Molecular control of the Endothelial to Hematopoietic Transition

Hanane Khoury

Sorbonne Université

Ecole Doctorale Complexité du vivant (ED515)

LBD UMR CNRS 7622, INSERM U1156

Equipe migration et différentiation des cellules souches hématopoïétiques

Molecular control of the Endothelial to Hematopoietic Transition

Contrôle moléculaire de la Transition Endothélio-Hématopoïétique

Par Hanane KHOURY

Thèse de doctorat de Biologie moléculaire et cellulaire, Biologie de développement et cellules souches

Dirigée par Thierry JAFFREDO

Présentée et soutenue publiquement le 24 Septembre 2019

Devant un jury composé de :

Mme. SOBCZAK-THEPOT Joëlle	Professeur	Président
Mme. BIGAS Anna	Group leader	Rapporteur
Mr. UZAN George	Directeur de recherche	Rapporteur
Mme. SCHMIDT Anne	Chargé de recherche	Examinateur
Mme. OBERLIN Estelle	Chargé de recherche	Examinateur
Mr. LACAUD Georges	Professeur	Examinateur
Mr. CHARBORD Pierre	Directeur de recherche émérite	Membre invité
Mr. JAFFREDO Thierry	Directeur de recherche	Directeur de thèse

« Do not follow where the path may lead.

Go instead where there is no path and leave a trail »

- Ralph Waldo Emerson

À vous mes chers parents je dédie cette thèse…

ACKNOWLEDGEMENT

I would like to acknowledge here the people who, from near and far, have contributed to the realization of this Ph.D. work.

First and foremost, I would like to express my sincere gratitude to all the jury members for accepting to evaluate my thesis and for their precious time especially during this time of the year. I would like to thank Dr. ***Anna BIGAS*** and Dr. ***Georges UZAN*** for evaluating my manuscripts, and for their precious comments and suggestions. My sincere thanks also go to Dr. ***Anne SCHMIDT***, Dr. ***Estelle OBERLIN***, Pr. ***Joelle SOBCZAK THEPOT***, for accepting to evaluate my thesis. And finally, I would like to thank Dr. ***Georges LACAUD*** for accepting to be part of my PhD committee before being part of my PhD jury today. Thank you so much for all the advices and support during these past four years and for all the scientific exchanges.

I would like to express my sincere gratitude to Dr. ***Thierry JAFFREDO*** who welcomed me in his lab, supervised and mentored my work, for his attention, his advices which were really useful and appreciated. For five years he helped me to progress and to learn a lot in embryology and hematopoiesis field. For all the time you awarded to me, Thierry, Thank you.

I am extremely thankful to Dr. ***Pierre CHARBORD*** for his guidance, his support and all he taught me during my thesis especially for my initiation to bioinformatics and monitoring all my analysis. I would never been able to accomplish this work without you, Pierre, Thank you.

My thanks also go to ***Rodolphe GAUTIER*** who helped me with the qPCR part and worked with me. Thank you Rodolphe for your help and your good mood.

I am also thankful to ***Charles DURAND*** who I met during my Master 1 and motivates me to enter the Stem Cell world especially the hematopoiesis world. Charles, thank you.

To ***Laurence PETIT***, ***Pierre Yves CANTO***, ***Isabelle CACCIAPUOTI*****,** thank you for everything I learned from you and for all the discussions we had during my stay in the lab.

You know you are never alone when you have in your life someone like ***Viviane PELTIER***, it is never easy to be far from home, but Viviane made it easy for me. Vivi you were like my mother during these 5 years, always here, taking care of me, discussing with me, every problem seems easy to solve when discussed with you. Thank you for literally everything, thank you for each lunch, each coffee, each discussion we shared.

Every person has a guardian angel somewhere, I am more than confident that ***Cecile DREVON*** were mine starting my very first day in the lab. Cecile, you believed in me and always pushed me forward, always told me to pursue my dream and to have faith in myself. It was hard to see you leaving 2 years ago but I knew that will not change the fact that you are always here for me. I promise you even if my journey in this lab has finished, we will always maintain this special relationship and you will always be my guardian angel.

To ***Gregoire STIK*** with whom I shared the office for almost 2 years and above all I shared great memories and great scientific exchanges, thank you Greg. I also owe a lot to all the people who were part of my daily life at one time or another and without whom these years wouldn't be the same, all the people of the Developmental Biology Laboratory especially ***Sylvie SCHNEIDER MANNOURY*** who welcomed me in the unit and was always ready to answer my needs. I am thinking also about all the master students I worked with especially ***Lea***, ***Marietta*** and ***Shako and Olivera*** with whom I shared the office for 6 months. As well as our neighbours of the 6th floor: ***Delphine, Marie Ange, Cedrine, Claire, Emmanuelle***, ***Estelle, Berbang*** and the two amazing friends who always gave me strength **Joanna** and **Sonya** thank you. My special thanks also go to ***Sophie GOURNET*** who initiated me to Photoshop and was always ready to help me presenting my results using her great graphic talent. I have no choice but to mention a special person who followed my evolution starting my Master 1, ***Sophie LOUVET***, thank you for every word you said to recomfort me, for every advice and every "Bonjour souriant", you are the kind of teacher and colleague that always stay in mind.

There is two persons who enlighten my stay in the lab and who truly became more than friends and whom I can call without hesitation "My Lebanese", ***Keirththana*** and ***Ludovic***. To all the selfies we took, to the Christmas trees we set up to all the

coffee breaks and the laugh we had, to each time we mutually supported each other, to the different culture we exchanged: I am more than grateful! Thank you Kiki et Ludo.

To all the ***PhD students*** I represented in the UMR 7622 laboratory council, I hope I was able to meet your expectations and demands. It was a pleasure for me to organize the «Club des Docs» of the UMR7622 and to meet each and every one of you.

If I could finish this thesis, it was also thanks to ***my friends*** and the outside activities that were there to change my ideas. I am thinking about all the great meetings and friendships I have made during these six years in Paris as well as my childhood and university friends back home. I will not list everyone but be sure that each one of you my friends helped me somehow to achieve my goal.

There are two peoples who left a special print during this journey, ***Lama*** meeting you was one of the best thing that happened to me, you were a source of inspiration and of strength thank you ! and ***Sandy*** you were always here when I needed you, listening to me, advising me, helping me every time you could, thank you!

Wassim, thank you for your creativity and your help in defining each chapter with a wonderful illustration.

Well now the most important part has arrived, I don't know if I will be able to express what I am feeling for my two families. ***Bassam*** and ***Rebecca*** without you I would never ever think about moving to France. Bassam you made everything easier, you believed in me and helped me achieve my dream. Rebecca my travel partner being around you always gave me kind of positive energy. Thank you so much.

What can I say to you? ***Rita, Remie*** and ***Nadine*** my three sweet sisters and my brothers in law (***Ziad, Rawad*** and ***Georges***). Sharing with you the happiest events in your life these past years (engagement, marriage, parenthood) made me the luckiest sister ever. Thank you for everything, but foremost thank you for giving me the best niece and nephews in the world, to ***Chris, Rhea and Tony***, I am the most delighted aunt ever, sometimes it was enough looking to some of your pictures or videos to cheer me up.

Maroun, my caring brother, the best thing that happened to me was you moving to Paris and living together. You have filled my life, we shared moments of joy,

happiness, craziness. Thank you for being my backbone, my support, my adviser. Thank you for supporting me these last days, I know I was terrible and always upset and you always knew how to comfort me up and to help me. Thank you!

Maman et ***papa***, je mets entre vos mains, le fruit de longues années d'études, de longues années de distance de votre amour de votre tendresse et de longs jours d'apprentissage. Loin de vous, votre soutien et votre encouragement m'ont toujours donné de la force pour persévérer et pour prospérer dans la vie. Chaque ligne de cette thèse chaque mot et chaque lettre vous exprime ma reconnaissance, mon respect, mon estime et le merci d'être mes parents. Je suis la fille la plus chanceuse au monde pour vous avoir comme parents, je vous aime très fort !

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION

AcLDL-DiI: 1,10 -dioctadecyl-3,3,30,30-tetramethyl indocarbocyanine perchlorate-labeled human acetylated low-density lipo- protein

AGM: Aorta Gonad Mesonephros

AML1: Acute Myeloid Leukemia 1

BFU-E: Burst Forming Unit- Erythroid

BL-CFC: Blast Colony-Forming Cell

BM: Bone Marrow

CA-FC: Cobblestone Area-Forming Cell

CBFA2: Core-Binding Factor subunit Alpha-2

CBP: Creb Binding Protein

CFC: Colony Forming Cell assay

CFU- E/Mk: Colony Forming Unit-Erythroid/ Megakaryocyte

CFU-B: Colony Forming Unit-Lymphocyte B

CFU-E: Colony Forming Unit-Erythroid

CFU-G: Colony Forming Unit-Granulocyte

CFU-GEMM: Colony-Forming Unit Granulocyte-Erythroid-Macrophage-Megakaryocyte

CFU-GM: Colony Forming Unit-Granulocyte/ Macrophage

CFU-M: Colony Forming Unit-Monocyte/Macrophage

CFU-Mk: Colony Forming Unit-Megakaryocyte

CFU-S: Colony Forming Unit in the Spleen

CHT: Caudal Hematopoietic Tissue

CLP: Common Lymphoid Progenitor

CMP: Common Myeloid Progenitor

CSFs: Colony Stimulating Factors

DEGs: Differentially expressed genes

DLP: Dorsal Lateral Plate

E: Embryonic day

EBs: Embryonic Bodies

ECM: Extracellular Matrix

ECs: Endothelial Cells

EHT: Endothelial-to-Hematopoietic Transition

ELTC-IC: Extended Long-Term Culture-Initiating Cell

EMP: Erythro-Myeloid Progenitors

EMT: Epithelial-mesenchymal transition

EryP-CFC: Primitive Erythroid- Colony Forming Cell

ESC: Embryonic Stem Cell

ESC: Embryonic Stem Cell

FACS: Fluorescent Activating Cell Sorting

FL: Fetal Liver

FLK: Fetal Liver Kinase

GMP: Granulocyte and Monocyte Progenitor

h.p.f: hours post fertilization

HC: Hematopoietic Cell

HCT: Hematopoietic Cell Transplantation

hESC: human ESC

HPP-CFCs: High Proliferative Potential-Colony Forming Cells

HSC: Hematopoietic Stem Cell

HSPC: Hematopoietic Stem and Progenitor Cell

IAHCs: Intra-Aortic Hematopoietic cell Clusters

ICM: Inner Cell Mass

IPSC: Induced Pluripotent Stem Cell

KI: Knock In

LDL: Low-Density Lipoprotein

LMO2: LIM Only domain 2

LP: Lateral Plate

LTC-IC: Long-Term Culture-Initiating Cell

LT-HSC: Long-Term HSC

MEP: Megakaryocyte and Erythrocyte Progenitor

MET: mesenchymal-Epithelial transition

MPP: Multipotent Progenitor

MT-MMPs: Membrane-Type Matrix Metalloproteinases

NK: Natural Killer Cells

POFUT: GDP-fucose protein O-fucosyltransferase

PSM: Pre-Somitic Mesoderm

P-Sp: Para-aortic-Splanchnopleura

RBI: Rostral Blood Island

SC: Stem Cell

Sca: Stem Cell Antigen

SDF: Stromal Cell-Derived Factor

SLAM: Signaling Lymphocyte Activation Molecule

Sp: Splanchnopleure

SRY: Sex Determining Region Y

ST-HSC: Short-Term HSC

TSR: Thrombospondin Repeats

VBI: Ventral Blood Island

VEGFR: Vascular Endothelial Growth Factor Receptor

YS: Yolk Sac

PREAMBLE

Hematopoietic Stem Cells (HSCs) are multipotent stem cells, sitting at the top of a complex hierarchy and are capable, through complex subtlety regulated pathways, of giving birth to the multi-lineage progenitors. Those progenitors have enhanced proliferative capacities and give rise to committed progenitors, which will differentiate into specific blood cell types. Normal hematopoiesis is required to constantly generate new mature blood cells, allowing, for example, the turnover of erythrocytes, and to preserve the HSC pool.

A small disruption of this subtle equilibrium leads to hematological disorders, like leukemia, multiple myeloma, or non-Hodgkin's lymphoma. In these diseases, the HSC pool homeostasis is severely affected, and the consequences can be lethal. Bone marrow and peripheral blood stem cell transplants are the gold standard when it comes to treating blood disorders. The injection of allogenic HSCs allows the patients to restore their hematopoietic system after radiation therapy or chemotherapy. As many as 50,000 transplants are performed each year, therefore, their improvement is of crucial significance to contemporary medicine. It is undeniable that this strategy has improved over the years, with the increase of donors, development of alternative sources to the bone marrow, such as the cord blood, or improvement of the conditioning methods. Nevertheless, HSCs transplant still presents flaws. HSCs are difficult to harvest (HSCs represent only 0,01 to 0,1% of the bone marrow cell population), and complications can occur, such as engraftment failure, graft-versus-host disease, and delayed reconstitution. Therefore, theoretical and applied researches need to be conducted to solve this public health matter problem.

Throughout an adult lifetime, HSCs reside in the bone marrow. Although several waves of hematopoietic emergence occur early and transiently in multiple sites during embryonic development, only one site produces bona fide HSCs. The first two waves of blood cells emergence are initiated in the extra-embryonic yolk sac and generate mainly primitive erythrocytes and then erythroid-myeloid progenitors. Adult blood, including HSCs, originates from a next wave that occurs in the aorta-gonad-mesonephros (AGM) region. It was shown that HSCs emerge from clusters in the ventral floor of the dorsal aorta, from a specific type of endothelium named the hemogenic endothelium HE. They lose their endothelial characteristics and acquire hematopoietic traits, in a process called the endothelial-to-hematopoietic transition (EHT). The molecular mechanisms controlling EHT and the specification and

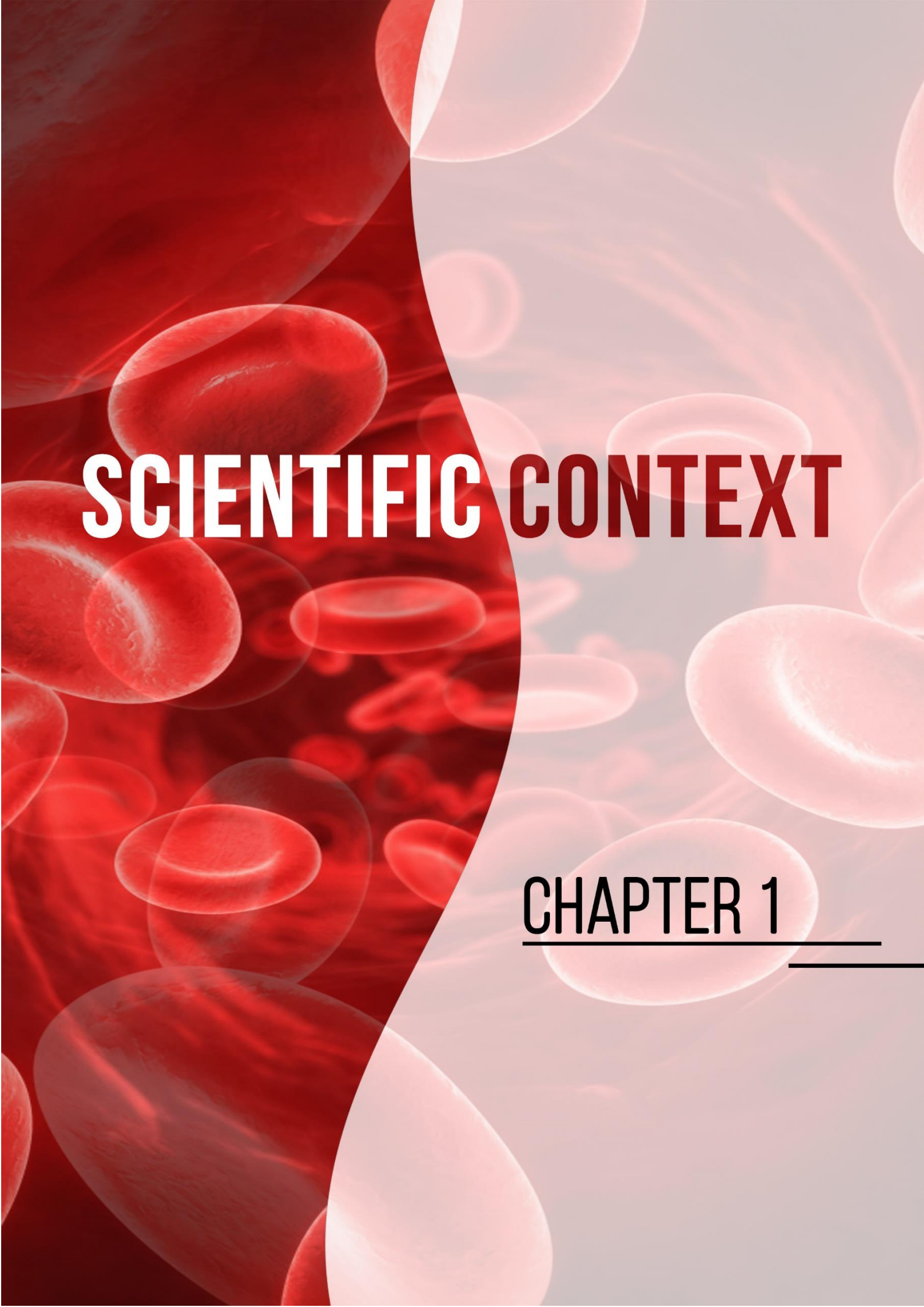
emergence of HSCs are still poorly understood. Different systems are utilized to disclose those factors. Embryonic (ESCs) and induced (IPSCs) pluripotent stem cells are widely used to study how to orient non-committed cells towards the hematopoietic lineage. One of the advantages of these cells is that they can be genetically modified to study the precise targets involved in hematopoietic development. On the other hand, a different approach consists of culturing immature cells in strictly defined culture media and trying to produce hematopoietic cells and eventually HSCs. Characterization of hematopoietic cells is achieved with surface markers, analysis of specific expression markers and functional tests. However, this identification is always retrospective.

