senserrich, J., Nerushev, O. and Medvinsky, A. (2014) 'Tracing the Origin of the HSC Hierarchy Reveals an SCF-Dependent, IL-3-Independent CD43− Embryonic Precursor', *Stem Cell Reports*, 3(3), pp. 489–501.

Rybtsov, S., Ivanovs, A., Zhao, S. and Medvinsky, A. (2016) 'Concealed expansion of immature precursors underpins acute burst of adult HSC activity in foetal liver', *Development*, 143, pp. 1284–1289.

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', *Journal of Experimental Medicine*, 208(6).

Sabbatini, P. and McCormick, F. (1999) 'Phosphoinositide 3-OH kinase (PI3K) and PKB/Akt delay the onset of p53- mediated, transcriptionally dependent apoptosis', *Journal of Biological Chemistry*, 274(34), pp. 24263–24269.

Sabin, F. R. (1917) 'Preliminary Note on the Differentiation of Angioblasts and the Method by Which They Produce Blood-Vessels', 13, pp. 5–7.

Samokhvalov, I. M., Samokhvalova, N. I. and Nishikawa, S. (2007) 'Cell tracing shows the contribution of the yolk sac to adult haematopoiesis', *Nature*, 446(7139), pp. 1056–1061.

Sandler, V. M., Lis, R., Liu, Y., Kedem, A., James, D., Elemento, O., Butler, J. M., Scandura, J. M. and Rafii, S. (2014) 'Reprogramming human endothelial cells to haematopoietic cells requires vascular induction.', *Nature*, 511(7509), pp. 312–8.

Schledzewski, K., Géraud, C., Arnold, B., Wang, S., Gröne, H. J., Kempf, T., Wollert, K. C., Straub, B. K., Schirmacher, P., Demory, A., Schönhaber, H., Gratchev, A., Dietz, L., Thierse, H. J., Kzhyshkowska, J. and Goerdt, S. (2011) 'Deficiency of liver sinusoidal scavenger receptors stabilin-1 and -2 in mice causes glomerulofibrotic nephropathy via impaired hepatic clearance of noxious blood factors', *Journal of Clinical Investigation*, 121(2), pp. 703–714.

Schmits, R., Filmus, J., Gerwin, N., Senaldi, G., Kiefer, F., Kundig, T., Wakeham, A., Shahinian, A., Catzavelos, C., Rak, J., Furlonger, C., Zakarian, A., Simard, J. J. L., Ohashi, P. S., Paige, C. J., Gutierrez-Ramos, J. C. and Mak, T. W. (1997) 'CD44 regulates hematopoietic progenitor distribution, granuloma formation, and tumorigenicity.', *Blood*, 90(6), pp. 2217–33.

Screaton, G. R., Bell, M. V, Jackson, D. G., Cornelis, F. B., Gerth, U. and Bell, J. I. (1992) 'Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons.', *Proceedings of the National Academy of Sciences*, 89(24), pp. 12160–12164.

Setälä, L. P., Tammi, M. I., Tammi, R. H., Eskelinen, M. J., Lipponen, P. K., Ågren, U. M., Parkkinen, J., Alhava, E. M. and Kosma, V. M. (1999) 'Hyaluronan expression in gastric cancer cells is associated with local and nodal spread and reduced survival rate', *British Journal of Cancer*, 79(7–8), pp. 1133–1138.

Shalaby, F., Ho, J., Stanford, W. L., Fischer, K. D., Schuh, A. C., Schwartz, L., Bernstein, A. and Rossant, J. (1997) 'A requirement for Flk1 in primitive and definitive hematopoiesis and vasculogenesis', *Cell*, 89(6), pp. 981–990.

Shalaby, F., Rossant, J., Yamaguchi, T. P., Gertsenstein, M., Wu, X. F., Breitman, M. L. and Schuh, a C. (1995) 'Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice.', *Nature*, pp. 62–66.

Shivdasanl, 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*, pp. 432–434.

Silver, L. and Palis, J. (1997) 'Initiation of murine embryonic erythropoiesis: a spatial analysis.', *Blood*, 89(4), pp. 1154–64.

Simsek, T., Kocabas, F., Zheng, J., Deberardinis, R. J., Mahmoud, A. I., Olson, E. N., Schneider, J. W., Zhang, C. C. and Sadek, H. A. (2010) 'The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche', *Cell Stem Cell*, 7(3), pp. 380–390.

Sleeman, J., Rudy, W., Hofmann, M., Moll, J., Herrlich, P. and Ponta, H. (1996) 'Regulated clustering of variant CD44 proteins increases their hyaluronate binding capacity', *Journal of Cell Biology*, 135(4), pp. 1139–1150.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. and Rogers, D. (1988) 'Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides', *Nature*, 336(6200), pp. 688–690.

Sørensen, K. K., Simon-Santamaria, J., McCuskey, R. S. and Smedsrød, B. (2015) 'Liver sinusoidal endothelial cells', *Comprehensive Physiology*, 5(4), pp. 1751–1774.

Solaimani Kartalaei, P., Yamada-Inagawa T., Vink, C. S., de Pater, E., van der Linden, R., Marks-Bluth, J., van der Sloot, A., van den Hout, M., Yokomizo, T., van Schaick-Solernó M. L., Delwel, R., Pimanda, J. E., van Ijcken W. F. J. and Dzierzak, E. (2015) 'Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation', *Journal of Experimental Medicine,* 212(1), pp. 93 - 106.

Souilhol, C., Gonneau, C., Lendinez, J. G., Batsivari, A., Rybtsov, S., Wilson, H., Morgado-Palacin, L., Hills, D., Taoudi, S., Antonchuk, J., Zhao, S. and Medvinsky, A. (2016) 'Inductive interactions mediated by interplay of asymmetric signalling underlie development of adult haematopoietic stem

cells', *Nature Communications*, 7, pp. 1–13.

Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V. and Lacaud, G. (2009) 'In Vitro Differentiation of Embryonic Stem Cells as a Model of Early Hematopoietic Development', *Methods of Molecular Biology*, pp. 317–334.

Su, W., Foster, S. C., Xing, R., Feistel, K., Olsen, R. H. J., Acevedo, S. F., Raber, J. and Sherman, L. S. (2017) 'CD44 transmembrane receptor and hyaluronan regulate adult hippocampal neural stem cell quiescence and differentiation', *Journal of Biological Chemistry*, 292(11), pp. 4434–4445.

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., Wong, I., Sousa, P., Zhu, T. N., Ditadi, A., Keller, G., Engelman, A. N., Snapper, S. B., Doulatov, S. and Daley, G. Q. (2017) 'Haematopoietic stem and progenitor cells from human pluripotent stem cells', *Nature*, 545(7655), pp. 432–438.

Sugiyama, T., Kohara, H., Noda, M. and Nagasawa, T. (2006) 'Maintenance of the Hematopoietic Stem Cell Pool by CXCL12-CXCR4 Chemokine Signaling in Bone Marrow Stromal Cell Niches', *Immunity*, 25(6), pp. 977–988.

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K. D., Yoshimoto, M., Yoder, M. C., Frampton, J., Schroeder, T., Enver, T., Göttgens, B. and de Bruijn, M. F. T. R. (2013) 'Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level.', *Nature communications*, 4, p. 2924.

Szabo, E., Rampalli, S., Risueño, R. M., Schnerch, A., Mitchell, R., Fiebig-Comyn, A., Levadoux-Martin, M. and Bhatia, M. (2010) 'Direct conversion of human fibroblasts to multilineage blood progenitors', *Nature*, 468(7323), pp. 521–526.

Takaishi, S., Okumura, T., Tu, S., Wang, S. S. W., Shibata, W., Vigneshwaran, R., Gordon, S. A. K., Shimada, Y. and Wang, T. C. (2009) 'Identification of gastric cancer stem cells using the cell surface marker CD44', *Stem Cells*, 27(5), pp. 1006–1020.

Takeda, M., Ogino, S., Umemoto, R., Sakakura, M., Kajiwara, M., Sugahara, K. N., Hayasaka, H., Miyasaka, M., Terasawa, H. and Shimada, I. (2006) 'Ligand-induced structural changes of the CD44 hyaluronan-binding domain revealed by NMR', *Journal of Biological Chemistry*, 281(52), pp. 40089–40095.

Takubo, K., Nagamatsu, G., Kobayashi, C. I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., Goda, N., Rahimi, Y., Johnson, R. S., Soga, T., Hirao, A., Suematsu, M. and Suda, T. (2013) 'Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells', *Cell Stem Cell*, 12(1), pp. 49–61.

Tamada, M., Nagano, O., Tateyama, S., Ohmura, M., Yae, T., Ishimoto, T.,

Sugihara, E., Onishi, N., Yamamoto, T., Yanagawa, H., Suematsu, M. and Saya, H. (2012) 'Modulation of glucose metabolism by CD44 contributes to antioxidant status and drug resistance in cancer cells', *Cancer Research*, 72(6), pp. 1438–1448.

Taoudi, S., Gonneau, C., Moore, K., Sheridan, J. M., Blackburn, C. C., Taylor, E. and Medvinsky, A. (2008) 'Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-Cadherin+CD45+pre-definitive HSCs', *Cell Stem Cell*, 3(1), pp. 99–108.

Taoudi, S. and Medvinsky, A. (2007) 'Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta.', *Proceedings of the National Academy of Sciences*, 104(22), pp. 9399–403.

Tavian, B. M., Coulombel, L., Luton, D., Clemente, H. S., Dieterlen-lievre, F. and Peault, B. (1996) 'Aorta-Associated CD34+ Hematopoietic Cells in the Early Human Embryo', *Blood*, 87(1), pp. 67–72.

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

Tavian, M., Robin, C., Coulombel, L. and Péault, B. (2001) 'The human embryo, but not its yolk sac, generates lympho-myeloid stem cells: Mapping multipotent hematopoietic cell fate in intraembryonic mesoderm', *Immunity*, 15(3), pp. 487–495.

Teriete, P., Banerji, S., Noble, M., Blundell, C. D., Wright, A. J., Pickford, A. R., Lowe, E., Mahoney, D. J., Tammi, M. I., Kahmann, J. D., Campbell, I. D., Day, A. J. and Jackson, D. G. (2004) 'Structure of the regulatory hyaluronan binding domain in the inflammatory leukocyte homing receptor CD44', *Molecular Cell*, 13(4), pp. 483–496.

Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Möröy, T., Robin, C., Miller, C., Cowley, S., Göttgens, B., Kousko, V. and Lacaud, G. (2015) 'GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1', *Nature Cell Biology*, 18(1), pp. 21-32.

Thomas, M. L. (1989) 'The leukocyte common antigen family', *Ann. Rev. Immunol.*, 7, pp. 339–369.

Thorne, R. F. (2003) 'The role of the CD44 transmembrane and cytoplasmic domains in co-ordinating adhesive and signalling events', *Journal of Cell Science*, 117(3), pp. 373–380.

Tian, X., Azpurua, J., Hine, C., Vaidya, A., Myakishev-Rempel, M., Ablaeva, J., Mao, Z., Nevo, E., Gorbunova, V. and Seluanov, A. (2013) 'High-molecular-mass hyaluronan mediates the cancer resistance of the naked

mole rat', *Nature*, 499(7458), pp. 346–349.

Tober, J., Koniski, A., McGrath, K. E., Vemishetti, R., Emerson, R., De Mesy-Bentley, K. K. L., Waugh, R. and Palis, J. (2007) 'The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis', *Blood*, 109(4), pp. 1433–1441.

Tober, J., Yzaguirre, A. D., Piwarzyk, E. and Speck, N. A. (2013) 'Distinct temporal requirements for Runx1 in hematopoietic progenitors and stem cells', *Development*, 140(18), pp. 3765–3776.

Tölg, C., Hofmann, M., Herrlich, P. and Ponta, H. (1993) 'Splicing choice from ten variant exons establishes CD44 variability', *Nucleic Acids Research*, 21(5), pp. 1225–1229.

Toole, B. P. (2004) 'Hyaluronan: From extracellular glue to pericellular cue', *Nature Reviews Cancer*, 4(7), pp. 528–539.

Travnickova, J., Tran Chau, V., Julien, E., Mateos-Langerak, J., Gonzalez, C., Lelièvre, E., Lutfalla, G., Tavian, M. and Kissa, K. (2015) 'Primitive macrophages control HSPC mobilization and definitive haematopoiesis.', *Nature communications*, 6, p. 6227.

Tremmel, M., Matzke, A., Albrecht, I., Laib, A. M., Olaku, V., Ballmer-Hofer, K., Christofori, G., Héroult, M., Augustin, H. G., Ponta, H. and Orian-Rousseau, V. (2009) 'A CD44v6 peptide reveals a role of CD44 in VEGFR-2 signaling and angiogenesis', *Blood*, 114(25), pp. 5236–5244.

Trochon, V., Mabilat, C., Bertrand, P., Legrand, Y., Smadja-Joffe, F., Soria, C., Delpech, B. and Lu, H. (1996) 'Evidence of involvement of CD44 in endothelial cell proliferation, migration and angiogenesis in vitro.', *International journal of cancer*, 66(5), pp. 664–8.

Trowbridge, I. S., Lesley, J., Schulte, R., Hyman, R. and Trotter, J. (1982) 'Biochemical characterization and cellular distribution of a polymorphic, murine cell-surface glycoprotein expressed on lymphoid tissues', *Immunogenetics*, 15(3), pp. 299–312.

Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W. and Orkin, S. H. (1994) 'An early haematopoietic defect in mice lacking the transcription factor GATA-2', *Nature*, pp. 221–226.

Uchida, H., Zhang, J. and Nimer, S. D. (1997) 'AML1A and AML1B can transactivate the human IL-3 promoter.', *The Journal of Immunology* , 158(5), pp. 2251–2258.

Udan, R. S., Vadakkan, T. J. and Dickinson, M. E. (2013) 'Dynamic responses of endothelial cells to changes in blood flow during vascular remodeling of the mouse yolk sac.', *Development*, 140(19), pp. 4041–50.

Ueno, H. and Weissman, I. L. (2006) ‘Clonal Analysis of Mouse Development Reveals a Polyclonal Origin for Yolk Sac Blood Islands’, *Developmental Cell*, 11(4), pp. 519–533.

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 TGF β pathway impairs endothelial to haematopoietic transition’, *Scientific Reports*, pp. 1–15.

Visvader, J. E., Fujiwara, Y. and Orkin, S. H. (1998) ‘Unsuspected role for the T-cell leukemia protein SCL / tal-1 in vascular development’, *Genes & Development*, 12, pp. 473–479.