Those approaches led to a better understanding of the EHT (e.g. role of Runx1, GFI1-GFI1B, GATA2…), but these models do not recapitulate all the steps of development, and the yield of HEC is very low. In spite of those studies, it is still impossible to produce HSCs de novo or amplify them ex vivo. Even though some notable advances were made in which cocktails of transcription factors are used to generate hemogenic ECs directly from differentiated cells through direct reprogramming (Sugimura et al., 2017, Lis et al., 2017). This brake needs to be overcome, and thus a better understanding of cellular and molecular mechanisms controlling HSCs emergence and expansion is required to develop future therapeutic approaches.

In the lab, we developed an ex vivo model recapitulating the different steps of mesoderm commitment towards hematopoietic production, including EHT. This model uses uncommitted pre-somitic quail mesoderm (PSM) that, following culture in a specific media, will orient the cells towards the endothelial and hemogenic endothelial commitment. Hemogenic endothelial cells (HECs) commitment is testified by the expression of the transcription factor Runx1. HECs undergo an endothelial-to-hematopoietic transition, losing their endothelial features and acquiring a hematopoietic phenotype, producing HSCs (Yvernogeau et al., 2016). The progression through the different phenotypes follows a precise calendar, characterized by the expression of specific genes and markers.

In this study, we investigated the molecular mechanisms controlling EHT in this culture system. To do so we generated transcriptomic data for each population subsets (mesodermal cells, endothelial cells, hemogenic endothelial cells, and

hematopoietic cells), and bioinformatics analyses were carried out to identify key genes and pathways regulating the transition from endothelial cells to hematopoietic production. For each population, differentially expressed genes were selected (up or down-regulated genes). WGCNA (data mining method) was performed using the matrix of specific up-regulated genes, it revealed the existence of 5 modules, 1 for EC, 1 for M, 1 for HC and 2 for HEC. In the present work, I focused on studying the two networks of HEC. Based on the list of genes positively correlated to the HEC state obtained following WGCNA analysis and the measure of gene connectivity, we selected several candidate genes shown as hubs in our networks and which we chose to functionally validate in our culture system. We could demonstrate a prominent role of two novel genes POFUT2 and CREB3L1 (on going) implicated in this transition. Furthermore, we validated our culture system by testing two pathways known to play a crucial role in the EHT Wnt and Notch. We showed that they are implicated in the regulation of the EHT in our system.



SCIENTIFIC CONTEXT

CHAPTER 1

I. Hematopoietic stem cells: intriguing yet remarkable cells

1. Generalities about hematopoiesis

Blood is a mesenchymal tissue composed of cells and an aqueous fluid known as plasma which allows cells to circulate continuously in all the vessels of the body. Two major functions of the blood include transporting substances to and from our cells and providing immunity and protection against infectious agents such as bacteria and viruses.

Blood consists of several elements and represents approximately 8% of total body weight. The major components include plasma (55%) and cellular elements (red blood cells, white blood cells, and platelets: 45%) Figure 1. Those cells named hematopoietic cells (HCs) are generated during a process called hematopoiesis and possess, depending on their type, special functions. The red blood cells (erythrocytes) ensure the transport of oxygen. The white blood cells (leukocytes) play an important role in the immune system and lymphatic system by defending the body against infection. The platelets (thrombocytes) circulate through the bloodstream and play a major role in clotting.

Mature blood cells have varying life spans. Red blood cells circulate for about 4 months, platelets for about 9 days, and white blood cells range from about a few hours to several days. When oxygen in tissues is low, the body responds by stimulating the bone marrow (BM) to produce more red blood cells. When the body is infected, more white blood cells are produced.

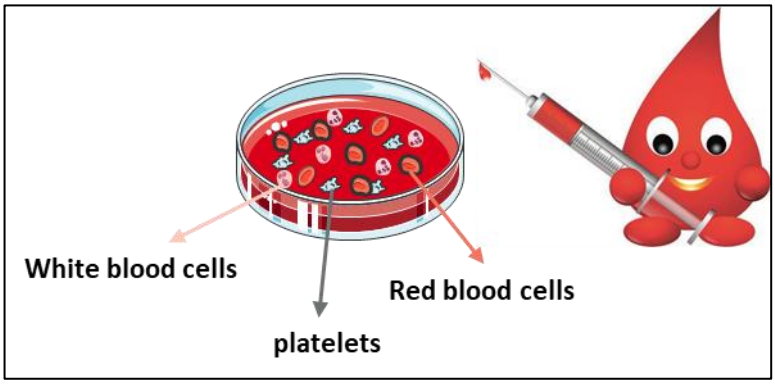


Figure 1: Schematic representation of blood cell constituents
The major blood components include plasma (55%) and cellular elements (red blood cells, white blood cells, and platelets*: 45%)*

The turnover of cells of the hematopoietic system may be estimated to be close to 1 trillion cells per day (Ogawa, 1993). This remarkable cell renewal process is supported by a small population of BM cells termed hematopoietic stem cells HSCs.

2. Stem Cells SCs and HSCs

A. Self-renewal and multipotency

There are four types of SCs: totipotent, pluripotent, multipotent and unipotent. They appear during development and some are also present in the adult. Totipotent SCs are perhaps the most versatile of the SC types. A totipotent fertilized egg has the potential to give rise to virtually all human cells, such as nerve, heart, skin... including gametes and all the embryonic appendages. So far totipotency comprises the zygote and extends to some early cleavage-stage blastomeres; the precise early-stage depends on the species. Shortly after the early cleavage stage, totipotency is lost and is replaced by a second SC state i.e., pluripotency. In the mouse embryo, pluripotency occurs during the first fourth days of development. Pluripotent SCs can do everything the totipotent ones except that they are not able to give rise to the placenta. Trophectoderm and inner cell mass fates separate early due to the expression of the transcription factor Cdx2. During development, the pluripotent SC state is limited to the inner cell mass stage. Multipotent SCs reside in most, if not all the tissues. They are limited in numbers and their role is to replenish the organs or tissues when cells are dying. Examples of multipotent SCs are neural SCs that give rise to different types of neural cells and glia, intestinal SCs that renew the intestinal crypts or hematopoietic SCs which can give rise to different blood cell types. Finally, a unipotent SC refers to a cell that can differentiate along only one lineage. The word 'uni' itself is derived from the Latin word 'unus,' meaning one. Found in adult tissues, a unipotent SC, in comparison with other types of SCs, has the lowest differentiation potential. A good example is the epidermis SC that gives rise only to keratinocytes.

Two modes of SC divisions are reported in the literature. The asymmetric division, in which a SC produces one differentiated cell and one SC. This suggests that the homeostatic control of the SC pool is maintained at the level of single cells. The symmetric division model also starts from a SC and produces two differentiated cells or two SCs. In this context, homeostasis is controlled at the population level, rather than at the individual cell level (Shahriyari and Komarova, 2013). Overall, the

number of adult SCs within a tissue is constant, because of the balance between symmetric and asymmetric divisions Figure 2.

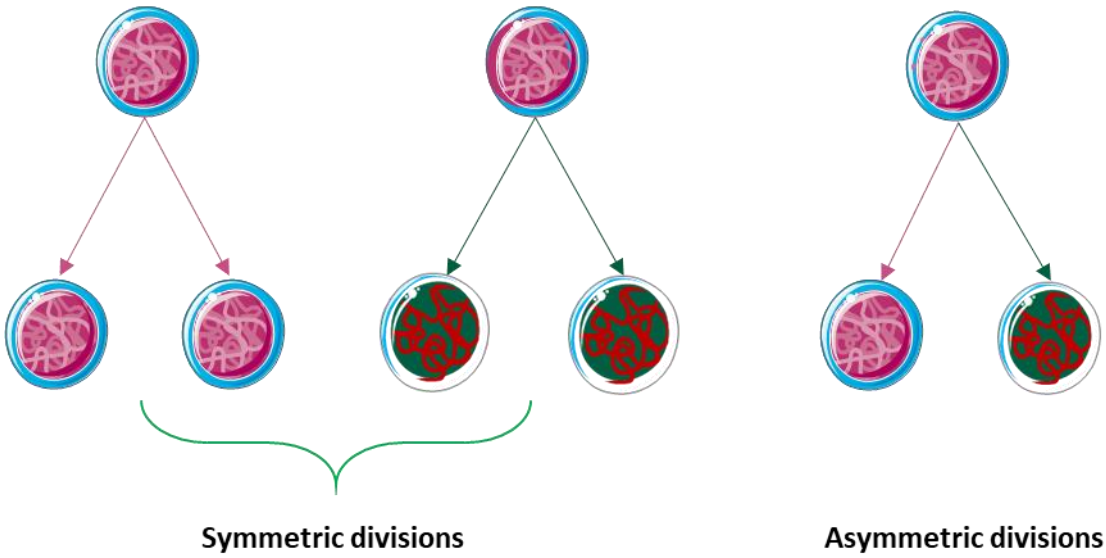


Figure 2: Symmetric and asymmetric divisions schematic representation
Symmetric divisions start from a SC and produces two differentiated cells or two SCs
Asymmetric division starts from a SC and produces one differentiated cell and one SC
Pink cells: Stem cells
Green cells: Differentiated cells

In most tissues, SCs are rare and reside in a specific microenvironment named niche. Niche cells regulate the signals sent to SCs by the surrounding tissues and maintain their two specific SC properties: self-renewal and multipotency. Like all SCs, HSCs are able to differentiate into a series of more and more committed progenitors that will proliferate to produce precursors, engaged into specific lineages to finally give rise to all mature blood cells such as lymphoid, myeloid, and erythroid cells. The hematopoietic system is a complex system formed of a progression of different cell types each defined by specific self-renewal, multilineage capacities, and maturity degrees. HCs are organized into different compartments according to a pyramidal hierarchy, starting from HSCs and ending in all differentiated mature blood cells, through progenitors and hematopoietic precursors. Progenitors are more specific than a SC and are pushed to differentiate into their "target" cell. The most important difference between SCs and progenitor cells is that SCs can replicate indefinitely, whereas progenitor cells can divide only a limited number of times. While precursors are a partially differentiated cell, usually referred to as a unipotent cell that has lost most of its SC properties.

HSCs sit at the apex of the tree with their capacity of long-term self-renewal and multilineage differentiation. They are very rare and represent 0.05% of the total mononucleated cells of the BM. Those cells are divided based on their capacity to self-renew into Long-Term HSCs (LT-HSCs) and Short-Term HSCs (ST-HSCs). The use of the classical technologies (in vitro colony-forming cell assays, other clonal assays and transient in vivo repopulation) led to the identification of multipotent progenitor MPP, which is still able to produce all mature blood cells but lost the capacity to self-renew (Adolfsson et al., 2001). MPP produces two types of progenitors, the Common Myeloid Progenitor CMP (Akashi et al., 2000) and the Common Lymphoid Progenitor CLP (Hao, 2001; Kondo et al., 1997). CMPs and CLPs lose the multipotency and generate successively the myeloid and lymphoid lineages. CMPs produce 2 types of progenitors: GMPs (granulocyte and monocyte progenitors) and MEPs (megakaryocyte and erythrocyte progenitors) whereas CLPs differentiate into B and T Lymphocytes and Natural Killer (NK) cells. Regarding the dendritic cells, it has been shown according to their phenotype, that they can derive from CMP or CLP as well (Manz, 2001; Manz et al., 2001; Traver et al., 2000). In the embryo as in the adult, HSCs are the only HCs able to self-renew. However, unlike the embryonic HSCs, adult HSCs have low proliferative potential (Abkowitz, 2002; Gordon, 2002) and the maintenance of mature blood cells is insured by the progenitors and hematopoietic precursors.

The study of HSC differentiation which allowed the generation of this hierarchical pyramid Figure 3, has relied mainly on ex vivo colony assays and the transplantation of hematopoietic stem and progenitor cells (HSPCs) into irradiated animals.

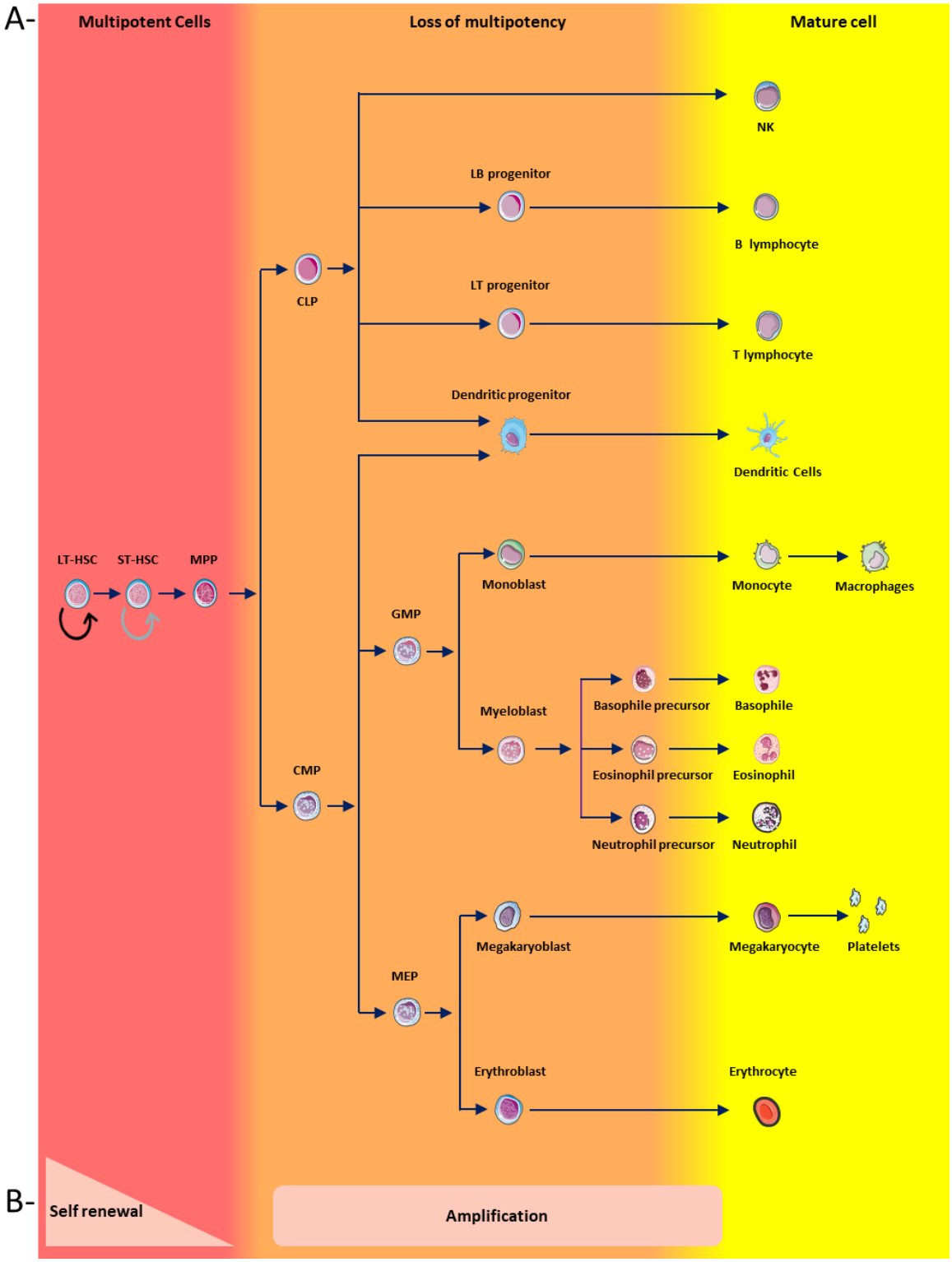


Figure 3: Schematic representation of the hierarchy of hematopoietic cells

A- Multipotency degree

B- Self-renewal and amplification degree

HSCs sit at the apex of the tree with their capacity of long-term self-renewal and multilineage differentiation. They give rise to multipotent progenitor MPP, which is still able to produce all mature blood cells but lost the capacity to self-renew. MPPs produce two types of progenitor CMP and CLP which lose the multipotency and generate successively the myeloid and lymphoid lineages.

Abbreviations: HSC: Hematopoietic Stem Cell, LT-HSC: Long Term HSC, ST-HSC: Short-Term HSC, MPP: Multipotent Progenitor, CMP: Common Myeloid Progenitor, CLP: Common Lymphoid Progenitor, GMP: granulocyte and monocyte progenitors, MEP: megakaryocyte and erythrocyte progenitors, LB: Lymphocyte B, LT: Lymphocyte T, NK: Natural Killer

In all proposed hierarchies, hematopoiesis originates in multipotent repopulating HSCs, and lineage development occurs via discrete and increasingly lineage-restricted progenitor populations. However, as predominantly purified cell populations were analyzed as opposed to individual cells, functional heterogeneity within the identified populations might have been underestimated. In addition, although transplantation of purified prospective HSC populations into irradiated recipient mice is a powerful assay to evaluate repopulating HSCs at the single-cell level, it may not be able to de-convolute HSC heterogeneity and function in steady-state hematopoiesis. Recently, several studies have explored emerging single-cell technologies that have the potential to overcome many of these limitations. Single-cell whole transcriptome analysis has allowed molecular heterogeneity to be systematically explored across the hematopoietic system. In addition, in vivo barcoding has emerged as a means to study the fate and contribution of individual stem and progenitor cells within larger populations, including their roles in unperturbed hematopoiesis. Together, these innovative single-cell technologies promise to considerably advance our understanding of hematopoiesis. However, as for the more classical single-cell technologies, recent single-cell approaches have technical and conceptual limits that need to be considered when interpreting the results, they generate (Jacobsen and Nerlov, 2019).