Vogeli, K. M., Jin, S. W., Martin, G. R. and Stainier, D. Y. R. (2006) ‘A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula’, *Nature*, 443(7109), pp. 337–339.

Wakimoto, K., Kobayashi, K., Kuro-o, M., Yao, A., Iwamoto, T., Yanaka, N., Kita, S., Nishida, A., Azuma, S., Toyoda, Y., Omori, K., Imahie, H., Oka, T., Kudoh, S., Kohmoto, O., Yazaki, Y., Shigekawa, M., Imai, Y., Nabeshima, Y. I. and Komuro, I. (2000) ‘Targeted disruption of Na+/Ca2+ exchanger gene leads to cardiomyocyte apoptosis and defects in heartbeat’, *Journal of Biological Chemistry*, 275(47), pp. 36991–36998.

Wang, H. U., Chen, Z. F. and Anderson, D. J. (1998) ‘Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4’, *Cell*, 93(5), pp. 741–753.

Wang, Q., Stacy, T., Binder, M., Marín-Padilla, M., Sharpe A. H., Speck, N. A. (1996) 'Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis', *PNAS,* 93, pp. 3444-3449.

Wareing, S., Eliades, A., Lacaud, G. and Kouskoff, V. (2012) ‘ETV2 expression marks blood and endothelium precursors, including hemogenic endothelium, at the onset of blood development’, *Developmental Dynamics*, 241(9), pp. 1454–1464.

Weber, G. F., Ashkar, S., Glimcher, M. J. and Cantor, H. (1996) ‘Receptor-Ligand Interaction Between CD44 and Osteopontin (Eta-1)’, *Science*, 271(January), pp. 509–513.

Weigel, P. H., Hascall, V. C. and Tammi, M. (1997) ‘Hyaluronan synthases’, *Journal of Biological Chemistry*, 272(22), pp. 13997–14000.

Weissmann, B. and Meyer, K. (1954) ‘The Structure of Hyalobiuronic Acid and of Hyaluronic Acid from Umbilical Cord’, *Journal of the American Chemical Society*, 76(7), pp. 1753–1757.

West, D. C., Hampson, I. N., Arnold, F. and Kumar, S. (1985) 'Angiogenesis

induced by degradation products of hyaluronic acid', *Science*, 228(4705), pp. 1324-1326.

Wettstein, R., Bodak, M. and Ciaudo, C. (2016) 'Generation of a Knockout Mouse Embryonic Stem Cell Line Using a Paired CRISPR/Cas9 Genome Engineering Tool', in Turksen, K. (ed.) *Embryonic Stem Cell Protocols*, pp. 321–343.

Wicha, M. S., Liu, S. and Dontu, G. (2006) 'Cancer stem cells: An old idea - A paradigm shift', *Cancer Research*, 66(4), pp. 1883–1890.

Wiles, M. V and Keller, G. (1991) 'Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture.', *Development*, 111(2), pp. 259–267.

Wilhelm, K., Happel, K., Eelen, G., Schoors, S., Oellerich, M. F., Lim, R., Zimmermann, B., Aspalter, I. M., Franco, C. A., Boettger, T., Braun, T., Fruttiger, M., Rajewsky, K., Keller, C., Brüning, J. C., Gerhardt, H., Carmeliet, P. and Potente, M. (2016) 'FOXO1 couples metabolic activity and growth state in the vascular endothelium', *Nature*, 529(7585), pp. 216–220.

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', *Developmental Cell*, 16(6), pp. 909–916.

Willert, K., Brown, J. D., Danenberg, E., Duncan, A. W., Weissman, I. L., Reya, T., Yates, J. R. and Nusse, R. (2003) 'Wnt proteins are lipid-modified and can act as stem cell growth factors', *Nature*, 423(6938), pp. 448–452.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. a, Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. a and Gough, N. M. (1988) 'Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.', *Nature*, 336(6200), pp. 684–687.

Wilson, N. K., Foster, S. D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P. M., Kinston, S., Ouwehand, W. H., Dzierzak, E., Pimanda, J. E., De Bruijn, M. F. T. R. and Göttgens, B. (2010) 'Combinatorial transcriptional control in blood stem/progenitor cells: Genome-wide analysis of ten major transcriptional regulators', *Cell Stem Cell*, 7(4), pp. 532–544.

Wolny, P. M., Banerji, S., Gounou, C., Brisson, A. R., Day, A. J., Jackson, D. G. and Richter, R. P. (2010) 'Analysis of CD44-hyaluronan interactions in an artificial membrane system: Insights into the distinct binding properties of high and low molecular weight hyaluronan', *Journal of Biological Chemistry*, 285(39), pp. 30170–30180.

Wong, T. Y., Chang, C. H., Yu, C. H. and Huang, L. L. H. (2017) 'Hyaluronan keeps mesenchymal stem cells quiescent and maintains the differentiation potential over time', *Aging Cell*, 16(3), pp. 451–460.

Xu M, J., Matsuoka, S., Yang, F. C., Ebihara, Y., Manabe, a, Tanaka, R., Eguchi, M., Asano, S., Nakahata, T. and Tsuji, K. (2001) 'Evidence for the presence of murine primitive megakaryocytopoiesis in the early yolk sac.', *Blood*, 97(7), pp. 2016–2022.

Yamada, Y., Warren, A. J., Dobson, C., Forster, A., Pannell, R. and Rabbitts, T. H. (1998) 'The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis.', *Proceedings of the National Academy of Sciences of the United States of America*, 95(7), pp. 3890–3895.

Yamaguchi, T. P., Dumont, D. J., Conlon, R. A., Breitman, M. L. and Rossant, J. (1993) 'Flk-1, an Flt-Related Receptor Tyrosine Kinase Is an Early Marker for Endothelial Cell Precursors', *Development*, 118(2), pp. 489–498.

Yang, B., Yang, B. L., Savani, R. C. and Turley, E. A. (1994) 'Identification of a common hyaluronan binding motif in the hyaluronan binding proteins RHAMM, CD44 and link protein', *Embo J.*, 1(2), pp. 286–296.

Yoder, M. C., Hiatt, K., Dutt, P., Mukherjee, P., Bodine, D. M. and Orlic, D. (1997) 'Characterization of definitive lymphohematopoietic stem cells in the day 9 murine yolk sac', *Immunity*, 7(3), pp. 335–344.

Yoder, M. C., Hiatt, K. and Mukherjee, P. (1997) 'In vivo repopulating hematopoietic stem cells are present in the murine yolk sac at day 9.0 postcoitus', *Proceedings of the National Academy of Sciences*, 94(13), pp. 6776–6780.

Yokomizo, T. and Dzierzak, E. (2010) 'Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos.', *Development*, 137(21), pp. 3651–3661.

Yokomizo, T., Ogawa, M., Osato, M., Kanno, T., Yoshida, H., Fujimoto, T., Fraser, S., Nishikawa, S., Okada, H., Satake, M., Noda, T., Nishikawa, S. and Ito, Y. (2001) 'Requirement of Runx1/AML1/PEBP2alphaB for the generation of haematopoietic cells from endothelial cells.', *Genes to cells*, 6(1), pp. 13–23.

Young, P. E., Baumhueter, S. and Lasky, L. A. (1995) 'The sialomucin CD34 is expressed on hematopoietic cells and blood vessels during murine development.', *Blood*, 85(1), pp. 96–105.

Yu, Q. and Stamenkovic, I. (1999) 'Localization of matrix metalloproteinase 9 to the cell surface provides a mechaism for CD44-mediated tumor invasion.', *Genes & Devel.*, 13, pp. 35–48.

Yu, Q. and Stamenkovic, I. (2000) 'Cell surface-localized matrix mealloproteinase-9 proteolytically activates TGF-beta and promotes tumor invasion and angiogenesis', *Genes and Development*, 14, pp. 163–176.

Yu, W., Woessner, J. F., Mcneish, J. D. and Stamenkovic, I. (2002) 'CD44 anchors the assembly of matrilysin/MMP-7 with heparin-binding epidermal growth factor precursor and ErbB4 and regulate female reproductive organ remodeling', *Genes & Development*, pp. 307–323.

Yvernogeau, L. and Robin, C. (2017) 'Restricted intra-embryonic origin of bona fide hematopoietic stem cells in the chicken', *Development*, 144(13), pp. 2352–2363.

Yzaguirre, A. D., Howell, E. D., Li, Y., Liu, Z. and Speck, N. A. (2018) 'Runx1 is sufficient for blood cell formation from non-hemogenic endothelial cells *in vivo* only during early embryogenesis', *Development*, 145(2)

Zhang, C. C., Kaba, M., Ge, G., Xie, K., Tong, W., Hug, C. and Lodish, H. F. (2006) 'Angiopoietin-like proteins stimulate ex vivo expansion of hematopoietic stem cells', *Nature Medicine*, 12(2), pp. 240–245.

Zhang, C. C. and Lodish, H. F. (2005) 'Murine hematopoietic stem cells change their surface phenotype during ex vivo expansion', *Leukemia and Lymphoma*, 105(11), pp. 4314–4320.

Zhang, J., Niu, C., Ye, L., Huang, H., He, X., Tong, W.-G., Ross, J., Haug, J., Johnson, T., Feng, J. Q., Harris, S., Wiedemann, L. M., Mishina, Y. and Li, L. (2003) 'Identification of the haematopoietic stem cell niche and control of the niche size', *Nature*, 425, pp. 836–841.

Zhang, P., Fu, C., Bai, H., Song, E. and Song, Y. (2014) 'CD44 variant, but not standard CD44 isoforms, mediate disassembly of endothelial VE-cadherin junction on metastatic melanoma cells', *FEBS Letters*, 588(24), pp. 4573–4582.

Zhong, T. P., Childs, S., Leu, J. P. and Fishman, M. C. (2001) 'Gridlock signalling pathway fashions the first embryonic artery', *Nature*, 414(6860), pp. 216–220.

Zhou, B., Weigel, J. A., Fauss, L. and Weigel, P. H. (2000) 'Identification of the hyaluronan receptor for endocytosis (HARE)', *Journal of Biological Chemistry*, 275(48), pp. 37733–37741.

Zhou, F., Li, X., Wang, W., Zhu, P., Zhou, J., He, W., Ding, M., Xiong, F., Zheng, X., Li, Z., Ni, Y., Mu, X., Wen, L., Cheng, T., Lan, Y., Yuan, W., Tang, F. and Liu, B. (2016) 'Tracing haematopoietic stem cell formation at single-cell resolution', *Nature*. Nature Publishing Group, 533(7604), pp. 487–492.

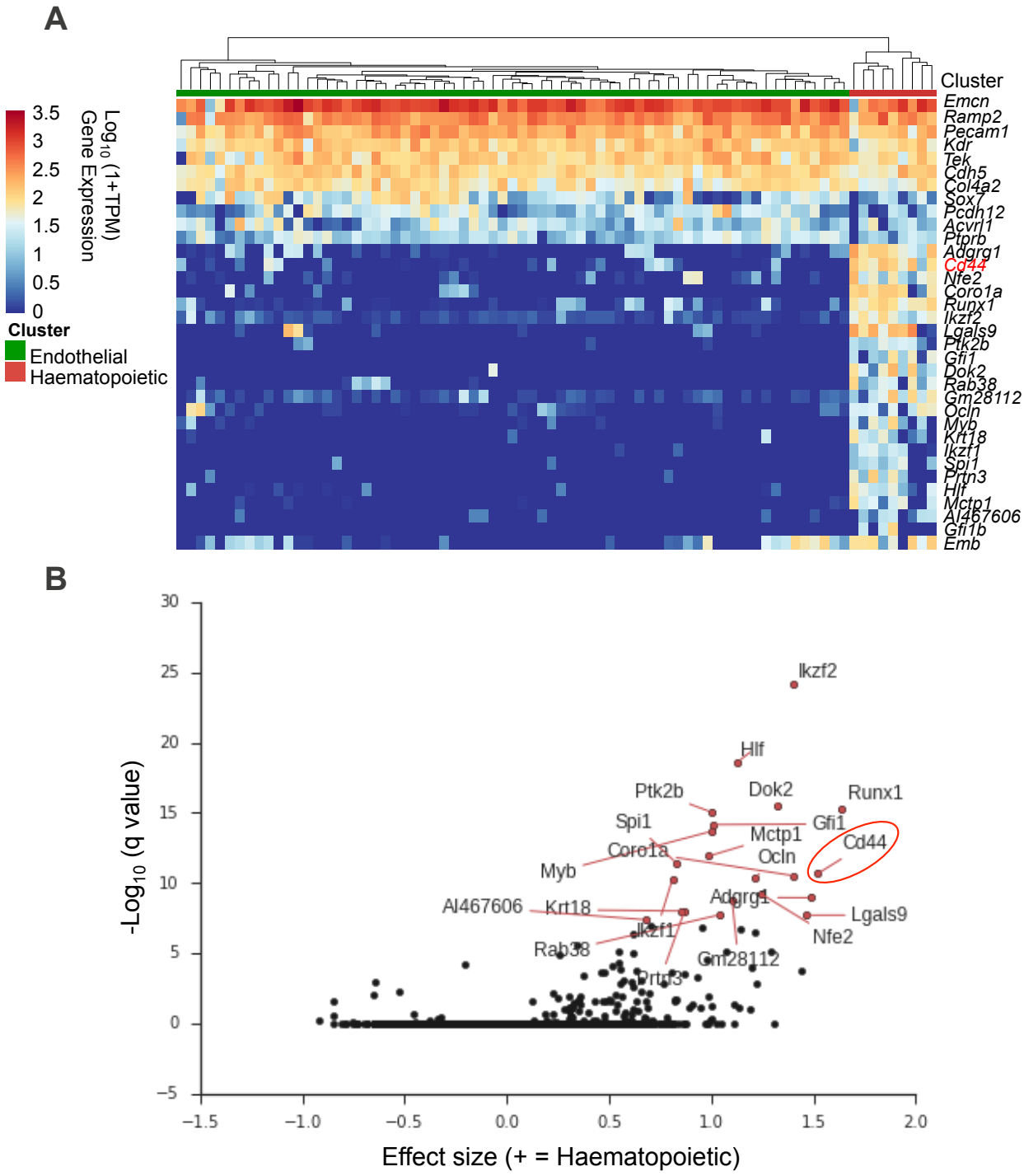
Zöller, M. (2011) 'CD44: Can a cancer-initiating cell profit from an abundantly expressed molecule?', *Nature Reviews Cancer*, 11(4), pp. 254–267.