Concerning new genetic fate mapping, clonal marking techniques and single-cell analysis, they shed light on hematopoiesis under physiological conditions. The inducible lineage tracing of cells emerging from HSCs is assessed by using mice expressing a dependent version of Cre recombinase from different loci: Scl, Runx1, Hoxb4, Fgd5… or by using Knock-In (KI) mouse expressing Tie2 locus Cre recombinase flanked on both ends by a modified estrogen receptor domain. This system combined with limiting dilution and mathematical modeling revealed quantitative properties of hematopoiesis from SCs in the BM (Busch and Rodewald, 2016; Busch et al., 2015; Henninger et al., 2017; Maetzig et al., 2017; Upadhaya et al., 2018).

Together the *ex vivo* and *in vivo* analysis of hematopoiesis allowed a new global and more precise overview of the implications of each component, HSCs progenitors and precursors, in producing the mature blood cells.

B. The balance of the number of HSCs

The number of HSCs in the BM is roughly constant (1-4 x 10^5 mononuclear cells, about 150 are solicited each day) however how the balance between self-renewal and quiescence is controlled is not totally understood. The number of HSCs is controlled by specific genetic programs defined by three major axes: the specification of HSCs, their self-renewal (expansion and maintenance) and finally their commitment, proliferation, and differentiation (Lessard et al., 2004).

The specification of HSCs occurs earlier during development and generates a pool present in the BM at birth. HSC amplification occurs in 2 defined embryonic sites: the placenta and the fetal liver. It is assumed that expansion of the HSC pool is mainly driven by symmetric division. The self-renewal capacity allows the conservation of a sufficient number of HSCs throughout the lifetime of a person. Those cells undergo asymmetric divisions in order to produce at the same time HSCs and MPPs. Thus, we talk about maintenance Figure 4.

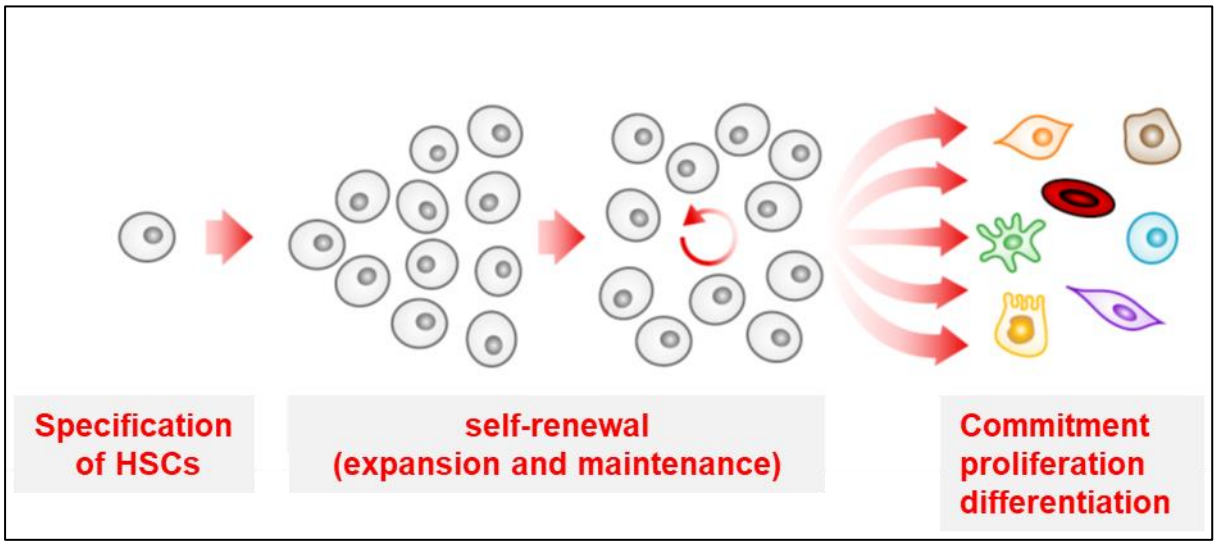


Figure 4: Genetic programs in hematopoiesis
At least three genetic programs are required for the development of the blood system.

(a) The specification of HSCs
(b) The self-renewal, including expansion and maintenance of HSCs
(c) The commitment, proliferation, and differentiation of HSCs

3. Importance of HSCs *ex vivo* in clinic

HSCs are the rare cells within the human BM and blood responsible for the life-long curative effects of hematopoietic cell transplantation (HCT) indicated for the treatment of blood cell disorders. These disorders, including leukemia, BM failure, and autoimmunity, are major causes of morbidity and mortality in France and in the developed countries. Moreover, HCT is a common procedure to cure BM failure resulting from the intensive radio/chemotherapy mandatory for the treatment of aggressive solid tumors. Although HCT has been used therapeutically for over 40 years, the demand for clinical-grade allogeneic HSCs has significantly increased over the past decades with no parallel increase of the HSC offers. A major goal in regenerative medicine is to derive HSCs from Embryonic Stem (ES) or induced Pluripotent Stem Cells (iPSC) or by direct reprogramming of specific cell types. Despite strong efforts over the past ten years, the directed generation of clinical-grade, transplantable human HSCs has not yet been achieved although the efficient generation of several more specialized blood cell types useful in human therapies was reported including erythroid cells (Lu et al., 2010) B and T lymphoid cells and megakaryocytes. There is, therefore, an urgent need to establish novel procedures to produce HSCs *de novo* since our lack of knowledge has a profound impact on therapies, weighs heavily on health budgets and is responsible for the poor quality of life for patients suffering from hematological diseases.

4. Functional and phenotypic identification of HSPCs

A. History of the concept of HSCs

In 1908, a Russian histologist named Alexander Maksimov, first put forward the existence of the SC (and coming up with the term himself) as part of his theory of hematopoiesis but, the first studies of the hematopoietic system started in 1914 when Franz Ernst Christian Neumann (1834-1918) described the BM as the organ of blood cell formation Figure 5.

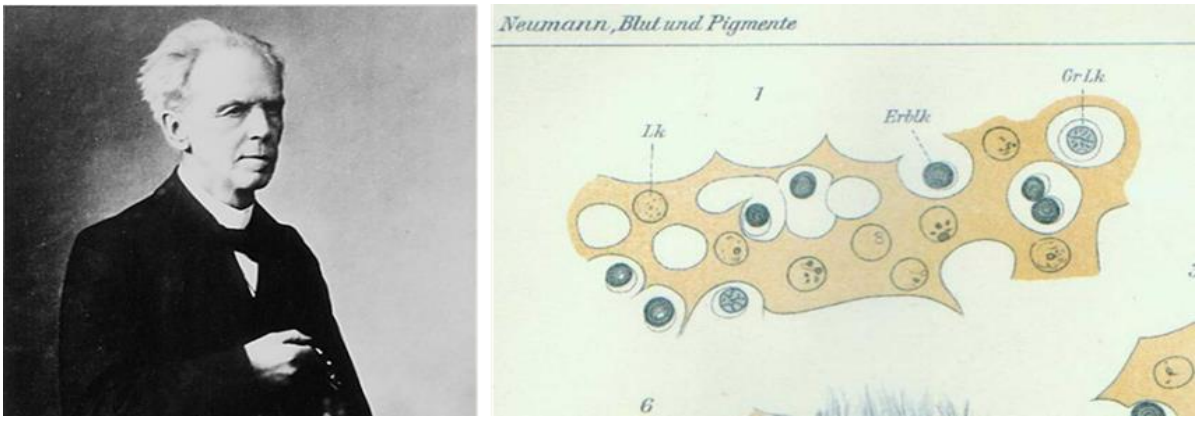


Figure 5: Ernst Neumann illustrating his "stem cell" 1914
Draft by Ernst Neumann himself, showing the "great-lymphozyt-stem-cell" (1912) or "great-lymphocyt" as stem cell for the postembryonic and embryonic development of erythropoiesis and leukopoiesis in the bone marrow and, as shown here, in the embryonic liver; GrLK: nucleus of great-lymphocyt-stem-cell; Erblk: Erythroblast; (Neumann 1914)

In 1945, the United States detonated two nuclear bombs over the cities of Hiroshima and Nagasaki. Many people who did not die immediately from the blast and heat died afterward because of a hematopoietic failure due to a lethal dose of irradiation. This was the starting point for several seminal discoveries concerning the field of hematopoietic cell transplantation. Jacobson and Lorenz in early 1951, trying to mimic radiation effects on mice, showed that spleen and BM cells of the adult mouse were able to reconstitute all hematopoietic cells of a lethally irradiated animal, proving the existence of cells responsible for the replenishment of hematopoietic cells (Jacobson et al 1951, Lorenz et al, 1951).

After those discoveries, Till & McCulloch (1960), revealed that the injection of BM cells into lethally irradiated mice led to the formation of hematopoietic nodules within the spleen. They proposed the term: Colony Forming Unit in the Spleen CFU-S (McCulloch and Till, 1960) for these nodules and showed that day 8 post-injection

nodules were essentially formed of unipotent HCs engaged in the myeloid, erythroid or megakaryocytic lineages, whereas day 12 post-injection nodules contained less differentiated HCs Figure 6. This difference between nodules suggests that the CFU-S giving rise to these 2 nodule types does not have the same degree of immaturity hence demonstrating a CFU-S hierarchy. Only day 12 CFU-S were able to reconstitute a secondary irradiated mice recipient and to form nodules within the spleen. This suggests that day 12 CFU-S possess a self-renewal potential and are more immature than day 8 CFU-S (Charbord, 1994).

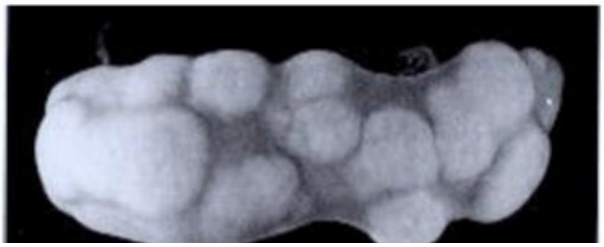


Figure 6: Colony Forming Unit in the Spleen: CFU-S

After injecting bone marrow cells into an irradiated animal, isolated colonies of hematopoietic cells develop on the surface of the spleen (Till 1981, from "Developmental Biology 6th edition" Scott F. Gilbert).

Several experiments showed that CFU-S are not able to maintain long-term hematopoiesis and are more mature than HSCs (Hodgson and Bradley, 1979; Jones et al., 1990). In order to better characterize CFU-S, Ploemacher and Brons introduced for the first time the notion of pre-CFU-S (Ploemacher and Brons 1988). Thirteen days after the primary injection, BM cells of an irradiated mouse were retrieved and injected in an another lethally irradiated recipient. Twelve days after the secondary injection, CFU-S were described in the spleen. Cells at the origin of this secondary day 12 CFU-S were named pre-CFU-S. They were shown to be able to migrate in the BM of the first mouse recipient and to be more immature than the primary day 12 CFU-S. Those pre-CFU-S were compared to the LT-HSCs and shown as cells able to reconstitute the hematopoietic system of the lethally irradiated mouse in the long-term. Nowadays the CFU-S test remains a classical way to assess the hematopoietic potential of certain embryonic and adult cell populations *in vivo*.

In the embryo, as well as in the adult, HSCs are rare and difficult to characterize, which limits considerably their fundamental study and their use in clinical therapies. The paradigm of HSCs is that it is still impossible to isolate HSCs at purity since no specific marker of these cells is highlighted. Due to the combination of membrane and intracellular markers, some populations of cells more or less enriched in HSCs can be isolated, and their respective hematopoietic potentials can be evaluated retrospectively through *in vivo* and *ex vivo tests.*

B. Ex vivo tests

a. Colony Forming Cell assay: CFC

The Colony Forming Cell assay is used in human and mouse to identify progenitors and hematopoietic precursors in a given population. It uses the property of committed progenitors to form colonies in a semi-solid medium in the presence of defined growth factors. Each colony is derived from a single progenitor, therefore counting the number of colonies provides the frequency and the total number of progenitors in the cell suspension. This assay also defines the quality of the committed progenitor which is assessed by looking at the size of the colony and its cell composition. MPPs will give rise to mixed colonies containing various types of differentiated HCs whereas, at the opposite, unipotent progenitors will give rise to colonies containing only one type of blood cells i.e., erythrocytes, granulocytes, macrophages… After 7 days of culture, colonies formed contained different lineages: Erythrocytes, granulocytes, macrophages, megakaryocytes (Johnson and Metcalf, 1977) Table 1, Figure 7.

Table 1: The different types of colonies in the CFC assay

CFU-E	Colony Forming Unit-Erythroid	Erythroid progenitors
CFU-Mk	Colony Forming Unit-Megakaryocyte	Megakaryocyte progenitors
CFU-G	Colony Forming Unit-Granulocyte	Granulocyte progenitors
CFU-M	Colony Forming Unit-Monocyte/Macrophage	Monocyte/macrophage progenitors
CFU-B	Colony Forming Unit-Lymphocyte B	Lymphocyte B progenitors
CFU-GM	Colony Forming Unit-granulocyte/macrophage	GMP progenitors
CFU- E/Mk	Colony Forming Unit-Erythroid/Megakaryocyte	MEP progenitors

However, the CFU test does not allow the identification of T lymphocyte progenitors since these cells need interaction with a permissive microenvironment.

Starting Day 10 we can find MPPs belonging to more than one lineage: CFU-GEMM colony-forming unit granulocyte-erythroid-macrophage-megakaryocyte.

In addition to these above-cited types of colonies, the CFC assay allows the formation of HPP-CFCs (High Proliferative Potential Colony Forming Cells) (Hodgson and Bradley, 1979; Srour et al., 1993). They are characterized by their ability to generate large colonies which, after 10-12 days of culture, contain more than 50,000 cells of the type granulocyte, erythrocyte, megakaryocyte and macrophage. Populations resistant to 5-fluorouracil (5-FU-used to destroy cells in cycle) are enriched with HPP-CFC, suggesting the quiescent state of these cells in vivo. HPP-CFCs are capable of producing new colonies once they are replated, demonstrating their ability to self-renew, a characteristic common to CFC and HSCs.

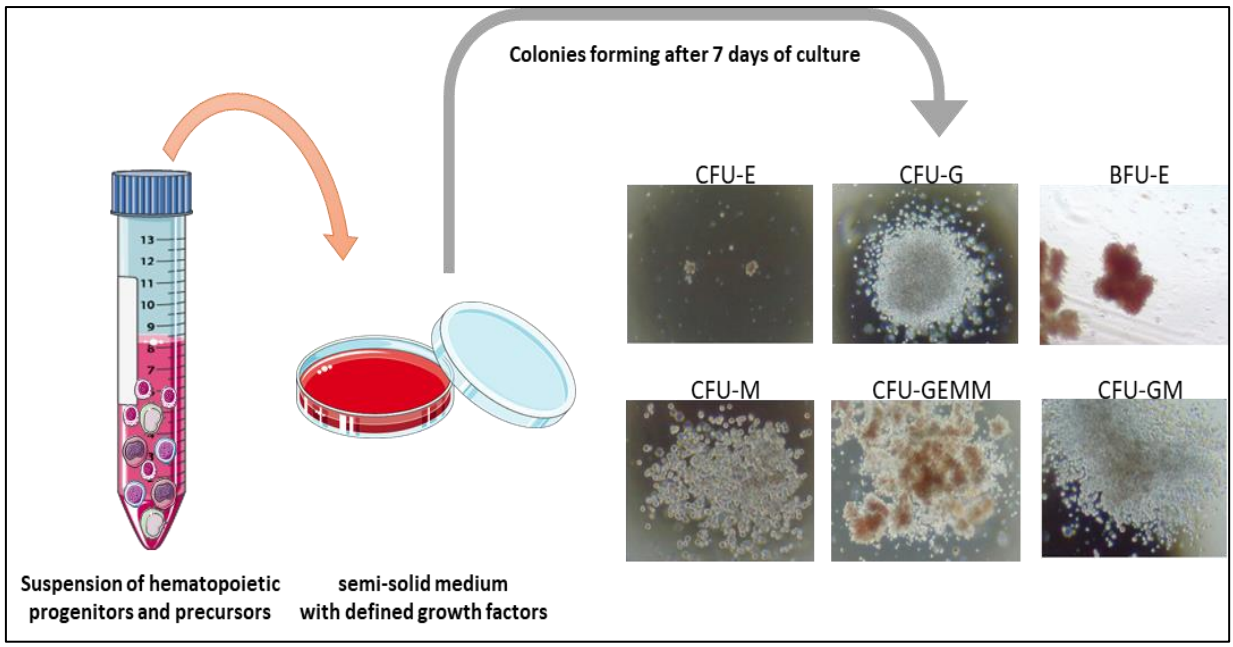


Figure 7: Colony-Forming Cell (CFC) assay

The suspension of hematopoietic progenitors and precursors is cultured in a semi-solid medium with defined growth factors for 7 days and then the number and the type of colonies formed are assessed.

Abbreviations: CFU Colony-Forming Unit, BFU Burst-Forming Unit, CFU - E Erythrocyte, CFU - E Erythrocyte, CFU - G Granulocyte, CFU - M Monocyte/Macrophage, CFU -GM Granulocyte/Macrophage, CFU - GEMM Granulocyte/Erythrocyte/Macrophage/Megakaryocyte

Other *ex vivo* tests have been used to identify more immature hematopoietic progenitors and stem compartment. Those tests are described in the upcoming paragraphs.

b. Co-culture systems

In vivo, HSCs are in tight contact with stromal cells in their niche. The interactions resulting from this contact are essential to regulate HSC biology. The long-term culture of HSCs is not possible without stromal support, showing the importance of the microenvironment promoting hematopoietic growth and differentiation especially through the production of hematopoietic growth factors, like Colony Stimulating Factors (CSFs), the existence of extracellular matrix proteins and receptors and also by the biosynthesis of microvesicles emitted by the stromal cells towards the HSCs. The longer the co-culture is, the more immature the HCs are.

c. The Cobblestone Area-Forming Cell CA-FC

The CA-FC test detects, in a given population, the presence of more immature HSCs than those detected by the clonogenic test. It consists of co-culturing with limiting dilution, an enriched population of immature HCs on a pre-established stromal cell layer supporting the hematopoiesis. The medium used promotes only the myelopoiesis and two types of cells arise: in the supernatant, round and refractive HCs and CA-FC cells that make paved colonies in close contact with the stromal supporting cells Figure 8. Depending on the duration of co-culture, observation of the CA-FC is correlated to a given compartment of the HCs hierarchy: after 7 days of co-culture, the presence of CA-FC allows to say that the initial population contains hematopoietic progenitors, while CA-FC at day 28 shows that the initial population contain more immature cells able to persist longer. This test was also used to determine the susceptibility of HCs to radiation (Ploemacher et al., 1992) and cytotoxic drugs (Down and Ploemacher, 1993).

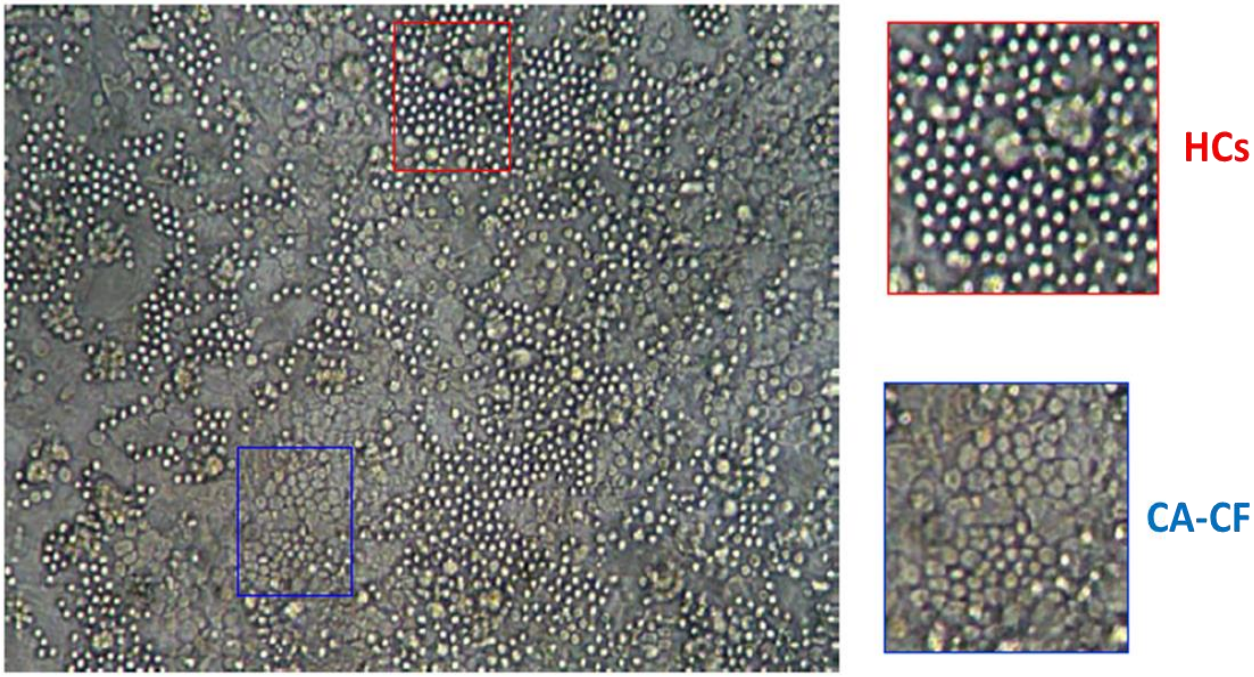


Figure 8: Illustration of a coculture of murine hematopoietic cells on a layer of MS5 stromal cells
Two types of cells arise: in the supernatant, round and refractive HCs and CA-FC cells that make paved colonies in close contact with the stromal supporting cells
Abbreviations: CA-FC Cobblestone Area-Forming Cell, HCs Hematopoietic Cells.

d. Long-Term Culture-Initiating Cell LTC-IC

In the direct continuity of the previously described co-culture, it is possible to go further in determining the degree of the immaturity of HCs contained in the population seeded on the stromal cell layer. Thus, after 28 to 35 days of culture, the cells of this layer, as well as the supernatant, can be recovered and grown in methylcellulose medium. If 7 to 14 days later, CFCs representative of all myeloid lineage are observable, it is possible to conclude retrospectively that the starting population initially seeded contained LTC-IC, HCs more immature than CA-FC D28. Originally, the growing conditions for the identification of LTC-IC in a population had been identified to study the myeloid differentiation of very immature human CHs *ex vivo* (Sutherland et al., 1989), but they turn out to be also relevant to the mouse. Some disadvantages to keep in mind is the absence of lymphoid differentiation and the loss of erythroid potential and heterogeneity of the progenitors Figure 9.

In order to estimate the lymphocyte potential, a variation of the LTC-IC test is set up. After 28 days of culture, the supernatant is totally replaced by a medium promoting the B lymphocyte production (Whitlock and Witte, 1982). The culture is maintained for 7 to 14 days. Cells of the supernatant are analyzed by flow cytometry.

However, this test doesn't reveal the T lymphocyte potential. It has been shown that the NOTCH pathway plays a crucial role during T cell development. And the co-culture of immature Progenitors with DLL1 expressing stromal cell allows a T cell differentiation (La Motte-Mohs, 2005). Another variation of the LTC-IC test is the Extended Long-Term Culture-Initiating Cell ELTC-IC. This test allows going for 60 days of culture which allows detecting the most immature HCs.

Thus, the current co-culture systems allow at different times of the culture, to determine the frequency of HCs and their degree of immaturity in a given population.

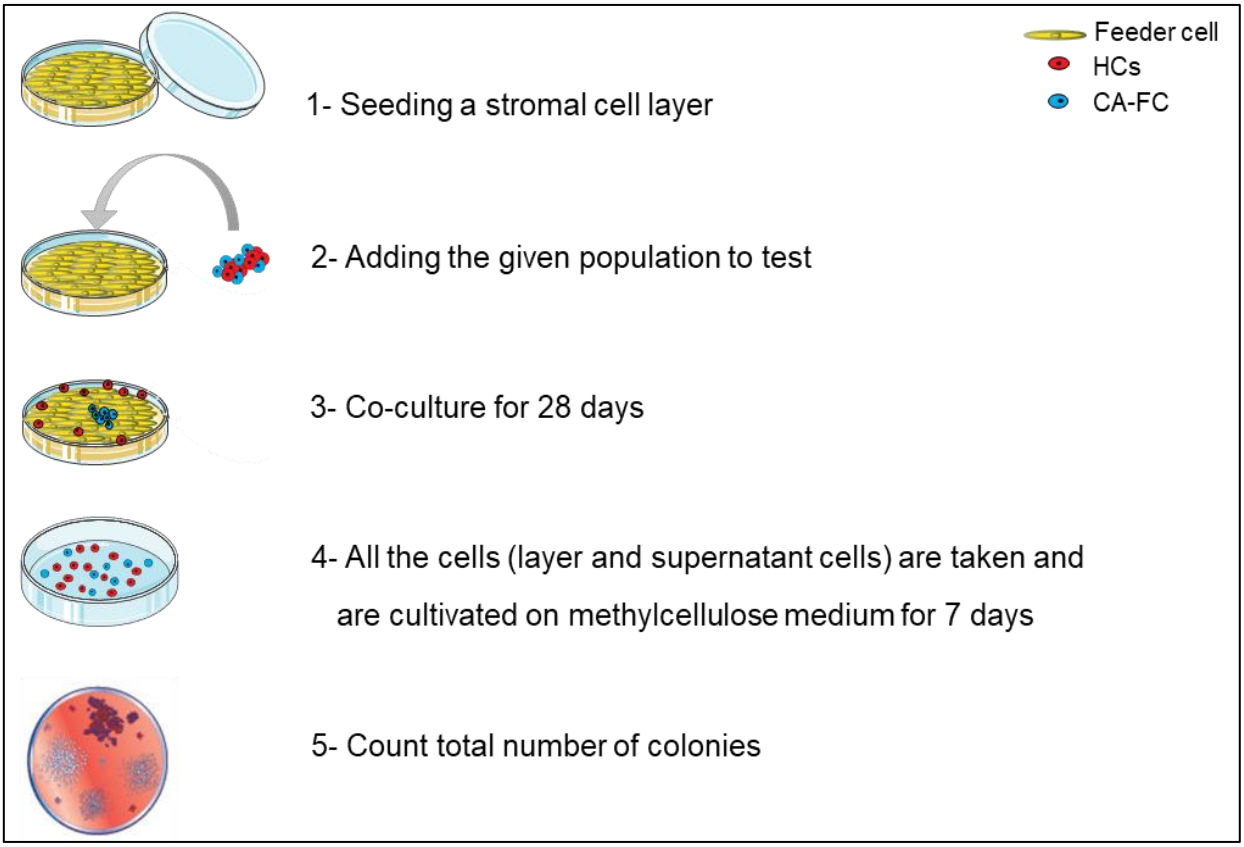


Figure 9: Test LTC-IC: Long-term Culture Initiating Cell assay
The population to test is added to a stromal cell layer and co-cultured for 28 days. Then all the cells are taken and cultivated on methylcellulose medium for 7 days and the number of colonies is assessed.
Abbreviation: CA-FC Cobblestone Area-Forming Cell, HCs: Hematopoietic Cells

C. In vivo tests

The worldwide recognized test to assess the potential of HSCs is their transplantation into lethally irradiated mice recipients. These experiments are based on Till&McChulloch assay in the 1960s. When irradiated, the mouse loses all its cycling HCs including progenitors and its survival depends on the existence of HSCs within the fraction during transplantation. The short-term survival depends on progenitor cells and injected cells are defined as ST-HSCs, while long-term survival depends on the presence of HSCs able to produce all the hematopoietic lineages of the mouse and defined as LT-HSCs.

The analysis is based on the expression of different markers especially the CD45 marker having different alleles in human and mouse and allowing the differentiation between donor and receiver cells Figure 10.

The first reconstitution is named the primary reconstitution and assesses the multipotency of the injected cells. This step is followed by secondary transplantation in order to assess the long-term self-renewal potential Figure 11.

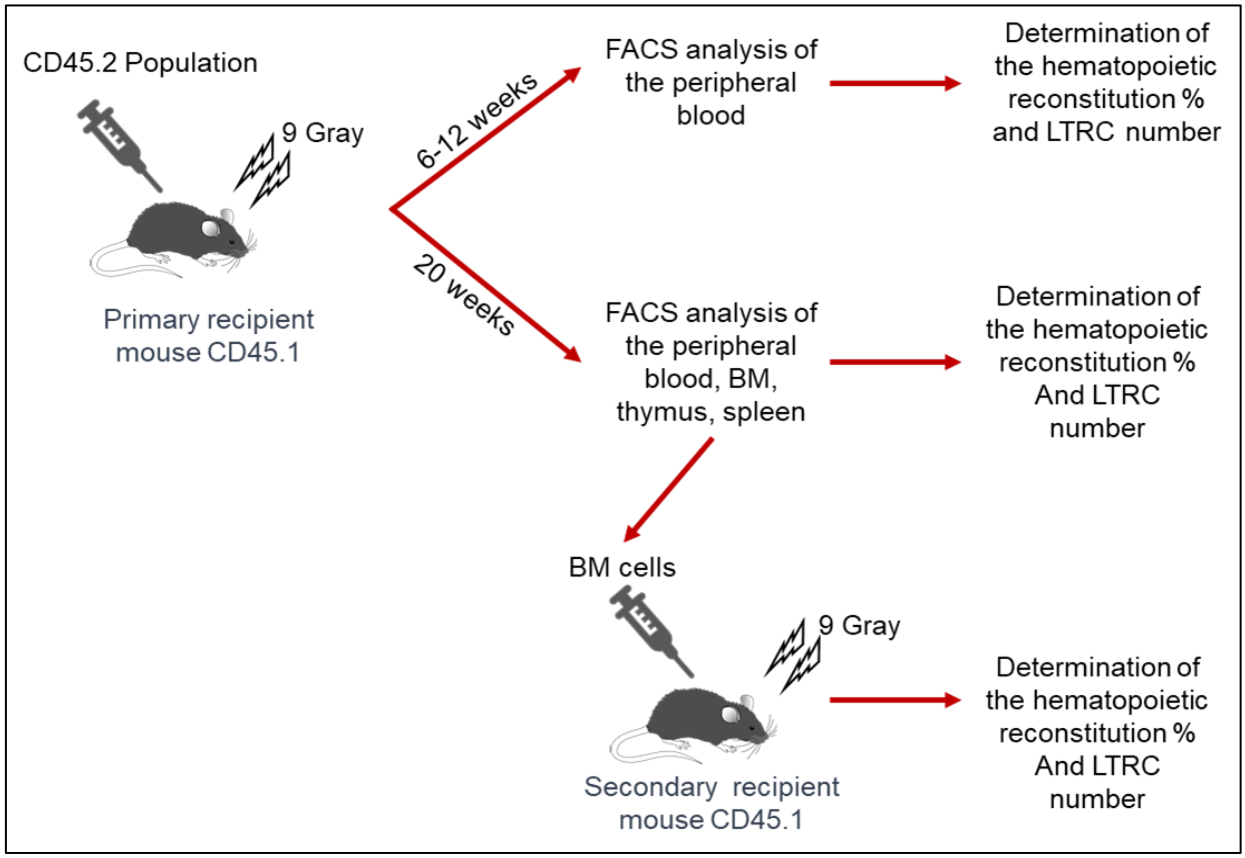


Figure 10: Primary and secondary hematopoietic reconstitution
Abbreviations: LTRC: Long-Term Repopulating Cell FACS: Fluorescence Activated Cell Sorting, BM: Bone Marrow

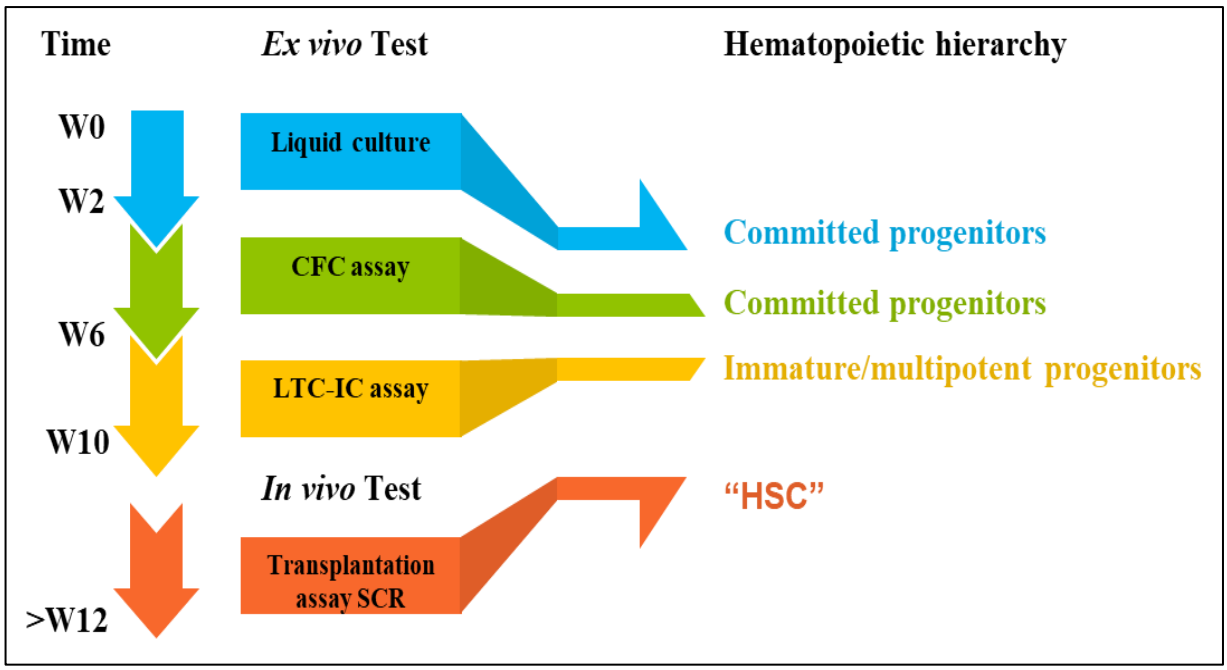


Figure 11: Recapitulative of Ex vivo and In vivo tests in parallel with the hematopoietic hierarchy
Abbreviations: HSC: Hematopoietic Stem Cell; CFC: Colony-Forming Cell; LTC-IC: Long-term Culture Initiating Cell assay

D. Phenotypic identification of HSPCs

Depending on the type, lineage and maturation degree, hematopoietic cells express specific cell-surface proteins. These proteins can mediate several cellular functions such as receptors, ligands, adhesion or migration… Together with the discovery of Fluorescent Activating Cell Sorting FACS and the production of monoclonal antibodies, those markers allowed the identification and the follow up of the hematopoietic system.