Zovein, A. C., Hofmann, J. J., Lynch, M., French, W. J., Turlo, K. a., Yang, Y., Becker, M. S., Zanetta, L., Dejana, E., Gasson, J. C., Tallquist, M. D. and Iruela-Arispe, M. L. (2008) 'Fate Tracing Reveals the Endothelial Origin of

Hematopoietic Stem Cells', *Cell Stem Cell*, 3(6), pp. 625–636.

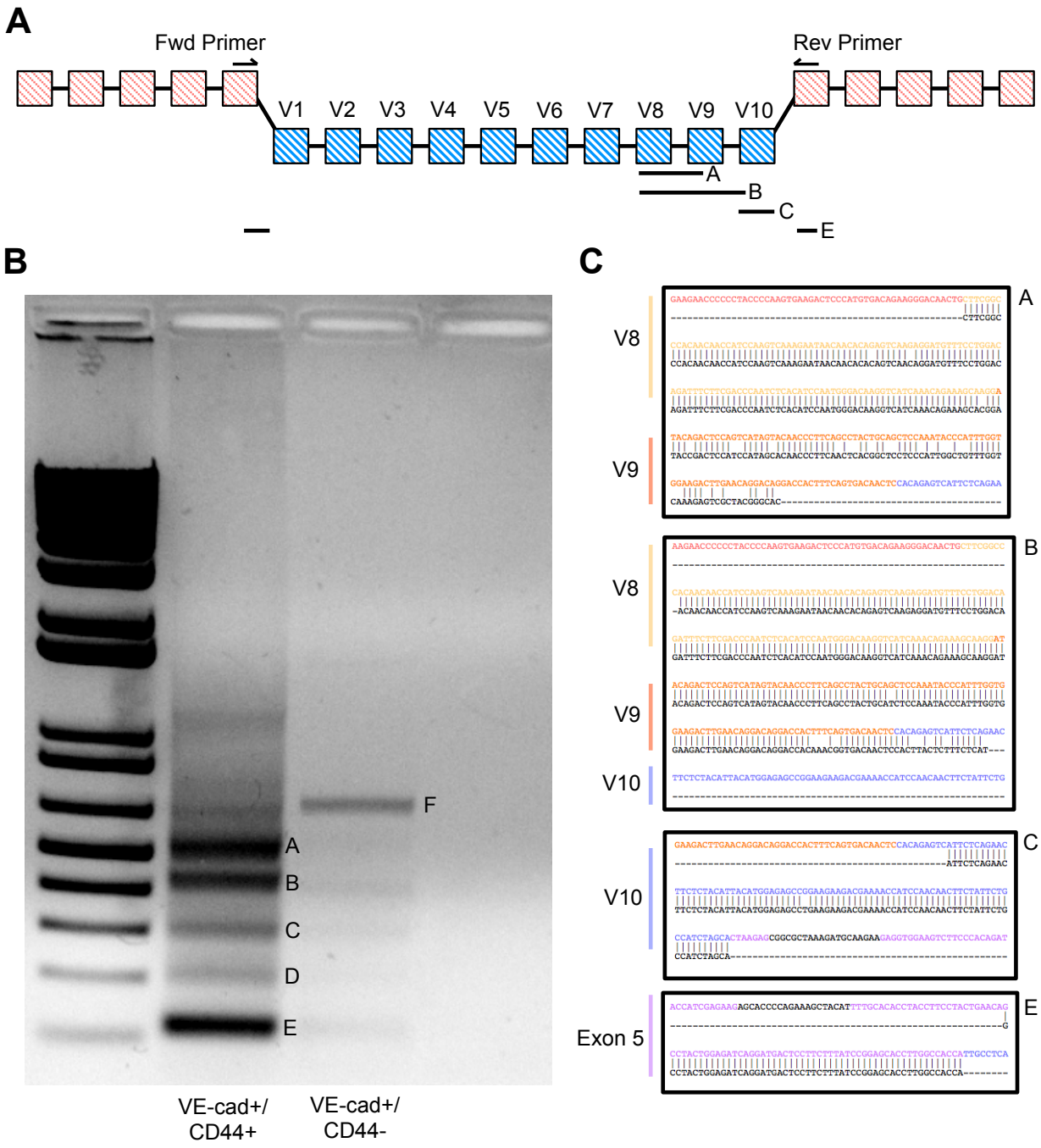
Zovein, A. C., Turlo, K. A., Ponec, R. M., Lynch, M. R., Chen, K. C., Hofmann, J. J., Cox, T. C., Gasson, J. C. and Iruela-Arispe, M. L. (2010) 'Vascular remodeling of the vitelline artery initiates extravascular emergence of hematopoietic clusters', *Blood*, 116(18), pp. 3435–3444.

Supplementary Material



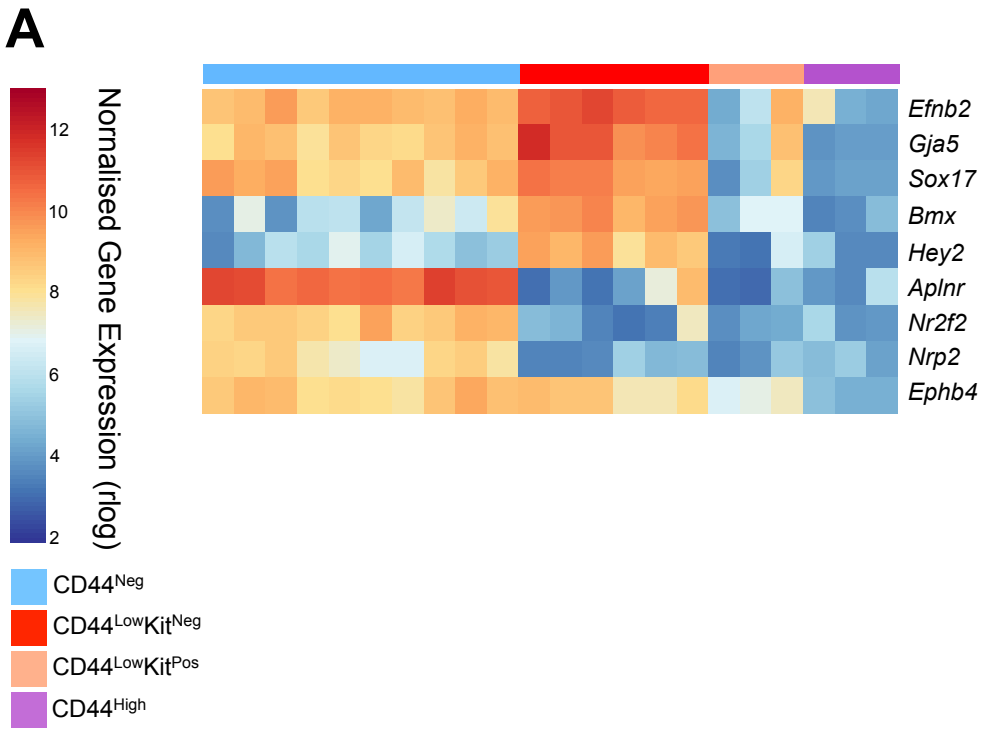
Supplementary Figure 1: Single cell RNA sequencing on 78 VE-Cadherin+ cells from the AGM

The transcriptional profile of 78 single cells derived from the VE-Cadherin+ fraction of the AGM (A). Bioinformatic analysis comparing the expression of the endothelial cluster and haematopoietic cluster identified CD44 as a significant marker gene (B). This experiment was performed by Ozge Vargel and the bioinformatics analysis was performed by Valentine Svensson.



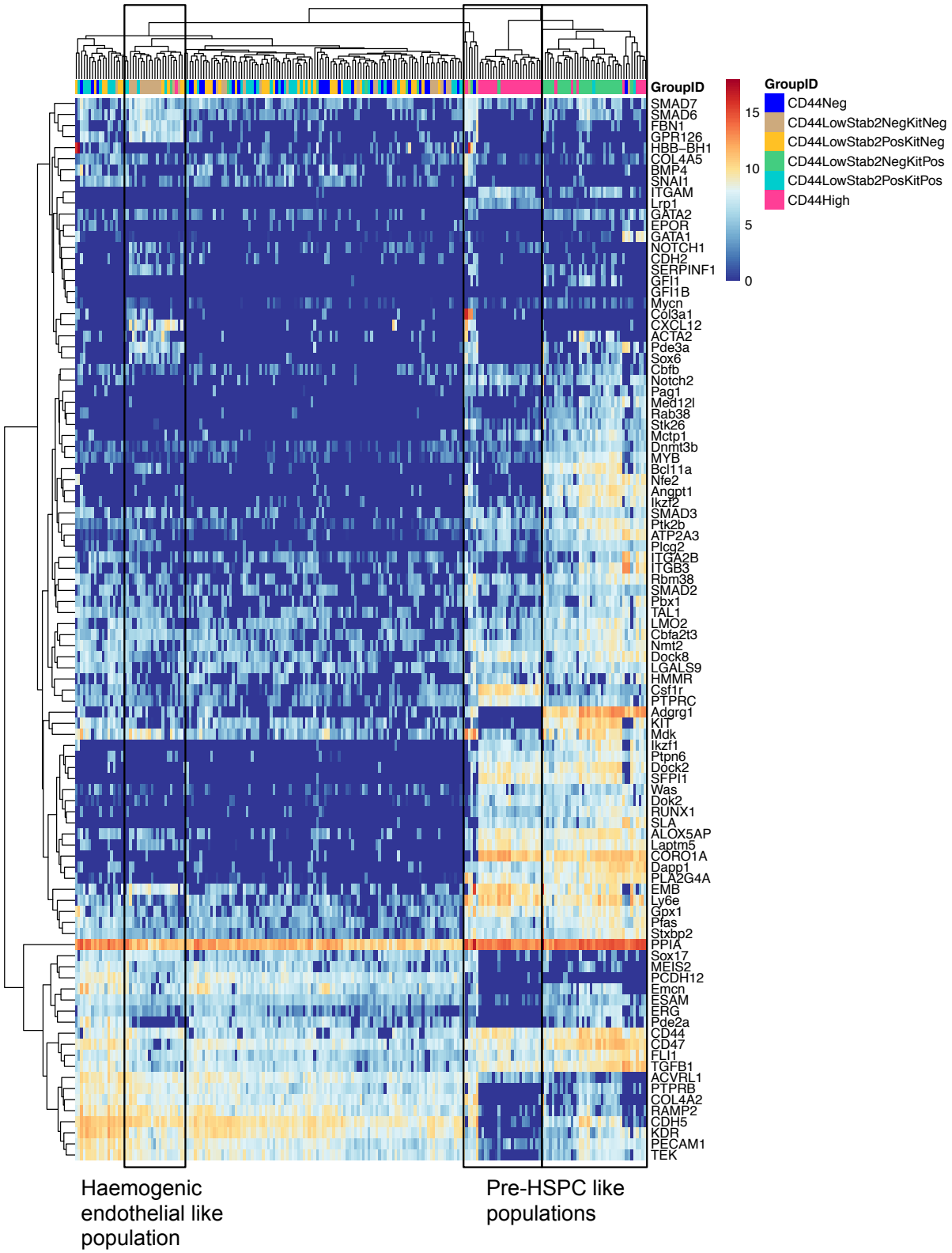
Supplementary Figure 2: CD44 splice variation analysis in VE-Cadherin+/CD44+ cells from the AGM

Depiction of CD44 constitutive exons (red) and variant exons (blue) indicating the placement of forward and reverse primers used to amplify alternative CD44 isoforms (A). Lines indicate where bands A, B, C and E aligned in relation to the CD44 transcript. Bands D and F did not align to the CD44 coding sequence. Agarose gel showing CD44 splice isoforms amplified from VE-Cadherin+/CD44+ and VE-Cadherin+/CD44- FACS sorted AGM cells (B). Letters indicate bands that were cut and sequenced. Sequence alignment between CD44 coding sequence and bands indicating the variant and constitutive exons to which they align (C).



Supplementary Figure 3: Expression of venous and arterial markers

25-bulk RNA sequencing reveals differences in arterial and venous gene expression between CD44Neg and CD44Low/Kit- cells (A).



Supplementary Figure 4: Heatmap of single cell qPCR of Stab2+ and Stab2- yolk sac endothelium

Single cell qPCR data of 212 cells sorted based on the cell surface markers VE-Cadherin, CD44, c-Kit and Stab2 from the yolk sac at embryonic day E10 and E11. Clusters of cells are marked that display a transcriptional profile similar to the pre-HSPCs and haemogenic endothelial populations previously identified in the AGM.