While many aspects of HSC biology are shared between species, the cell surface markers between mice and humans appear to show a lack of congruence.

In the mouse BM, two cell populations can be distinguished: cells expressing markers of mature lineages (CD3, CD4, CD8, CD19, NK1.1, Gr1, CD11b and Ter119, Lin^+), and cells negative for these markers. This population called lineage minus (Lin^-) was shown to contain hematopoietic stem and progenitor cells. HSCs and MPPs can be fractionated according to a high level of Stem Cell Antigen (Sca)-1 and c-Kit tyrosine kinase receptor ($Lin^-Sca\text{-}1^{hi}$ $c\text{-}Kit^{hi}$, LSK) (Okada et al.). Using CD34 and FLT3 markers, LT-HSCs can be distinguished from ST-HSCs and MPPs. Within the LSK $CD34^-FLT3^-$ subset, around 35% are HSCs with LT repopulation activity. In contrast, LSK $CD34^+FLT3^-$ cells are multipotent but have a limited repopulation activity in lethally irradiated mice (Yang, 2005). The most commonly used method nowadays to purify mouse HSCs is the use of cell surface receptors of the SLAM (Signaling Lymphocyte Activation Molecule) family. The SLAM family is a group of 10–11 cell surface receptors that are tandemly arrayed at a single locus on chromosome 1 (Engel et al., 2003). Using the SLAM code, HSCs were highly purified as $CD150^+CD244^-CD48^-$ cells while MPPs were $CD244^+CD150^-CD48^-$ and most restricted progenitors were $CD48^+CD244^+CD150^-$ (Kiel et al., 2005). Among the LSK $CD150^+$ $CD48^-$ cell population, 30% of cells do not express CD34 and this population might be highly enriched in HSCs.

In human, HSCs express several surface markers, but their number is not enough to specify their determination/self-renewal potential. The most used surface marker is the CD34 adhesion mediated glycoprotein. $CD34^+$ cells account about 1% of the BM cells and are enriched in progenitors and HSCs. CD34 expression is not

exclusive to HSCs as some endothelial cells (ECs), some stromal BM cells and many others also carry the CD34 marker (Sidney et al., 2014) Figure 12.

Several works showed that HSCs can also be found in $CD34^{-}Lin^{-}$ cells. Other studies conclude that $CD34^{-}$ cells could be the ancestors of $CD34^{+}$ cells (Anjos-Afonso et al., 2013). Since CD34 also marks committed progenitors, other markers are to be combined to enrich in more immature cells. The most used are CD38, Thy1/CD90, and CD133/prominim1.

These markers can be combined with the ability of HSCs to efflux the Rhodamine or Hoechst dye to assess the ability of HSCs at the clonal level.

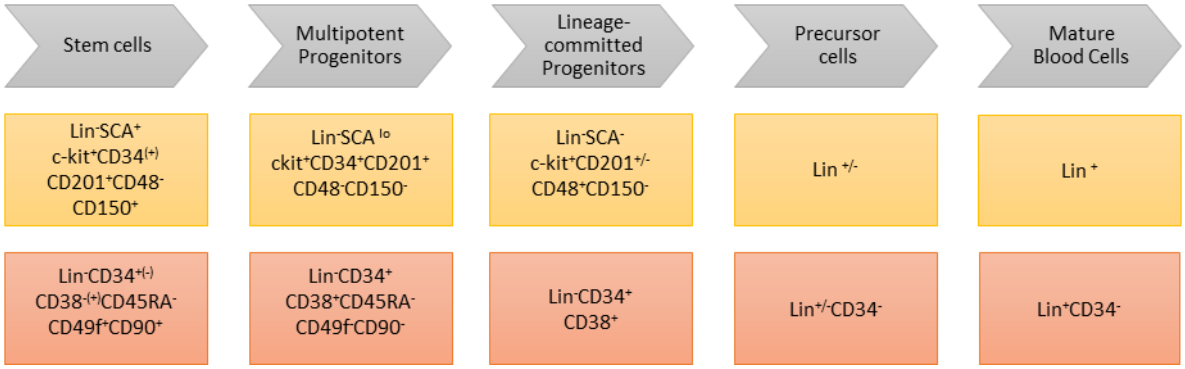


Figure 12: Parallel between human and mouse most definitive markers used to identify the various types of hematopoietic cells.
In grey cell types
In yellow mouse markers
In orange human markers

II. On the origin of blood

The origins of HSCs have been the subject of intense debate. Development of the hematopoietic system is characterized by the existence of multiple waves, in multiple origins and changing locations. There are also specificities related to the species. The chick embryo was instrumental in decoding many rules of developmental hematopoiesis. In order to investigate how the blood system forms during development, Moore and Owen chose the chicken embryo, already the subject of many descriptive studies since the end of the XIXth century, because it is directly visible after shell opening and develops independently from a maternal circulation. They applied an array of experimental approaches: irradiation/restoration, parabiosis, grafting of rudiments (thymus, spleen, bursa of Fabricius) on the chorioallantoic membrane (Moore and Metcalf, 1970 For review) To follow cell origins in these different designs, Moore and Owen used the pair of sexual chromosomes as a cell marker. Despite several drawbacks inherent to this labelling system (the chromosomes can be observed only in 5-10% of cells arrested at the metaphase stage of mitosis after a colcemid treatment which elicits high mortality of the embryos; furthermore the adequate opposite sex combination is left to chance), these investigators discovered an important trait which characterizes the development of the blood system: HSCs colonize all rudiments to the exception of the yolk sac (YS). They also grafted mouse embryo rudiments on the chicken chorioallantois and observed that, in contrast with cell type-specific recognition capacities displayed by most other embryonic cell types across different classes of vertebrates, attraction or recognition did not work between mouse stromal rudiments and chicken HSCs. Retrieved at an early stage and grafted on the chorioallantois, the mouse thymic rudiment remained lymphoid, allowing Moore and Owen to conclude in 1970 that, in mammalian developmental hematopoiesis also, HSC emerge in one location and differentiate in another (Moore and Metcalf, 1970).

Benefiting from accurate functional testing, the mouse embryo took the lead and was instrumental in building a precise, comprehensive picture of hematopoietic development in mammals. Developmental hematopoiesis is organized into 3 separated waves. The first one takes place in the YS at embryonic day (E)7.5 in the mouse embryo (Palis et al. 1999) and at 19-22 hours in the chick embryo (Minko et al., 2003). This wave gives rise to primitive erythroid cells, macrophages, and

megakaryocytes (Palis et al 1999). Shortly after the onset of this primitive hematopoiesis, the first definitive erythro-myeloid progenitors and immune-restricted progenitors emerge from the YS at E8.5 in the mouse and E3 in the chick embryo. The third wave occurs in the embryo proper in a specific region named aorta gonad mesonephros (AGM) and is characterized by the production of HSCs and adult type lympho-myeloid progenitors (Drevon and Jaffredo, 2014; Gritz and Hirschi, 2016). HSCs detectable by direct transplantation in adult recipients are not present until E10.5 in the AGM but hematopoietic emergence is detected from E9.5 closely associated with vitelline and umbilical arteries (de Bruijn, 2000; Tavian et al., 2001). Subsequently, transplantable HSCs are found in the YS, the placenta and the head (Li et al 2018, Alvarez-Silva et al., 2003) but it is not clear whether HSCs emerge *de novo* in these sites. Shortly after this period, HSCs migrate and either colonize the fetal liver in mammals or reach the sub-aortic mesenchyme in birds. In both situations, they undergo multiplication and amplification and eventually migrate to the definitive hematopoietic organs (BM, Thymus, spleen…) where they will reside during the lifetime of the organism Figure 13.

Figure 13: The Journey of HSC from in different species: Human, Mouse, Chick and Zebrafish.
Starting with fertilization, circulation, birth and larvae and finishing with the Adult life.
The four species share two major waves of hematopoiesis: primitive and definitive waves.
Human, Mouse and Chick have the same hematopoietic organs, however in Chick the Fetal Liver is replaced by the para-aortic foci.
Concerning the Zebrafish, it shares the AGM and the Thymus with the other species. It got the ICM as an equivalent of the YS, the CHT as equivalent of the Fetal Liver and the Kidney marrow as equivalent of the bone marrow.
Abbreviations: YS: Yolk Sac, ICM: Inner Cell Mass, CHT: Caudal Hematopoietic Tissue.
d: day, h: hour, W: Week, hpf: hours post fertilization, Mo: month
Adapted from *Rowe et al 2016*

1. The different hematopoietic sites

A. The yolk sac (YS)

The YS is a bilayer organ composed of extraembryonic mesoderm cells opposed to visceral endoderm cells. The association of cells derived from mesoderm and endoderm germ layers is termed splanchnopleure. In mouse and human, normal embryonic development is critically dependent on the proper functioning of the YS since its endoderm layer transports and metabolizes maternally derived macromolecules and synthesizes serum proteins, and its mesoderm layer produces the first blood cells within blood islands. Furthermore, the visceral endoderm is thought to serve as a source of inductive signals important for the formation of blood cells and endothelial networks in the YS (Palis and Yoder, 2001) Figure 14.

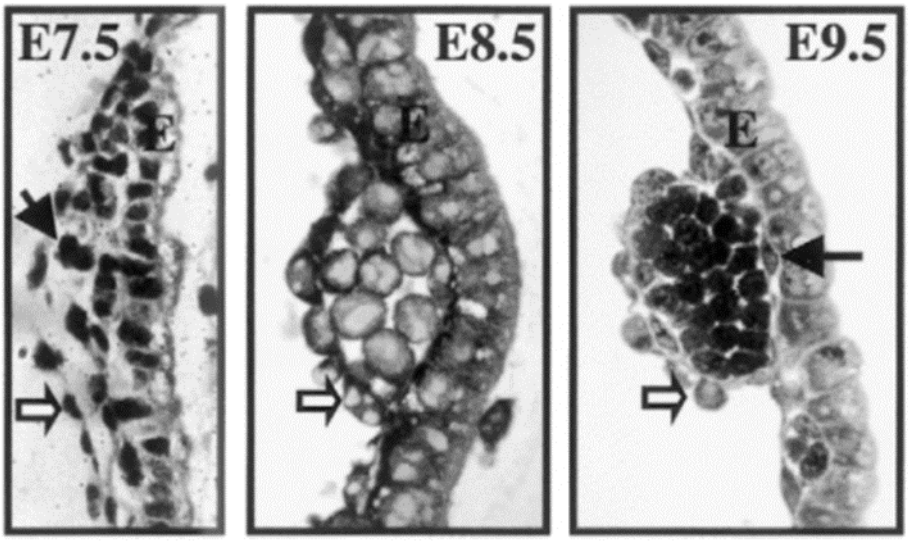


Figure 14: Maturation of the blood island in mice embryo between E7,5 and E9,5
Development of yolk sac blood islands in the mouse embryo between E7.5 and E9.5. Yolk sac blood islands form between the single-cell layer of endoderm cells (E) and mesothelial cells (open arrow). Blood islands develop from undifferentiated mesoderm cells (E7.5, closed arrow) that give rise to inner blood cells surrounded by an outer endothelial lining (E9.5, closed arrow). (Palis and Yoder, 2001)

a. The first primitive hematopoietic wave.

The first morphological evidence of hematopoietic activity in multiple mammalian and non-mammalian embryos is the appearance of cohorts of immature erythroid cells within the YS mesoderm beginning at E7.5 in the mouse embryo. These “blood islands” consist of immature erythroid cells wrapped by a monolayer of ECs emerging from seemingly identical mesoderm cells. This close spatial and temporal relationship between the two cell types suggested to Maximow (1909), Vera Dantschakoff (1909) and Florence Sabin (1920) that these cells arise from a common mesoderm precursor designated as the hemangioblast (Murray, 1932). See part III page 55 for a description of the hemangioblast.

The fact that primitive erythroid cells emerge from the YS blood islands and constitute the first circulating cells of the embryo was first recognized by Maximov (1909). The large, nucleated erythroid cells contrasted with small, enucleated, red blood cells that circulate during the life span. These primitive erythrocytes have some specific features that distinguish them from definitive erythrocytes: 1° they are several folds larger and contain six-fold more hemoglobin when mature. Palis and co-workers showed that, contrary to expectations, primitive erythroblasts progressively enucleate between embryonic days 12.5 and 16.5, generating mature primitive erythrocytes that are similar in size to their nucleated counterparts. These enucleated primitive erythrocytes circulate as late as 5 days after birth (Kingsley, 2004). 2° primitive and definitive red blood cells express different combinations of globin chains. Primitive erythroblasts express both embryonic (Zeta, ßH1, and ey) and adult (alpha1, 2, **ß**1, and ß2) globins during cell maturation. z- and alpha-globin mRNAs have been detected in YS blood islands in pre-somite stage embryos. In functional *ex vivo* tests, primitive erythrocytes arise from a unique progenitor termed EryP-CFC that generates within 5 days compact erythroid colonies comprised of large erythroid cells that express both embryonic and adult globin chains (Palis, 1999a; Wong et al., 1986).

In addition to primitive erythroid cells, the early YS also generates colonies of macrophages that emerge temporarily and spatially with the primitive erythroid cells at E7.5. These colonies are generated from a unipotent progenitor and are not a component of a bi- or multipotent progenitors until E8.5. Most of these macrophages give rise to resident macrophages involved in cell debris removal during tissue remodeling but it was shown recently that a cohort enters the central nervous system

and differentiate into microglia (Ginhoux et al., 2010). The first experimental evidence was obtained on the avian model by constructing YS chimeras where a chick embryo is grafted onto a quail YS or vice versa and immunostain them with an antibody recognizing the hemangioblastic lineage QH1 or MB1. This experimental design showed that a phagocytic lineage originating from the YS seeds the embryo body and invade the tissues including the central nervous system (Cuadros et al., 1992). Slightly later on, the existence of a microglial progenitor was shown from E8 onward in the mouse embryo. Later on, using inducible lineage or gene expression mapping experiments using Runx1, c-Kit, CSFR1, Tie2 or Flt3 reporter mice, tracing of microglial progenitors was achieved, and their YS origin formerly demonstrated (McGrath et al., 2015)

Megakaryocyte progenitors have also been isolated from the E7.5 YS of the mouse embryo. Mk-CFCs increase with time from E7.5 to E9.5 indicating that megakaryocyte formation overlaps between primitive and definitive YS hematopoiesis. However, megakaryocyte precursors revealed with anti Gp1bß are not detected until E9.5 in the YS. Interestingly, if in the BM, a bipotent erythro-megakaryocyte progenitor has been revealed, most of the erythroid and megakaryocyte colonies were composed of a single cell type indicating a unipotent progenitor. However, some bipotential progenitors were detected that are biased towards the erythroid lineage (Tober et al., 2007).

b. The second definitive hematopoietic wave.

Interestingly, the first definitive erythroid progenitors Burst Forming Unit-Erythroid (BFU-E) were detected from E8.5 in the mouse YS. They will massively expand over the next 48 hours and will be found in increasing numbers in the bloodstream between E9.5 and 10.5. As the liver emerges as a hematopoietic organ, they will be massively amplified from E10. Small numbers of BFU-E are detected within the YS after the appearance of CFU-E. Large numbers of CFU-E are detected in the liver, prior to the colonization by adult-repopulating HSCs, indicating a major role in definitive erythropoiesis. Definitive erythrocytes enter the bloodstream between E11.5 and E12.5.

In addition to definitive erythrocytes, several additional hematopoietic progenitors are detected in the YS. These are unipotent mast cell progenitor and a

bipotential granulocyte-macrophage progenitor and multilineage, high proliferative potential colony-forming cells (HPP-CFC) which give rise *ex vivo* to macrophages, mastocytes, and granulocytes. This myeloid potential originates from erythro-myeloid progenitors that express high levels of c-kit and CD41 at their surface (Ferkowicz, 2003; Mikkola, 2003). Single E9.5 Erythro-Myeloid Progenitors (EMPs) have the capacity to generate multiple myeloid lineages ex vivo as well as definitive erythroid potential (McGrath et al., 2015). However, EMP does not have B-lymphoid potential. Of note, EMP arises via an endothelial-to-hematopoietic transition (EHT) in a Runx1-dependent manner (Chen et al., 2011; Frame et al., 2016; Miller et al., 2002; North et al., 1999). Macrophage and megakaryocyte progenitors persist after primitive erythropoiesis indicating that these progenitors are component of the two waves.

Finally, small numbers of lymphoid progenitors have been detected around E9 in the YS and the embryo proper. However, it is not clear whether these lymphoid progenitors arise from the YS or seed the YS from the embryo. Thorough examination of the YS and para-aortic splanchnopleura lymphoid potential indicates that this potential is not present in the YS until the onset of circulation but is detected in the future aortic region as early as E7.5 days (Cumano et al., 1996). However, YS chimeras indicate that the generation of B-lymphoid cells from the YS is possible but at very low frequency (Sugiyama et al., 2007). Another study indicates that B1 progenitor cells, that preferentially differentiate into innate type-B1 and marginal zone B cells, but not into B2 cells upon transplantation, arise directly from the YS through an Endothelial to Hematopoietic Transition (EHT) independently of HSCs (Yoshimoto et al., 2011). A similar YS-derived immune-restricted and lymphoid-primed progenitor was also described prior to the detection of definitive HSCs and contribute to giving rise to lymphoid as well as myeloid components of the immune system. Dissection of the YS and the Para-aortic-Splanchnopleura (P-Sp) prior to circulation followed by a short culture period show that these cells emerge autonomously from the YS and are detected based on Rag1 expression as early as E9.5 (Böiers et al., 2013).

B. The allantois/placenta

A hematopoietic function for the placenta was considered a long time ago. While the presence of reconstituting HSC (McCulloch and Till, 1960 ; Dancis et al., 1968) and of B lymphoid progenitors (Melchers, 1979) in the placenta has been reported many years ago, a contribution of the placenta to fetal hematopoiesis was not seriously considered then.

It started with the quail chicken system that revealed an additional hematopoietic organ starting at E3.5, the allantois, an embryonic appendage situated at the caudal part of the embryo and developing from the intestine. The allantois is basically formed of endoderm associated with mesoderm, allowing it to produce ECs and HCs. Quail allantois grafted in the chicken host could produce ECs and HCs that colonize the BM. Moreover, HCs produced by the allantois express different globin isoforms including the adult ß globin (Caprioli et al., 1998). Detailed examination of the earliest phases of allantois development showed that this structure develops blood islands harboring red blood cells and CD45+ cells. *Ex vivo* culture of the allantois indicated that the red blood cells were of definitive lineage in keeping with Caprioli et al., 1998 (Caprioli et al., 2001) Figure 15.

Like the avian chorioallantoic membrane, the placenta is an extra-embryonic appendage richly vascularized. It is formed in the mouse, from E8,5 and provides gas and nutrient exchange necessary for the survival of the embryo. It is formed by the fusion of the allantois with the ecto-placental cone at E8.5 and is composed of these two structures, the first of fetal origin and the second of maternal origin. The hematopoietic potential of the placenta was detected in both allantois and pre-fusion chorion (Corbel et al., 2007) Figure 16. The presence of clonogenic hematopoietic progenitors was investigated in the E8/E9 placenta and compared to the fetal liver of YS at the same stage. HPP-CFC were 2-4 times more frequent in the placenta (Alvarez-Silva, 2003). The placenta was also shown to be an amplification niche for HSCs between E12.5 to E13.5 (Gekas et al., 2005) in the labyrinth region (Ottersbach and Dzierzak, 2005). However, it was not clear whether the allantois/placenta was a site of maturation or a source of hematopoietic cells. The allantois isolated before the establishment of the circulation is capable of giving rise to myeloid and definitive erythroid cells (Ziegler et al., 2006). In contrast to the fetal liver, the placenta is

enriched in immature progenitors and few progenitors involved in specific lineages are observed. In addition to being the emergence site of HSCs, this extraembryonic structure plays a temporary niche role in the maturation and expansion of HSCs (Alvarez-Silva, 2003; Gekas et al., 2005) Using mouse embryos deficient for NCX1 and lacking heartbeat suggest that placenta is able to generate B, T and erythroid cells suggesting the in-situ emergence of hematopoietic progenitors (Rhodes et al., 2008). The approach is controversial as the blood flow itself plays a role in maintaining the endothelial structure and the appearance of HSCs. Indeed, recent experiments in Zebrafish and Mouse show that the absence of blood circulation causes a severe decrease in the number of HSCs in the embryo (North et al., 2009). Additional in vivo and in vitro experiments were performed, showing that blood flow has a significant influence on the level of hematopoietic gene expression, such as runx1 (Adamo et al., 2009).

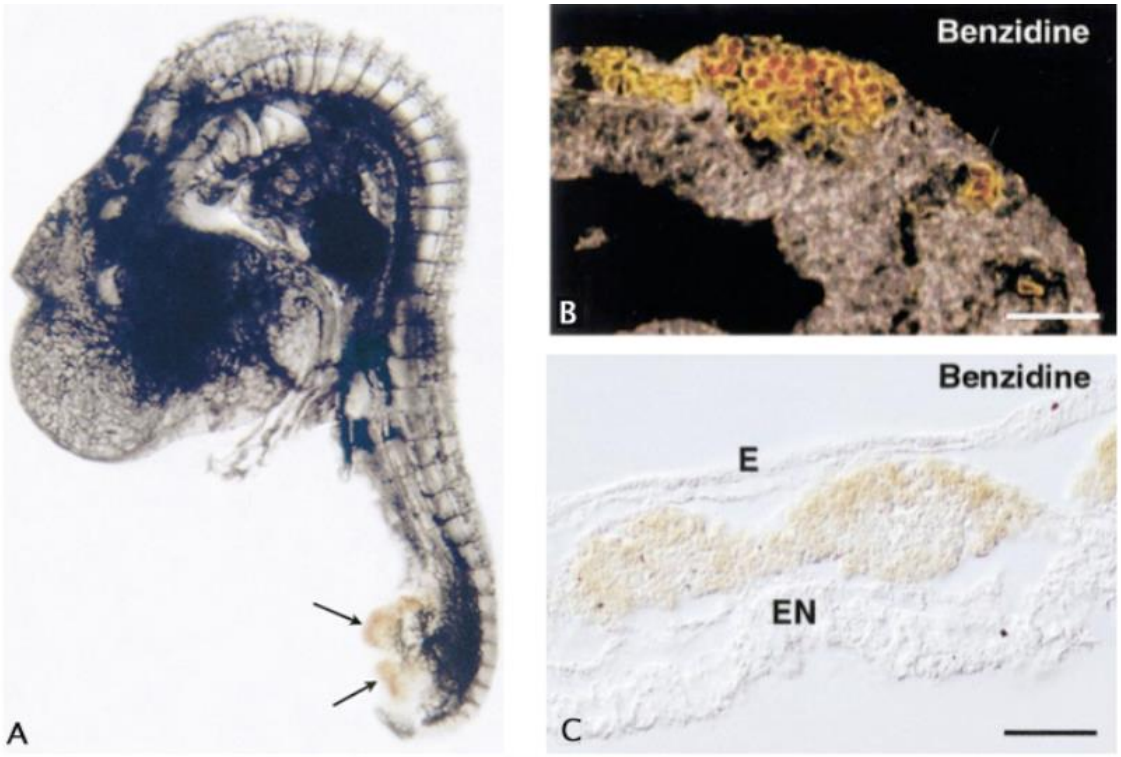


Figure 15: The Allantois a hematopoietic emergence site
A- *HH20 chick embryo. India ink inoculation followed by methyl salicylate clearing. The whole vascular network including plexuses in the brain, neural tube, and posterior intestinal portal are filled with ink. Only the base of the allantois is vascularized, while the large aggregates of red blood cells at the apex remain free of ink (arrows).*
B- *Benzidine staining 1 phase contrast, cross section. Bar 5 20 mm*
C- *Benzidine staining 1 Nomarski's interference contrast of yolk sac blood islands from an E2 chicken embryo. Allantoic and yolk sac blood islands are very similar.*
Abbreviations: E, ectoderm; En,allantoic endoderm.
Adapted from Caprioli et al, 2001

Similar investigations were undertaken in the human embryo. Multiple types of progenitors i.e., myeloid, erythroid, NK and B cells were detected in the labyrinth of the human placenta (Barcena et al., 2009). In addition, progenitors and HSC were detected in the placenta all along with development. However, no amplification of HSC or progenitor could be detected indicating a slightly different role than the mouse placenta (Robin et al., 2009) Figure 16

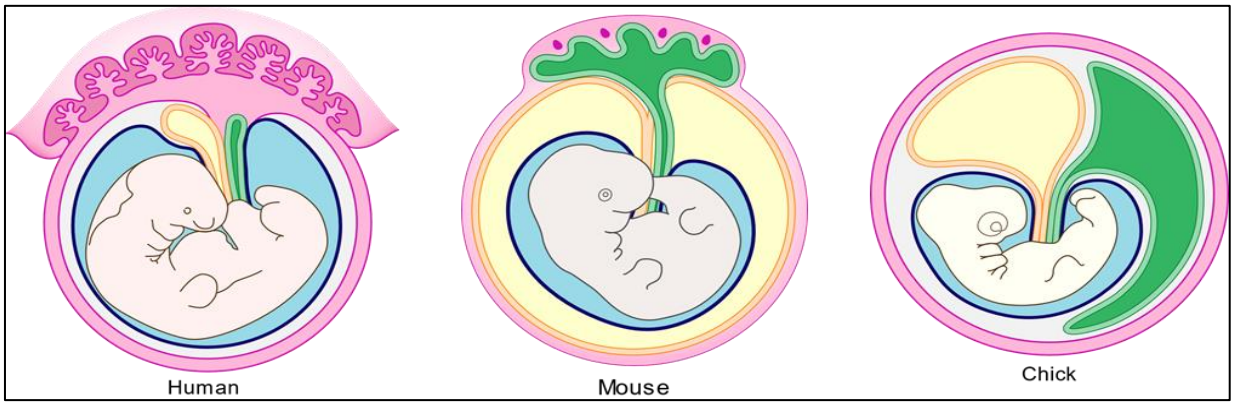


Figure 16: The extraembryonic membranes in human, mouse and chick embryos.

In the middle: embryo
In yellow: Yolk Sac
In green: The allantois
In blue: The Amnion
In pink: Chorion
The placenta of human and mouse embryos is shown on the upper part in pink.

C. The aortic region

At the dawn of embryology, scientists described the seemingly coincident emergence of endothelial and hematopoietic cells in the YS giving rise to the concept of hemangioblast (Murray, 1932). It is not difficult to imagine why the YS was the first hematopoietic site to be investigated. Indeed, the YS shows the first signs of hematopoiesis testified by the presence of the red blood islands and was naturally proposed to be the site of HSC production. In the mid and late 60s, Moore and Owen reported that the YS was the site where hematopoietic stem cells emerge in a unique, early, developmental event and subsequently colonize the hematopoietic organs (Moore and Owen, 1967). The experimental design used a series of graft and parabiosis experiments relying on the identification of sex chromosomes Figure 17. In addition, they also showed that HSPCs were of extrinsic origin (see above). These two pillars made the core of the hematogenous theory. During the 70s the relevance of this hypothesis was investigated in mammalian species by culturing either E7 whole mouse embryo or separated YS and embryo. The conclusion was that the YS was the only site where HSC were formed reinforcing the central role of this site in HSC generation (Moore and Metcalf, 1970). Based on this and on other reports (Weissman et al., 1977), the hematogenous theory became a dogma and was widely accepted among the scientific community.

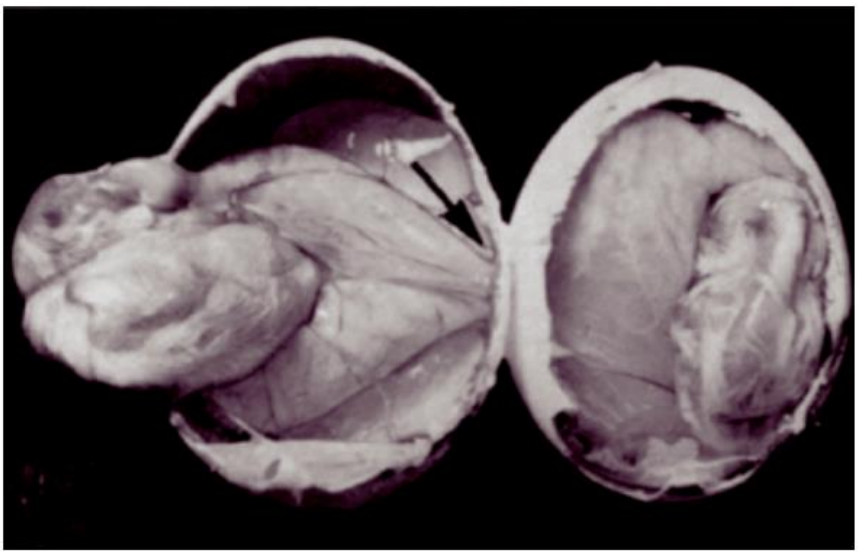


Figure 17: Parabiosis procedures between the YS of two chicken embryos
Embryos joined by a yolk sac anastomosis (arrowed)
Image from Moore&Owen 1967

In 1972, a surgical technique to construct a unique type of quail/chick chimera was devised by Martin (1972). The flatness of the avian embryo as it lies on the surface of the spherical YS facilitates its dissection, allowing it to be replaced with the embryo of another species *in ovo*. When such manipulations were carried out, it was noted that there were a rapid reconstitution and joining of blood vessels between the grafted embryo body and the host YS and that the free circulation of blood cells was established between the YS and the body, without any disruption to normal development in the resulting chimeras. The quail/chick chimera consisted in grafting a quail embryo on a chick YS at a very early stage of development, before the onset of heart beating Figure 18 The embryos resulting from this operation possessed hemopoietic organs in which the cells belonging to blood lineages in the thymus, spleen, bursa, and BM were all quail while stromal cells were chicken (Dieterlen-Lievre, 1975 ; Martin et al., 1978). Circulating blood was chicken (derived from YS progenitors) until day 5 of incubation, then became mixed, and then richer and richer in quail erythrocytes (Beaupain et al., 1979). The replacement of erythrocytes derived from YS progenitors by these derived from intraembryonic HSC was strikingly confirmed when the chimeras were built between congenic strains of chickens, differing in their immunoglobulin allotypes or in their major histocompatibility antigens (Lassila et al., 1978). These results led to the indisputable conclusion that YS progenitors give rise to a short-lived progeny, hence cannot be considered as long-life hematopoietic stem cells, since this term specifies the property of long-term self-renewal. The extinction of YS progenitors may result from the differentiation pressure resulting from the exponential need for red cells in the rapidly developing embryo. Indeed, associated in vitro with an attractive thymic rudiment, the extraembryonic area (i.e. future YS) was capable of providing lymphoid progenitors.

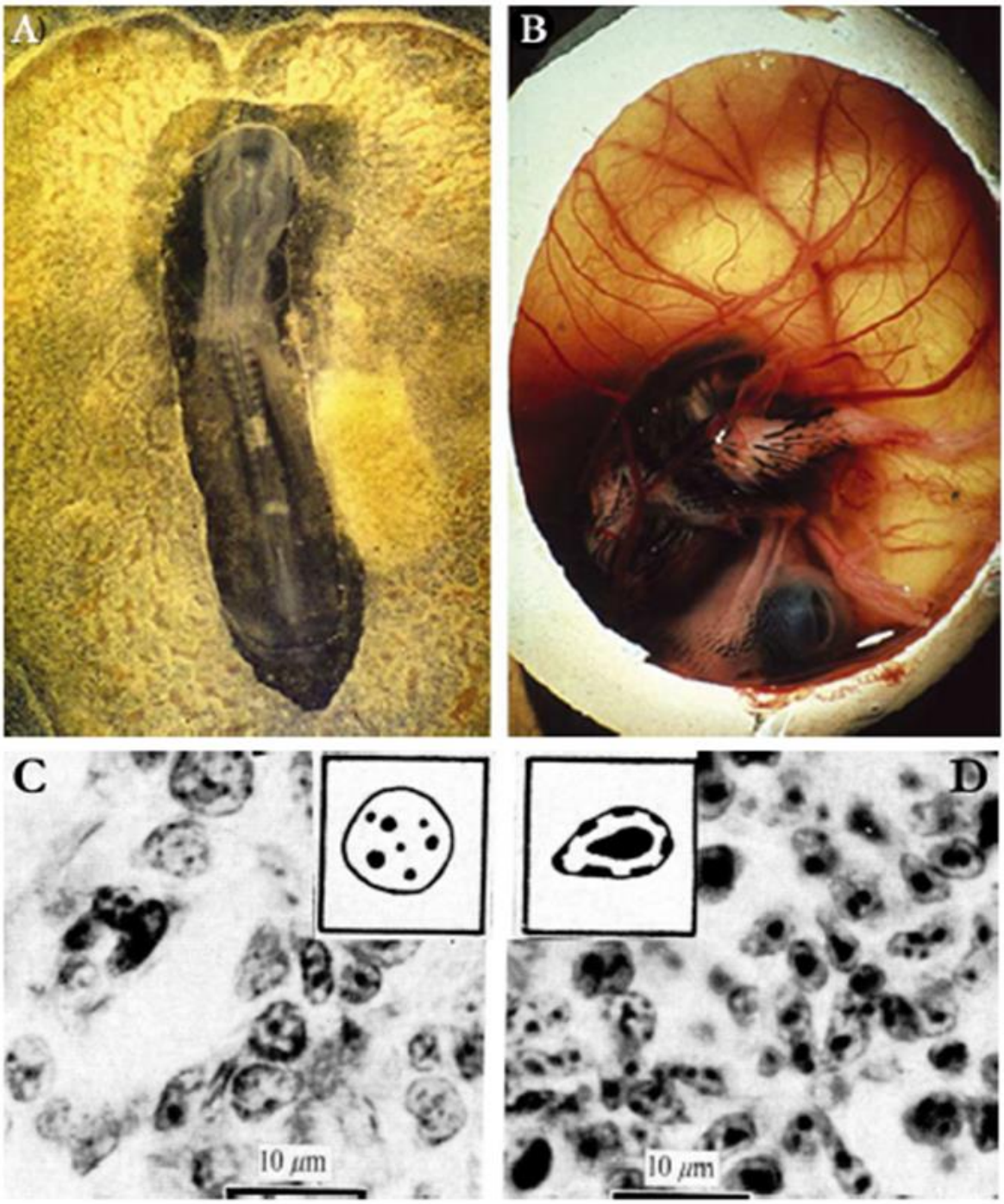


Figure 18: Quail–chicken yolk sac chimera
A*- Immediately after surgical construction*
B*- E13 chimera developed in ovo. The quail embryo (with typical pigmentation) is associated to a chicken yolk sac.*
C*- Erythropoiesis in the chick embryo spleen, 13 days. Early erythroblast nuclei show fine lumps of chromatin, regularly distributed, as schematized in the frame*
D*- Early erythroblasts have Feulgen-positive material in a heavy central mass and smaller peripheral patches in their nuclei (schematized in the frame)*
(Images from Dieterlen-Lièvre,1975 and Dieterlen-Lièvre & Le Douarin 2004)

It was clear that another source was responsible for the production of hematopoietic progenitors that seeded the hematopoietic organs as well as for HSCs. Intra-embryonic clusters had been described from the beginning of the 20th century in both avian and mammalian species as groups of hematopoietic cells, located on the aortic floor, protruding into the aortic lumen (Dieterlen-Lièvre et al., 2006) Figure 19. This region thus appeared as an interesting candidate to produce the cells that seeded the definitive hematopoietic organs. In quail/chick chimeras, these intra-aortic clusters are quail, i.e. they have an *in situ* intra-embryonic origin (Dieterlen-Livre and Martin, 1981). Moreover, their affinity for anti-ITAG2b integrin (also known as CD41) antibody authenticates their hematopoietic progenitor nature (Corbel, 2002). Indeed the peri-aortic region of the chicken embryo, dissociated into single cells and seeded in a semi-solid medium, gave rise to 3–4 times more colonies (erythroid or macrophages) than the BM from a newly hatched animal (Cormier and Dieterlen-Lievre, 1988).