Supplementary Table 1: Primers designed against 95 genes for single-cell qPCR

Gene Name	Outer Forward Primer	Outer Reverse Primer	Inner Forward Primer	Inner Reverse Primer
Acta2	AGGCACCACTGA ACCCTAAG	CACAGCCTGAATA GCCACAT	CCAACCGGGAGA AAATGAC	ATGGCGGGGACA TTGAAG
Acvrl1	CGAATTGCCCATC GTGACCTCAA	CGTGGTTGTTGC CGATATCCAGGTA	CGAAGTCGCAAT GTGCTGGTCAA	CGAAGTCGCAAT GTGCTGGTCAA
Adgrg1	GTGGTGGAGGTC TTCGGTAC	GGGCCGTAGTTA TTCACATCCA	CTATGTGCCCGG CTATCTG	CAACGCCACCAG AGTGA
Alox5ap	TGCGTACCCCAC TTTCCTTGT	TCTCTCTCCCAGA TAGCCGACA	GCAGGACTACTTT GCAGCCA	AGGTACATCAGTC CGGCGAAGG
Angpt1	CAGCACGAAGGA TGCTGATAA	CCCGCAGTGTAG AACATTCC	CTGTATGTGCAAA TGCGCTCTC	ATTTAGATTGGAA GGGCCACAGG
Atp2a3	CGACCATTGTGG CTGCAGTAGAA	CGTCCCAGAATT GCTGTGAGGAA	CGAGAGGGCAGG GCCATCTACA	CGAGAGGGCAGG GCCATCTACA
Bcl11a	CCCGGGATGAGT GCAGAATA	CTGTGCGTGTTG CAAGAGAA	GCCCCGCAGGGT ATTTGTAAA	ATGCACTGGTGA ATGGCTGTT
Bmp4	CAGCCGAGCCAA CACTGTGA	TGGGATGCTGCT GAGGTTGA	AGTTTCCATCACG AAGAACATCTGG	GAGGAAACGAAA AGCAGAGC
Cbfa2t3	AACCAGCAAGAG GACTCCA	CTCCCAGTCTTTG TGCTGAC	GCGAGAGCTGCT GGAAC	AGAAGGACCCGC AGTAGC
Cbfb	CCCGGACCAGAG GAGCAA	CGGCAGGCGTTC TGGAA	TCGAGAACGAGG AGTTCTTCAG	CGTGTCTGGCGC TCCT
Cd44	CCCCTCCTGAAG AAGACTGTACA	GCTGTAGCGAGT ACCATCACG	GTCACAGACCTA CCCAATTCCTTCG A	GCTGTAGCGAGT ACCATCACG
Cd47	CTTCTGGACTTGG CCTCATTGT	ACGTAGCCCAGC ACTTGAGT	TCTCTACGGGGA TATTAATACTACT TCAG	TGGCAATGGTGA AAGAGGTCA
Cdh2	ATCAACAATGAGA CTGGGGACATCA	CTTCCATGTCTGT GGCTTGAA	CACTGTGGCAGC TGGTCTGG	ATTAACGTATACT GTTGCACTTTCTC TCG
Cdh5	CGACACCATCGC CAAAAGAGAGAC	CGTCTTAGCATTC TGGCGGTTCAC	CGAGGATTTGGA ATCAAATGCACAT CG	CGAGGATTTGGA ATCAAATGCACAT CG
Col3a1	TCTTCTCACCCTT CTTCATCCC	TGACATGGTTCTG GCTTCCA	ACTCTTATTTTGG CACAGCAGTC	TCTCTAGACTCAT AGGACTGACCAA
Col4a2	CGACCCTGGAAG CCCTGGATTTA	CGTTGTCTTCCCT TAAGTCCCAACA	CGATGGCAGGGA TGCCTGGT	CGATGGCAGGGA TGCCTGGT
Col4a5	CGAGGACTCTCT GTGGATTGGCTA	CGTATGAAGGGA GCGGAACGAA	CGATTCATGATGC ATACAAGTGCAG GA	CGATTCATGATGC ATACAAGTGCAG GA
Coro1a	TCTGCGCTGTCAA CCCCAAGT	ACCAGGGGCACG TTCTTGTCT	GCTCTGATCTGTG AGGCCAG	ACTCGTCCAGTCT TGCCTAG
Csf1r	ACCCTACTCAGTT GCCCTACA	GCCCAGACCAAA GGCTGTA	GTGGGAGTTCCC TCGGAACAA	GCCTCCACCACC TTCCCAAA
Cxcl12	TCTTCGAGAGCC ACATCGCC	TGTCTGTTGTTGT TCTTCAGCCGT	AGAGCCAACGTC AAGCATCTGA	CTGAAGGGCACA GTTTGGAGT
Dapp1	ACAGAGTGCTCA GCTGTTCAA	CCATTCATCGGCT TCAACTCC	TTTGATTACTCAC AGGAACGAGTAA	CGTCTTTGCACAG AGATAAAATGT
Dnmt3b	TTGGTGGAAGCC CATGCA	CCTTGGGGCGGG TATAATTCA	ATGATCTCTCTAA CGTCAATCCTG	GCAAGTGGTAAA ACTCGAAGAAG
Dock2	GTACCACTCAGTC GTCTACTACC	GAGCGGTGTCGA AACATGAA	AAGTCAAACAGC CCCGATG	CCTCAAATGGATC CTCTGCAT
Dock8	GGATTCTGGACC TGCTGTTCA	TCTTCCAGCTTGG CCTTGAC	TCTGTGTCTCCTG CTTTGAATA	TCTCTTGACTTCT GCAGGAC
Dok2	AGCCAGGCACAC AGCTATA	ACTCAAAGTTGCC CTCTCCA	TGACTGGCCCTA CAGGTTTCT	GTCTGCCAGCCT CAAAGGAAA
Emb	CCTGGTTACCCG CTCCTCCTTCT	CGCCATCATTTCT TCTCGAACAGGT	ATGTCTTCTGGCG GCGAC	ACTTGTAAAAGTT GGATCTGTGGGA
Emcn	TAGGAAGATTGCA	ATTCGATACAAAC	CAACCACCCCCT	GTTATAACAACCA

	ACCACTCCA	CCACCAGAGTA	CCTATTCCA	GCGCGATAACCA
Epor	TGAAGTGGACGT GTCGGCAG	ACAGCGAAGGTG TAGCGCGT	CAACCGGGCAGG AGGGACA	CCGCCCCGCAGG TTGCTCAGAA
Erg	TCCCGAAGCTAC GCAAAGAA	TTTGGACTGAGG GGTGAGGTG	TACAACTAGGCCA GATTTACCTTATG A	TGTGGCCGGTCC AGGCTGAT
Esam	CGAAGTCTATGTC TGCAAGGCTCAA	CGTCCCAACAAAA GTGCCCACAA	CGACAGAGTGGG CTTTGCCAAGT	CGACAGAGTGGG CTTTGCCAAGT
Fbn1	ACGTGGCGGGGA ATGTACAAACA	CGTCAGAGCTGT GTAGCAGTAACC A	CGACTGTCAGCA GCTACTTCTGCAA AT	CGACTGTCAGCA GCTACTTCTGCAA AT
Fli1	TGCTGTTGTCGCA CCTCAGTT	TTCCTTGACATTC AGTCGTGAGGA	CTCAGGGAAAGT TCACTGCTGGCC TA	TGGTCTGTATGG GAGGTTGTG
Gata1	CCTGTGCAATGC CTGTGGCT	TGCCTGCCCGTTT GCTGACAA	GTATCACAAGATG AATGGTCAGAAC C	CATTCGCTTCTTG GGCCGGATG
Gata2	AAGCAAGGCTCG CTCCTG	CACAGGCATTGC ACAGGTAGT	CAGAAGGCCGGG AGTGTGTC	GCCCGTGCCATC TCGT
Gfi1	CGAGAGATGTGC GGCAAGACC	CGTAGCGTGGAT GACCTCTTGAA	CGAGTGAGCCTG GAGCAACACAA	CGAGTGAGCCTG GAGCAACACAA
Gfi1b	CGATGGACACTTA CCACTGTGTCA	CGTAGGTTTTGCC ACAGACATCAC	CGAAGTGCAACA AGGTGTTCTCC	CGAAGTGCAACA AGGTGTTCTCC
Gpr126	CGACTGTGCAGC CACTTCACTCA	CGTGGCAGATATT CCGCACCCAATA	CGATGGAGTTCT GATGGATCTTCC	CGATGGAGTTCT GATGGATCTTCC
Gpx1	CGGGACTACACC GAGATGAA	CGGACGTACTTG AGGGAATTCA	CAGAAGCGTCTG GGACCT	TCTTCATTCTTGC CATTCTCCTG
Hbb-bh1	GAGCTGCACTGT GACAAGCTTCA	GGGGTGAATTCC TTGGCAAAA	TGGATCCTGAGA ACTTCAAGC	GAGTAGAAAGGA CAATCACCAACA
Hmmr	TGCACAGCTACTT GGTCACCA	AGCTGAGATCGG AGTTTTGACACCT	GGTCACCAGAAC CTAAAGCAAA	CACCTCCGATTTG AGTTGGCT
Ikzf1	GCAAGCAATGTC GCCAAAC	TCCATCACGTGG GATGTCA	GAGACAAGTGCC TGTCAGA	TCATATCCTCCTT CTCATAGTTGG
Ikzf2	AAGAAGGGACGC TCTCACA	ACAGCGTTCCTTG TGTTCC	GACACCTCAGGA CCCATTCTG	TGAGCTGCGCTG CTTGTA
Itga2b	TTCCAACCAGCG CTTCACCT	TGCTCGGATCCC CATCAAAC	CGACAACAGCAA CCCAGTGTTT	GCCCACGGCTAC CGAATATC
Itgam	AGCAGGGGTCAT TCGCTACG	CAGCTGGCTTAG ATGCGATGG	ATTGGGGTGGGA AATGCCTTC	GTCGAGCTCTCT GCGGGACT
Itgb3	TCCTCCAGCTCAT TGTTGATGC	AGGCAGGTGGCA TTGAAGGA	ACGGGAAAATCC GCTCTAAA	AGTGACAGTTCTT CCGGCAGGT
Kdr	TGTGGGGCTTGA TTTCACCTG	TCGCCACAGTCC CAGGAAAG	CACTCTCCACCTT CAAAGTCTCATCA	TTTCACATCCCGG TTTACAATCTTC
Kit	CTGGCTCTGGAC CTGGATGA	CCTGGCTGCCAA ATCTCTGTG	TGCTGAGCTTCTC CTACCAGGTG	ATACAATTCTTGG AGGCGAGGAA
Laptm5	ATGACACTGCAG CTGCTAGAC	CCTCCTGGCTTG GGAGGTAA	CTGTTTGAGTATC CTGACCCTGTG	CATGTGGTTCATG GACTTGAAGTTG
Lgals9	TGGTGCAGAGGT CAGAGTTCA	AAGCAGCCGGAG ACAGCGAT	AGGTGATGGTGA ACAAGAAATTCT	GGTAGGGTACGC GGTGTTGG
Lmo2	TCGGCCATCGAA AGGAAGAG	GCGGTCCCCTAT GTTCTGCT	CTGGACCCGTCT GAGGAACC	GCAGCCACCACA TGTCAGCA
Lrp1	TGGGTGGATGCC TTCTATGAC	AGGTAGTTGCCAT GGTGACA	AATTGAGACCATA CTGCTCAATGG	CAGGCCGAAGGC ATGATTC
Ly6e	CTTTGGCTTGCGA ACCTTCA	ACATGAGAAGCA CATCAGGGAA	GCAGATGTCTGC CACTTCCAA	GAACTTGCTCCAT GCCCAGAA
Mctp1	CCTAAATCCTGTG TGGGAGGAA	AAGGCTGAGCCC ATAAAGTCA	AAGGCTTGCGTG CTCATT	TCTTGCAGTCCAA AATCATAGTCA
Mdk	CCTGTCAGCTCT GTCAATCAC	TCAGGGTGGGGA GAACAAAA	TGTCCTCTCACGC CCACAC	AGGGTATGGGGA GGCTCACT
Med12l	CTTTGTGTGCAAC ACCCTCA	CCACTCCGAACT CAGAACTGTA	AATGTGTGCATG GGACACCAG	CAGGCAGTTAGC TCAGAGGAGA
Meis2	CGACCCGTACCC TTCAGAAGAACAG AA	CGTTGGTCAATCA TGGGCTGCACTA	CGAGAAACAGTTA GCGCAAGACAC	CGAGAAACAGTTA GCGCAAGACAC

Myb	CGAGTGGCAGAAAGTGCTGAACC	CGTTGCTTGGCAATAACAGACCAAC	CGACATCAAAGGTCCCTGGACCAAA	CGACATCAAAGGTCCCTGGACCAAA
Mycn	GAGCCTTCGAATTGGGCTAC	CATGCAGTCCTGAAGGATGAC	AGATGCTGCTGCCGGAGGCCGAC	AGTGAGGCCACCCAGGCCCCC
Nfe2	TCCTCAGCAGAACAGGAACA	TGAGGCTCAAAAGATGTCTCAC	GGTTATCACAGCTGCCTGTT	TTTAGACCCTGCAGCTCAGTA
Nmt2	GCCACAAAACGCAGATACCA	AAGAATATGGTTCCTGGCGGATA	TGGGACACACAGCCAGT	TTGTCTTTGTCTGGTTCAATTGC
Notch1	CGACCAACCCTGTCAACGGCAAA	CGTATTTGCCTGCGTGCTCACAA	CGATGCCCCTCGGGGTACA	CGATGCCCCTCGGGGTACA
Notch2	AGCTGCTACTCACAGGTGAA	CCTCACAGTTGACACCAACC	CGAGTGCCTGAGCAATCC	CATCGCAGAGGCACTTATAGC
Pag1	ACAAGTCTCGGGAAGAAGACC	GCAGATTCTGGCCCTTTCATAC	CGACTCTTACAGAAGAGGAGATCTCAG	CCAGGTTTGTGTGCCGACTG
Pbx1	CAGGACATCGGGGACATTTTACA	CACATTAAACAAGGCAGGCTTCA	AATTATGACCATCACAGACCAGAGTT	TTCTGTGGCAGTTTAAAGCATGTT
Pcdh12	CGATGGCTGCTTTTGCGGAAC	TCGTGGTTTGGTTTGGGCTGGAA	CGAGGAACCCGGTGGAGGA	CGAGGAACCCGGTGGAGGA
Pde2a	CTGTCCTGCTACAGGAGATCATCA	CCACAACGCCACCATCGAA	CGGAAGCCAGAAACCTCA	CCACCAGCTCGTTCTGAT
Pde3a	TACACGCTGTGTGGTATCTCA	ATCCCATGTGTCCGTGTGTA	ATTCCTGGCCTCCCAAGT	TCCACTGTCAGAATCGGAGT
Pecam1	TGCGGTGGTTGTCATTGGAG	CTGGACATCTCCACGGGTTT	GTCATCGCCACCTTAATAGTTGCAG	TGTTTGGCCTTGGCTTTCCTC
Pfas	CAAGAGCTGCAGCGGAAC	CGCTAGCTTCTTCCCATCCA	CCAGTACAGTGGAGGTCTTTGA	CATGGAGCTGGCCCTTGA
Pla2g4a	CGTGTGGAATGAGACCTTTGAGT	AGGGAATGTAGCTGTGCCTAGGGT	ACCTTTGAGTTCATTTTGGATCCT	AGTGTGATCTCCAAAACATTTTCCTGA
Plcg2	AGCATGCGCTCTGAGAAGTA	GGCAATACTCCGCCCTAGTTTA	TGATCCGATGCCCCTGGAG	GGTGGCGTGCACCAAGAA
Ppia	CGACCGACTGTGGACAGCTCTAA	CGTAGTGAGAGCAGAGATTACAGGAC	CGATTTCTTTTGACTTGCGGGCATT	CGATTTCTTTTGACTTGCGGGCATT
Ptk2b	AGAAAGACTGTACCCAGGACAAC	TCCAGGTGGGTTCCTCTTCA	AAGGAGAAGTTCATGAGTGAGG	TGCCAATCAGCTTCACGAT
Ptprb	ACGCCAAGAGCGGCAATTATGCA	CGTTGCACCCAGGACACCTTTAA	CGACCACTCCTTCACCGAGGAA	CGACCACTCCTTCACCGAGGAA
Ptprc	GGCTTCAAGGAACCCAGGAAATA	TGACAATAACTGTGGCCTTTTGCTC	ATTGCTGCACAAGGGCCCCGGGATG	CAGATCATCCTCCAGAAGTCATCAA
Rab38	GCGCTGAAGGTGCTCCA	CCCCATAGCTTCCCGGTAATAA	GTGGTGCGCTTGCAGCTCTG	ACTCTTGTCATGTTTCCAAATC
Ramp2	TCCCACTGAGGACAGCCTTG	TCCTTGACAGAGTCCATGCAA	TCAAAAGGGAAGATGGAAGACTACGA	TCTTGTACTCATACCAGCAAGGTAGGACA
Rbm38	GCTGATAGGGCTTGCAAAGAC	GCTGCACACCAACAGCAA	CCTATCATCGATGGTCGCAAGG	AGCCCGTCTGTAAGCTCCTAG
Runx1	CGAACTACTCGGCAGAACTGAGAA	CGTACGGTGATGGTCAGAGTGAA	CGAATGCTACCGCGGCCATG	CGAATGCTACCGCGGCCATG
Serpinf1	AGAACGTCCCCAGCAGCTCT	TGCCAGCTTGTTCACAGGGACC	AGCAGCTCTGAGGGCTCCC	TCCTCCTCCTCCACGGGCT
Sla	CGACGAATCTTCCGTCTTCCCAAC	TCGGGTGAGCACACAGCATAGAC	CGAACTGGTACTACATCTCACCAAGG	CGAACTGGTACTACATCTCACCAAGG
Smad2	TGCTCTCCAACGTTAACCGAAA	TCAGCAAACACTTCCCCACCT	GCCACTGTAGAAATGACAAGAAGACA	TGTAATACAAGCGCACTCCCCTTC
Smad3	CCAATGTCAACCGGAATGCAG	TGAGGCACTCCGCAAAGACC	CGTGGAACTTACAAGGCGACA	CCCCTCCGATGTAGTAGAGC
Smad6	TTCTCGGCTGTCTCCTCCTGAC	TTCACCCGGAGCAGTGATGA	GTACAAGCCACTGGATCTGTCCGA	GGAGTTGGTGGCCTCGGTTT

			TT	
Smad7	GGAAGATCAACC CCGAGCTG	TGAGAAAATCCAT TGGGTATCTGGA	TGTGCTGCAACC CCCATCAC	AAGGAGGAGGGG GAGACTCTA
Snai1	CGATCTGCACGA CCTGTGGAAA	CGTGAGCGGTCA GCAAAAGCA	CGACTCTAGGCC CTGGCTGCTT	CGACTCTAGGCC CTGGCTGCTT
Sox17	CGAGATGGGTCT TCCCTACC	TCGGGGTAGTTG CAATAGTAGAC	GGACACGACTGC GGAGTGAA	GTCGGACACCAC GGAGGAAAT
Sox6	TCAACCTGCCAAA CAAAAGCA	CTGTGTCCTGGT CCCCAAA	GCATCCCCAGCC CCATTG	CAGGGCAGGAGA GTTGAGACT
Spi1	GTGGGCAGCGAT GGAGAAAG	TGCAGCTCTGTG AAGTGGTTCTC	ATAGCGATCACTA CTGGGATTTCTCC	GGGAAGTTCTCA AACTCGTTGTTG
Stk26	AGCAGCGCGAAG TAGCA	GCCTTTTCCAATG CGCTCTAA	CCACCATGGCCC ACTCA	TTTGTGAACAGTT CTTCTGGATCTG
Stxbp2	AAACGGAGAGAA CCCATTCC	GGGCTGCTTTGT AGGTGAA	GCTTGGAGGCAA TTTATTTGCTGAG	GGTTGGTGTTCC CTGGAAGTC
Tal1	ACCGGATGCCTT CCCCATGTT	GCGCCGCACTAC TTTGGTGT	CCAACAACAACC GGGTGAAGA	AGGACCATCAGA AATCTCCATCTCA
Tek	TCCAAAGGAGAAT GGCTCAGG	TCCGGATTGTTTT TGGCCTTC	TTCCAGAACGTGA GAGAAGAACCA	TGTTAAGGGCCA GAGTTCCTGA
Tgfb1	ACCCCCACTGATA CGCCTGA	GCAGTGAGCGCT GAATCGAA	TGGCTGTCTTTTG ACGTCACTG	GCCCTGTATTCC GTCTCCTTGG
Was	CGTTCAGCAGAA CATTCCTTCC	AGTGCCAGGTAG AGCTGAAC	AACCTCCTCCAG GACCATGAA	CTGTGGTAGCCA GTGTCCAG

Supplementary Table 2: Differentially expressed genes between CD44Negative and CD44Low/Kit- populations

	Base Mean	log2 Fold Change	Adjusted p-value
Ankrd1	25.50618636	30.80488201	7.14E-15
Tgfbi	161.3901095	10.51636231	1.03E-07
Apln	63.40484308	9.861703974	9.04E-18
Lhx6	52.64044743	9.603183946	1.32E-17
Gas7	93.98741793	9.266688291	1.18E-13
Fam167b	31.4403707	8.848524118	8.40E-08
Kbtbd11	31.29723969	8.788236531	3.24E-12
Prl3b1	230.2757472	7.816149948	1.13E-09
A730020E08Rik	17.79762791	7.723318885	2.68E-06
Mctp2	17.0317109	7.572013538	0.00270569
Dfna5	15.71572309	7.548019237	4.95E-05
Rasd1	13.33515905	7.517530036	1.49E-05
Slc35f2	14.59325379	7.509254061	0.000174517
Ccser1	11.41392086	7.306562694	0.000505947
Fut4	13.22931601	7.284141351	2.17E-06
Mlkl	21.28689963	7.263506449	1.86E-07
Slc2a12	10.9733371	7.229540552	0.009107188
Tnik	15.0499464	7.224897962	0.000364522
Pip5k1b	11.17532413	7.159959334	0.008066971
Tnfrsf9	13.30384109	7.11462124	0.00062631
Fam19a2	9.357954411	7.110944663	0.003798265
Nr2f1	9.541781298	7.093658504	0.000447407
6430573F11Rik	9.128588977	7.056586581	4.56E-05
D330045A20Rik	23.55768296	7.000742054	8.48E-05
Elovl4	11.57549775	6.933894172	0.003861909
Fancf	11.66122717	6.878240253	0.000376731
Pcx	8.721972385	6.802060229	0.000260268
Cmah	35.43330851	6.791001024	1.31E-05
Thsd4	10.80749624	6.769092364	0.007731119
Gm826	8.099302714	6.664694074	0.000645957
Fbxo48	9.508180036	6.626351963	0.001796723
Fam212b	6.771131773	6.540923213	0.006616589
Shisa2	8.1515097	6.513069612	0.006730245
Cnr2	21.48756544	6.473032871	3.57E-06
Nlrc3	26.22581678	6.439929129	1.53E-06
Prkcq	26.99969101	6.352413377	0.000532665
Gprasp2	11.76309195	6.232166078	0.005910292
3300005D01Rik	8.973928472	6.069259786	0.00093867
Kcne3	542.7033245	6.06867804	2.55E-08
Nkd1	29.64404099	6.033836205	3.60E-06
Zfp619	20.52222821	5.900063169	8.13E-06

Best1	6.700381274	5.820134532	0.007058308
Sergef	23.44004177	5.626848995	0.002595308
Ccl6	120.1257075	5.570351112	0.002777379
Mt1	330.0015636	5.520101233	7.11E-14
Bik	42.94178703	5.448640006	2.49E-08
Mafb	35.83651387	5.415497371	2.81E-06
Spry1	47.98470795	5.331786413	1.04E-09
Nrp2	221.354704	5.304305382	4.12E-07
Hgf	40.43768824	5.225349597	0.006538649
Hps6	9.713434711	5.116648795	0.00239975
Pde2a	383.1668234	5.077376892	9.04E-06
Fzd10	26.85566228	5.03765585	1.76E-05
Mylpf	12.19876539	5.01017655	0.000255187
Epb41l3	16.14208692	4.96726305	0.003068604
St8sia1	66.62618437	4.961531972	9.09E-09
Gpr183	419.6854705	4.906296555	8.60E-06
Stab2	115.2685186	4.77357517	0.005418404
Srl	9.838735419	4.772621086	0.008313103
Ackr1	35.26158033	4.673118011	0.001894871
Lama1	44.02834481	4.668218739	1.20E-05
Lalba	13.01679218	4.642384763	0.003056338
Adamtsl2	29.06907061	4.603317373	5.84E-05
Tmem184c	118.6019203	4.586103589	8.49E-08
Colgalt2	55.1848555	4.537715687	0.000700391
Piezo2	249.2624354	4.519063563	5.08E-09
Kcnk5	20.8338197	4.498626032	0.002146742
Aplnr	1318.543925	4.431561993	0.002409416
Adamts9	51.69351808	4.395534174	0.00023005
Etv2	33.62657579	4.316847355	0.002838678
Ptpn7	93.77173225	4.298106231	3.87E-07
Orai2	25.97943456	4.270158623	0.002635872
Otud3	8.664034829	4.258012489	0.006102439
Tmem108	9.336023979	4.116789524	0.00160165
Mterf1b	8.288838383	4.095992307	0.004716617
Cdkl1	46.15810993	4.066157262	0.00111066
Fuom	14.41432652	4.059354061	0.000800277
3110035E14Rik	6.753435517	4.026324158	0.002949328
Cadm1	54.56282319	4.018209876	0.000162956
Nostrin	50.30814367	4.009328514	1.24E-05
Nr2f2	347.3381797	3.964062993	0.002875362
Prtg	407.954527	3.786525846	0.007471836
Tfec	163.2517233	3.768285495	4.72E-05
Bcat1	146.0503918	3.755350486	1.02E-05
Adgrg3	361.8876046	3.750773521	4.55E-06
Clec1b	532.6457321	3.67106534	0.004910519

Aunip	41.04170901	3.663970997	3.84E-05
Lrp8	37.58449465	3.586735407	0.000849563
Ak4	45.74298584	3.554337962	0.000282524
Mcc	61.53054348	3.530052601	0.005528985
Ramp3	24.93878935	3.525241566	0.009624378
Ttyh2	50.98333011	3.461134012	0.003780898
Snai1	84.34270284	3.418124761	0.000281456
Apold1	182.5033905	3.402928614	9.11E-05
Cmbl	34.74661164	3.343351781	0.001002703
Ydjc	15.49640277	3.281997459	0.006201782
Ciart	23.50574211	3.20648145	0.004063743
Sgk1	588.0161194	3.174327357	2.32E-05
Bok	180.1083177	3.084951568	1.35E-08
Ptges3l	19.74439998	3.015138261	0.000937891
Rgs7bp	32.29205105	3.014673925	0.008983211
Gng2	566.2973422	2.985285866	1.90E-09
Pcbp3	65.48101262	2.968402904	0.00043792
Lef1	99.12002007	2.948983775	3.24E-12
Ccnd1	531.5793752	2.93956831	4.26E-14
Phlda1	50.50855921	2.930155994	6.48E-07
Slc16a6	28.25647203	2.913338502	0.007127371
Sh2d3c	235.2968885	2.912104432	7.61E-10
Zfp41	108.7591782	2.844744034	0.001042896
Tln2	122.688666	2.837284952	0.004211829
Lcmt2	42.31211951	2.697843322	0.006242368
Dok4	191.312712	2.694139264	5.03E-07
Tiam1	35.52523816	2.690899886	0.003135874
Ptpn4	26.28700065	2.685462044	0.00270877
Lhpp	98.16169232	2.673170778	4.52E-06
2610203C22Rik	125.4281361	2.665946049	0.000452092
Fam102b	177.9650743	2.648061088	0.002513408
Arhgap10	35.43605036	2.629608838	0.000135465
Hlx	44.84965193	2.614098889	0.001464693
Lgr4	119.8794716	2.59014538	0.00347659
Mmrn1	1353.872461	2.589564232	4.30E-10
Asns	432.031957	2.587583511	4.55E-10
Lck	167.9620297	2.57940835	1.24E-06
Trib1	127.3071712	2.577190873	2.39E-06
Pfkp	141.3525152	2.563075919	6.81E-07
Rpph1	26.84442114	2.546609846	0.005323279
Ppp1r14a	22.12636926	2.542543733	0.007297326
Plau	126.5927648	2.540495059	5.31E-05
Bmp4	182.1357205	2.52712181	0.000173932
Ada	91.83534223	2.518627373	0.00077738
Gem	112.4974042	2.508378804	0.00865646

Peg13	67.58738143	2.506492362	0.001322676
Fam132a	44.21702681	2.506111959	0.003103894
Mns1	64.0331474	2.504644656	0.001809767
Rhobtb1	28.75688235	2.495552069	0.003617269
Fam174b	80.13067739	2.482717182	0.003326631
Slc16a3	163.465678	2.447539376	9.18E-07
Dmtn	102.4715512	2.427184483	0.001045852
Dnph1	146.1811871	2.415319699	0.006106394
2010204K13Rik	55.05728442	2.391924196	0.006656831
Abhd8	64.55880258	2.350006183	0.000387277
Stx2	113.7803836	2.272821442	0.007644503
Ripply3	218.5103965	2.253860526	0.000120111
Cnrip1	186.0186162	2.252050612	0.00370756
Gemin4	198.4461414	2.243080625	3.13E-13
Lat	85.43576622	2.230546252	0.00047926
Crmp1	532.1941855	2.201081263	3.46E-06
Mcm10	209.349248	2.19332879	1.21E-06
Pcdh12	125.1909264	2.172987949	0.004148032
Slc29a1	812.4106385	2.156677328	1.34E-11
Tubb6	1677.868899	2.127057195	8.14E-06
Lipt2	268.5149927	2.126937491	7.73E-10
Chaf1b	427.8517934	2.098338699	3.05E-16
Trps1	71.04208794	2.097537754	0.008286358
Pkm	5878.223495	2.093383049	1.76E-28
Asf1b	591.8978876	2.032363005	0.000263584
Eno1	16.25189932	2.027577027	0.00239975
Eno1b	333.3380746	2.009950924	2.21E-05
Mthfd2	183.9073635	2.004672066	0.001998631
Hspbap1	97.26241189	2.001488452	0.004166933
Txnrd2	124.6937997	2.001359405	4.46E-05
Zfp697	48.0247465	-2.00239518	0.001456976
Fndc3b	384.7662243	-2.011481223	1.01E-05
Cyp26b1	232.401227	-2.026172164	2.91E-05
Tanc2	148.3280735	-2.027383926	0.001925265
Insr	216.1669914	-2.030687775	4.49E-08
Mylip	160.5543537	-2.038397771	2.06E-05
Gja4	3607.144179	-2.04140761	0.000197682
Cpq	185.0160442	-2.045098576	3.51E-05
Cdk19	197.5256576	-2.051885101	0.003033104
2610008E11Rik	343.5634581	-2.053446593	1.07E-07
Ghr	319.0665962	-2.089524481	8.68E-08
Efnb2	1078.208725	-2.091005095	0.000878689
Dnajb9	1072.263465	-2.09202724	7.77E-11
Bcl9l	20.71942381	-2.092553823	0.005861375
Ccpg1	212.6461943	-2.096878144	1.57E-05