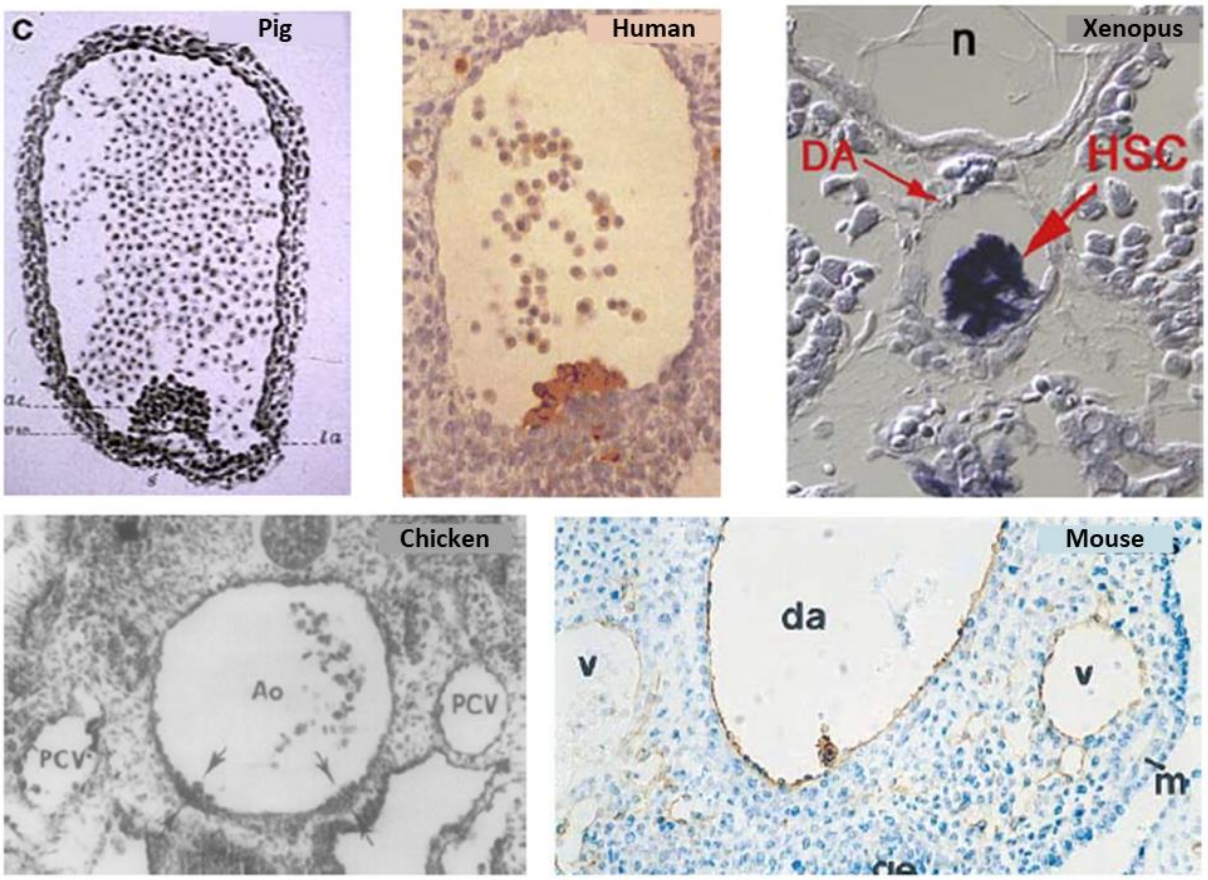


Figure 19: Hematopoietic clusters in the aorta of different species
Illustrations respectively from : Emmel et al 1916, Tavian et al 1996, Ciau-Uitz et al 2010, Dieterlen-Lièvre et al 1981 and Wood et al 1997

a. Mouse Aorta

In 1993, two groups, based on bird results, identified the aorta as a major site of HSC emergence in mice (Medvinsky et al., 1993 ; Godin et al., 1993). Prior observations had made it possible to highlight the embryological origin of the aorta: between E7 and E8.5, the presumptive territory of the aorta corresponds to the splanchnopleural mesoderm located in the caudal region of the embryo (Cumano et al., 1996). From E8.5, the splanchnopleure (Sp), located in the splanchnopleural mesenchyme and associated endoderm, evolves into a trunk and abdominal structure including the paired aortas, the subaortic mesenchyme, and the endoderm of the digestive tract, this structure is called Para-aortic Splanchnopleura (P-Sp) (Godin et al., 1995).

This P-Sp develops at E10.5- E11.5 to form the Aorte-Gonad-Mesonephros (AGM) region comprising the dorsal aorta (resulting of the merge of the paired aortas), its underlying mesenchyme, and the urogenital system (Medvinsky et al., 1993). It is within this AGM that we can observe between E10 and E11.5 the budding of the clusters of round cells in the light of the aorta from the endothelium Figure 20.

These structures are similar to those observed in Chickens and are mainly observed at the ventral wall of the aorta. On rare occasions, however, hematopoietic buds are seen in the vitelline arteries or in the dorsal position of the aorta, the latter marking here a difference from the situation observed in the bird (de Bruijn, 2000). Evidence of the aortic origin of the final HSCs was provided through experiments of P-Sp or YS transplants taken prior to the establishment of circulation between the YS and the embryo under the SCID mouse renal capsule (Godin et al., 1993).

The authors observed that only the aortic region allowed a reconstruction of the lymphoid lineage of the host. At the same time, another group showed that the AGM carries a CFU-S activity initiated by E9 and which peaks at E10 and then declines. Concurrently, CFU-S activity increases in the fetal liver, suggesting that hematopoietic progenitors emerging at the aorta subsequently colonize the fetal liver (Medvinsky et al., 1993).

The first progenitors are detected in the Sp while the circulation between the YS and the embryo is already established, which does not allow us to conclude on the

true site of emergence of the aortic HSC. The experiments conducted by Cumano et al. 1996 were able to respond to this question: the authors took before the establishment of the blood circulation, the YS, and the Sp and separately placed them in organotypic culture for 2 to 4 days before testing their hematopoietic potential (Cumano et al., 1996). The authors show that before establishing blood flow between the YS and the embryo, Sp contains progenitors capable of producing B and T lymphocytes. Experiments conducted by E. Dzierzak and A. Medvinsky in the same year complement these results (Medvinsky and Dzierzak, 1996). The authors show that cells from the AGM of mice at E11 are able to reconstitute the different hematopoietic blood lineages in irradiated adult mice, and are therefore capable of self-renewal. Similar experiments conducted prior to the establishment of blood flow between the embryo and the YS were conducted in 2001 and confirmed that HSCs of the AGM are produced in this tissue (Cumano et al., 2001). Whereas the culture of YS alone gives rise only to erythromyeloid progenitors that have reduced self-renew potential.

The embryo grown in the absence of YS and Sp has no hematopoietic activity: the intra-embryonic hematopoietic potential is therefore specifically restricted to the aortic region. Recent experiments with genetically marked YS transplants in mice showed that this tissue is capable of generating lymphocyte B progenitors, in very limited quantities: less than 1% (Sugiyama et al., 2007). However, no contribution of the YS to the T lymphoid population has been demonstrated: the contribution of the YS to adult hematopoiesis in the mouse, if it exists, is very limited.

The aortic hematopoiesis was subsequently slightly better characterized: 500 to 1000 progenitors of which only a dozen HSCs are generated between E8.5 and E13 in the AGM. The peak production is detected at E10.5-E11, and the quantity in the aortic region then decreases to E12.5-E13 (Godin and Cumano, 2002; Godin et al., 1995; Godin et al., 1999). The aortic region is mainly dedicated to the production of HSCs, multipotent progenitors capable of long-term self-renewal, the expansion and differentiation of these progenitors taking place in other hematopoietic sites, including

the fetal liver that becomes a major organ of hematopoietic expansion and differentiation (Ema and Nakauchi, 2000; Godin et al., 1999).

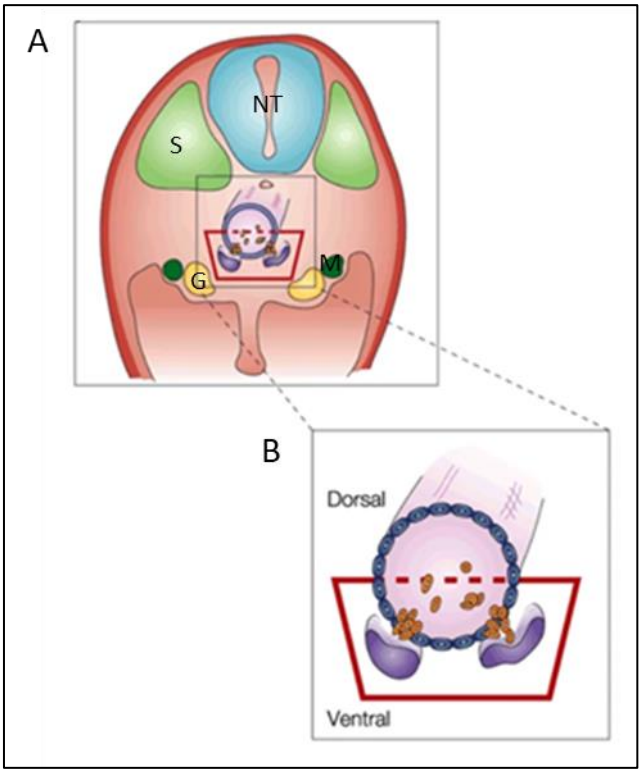


Figure 20: Location of intra-embryonic HSCs.

A- schematic of a transverse section through the embryo, showing the internal structure of the embryo at the level of the truncal AGM.

B- An enlargement of the aortic region, schematically showing the intra-aortic clusters, which are restricted to the ventral part (floor) of the vessel, and the sub-aortic patches. The area that has hemogenic activity is shown in a red box.

Abbreviations: AGM, aorta–gonad–mesonephros; NT: Neural Tube, S: Somite, M: Mesonephros, G: Gonad

Adapted from Godin & Cumano 2002

b. Amphibians aorta

In amphibians, removal of the ventral blood island resulted in the absence of circulating red blood cells. The origin of adult progenitors was hence first exclusively attributed to ventral blood island (VBI) (Federici, 1926). This hypothesis was then invalidated in the amphibian Rana pipiens by transplanting various territories subsequently identified according to their ploidy difference (Hollyfeild, 1966). These experiments first showed the existence of an alternative source of progenitors. When the cytogenetically labelled VBI was transplanted to the neurula stage, no larval erythrocyte came from the graft. This approach then made it possible to locate an alternative source of progenitors: if the dorsal mesodermal territory containing the pronephros is transplanted, erythrocytes from the graft are detected in the larva and in adults indicating that this region could be at the origin of the adult hematopoietic system. In agreement with this hypothesis, the lymphoid population of the larva is also derived from the dorsal region (Turpen et al., 1981; Turpen et al., 1983). The emergence of HSCs from the dorsal region was then demonstrated in Xenopus (Tompkins et al., 1980). The activity of the hematopoietic region is characterized by the presence of HCs in the pronephric tubules, the aorta and the cardinal veins (Turpen et al., 1981). The origin of these cells was precisely located in the Dorsal Lateral Plate (DLP) (Turpen & Knudson, 1982; Kau & Turpen, 1983). Data indicate that the DLP produces SCs contributing to larval and adult hematopoiesis. In particular, this territory participates to definitive erythropoiesis, thymic and hepatic hematopoietic activities (Bechtold et al., 1992 ; Kau & Turpen, 1983; Maneo et al., 1985a, b; Chen & Turpen, 1995). These results were later confirmed by two groups using dye injection (rhodamine dextran, GFP or βgalactosidase) into blastomeres and the follow up of the blastomere progeny (Ciau-Uitz et al., 2000; Lane and Smith, 1999). In particular, R. Patient's group was able to map the origin of different hematopoietic progenitors in stage 32 blastula cells. Regarding the ISV, the authors showed that the anterior part of the structure comes from blastomeres C1 and D1 while the posterior part comes from blastomere D4 (Ciau-Uitz et al., 2000). This study also highlighted that the dorsolateral plate (DLP) and the VBI lineages are already separated at a very early stage: only the C3 dorsal blastomere contributes to the formation of the DLP and the emergence of HCs in the aortic region of the embryo Figure 21. Nevertheless, a

more recent study suggests that embryonic and adult lineages could have a common origin in Xenopus (Lane and Sheets, 2002).

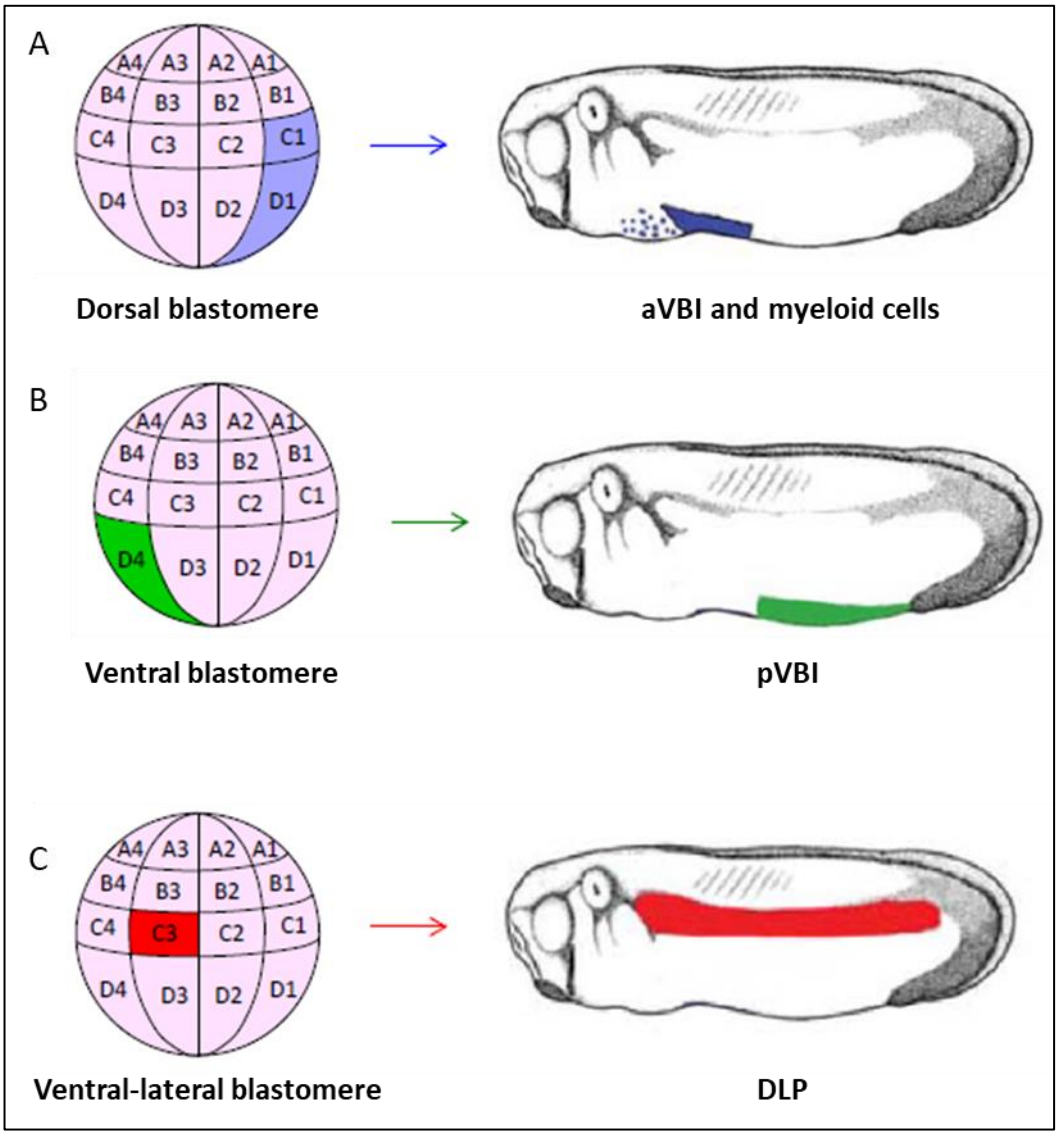


Figure 21: Ontogeny and hematopoietic potential of the different regions of Xenopus
A*-The aVBI and its precursor, the embryonic hemangioblasts, derive from dorsal blastomeres of the 32-cell stage embryo*
B*- Ventral blastomere D4 gives rise to the pVBI.*
C*-The C3 blastomere of the 32-cell stage embryo gives rise to the DLP which later generates HSCs in the ventral wall of the dorsal aorta.*
Abbreviations: VBI: Ventral Blood Island, DLP: Dorsal Lateral Plate, aVBI: anterior VBI, pVBI: posterior VBI.
Adapted from Ciau-Uitz et al 2010.

c. Zebrafish aorta

In developmental biology, and for the past two decades, the zebrafish model (Danio rerio) has been increasingly used for the study of HSCs due to its many advantages. In the same way as the Xenopus embryo, the fertilization of the zebrafish is external and almost 200 eggs are obtained each spawning. The body of these embryos is translucent, which facilitates phenotypic observations, especially when they are coupled with fluorescent markers. Finally, this model is particularly powerful for conducting genetic analyses, notably through the realization of large-scale screens and obtaining mutating lines.