Tob1	65.9974918	-2.102942757	0.000293692
Sat1	5786.321167	-2.108170217	1.89E-07
Smim3	167.9532155	-2.114904783	0.0016516
Xpr1	175.0623769	-2.11773641	0.000152169
Mgst3	236.4399353	-2.122949816	1.04E-07
Khdrbs3	92.81144596	-2.126084516	0.001093751
Diaph2	174.2756952	-2.12901338	1.53E-06
Gabbr1	59.25078487	-2.130276741	0.00270877
Fam101b	64.93688331	-2.149021013	0.000741716
Mfap2	938.0850374	-2.150106301	1.62E-06
Spata1	30.91092925	-2.153493069	0.004467802
Smad7	161.8905493	-2.15608351	2.62E-05
Asah2	76.55631038	-2.161335691	0.005177279
Mum1l1	425.3095902	-2.161376864	0.000260268
Nfkbiz	78.47104742	-2.164090677	0.002072117
Firre	111.6090085	-2.167895543	0.000331737
Itm2a	1782.061766	-2.172156365	4.56E-12
Slc25a40	114.7276806	-2.17588334	0.000261872
Mir99ahg	180.2345547	-2.212864064	0.000499024
Trim47	302.5256951	-2.218397032	2.32E-08
Dst	223.1978802	-2.24382604	5.16E-06
Ccdc80	238.4003441	-2.246684808	0.000117708
Npr3	494.8244693	-2.264198292	0.006061632
Tapbp	183.8310852	-2.264712582	4.81E-07
Emp3	392.7862493	-2.270409335	2.92E-08
Tox	112.4869272	-2.274080721	1.89E-07
Pmepa1	176.9453091	-2.300774507	3.54E-05
Gm14420	71.76150356	-2.301041817	0.009040874
Meis2	252.327756	-2.305849462	0.000322625
Camk2d	439.7186277	-2.31129653	3.64E-08
Yy2	212.722359	-2.313377624	0.000115196
Tspan8	174.6695876	-2.33468579	6.80E-06
Ifitm1	203.0983107	-2.337077573	0.0030486
Edn1	554.5931726	-2.337680279	0.000162291
Prdm16	99.05991569	-2.348027813	0.000503655
Stk38l	121.4119394	-2.348176317	3.66E-08
Laptm5	664.5691098	-2.35049818	1.46E-09
Gm4944	170.3473918	-2.354854452	0.006248635
Klhl13	615.4932674	-2.375564244	1.15E-13
Pde9a	140.6664275	-2.375798962	2.54E-06
Setbp1	50.35822821	-2.377380244	0.008798509
Nadk2	267.4012478	-2.390123642	1.29E-08
Acer2	474.7504486	-2.390745584	4.30E-07
Kank1	133.7521537	-2.406432786	0.002215898
Snhg14	100.2087273	-2.415460322	0.007888183

Scai	237.7870519	-2.428704528	2.17E-05
Prkcd	152.435323	-2.429452225	0.007084341
Tmx4	297.0814445	-2.441935768	7.68E-12
Fam198b	892.8470702	-2.445693944	8.39E-10
Zfp667	113.5986569	-2.446819119	0.002840966
Lpin1	52.80732225	-2.453053179	0.004629388
Pde3b	249.1256989	-2.456901943	1.73E-06
Pdcd4	421.403935	-2.459878723	5.76E-16
Cdk14	151.9795738	-2.461081481	9.16E-06
Zbtb4	47.56615494	-2.464782482	0.005079212
Ptchd1	527.1282626	-2.493047998	1.37E-09
Aqp1	60.15802138	-2.507683462	0.003592063
Senp7	178.6571111	-2.509966518	1.58E-10
Pld1	318.6574308	-2.520752502	1.02E-07
Fam181b	83.24531905	-2.522185116	0.002481606
Serpinf1	1173.297443	-2.525318969	5.52E-16
Thsd7a	123.9001816	-2.535879428	0.003258933
Aes	144.3662628	-2.546156229	1.31E-10
Stmn2	681.9122616	-2.551142191	0.000811735
Nfat5	420.6209037	-2.555868611	1.56E-07
Gja5	1022.975322	-2.561364945	4.15E-05
Il15	191.3510805	-2.563015325	5.30E-06
Col5a2	835.5105201	-2.564588528	4.56E-12
Igsf3	228.5166715	-2.569349426	5.15E-12
Sost	39.99614021	-2.578664215	0.000345192
Meis1	216.5926366	-2.579594011	1.45E-06
Slc9a3r1	139.8253626	-2.582919538	2.94E-06
Mecom	1071.132398	-2.611771496	1.00E-17
Trim34a	90.03978047	-2.61607329	0.007698747
Unc13b	76.50883651	-2.627542032	0.003750911
Gata6	102.9752231	-2.635742846	9.78E-05
Zfp97	51.94495484	-2.64416592	0.00585697
Hs3st1	156.4178276	-2.645065468	7.87E-06
Fn1	5985.622465	-2.675337291	3.42E-28
Sncaip	64.58688608	-2.690394777	0.005981093
Timp2	140.2697289	-2.692241973	2.11E-08
Notch3	71.82103636	-2.716639469	0.001155911
Calcrl	4603.27198	-2.720760135	2.15E-24
Nipal2	79.72927297	-2.747410085	0.003054315
Tmem176b	720.0868452	-2.748968116	1.17E-09
Fam84b	177.4557529	-2.757428754	0.00014231
Ssfa2	623.1719156	-2.759015082	1.43E-16
Plcg2	309.8345006	-2.793749742	0.000101095
Slc16a4	62.05283257	-2.799637079	6.45E-06
Gbp7	149.3667069	-2.80586988	0.000442463

Fgd4	147.1824455	-2.81097205	0.000638745
Ptgis	181.9398434	-2.812924918	3.55E-08
Lmo7	62.66091025	-2.832201053	0.004305355
Alpk3	45.19896139	-2.844537008	0.002157956
Ctsh	1401.109937	-2.86772808	6.18E-06
Sox6	162.1639305	-2.91816727	0.001364567
BC064078	52.3817415	-2.956683476	0.001018225
Neto2	125.6935462	-2.987457856	0.000410936
Enpp4	113.270224	-2.995746	0.000415199
Slfn2	146.8819422	-3.002569696	0.002252592
Mtap7d3	34.70340077	-3.003235697	0.008642032
Fermt3	232.347294	-3.003377117	0.005420852
Slc18a2	1180.551238	-3.004173686	4.96E-15
Ocln	284.9178802	-3.022406905	8.39E-09
Lama5	140.7140661	-3.023547261	1.93E-08
Smpdl3a	394.3377886	-3.042367317	2.10E-08
Fjx1	35.60743188	-3.046265094	0.000297076
Neurl3	293.6950791	-3.057968425	0.004260675
Krt18	177.2449094	-3.061396679	0.006656831
Pdlim1	451.2628355	-3.078267654	3.34E-14
Vegfa	57.60109553	-3.0837278	0.000139904
Capn5	31.8277457	-3.088913259	0.005797405
Dmd	190.4620428	-3.095095329	2.32E-10
Pla2g4a	622.4318404	-3.100313503	0.008679177
Crebrf	42.87213035	-3.107512312	0.002250157
Nkx2-3	106.1645302	-3.131720852	0.000139511
Procr	363.3611547	-3.138509913	1.67E-09
Srgn	604.7919667	-3.143177768	2.24E-05
Ctsc	1919.960507	-3.149736412	4.67E-08
Cd44	890.8903188	-3.152258915	9.71E-13
Dhh	86.24882163	-3.168188234	0.00112965
AU021092	328.3931668	-3.174369271	1.19E-09
Emb	2027.975447	-3.178579222	3.67E-18
Smad6	142.0697375	-3.180672146	1.28E-07
Mllt3	289.3727827	-3.224110034	5.69E-11
Pcsk5	453.7528369	-3.229863171	5.19E-08
Tm4sf1	1187.439505	-3.237339439	8.05E-12
Egfl8	51.52588808	-3.245861182	0.001874349
Jag1	1778.331273	-3.253399908	5.58E-06
Itga4	256.2172439	-3.276985812	0.000796385
Tmem100	1611.699263	-3.284758069	0.002704022
Gpr85	22.0927988	-3.28483231	0.009005113
Smtnl2	155.6631868	-3.342930901	0.000108383
Tnfrsf26	196.233847	-3.346020352	5.26E-05
Adgrg6	213.8991108	-3.352123622	7.08E-07

Tmem221	59.13493784	-3.372574835	0.002698392
Crispld1	104.3196103	-3.375935239	0.000492617
Sp110	61.14902254	-3.378180854	4.82E-05
Gm20939	89.99409159	-3.390713393	0.004367915
C530008M17Rik	144.2685862	-3.411073592	4.52E-06
Isl1	180.7125252	-3.415247376	0.000262179
Gpm6b	267.5987719	-3.420775509	0.00074013
Fas	218.5046748	-3.429257795	1.24E-06
Sfrp1	42.09572157	-3.461726617	0.005735329
Ankrd29	135.7451404	-3.468421351	2.64E-06
Lama3	298.3706902	-3.470725045	2.29E-07
Frk	146.2696365	-3.516116733	3.45E-06
Bcl11a	221.5621111	-3.519152267	1.02E-14
Xk	69.47002093	-3.60433993	0.001333542
Hacd4	419.1244511	-3.621811213	7.52E-12
Vwf	1543.343578	-3.622195129	1.60E-10
Synpo	186.5357296	-3.654890838	1.06E-05
Sdpr	996.1265572	-3.674639554	0.001940231
Was	45.48249758	-3.743889227	0.008521852
Npnt	85.94979989	-3.749120788	0.002520529
Adgrg1	397.4560578	-3.764534564	7.48E-05
Dcaf12l1	79.26006413	-3.766468952	0.002361003
Gas1	24.63902761	-3.771362018	0.003456357
Ltbp4	524.8986434	-3.782275174	1.16E-05
Wnk4	47.43961326	-3.785468914	0.000127433
Naip5	86.54078872	-3.786806454	0.001910655
Alox12	204.0352482	-3.790211478	0.002215898
Bcam	35.28086475	-3.79098106	0.001224602
Fbn1	1328.807222	-3.794576845	3.05E-16
AI467606	87.1251668	-3.81652144	0.001580202
Bmx	480.0257825	-3.83832336	9.47E-06
Ppargc1a	21.53886161	-3.843924835	0.009889554
Ddb2	18.82384718	-3.855791735	0.002156538
Ltbp1	357.4724101	-3.873083349	1.18E-14
Rps6ka5	162.409676	-3.877001758	3.12E-10
Wipf3	49.49306613	-3.896330049	0.001107964
Gcnt1	106.5783666	-3.911424284	2.06E-05
Cyp4b1	36.44988202	-3.928411273	0.006180715
4930412O13Rik	42.33118633	-3.995332324	0.001874349
Serpinb8	63.20614008	-4.008865931	0.000122099
Kcnn4	29.33351762	-4.023341203	0.000718934
Scarletltr	4.287175301	-4.033958253	0.003155112
Nfe2	127.8995985	-4.114917568	6.04E-05
Hey2	335.1618049	-4.190503806	7.97E-05
Psmb8	76.95527611	-4.192134905	0.001771326

Dkk2	242.0562233	-4.21312208	0.002397197
Mr1	61.54938681	-4.25230154	0.003406739
Plek2	33.2153158	-4.262347762	0.004152759
Ikzf2	163.8263323	-4.288764745	4.94E-08
Sult1a1	69.60759284	-4.31299754	5.24E-06
Pde3a	82.37867245	-4.36573932	6.81E-05
Pdzk1ip1	60.71222515	-4.388962016	0.000594807
Mfap5	47.47202926	-4.419359181	0.005992706
Pstpip1	16.64815545	-4.469141588	0.00796169
Grik5	14.78589095	-4.476214081	2.28E-06
Gp1ba	33.89053099	-4.488660505	0.00019326
Tnni2	22.64471465	-4.488953199	0.004546222
F10	104.5735738	-4.503076863	1.03E-06
Ms4a6d	344.8554883	-4.535570295	0.000694591
Klhl35	19.47662844	-4.567564877	0.004153971
Stat4	147.6705734	-4.574453513	0.001333542
Lrrc9	23.9465403	-4.583362593	0.006248635
Smtnl1	18.98691679	-4.624930114	0.006183205
Spi1	419.4956989	-4.641969638	0.000353015
Bmper	74.71196915	-4.66430177	0.006315645
5930403L14Rik	65.9032584	-4.67939998	0.000110954
Akr1c14	573.5535666	-4.691029526	5.52E-05
Tnfrsf23	53.5718724	-4.731559281	2.51E-06
Col4a5	101.1502735	-4.736695126	4.78E-07
Hlf	71.04738958	-4.758502325	0.000107682
Slc8a1	134.2257614	-4.761549625	0.00011505
Erich2	5.319713805	-4.775262561	0.009865634
Nfix	22.22833281	-4.781904636	7.24E-06
Calml4	26.27866243	-4.846395303	0.000827363
Isoc2b	13.29986744	-4.894334829	0.000140172
Lrp1	33.41498845	-4.964992877	0.005506852
Syne4	18.74216505	-4.974897841	3.96E-05
Psmb9	29.06089031	-4.975976946	0.001585076
Lrrc16a	32.75623191	-4.980354993	2.74E-06
Gfi1	61.39080196	-4.990739639	0.000309948
Rragd	23.45893981	-4.999731884	0.001988171
Nmnat3	20.36278187	-5.043358316	0.009889275
Anpep	24.79727263	-5.055674976	9.39E-06
Dsp	43.51339786	-5.10113372	1.95E-05
Pmaip1	47.16894906	-5.131386174	0.001876941
Pak7	32.77866077	-5.138943462	0.000329005
Met	151.2609249	-5.149681829	3.41E-05
Cytl1	50.18970608	-5.195519906	3.81E-05
Irf6	109.378421	-5.277840743	1.21E-08
Rhof	31.33032443	-5.291941806	2.12E-05

Selp	16.78962727	-5.364284262	0.004775781
Mctp1	131.8708823	-5.386165582	1.00E-07
Psd4	26.45007951	-5.397226565	0.002393873
Lmo1	14.37302096	-5.413149267	0.005300239
Mt3	19.84340166	-5.522226678	0.000180803
P4ha3	205.9657655	-5.546860417	0.000166442
Chil5	17.7634245	-5.572077357	0.003151683
Syt15	79.93034677	-5.634248142	7.03E-07
Acaa1b	10.98968606	-5.705877685	0.000141501
P2rx7	17.13769739	-5.712104747	0.006916767
Fgfr3	11.02227834	-5.911237503	7.24E-06
Gpr84	22.52893241	-5.914701165	6.21E-05
Ctse	14.21116783	-5.946404861	0.007936364
Efemp1	144.2600701	-5.952631051	0.000152809
Nupr1	116.8706155	-5.971952377	4.89E-09
Gucy1a3	55.1915894	-6.036565647	0.00269354
Has2	37.98998807	-6.068936712	4.59E-10
Olfml2a	18.92743865	-6.072841895	0.000924955
Ivl	7.507074635	-6.084293968	0.007008491
Olfr65	12.48187221	-6.19464371	0.00072814
Spaca3	9.945287611	-6.267494712	0.002843607
Htra1	354.0196329	-6.3347494	8.14E-11
Eln	84.34818961	-6.352916902	2.10E-08
Nox1	9.362588681	-6.402243085	0.008685178
Rprm	9.027051747	-6.454293908	0.000140172
B4galnt2	15.82020492	-6.462112453	1.16E-05
Atp2a3	36.95610278	-6.506208072	1.04E-06
Scx	7.501716724	-6.55619466	0.002422051
Syt10	9.610849779	-6.565298696	0.000800277
Pitpnm1	6.971062106	-6.631538221	0.000815258
N4bp2l1	40.60930877	-6.638796432	3.35E-07
B3galt2	13.17993668	-6.656381035	0.007258845
Dok2	245.255472	-6.766075464	6.71E-10
Gm7694	23.35661877	-6.834020521	0.000252112
Stk26	55.63633961	-6.958570306	0.002115685
Tmem71	21.92599808	-6.977893638	4.07E-05
Gria3	36.23682964	-7.183381614	4.40E-11
Ikzf1	110.3957947	-7.195718005	3.87E-10
Zcchc16	10.43416971	-7.490146108	0.003889647
Clec4e	123.8182189	-7.512868091	4.13E-09
Pcdh11x	11.09887025	-7.547919412	0.003717358
Casp12	17.83300622	-7.590816992	0.000153776
Cpne7	14.51037015	-7.869964983	0.000329284
Gkn3	89.08396995	-7.977462263	7.83E-07
S100a4	43.58391797	-8.657209255	2.63E-13

Mgp	36.20937276	-9.196189016	7.23E-05
Fbln5	410.100214	-9.351894972	0.000322495
Lox	43.80010341	-9.377415773	8.63E-05
Cfi	92.57464584	-10.21368425	1.93E-08
Sla	196.3674958	-22.33554716	1.90E-17
Chst9	8.208234389	-24.95145428	4.41E-12

Supplementary Table 3: Marker genes identified by single-cell RNA sequencing analysis

Cluster 1: VE-Cadherin- AGM					
	p-value	**Average logFC**	**Percentage 1**	**Percentage 2**	**Adjusted p-value**
Crabp1	4.52E-25	1.910126326	0.674	0	5.31E-21
Mdk	7.26E-19	1.675144325	0.93	0.341	8.54E-15
Gpc6	8.97E-17	1.737179173	0.86	0.244	1.06E-12
Peg3	9.20E-16	1.707464536	0.953	0.496	1.08E-11
Col4a1	7.70E-10	-1.506028093	0.395	0.8	9.06E-06
Afp	1.78E-09	-2.39757156	0.023	0.57	2.09E-05
Calcrl	9.21E-09	-2.253097582	0.023	0.533	0.00010841
Cdh5	1.16E-08	-1.884704132	0.023	0.526	0.000136971
Ttr	5.01E-08	-1.639174794	0.023	0.496	0.000589963
Flt1	2.00E-07	-1.777476838	0.047	0.496	0.002347822
Ptprm	2.98E-07	-1.75771161	0.209	0.6	0.003505492
Kdr	9.24E-07	-1.582539288	0	0.407	0.010876558
Elmo1	9.44E-07	-1.667179364	0.047	0.452	0.011105373
Stab2	3.05E-06	-2.767205077	0	0.378	0.035903345
Cluster 2: VE-Cadherin+ AGM					
	p-value	**Average logFC**	**Percentage 1**	**Percentage 2**	**Adjusted p-value**
Emcn	2.47E-23	1.589077525	0.745	0.038	2.91E-19
Hba-a1	1.11E-20	-3.437116576	0.745	0.992	1.30E-16
Hba-x	2.71E-20	-3.465859872	0.851	0.985	3.19E-16
Hbb-bh1	5.59E-20	-3.55073909	0.766	0.985	6.58E-16
Hbb-y	3.19E-19	-3.261473361	1	0.992	3.75E-15
Hba-a2	1.01E-15	-3.051395255	0.574	0.908	1.19E-11
Hbb-b2	3.25E-12	-2.539147613	0.128	0.725	3.82E-08
Mecom	3.33E-12	1.553983605	0.745	0.221	3.91E-08
Hbb-b1	1.03E-10	-1.807832267	0.021	0.595	1.21E-06
Afp	1.14E-10	-2.454551483	0.021	0.588	1.35E-06
Ttr	1.63E-09	-1.696949333	0	0.519	1.92E-05
Mt1	2.16E-09	-1.888451876	0.255	0.702	2.54E-05
Col1a1	2.41E-09	-2.64667958	0	0.511	2.84E-05
Apoa1	3.30E-08	-1.698690798	0	0.458	0.000388075
Col3a1	4.41E-08	-1.849208399	0.021	0.489	0.000518502
Car2	4.99E-08	-1.616543299	0.043	0.496	0.000586596
Col1a2	1.00E-07	-1.872684794	0.021	0.458	0.001176236
Mt2	6.83E-07	-1.505230803	0.085	0.496	0.008032469
Cluster 3: VE-Cadherin- YS					
	p-value	**Average logFC**	**Percentage 1**	**Percentage 2**	**Adjusted p-value**

Col1a2	3.14E-25	2.207998754	0.953	0.148	3.69E-21
Col1a1	6.76E-25	2.535544321	0.977	0.185	7.95E-21
Col3a1	5.72E-23	2.271305297	0.953	0.178	6.73E-19
Hba-a1	4.15E-21	2.308475377	1	0.904	4.88E-17
Hbb-y	4.36E-21	2.314894285	1	0.993	5.13E-17
Hba-x	5.44E-21	2.238500133	1	0.933	6.40E-17
Hbb-bh1	6.53E-21	2.186465982	1	0.904	7.68E-17
Hba-a2	1.35E-20	2.182455257	1	0.763	1.59E-16
Hbb-b2	3.50E-20	1.735779634	1	0.43	4.12E-16
Car2	3.74E-20	1.506399067	0.953	0.193	4.41E-16
Col5a1	2.40E-19	1.512775262	0.744	0.081	2.82E-15
Hbb-b1	2.27E-18	1.690789409	0.953	0.281	2.67E-14
Mt2	3.21E-18	1.756554941	0.884	0.23	3.78E-14
Acta2	2.00E-13	1.876104421	0.744	0.23	2.36E-09
Ptprm	5.17E-10	-2.043837539	0.093	0.637	6.08E-06
Ldb2	6.69E-10	-1.949390326	0.047	0.607	7.87E-06
Calcrl	8.05E-09	-2.289136435	0.023	0.533	9.47E-05
Flt1	9.37E-09	-1.950238015	0	0.511	0.000110211
Zfpm2	7.40E-08	-1.752546398	0.047	0.511	0.000870795
Cdh5	1.40E-07	-1.804651705	0.07	0.511	0.001650544
Mecom	2.21E-07	-1.817886893	0.023	0.467	0.002603218
Elmo1	3.44E-07	-1.671830876	0.023	0.459	0.004042094
Prex2	4.66E-07	-1.515435448	0.07	0.496	0.005479366
Kdr	9.24E-07	-1.582539288	0	0.407	0.010876558
Cluster 4: VE-Cadherin+ YS					
	p-value	**Average logFC**	**Percentage 1**	**Percentage 2**	**Adjusted p-value**
Stab2	6.05E-31	3.609106539	0.933	0.068	7.12E-27
Lyve1	3.17E-19	1.770801443	0.556	0.008	3.72E-15
Afp	3.24E-18	1.878811178	0.889	0.286	3.82E-14
Mrc1	7.84E-17	1.920614638	0.689	0.105	9.23E-13
Calcrl	2.28E-15	1.742748433	0.844	0.263	2.68E-11
Maf	1.27E-13	1.88838544	0.689	0.165	1.49E-09
Myh10	4.50E-12	-1.584283496	0.111	0.737	5.29E-08
Colec12	9.68E-11	1.769468469	0.644	0.218	1.14E-06
Ldb2	3.49E-10	1.63076388	0.756	0.376	4.10E-06
Dab2	8.10E-10	1.692139884	0.667	0.271	9.53E-06
Ralgapa2	1.06E-07	1.562843321	0.556	0.203	0.00124728
Gpc6	4.17E-06	-1.500072019	0.111	0.489	0.0490495

Appendix I: R code used for differential gene expression analysis of bulk RNA sequencing samples

```
## RNAseq analysis for thesis ##
setwd("/Users/morgan/Desktop/R_Analysis/Thesis_25RNA")
getwd()

## Your output files will start with "output" ##
output = "CD44Low"

if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("DESeq2", version = "3.8")
## Libraries ##
library(dplyr)
library(ggplot2)
library("pheatmap")
library("RColorBrewer")
library("DESeq2")

## Upload your RNAseq read matrix to R ##
RNAseq_Data <- read.table("all_data_together.tabular", header = TRUE, sep = "\t",
row.names = "gene")
dim(RNAseq_Data)

## Removed samples that I don't want to use for the analysis ##
RNAseq_Data2 <- RNAseq_Data[ -c(23:25) ]
dim(RNAseq_Data2)

## loaded metadata describing the conditions of the experiment ##
colNames <- read.table("colNames2.txt", header = TRUE, sep = "\t")


### look at summary of data with raw counts ###
(summary_table <- data.frame(summary(RNAseq_Data2)))
summary_table <- dplyr::select(summary_table, Var2, Freq)

## pre-filtering to remove zeros from the dataset ##
## Remove rows with only zeros ##
nrow(RNAseq_Data2)
RNAseq_Data2 <- RNAseq_Data2[rowSums(RNAseq_Data2) > 0,]
nrow(RNAseq_Data2)

## Remove rows with rowsum less or equal than number of samples ##
RNAseq_Data2 <- RNAseq_Data2[rowSums(RNAseq_Data2) >
ncol(RNAseq_Data2),]
nrow(RNAseq_Data2)

## look at summary of data with raw counts again. Medians changed noticeably ##
dim(RNAseq_Data2)
(summary_table = data.frame(summary(RNAseq_Data2)))

## create DESeqDataSet from our matrix with raw counts ##
```

```
dds <- DESeqDataSetFromMatrix(countData = RNAseq_Data2, colData = colNames,
design =~Group)

## to resolve issues with variability and stabilise the variance accross the mean,
there are two options ##
## rlog (regularised logarithm transformation) or vst (variance stabilising
transformation)##
rlog_RNAseq2 <- rlog(dds, blind = FALSE)
head(assay(rlog_RNAseq2), 3)


## Generate a table with normalised log expression values = expression matrix!##
write.table(assay(rlog_RNAseq2), paste0(output, "_rlog_normalized_counts.tsv"),
          quote = FALSE, sep = "\t")
exprMatrix <- as.data.frame(assay(rlog_RNAseq2))

## We can then run the differential expression analysis with the DESeq function ##
## The function will the estimate of size factors controlling for differences in the
sequencing depth of the samples ##
## Then estimate the dispersion values for each gene, and fit a generalized linear
model ##
dds <- DESeq(dds, minReplicatesForReplace = 3)

### plot a PCA graph with DESeq2 package build-in function plotPCA() ###
plotPCA(rlog_RNAseq2, intgroup = c("Group"), ntop = nrow(RNAseq_Data2))
PCAdata <- plotPCA(rlog_RNAseq2, intgroup = c("Group"), ntop =
nrow(RNAseq_Data2), returnData = TRUE)

## Alternative PCA ##
RNAseq_pca <- prcomp(t(exprMatrix))
scores = as.data.frame(RNAseq_pca$x)
scores <- cbind(scores, colNames)
install.packages("ggfortify")
library(ggfortify)
autoplot(RNAseq_pca, data = scores, colour = "Group")
pca3d(RNAseq_pca,group=colNames$Group)

## Tsne analysis ##
install.packages("Rtsne")
library(Rtsne)
RNAseq_tsne <- Rtsne(t(exprMatrix), perplexity = 2)
plot(RNAseq_tsne$Y,col=colNames$Group)
tsne_plot <- data.frame(x = RNAseq_tsne$Y[,1], y = RNAseq_tsne$Y[,2], col =
colNames$Group)
ggplot(tsne_plot) + geom_point(aes(x=x, y=y, color=col))

## Generate Tsne Plot ##
plotTSNE(exprMatrix, scale_features = TRUE, perplexity = 10,
        rand_seed = 7 #this function fixes the TSNE#
        , theme_size = 13) +
  geom_point(aes(colour = colNames$Group), size = I(3)) +
  scale_color_manual(values = c("blue","burlywood3", "cyan2","seagreen3",
"violetred1", "red"))
```

```
## Compare CD44Low/KitNeg Group with CD44Negative for differential expression
analysis ##
res2 <- results(dds, contrast = c("Group", "CD44Low", "CD44Negative"))
mcols(res2, use.names = TRUE)

## Filter out NA results ##
resFilt2 <- res2[!is.na(res2$padj),]
head(resFilt2)
dim(resFilt2)

## Create list of genes with an adjusted p-value of > 0.01 and generate text file ##
resSig2 <- subset(resFilt2, padj < 0.01)
dim(resSig2)
resSig2 <- as.data.frame(resSig2[ order(resSig2$log2FoldChange, decreasing =
TRUE), ])
write.table(resSig2, "~DESeq\\Sig2.txt")

## Mark genes within the data set that are significant so they can be displayed on
plot ##
data2 <- data.frame(gene = row.names(res2), pvalue = -log10(res2$padj), lfc =
(res2$log2FoldChange))
data2 <- na.omit(data2)
data3 <- data2 %>% mutate(significant = ifelse(pvalue > 2, "yes", "no"))
number_yes <- table(data3$significant)
table <- as.data.frame(number_yes)

## Create volcano plot of differentially expressed genes, plotting adjusted p-value
and log fold change ##
volcanoplot2 <- ggplot(data3, aes(x = lfc, y = pvalue, color = significant)) +
  geom_point(size = 2, alpha = 0.7, na.rm = T)
volcanoplot2 + scale_color_manual(values = c("black","goldenrod")) +
  geom_vline(aes(xintercept = 2), linetype = "dashed") +
  geom_vline(aes(xintercept = -2), linetype = "dashed") +
  theme(axis.text = element_text(size = 11, colour = "black"),
        axis.title = element_blank(),
        panel.background = element_rect(fill = NA),
        axis.line = element_line(size = 1, colour = "black"),
        legend.position = "none") +
  scale_x_continuous(limits = c(-15,15), breaks = c(-15, -10, -5, 0, 5, 10, 15)) +
  scale_y_continuous(limits = c(0, 30), breaks = c(0, 5, 10, 15, 20, 25))
```

Appendix II: R code used for single cell RNA sequencing analysis

```
setwd("/Users/morgan/Desktop/R_Analysis/Stab2_Countmat/Seurat")
getwd()

install.packages('Seurat')
library(Seurat)
library(dplyr)
library(gdata)

## Load matrix with count data and metadata file ##
stab2_counts <- as.matrix(read.table("MorganFirst_counts_cells_all.tsv", header =
TRUE, sep = "\t"))
stab2_meta <- read.table("MorganFirst_coldata.tsv", header = TRUE, sep = "\t")

## Look at the parameters of the matrix ##
## We have 200 cells and 15777 detected genes ##
dim(stab2_counts)
summary(colSums(stab2_counts))

## Check the number of genes detected in three or more cells ##
## We have 11,775 genes detected in more than three cells ##
Stab2_counts3 <- apply(stab2_counts, 1, function(x) sum(x>0))
table(Stab2_counts3>=3)

## Create Seurat object ##
stab2_sc <- CreateSeuratObject(raw.data = stab2_counts,
                               min.cells = 3,
                               min.genes = 200,
                               project = "stab2")

## Add meta-data cell group (VE- AGM, VE+ AGM, VE- YS and VE+ YS) to each cell
##
cell_group <- stab2_meta[c("state")]
stab2_sc <- AddMetaData(object = stab2_sc, metadata = cell_group, col.name =
"state" )
head(stab2_sc@meta.data)

## Identify mitochondrial genes in the matrix & remove cells with high percentage ##
## List of mitochondrial genes obtained from mouse mito.carta 2.0 ##
mito <- read.table("Mito_genes.txt", header = TRUE)
mito.genes <- intersect(mito$genes, rownames(x = stab2_sc@data))
length(mito.genes)
percent.mito <- Matrix::colSums(stab2_sc@raw.data[mito.genes, ]) /
Matrix::colSums(stab2_sc@raw.data)
class(percent.mito)
head(stab2_sc@meta.data)
stab2_sc <- AddMetaData(object = stab2_sc,
                        metadata = percent.mito,
                        col.name = "percent.mito")
head(stab2_sc@meta.data)
VlnPlot(object = stab2_sc, features.plot = c("nGene", "nUMI", "percent.mito"), nCol =
3)
```

```
## Look at the relationship between gene numbers and UMI ##
par(mfrow = c(1, 2))
GenePlot(object = stab2_sc, gene1 = "nUMI", gene2 = "percent.mito")
GenePlot(object = stab2_sc, gene1 = "nUMI", gene2 = "nGene")
table(stab2_sc@meta.data$percent.mito < 0.10 &
stab2_sc@meta.data$nGene<3000)

## Filter out the cells that appear to be doublets and with too high percentage of
mitochondrial genes ##
stab2_filter <- FilterCells(object = stab2_sc,
                    subset.names = c("nGene", "percent.mito"),
                    low.thresholds = c(200, -Inf),
                    high.thresholds = c(3000, 0.10))
stab2_filter
## We are left with 178 cells and 11767 genes ##

## Histogram of gene expression before normalisation ##
hist(colSums(stab2_filter@data),
    breaks = 100,
    main = "Total expression before normalisation",
    xlab = "Sum of expression")

## Normalise filtered data with a global scaling method ##
## This normalises the gene expression measurements for each cell by the total
expression ##
## Then we multiply this by a scale factor & log-transform the result ##
stab2_norm <- NormalizeData(object = stab2_filter,
                    normalization.method = "LogNormalize",
                    scale.factor = 10000)

## Histogram of gene expression after normalisation ##
hist(colSums(stab2_norm@data),
    breaks = 100,
    main = "Total expression after normalisation",
    xlab = "Sum of expression")

## Now we identify the highly variable genes ##
## FindVariableGenes calculates the average expression & dispersion for each gene
& places these genes into bins ##
## We can then calculate the z-score fr dispersion within each bin ##
## Macosko et al., ##
stab2_norm <- FindVariableGenes(object = stab2_norm,
                        mean.function = ExpMean,
                        dispersion.function = LogVMR,
                        x.low.cutoff = 0.0125,
                        x.high.cutoff = 3,
                        y.cutoff = 0.5)
head(stab2_norm@var.genes, 20)
length(stab2_norm@var.genes)
## 2848 highly variable genes were identified ##

## We can then scale the data to remove unwanted sources of variation such as
batch effects, cell cycle effects ect.. ##
stab2_norm <- ScaleData(object = stab2_norm,
```

```
                vars.to.regress = c("nUMI", "percent.mito"))


## We can then run PCA analysis on the highly variable genes ##
stab2_PCA <- RunPCA(object = stab2_norm,
                pc.genes = stab2_norm@var.genes,
                do.print = TRUE,
                pcs.print = 1:5,
                genes.print = 5, rev.pca = TRUE, weight.by.var = TRUE)
PCAPlot(object = stab2_PCA, dim.1 = 1, dim.2 = 2)
PCHeatmap(object = stab2_PCA,
        pc.use = 1:12,
        do.balanced = TRUE,
        label.columns = TRUE)

## To determine which PCAs are significant or relevant to feed into down-stream
analysis ##
PCElbowPlot(object = stab2_PCA)
stab2_PCA <- JackStraw(object = stab2_PCA, num.replicate = 100, display.progress
= FALSE)
JackStrawPlot(object = stab2_PCA, PCs = 1:20)

## To generate clusters based on the signficant PCA components and changing the
resolution to alter the number of clusters ##
stab2_clusters <- FindClusters(object = stab2_PCA, reduction.type = "pca", dims.use
= 1:9,
                resolution = 0.2, print.output = 0, save.SNN = TRUE)

## This sets the identity to cell state so that the Tsne plot will be coloured based on
your groups ##
stab2_clusters <- SetAllIdent(object = stab2_clusters, id = "state")

## Running the Tsne algorithm, you can change the perplexity ##
stab2_tsne <- RunTSNE(object = stab2_clusters, seed.use = 42, add.iter = 5000,
dims.use = 1:9, perplexity = 12)

## Plotting the Tsne ##
TSNEPlot(object = stab2_tsne, pt.size = 4,
       colors.use = (c("VE- AGM" = "cyan3", "VE- YS" = "peachpuff", "VE+ AGM" =
"cornflowerblue", "VE+ YS" = "salmon"))) +
  theme(axis.text = element_text(size = 11, colour = "black")) + labs(x="Dimension 1",
y="Dimension 2")
save_plot("Seurat_TSNE.pdf", p1, base_height = 10, base_width = 10)
## You can visualise the number of groups in your data ##
print(stab2_clusters@meta.data)

## Tsne plots with individual gene expression ##
FeaturePlot(object = stab2_tsne,
          features.plot = c("Cdh5"),
          cols.use = c("grey", "blue"),
          reduction.use = "tsne", pt.size = 4, no.legend = FALSE, no.axes = FALSE) +
background_grid()

## Finding marker genes for each cluster ##
```

```
cluster1.markers <- FindMarkers(object = stab2_clusters, ident.1 = "VE- AGM",
min.pct = 0.1, logfc.threshold = 1.5)
cluster2.markers <- FindMarkers(object = stab2_clusters, ident.1 = "VE+ AGM",
min.pct = 0.1, logfc.threshold = 1.5)
cluster3.markers <- FindMarkers(object = stab2_clusters, ident.1 = "VE- YS", min.pct
= 0.1, logfc.threshold = 1.5)
cluster4.markers <- FindMarkers(object = stab2_clusters, ident.1 = "VE+ YS", min.pct
= 0.1, logfc.threshold = 1.5)

## Writing the files ##
write.csv(cluster1.markers, file = "VE-AGM_markers_logfc1.5.csv")
write.csv(cluster2.markers, file = "VE+AGM_markers_logfc1.5.csv")
write.csv(cluster3.markers, file = "VE-YS_markers_logfc1.5.csv")
write.csv(cluster4.markers, file = "VE+YS_markers_logfc1.5.csv")

## I combined the files manually and uploaded the marker gene list ##
Marker_genes <- read.table(as.matrix("Marker_genes.txt"))

## Extracting the expression matrix from the Seurat object and creating a csv file ##
data_to_write_out <- as.data.frame(as.matrix(stab2_norm@data))
write.csv(data_to_write_out, file = "Seurat_expressionMat.csv")

## Extracting the metadata eg: Cell state for your set of filtered and normalised cells
##
meta_data_norm <- as.data.frame(stab2_norm@meta.data)
write.csv(meta_data_norm, file = "Seurat_metadata.csv")

## Create Heatmap of Marker genes ##
expressionTable <- read.xls("Seurat_Markers_HM.xlsx",sheet=1,
stringsAsFactors=FALSE)
exprID <- expressionTable[,1]
exprMatrix <- as.matrix(expressionTable[,-1])
rownames(exprMatrix) <- exprID
samples <- read.xls("Seurat_Markers_HM.xlsx", sheet=2, stringsAsFactors=FALSE)
rownames_samp = make.names(samples$SampleID, unique=TRUE)
samples = data.frame(GroupID = samples$GroupID)
rownames(samples) = rownames_samp
rm(rownames_samp)
stopifnot( all(rownames(samples) == colnames(exprMatrix))  )
genes <- read.xls("Seurat_Markers_HM.xlsx", sheet=3, stringsAsFactors=FALSE)
rownames(genes) <- genes$GeneID
stopifnot( all( exprID == rownames(genes) ) )
phenoData <- new("AnnotatedDataFrame", samples)
featureData <- new("AnnotatedDataFrame", genes)
eS <- new("ExpressionSet", expr=exprMatrix,featureData=featureData,
phenoData=phenoData)
library(pheatmap)
library(RColorBrewer)
hmcol <- colorRampPalette(brewer.pal(10, "RdYlBu"))(100)
ann_colors = list (GroupID = c("VE- AGM" = "cyan3", "VE- YS" = "peachpuff", "VE+
AGM" = "cornflowerblue", "VE+ YS" = "salmon"))
pheatmap(mat=exprs(eS),annotation_col=pData(eS), color= rev(hmcol),
annotation_colors = ann_colors, border_color = "white", cluster_cols = TRUE,
cluster_rows = TRUE, cellwidth = NA, cellheight = NA, fontsize = 11, show_colnames
= FALSE)
```

Appendix III: Title and abstract for manuscript available on BioRxiv

Single-cell transcriptomics identifies CD44 as a new marker and regulator of haematopoietic stem cell development

Morgan Oatley[1*], Özge Vargel Bölükbası[1,2*], Valentine Svensson[3,4,5], Maya Shvartsman[1], Kerstin Ganter[1], Katharina Zirngibl[6], Polina V. Pavlovich[1,7], Vladislava Milchevskaya[6,8], Vladimira Foteva[1], Kedar N. Natarajan[3,9], Bianka Baying[10], Vladimir Benes[10], Kiran R. Patil[6], Sarah A. Teichmann[3] & Christophe Lancrin[1,11]

[1] European Molecular Biology Laboratory, EMBL Rome - Epigenetics and Neurobiology Unit, Monterotondo, Italy.

[2] Current address: Boston Children's Hospital/Harvard Medical School, Boston, USA.

[3] Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, UK.

[4] European Molecular Biology Laboratory, EMBL-EBI, Wellcome Genome Campus, Hinxton, UK.

[5] Current address: Pachter Lab, California Institute of Technology, California, USA.

[6] European Molecular Biology Laboratory, Structural and Computational Biology Unit, Heidelberg, Germany

[7] Moscow Institute of Physics and Technology, Institutskii Per. 9, Moscow Region, Dolgoprudny 141700, Russia.

[8] Current address: Institut für Medizinische Statistik und Bioinformatik, Köln, Germany

[9] Current address: The University of Southern Denmark, Danish Institute for Advanced Study, Department of Biochemistry and Molecular Biology, Odense, Denmark.

[10] European Molecular Biology Laboratory, Genomics Core Facility, Heidelberg, Germany.

[11] Correspondence: christophe.lancrin@embl.it

[*] Co-first authors

Abstract

The endothelial to haematopoietic transition (EHT) is the process whereby haemogenic endothelium differentiates into haematopoietic stem and progenitor cells (HSPCs). The intermediary steps of this process are unclear, in particular the identity of endothelial cells that give rise to HSPCs is unknown. Using single-cell transcriptome analysis and antibody screening we identified CD44 as a new marker of EHT enabling us to isolate robustly the different stages of EHT in the aorta gonad mesonephros (AGM) region. This allowed us to provide a very detailed phenotypical and transcriptional profile for haemogenic endothelial cells, characterising them with high expression of genes related to Notch signalling, TGFbeta/BMP antagonists (Smad6, Smad7 and Bmper) and a downregulation of genes related to glycolysis and the TCA cycle. Moreover, we demonstrated that by inhibiting the interaction between CD44 and its ligand hyaluronan we could block EHT, identifying a new regulator of HSPC development.