The establishment of the hematopoietic system in the zebrafish is accomplished as in the avian model, via two waves, one primitive and the other definitive emerging from two distinct sites. The primitive wave occurs in the Inner Cell Mass (ICM) equivalent to the YS of birds. When the heart begins to beat (26 to 28 hours post-fertilization) erythrocytes and endothelial precursors within the ICM depart in circulation. At the same time, macrophages and granulocytes differentiate in an anterior area named Rostral Blood Island RBI (Herbomel et al., 1999; Lieschke et al., 2002).

Starting 26 hours post-fertilization, definitive hematopoiesis is initiated in the aortic region. Hematopoietic genes (runx1, c-myb) start to be expressed in the aortic endothelium reminding the situation observed in the bird (Thompson et al., 1998)

In 2002, runx1, the key gene in HSCs specification in vertebrates was described in zebrafish (Kalev-Zylinska, 2002). The hematopoietic potential of the aortic region was revealed in 2006. The group of P. Herbomel was able to establish that progenitors from the aortic region migrated via blood flow to an area called Caudal Hematopoietic Tissue (CHT), a probable equivalent of the para-aortic foci of birds, or the fetal liver of the mouse before colonizing the final hematopoietic organs (Murayama et al., 2006). The first lymphoid progenitors were detected at E3 in the thymus. At E5, they are found in the pronephros, equivalent of the BM (Zapata, 1979) Figure 22, Figure 23.

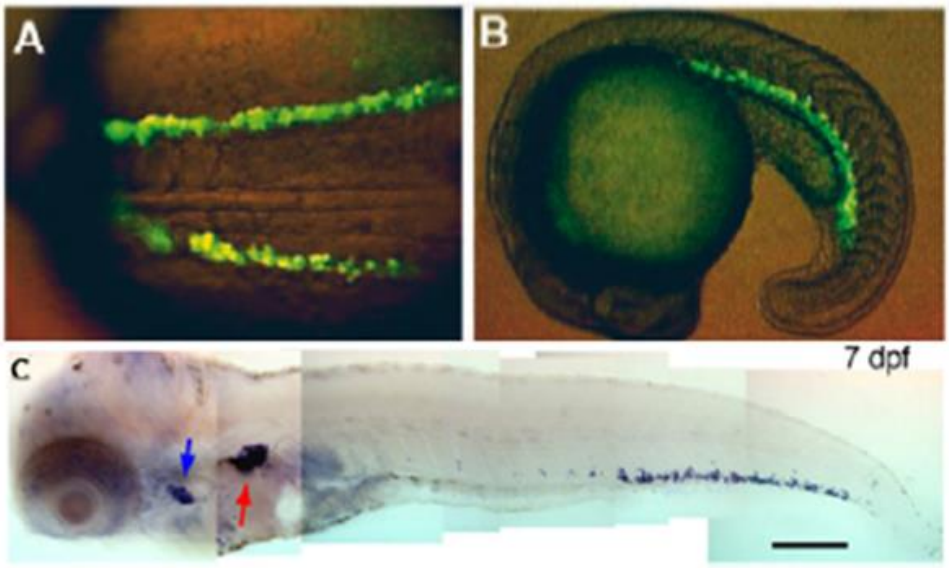


Figure 22: Hematopoietic sites in zebrafish
A,B*-.GFP expression driven by the gata 1 promoter in a 12 hour and a 22 hour transgenic zebrafish embryo, respectively reporting the hematopoietic activity of the ICM.*
C*- c-myb Expression Reveals the Sites of Larval and Definitive Hematopoiesis.*
Blue arrows, thymus; red arrows, pronephros
Adapted from Long et al 1997 and Murayama et al 2006.

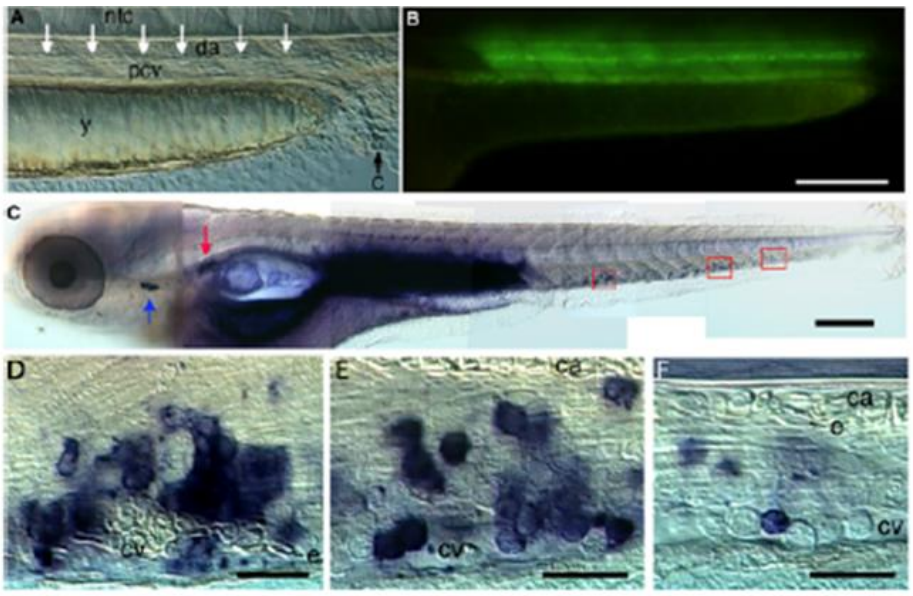


Figure 23: The aorta a source of hematopoietic progenitors
A*-arrows point at the DA-PCV joint targeted with the laser.*
B*- Fluorescence of fluorescein just laser-uncaged all along the DA-PCV joint*
C*-Cells marked by laser-uncaged migrate to the thymus (blue arrow) and to the pronephros (red arrow) and form cell aggregates in the CHT*
D-E-F*: Magnifying of C showing the progenitors in the CHT*
Abbreviations: CHT: caudal hematopoietic tissue, PCV: posterior cardinal vein, DA: Dorsal aorta, cv: caudal vein, ca: caudal artery, Y: yolk sac, c: cloaca, ntc: notochord.
Adapted from Murayama et al 2006

d. The aortic endothelium

i. Splanchnopleural origin of the hemogenic endothelium

During gastrulation, the mesodermal layer deposits between the ectoderm and the endoderm. It then regionalizes according to the embryonic axes, each region having different potentials and fates. Medio-lateral regionalization of the mesoderm takes place resulting in the formation of different compartments: notochord, somites, intermediate mesoderm, and lateral plate mesoderm. The latter therefore comprises the embryonic mesoderm as well as the extraembryonic mesoderm.

During its formation, the coelom divides the lateral mesoderm into two layers: a dorsal one designated as the somatopleure comprises the somatopleural mesoderm associated with the ectoderm; a ventral one designated as the splanchnopleure comprises the splanchnopleural mesoderm associated with the endoderm. Hematopoietic cells and a cohort of endothelial progenitors arise from the splanchnopleure. The blood islands are formed from the extra-embryonic splanchnopleural mesoderm and associate endothelial and hematopoietic progenitors. In the embryo proper, the splanchnopleural mesoderm gives rise to angioblasts which form capillaries that will merge to form a vascular network closely associated with the endoderm. The vascular system of the viscera and the aorta are formed this way. This origin of blood and endothelium in the different embryonic tissues was demonstrated by quail/chicken grafts experiments followed by QH1 (a marker of endothelial and hematopoietic quail cells) staining Figure 24. Results differed according to the germ layer constitution of the grafted rudiments. In the case of the limb buds and body wall, ECs from the host invaded the graft through an angiogenic process. Hematopoietic progenitors from the host also colonized the grafted BM. In contrast, rudiments of internal organs provided their own contingent of endothelial precursors, a process termed vasculogenesis. Nevertheless, HCs in these organs were all derived from the host (Pardanaud et al., 1989).

Seeking the origins of ECs Pardanaud and co-workers traced the distribution of $QH1^+$ cells after replacement of the last somite or heterotopic grafts of lateral plate mesoderm. When the authors grafted quail somitic tissue on a chicken host, they found $QH1^+$ cells in the endothelium of somatopleural derivatives, i.e the body of the

embryo, limbs and renal endothelium. When the splanchnopleural mesoderm was grafted, quail QH1^{+} cells contributed to the formation of the vascular system of the viscera as well as to the ventral endothelium of the aorta (Pardanaud et al., 1996).

The results lead to the conclusion that the embryo becomes vascularized by endothelial precursors from two distinct regions, splanchnopleural mesoderm and paraxial mesoderm i.e., somites. The territories respectively vascularized are complementary, precursors from the paraxial mesoderm occupy the body wall and kidney, i.e., they settle along with the other paraxial mesoderm derivatives and colonize the somatopleure. Interestingly, ECs originating from the somites are barred to contribute to vascularization of the internal organs and from initially integrating the aortic floor thus never giving rise to hematopoiesis. ECs originating from the splanchnopleural mesoderm colonize internal organs and contribute to hematopoiesis (Pardanaud et al., 1996). Finally, the endothelial and hematopoietic potential of these two mesodermic tissues can be flip-flopped by prior culturing them with endoderm or ectoderm or by treatment with several growth factors (Pardanaud and Dieterlen-Lievre, 1999). The endoderm induced a hemangiopoietic potential in the associated mesoderm. Indeed, the association of somatopleural mesoderm with endoderm promoted the ‘ventral homing’ and the production of hemopoietic cells from mesoderm not normally endowed with this potential. The hemangiopoietic induction by endoderm could be mimicked by VEGF, bFGF, and TGFβ1. In contrast, contact with ectoderm or EGF/TGFα treatments totally abrogated the hemangiopoietic capacity of the splanchnopleural mesoderm, which produced pure angioblasts with no ‘ventral homing’ behavior.

Given the difficulties encountered during in vitro mouse embryo cultures, an interspecific grafting approach has been used to demonstrate the endothelial potential of different mesodermal regions in mice (Fontaine-Perus and Jarno, 1995). Mouse somites transplanted orthotopically in a chicken embryo produce ECs that colonize the neural tube, limbs, mesonephros, and roof of the aorta (Ambler et al., 2001).

The splanchnopleural origin of hematopoietic precursors may extend to other sites of embryonic hematopoiesis. In fact, the yolk sac progenitors emerge from the splanchnopleural layer of the lateral plate, and the mesoderm involved in the formation of the allantois also corresponds to mesoderm associated with endoderm, therefore having a splanchnopleural identity (Caprioli et al., 1998; Caprioli et al., 2001).

Interspecific grafting experiments have shown that, as soon as the coelom is formed, somatopleural and splanchnopleural mesoderms have different potentials (Pardanaud and Dieterlen-Lievre, 1993). The hematopoietic potential would, therefore, be restricted to ECs of splanchnopleural origin.

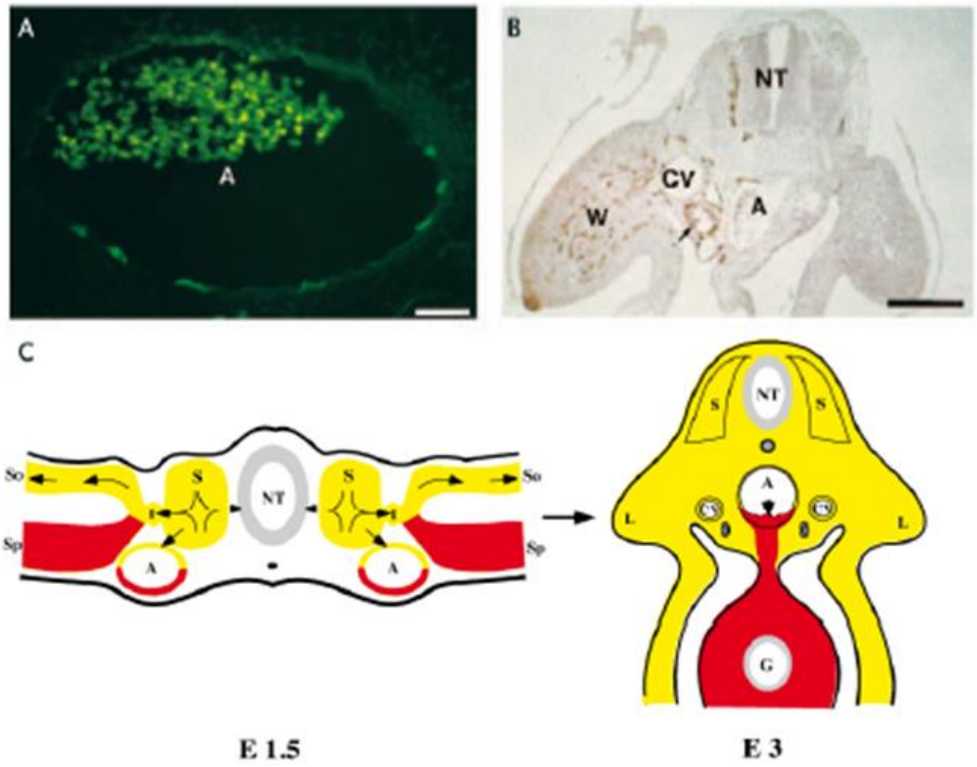


Figure 24: Mapping of the two endothelial lineages
A- *3-day chick host, 2 days after quail splanchnopleural mesoderm grafting. QH1+EC are integrated in the floor of the aorta*
B- *3-day chick host, 2 days after the graft of quail rostral segmental plate. QH1+EC have colonized the ipsilateral body wall of the chick host; they participate in the vascularization of the neural tube, somitic derivatives, wing bud , mesonephros (arrow) and roof of the aorta*
C- *The aorta appears to have a mosaic origin, roof and sides from the somite, floor from splanchnopleural mesoderm.*
Abbreviations: CV, cardinal vein; G, gut; L, limb bud, NT: Neural Tube, W: Wing bud, S: Somite, So: somatopleural mesoderm, Sp: splanchnopleural mesoderm.
Illustartions from Pardonaud et al 1996

ii. The aortic endothelium remodeling by cells of somitic origin

Experimental approaches by Pardanaud and collaborators indicate a double origin for the ECs that form the dorsal embryonic aorta (Pardanaud et al., 1996). Ventral cells with a hemogenic potential are of splanchnopleural origin whereas those in the dorsal position have a somitic origin. This demonstration was confirmed and completed in the team (Pouget et al., 2006; Pouget et al., 2008). While the work by Pardanaud et al., involved transplantation of a single somite, Pouget et al. performed grafts of the pre-somitic mesoderm (PSM) (material corresponding to a length of 10 somites) and analyzed the contribution of the graft to the formation of the aortic endothelium at different stages (Pouget et al., 2006). In sum, during their formation, the paired aortas are uniquely composed of ECs of splanchnopleural origin. From E2 in the bird embryo, while the aortas are still paired, somitic ECs migrate into the roof and sides of the aorta and replace the initial roof and sides of splanchnopleural origin Figure 25. Thus, when the aortas fuse into a single vessel, the ECs derived from the splanchnopleura are restricted to the aortic floor hence restricting the blood-forming ability to the ventral side of the aorta. The second phase of aortic remodeling occurs during and after aortic hematopoiesis. At the time of hematopoietic cluster formation, ECs from the somite begin to migrate and locate immediately underneath the clusters. Since hematopoietic cells are produced at the expense of splanchnopleural ECs, cells of somitic origin progressively replace the ECs of the floor ensuring the integrity of the vessel during the phase of HSC emergence. At the completion of hematopoiesis, the aortic floor of splachnopleural origin has been replaced by ECs originating from the somite. Hence within a short period of time, the origin of the aortic endothelium has changed from splanchnopleural to somatopleural (somite). In addition to replacing the aortic endothelium, the somite was also shown to give rise to the aorta smooth muscle cells. Interestingly, this aortic cell layer derives from cells of the ventral side of the sclerotome that turn their sclerotome fate and acquire a smooth muscle cell fate. In particular two sclerotome-specific markers Pax1 and FOXC2 are shortly expressed by the forming smooth muscle cells and are down-regulated as the cell layer forms (Pouget et al., 2008; Wiegreffe et al., 2007). In the mouse, the study of FLK1 expression shows that in the dorsolateral position, a cell population exists expressing this characteristic marker of ECs (Ema et al., 2006).

More recently, during orthotopic grafts of mouse somites in a chicken embryo followed by tracing the fate of ECs, there is also a gradual and complete replacement of the aortic endothelium starting with the roof of the vessel (Yvernogeau et al., 2012). Smooth muscle cells forming the tunica of the aorta are also established in two successive stages: cells originating from the lateral plate and placed under the aorta differentiate under the vessel, then a second generation of smooth muscle cells of sclerotomous origin forms the final tunica around the aorta concomitantly with the replacement of ECs (Wasteson et al., 2008; Wiegreffe et al., 2009).

This set of experiments shows that it exists at the time of hematopoiesis a complete replacement of the aortic endothelium of splanchnopleural origin by ECs of somitic origin. This total replacement suggests that all the aortic ECs of splanchnopleural origin have a hematopoietic destiny: the tissue origin seems to play a fundamental role in the hemogenic potential of the ECs.

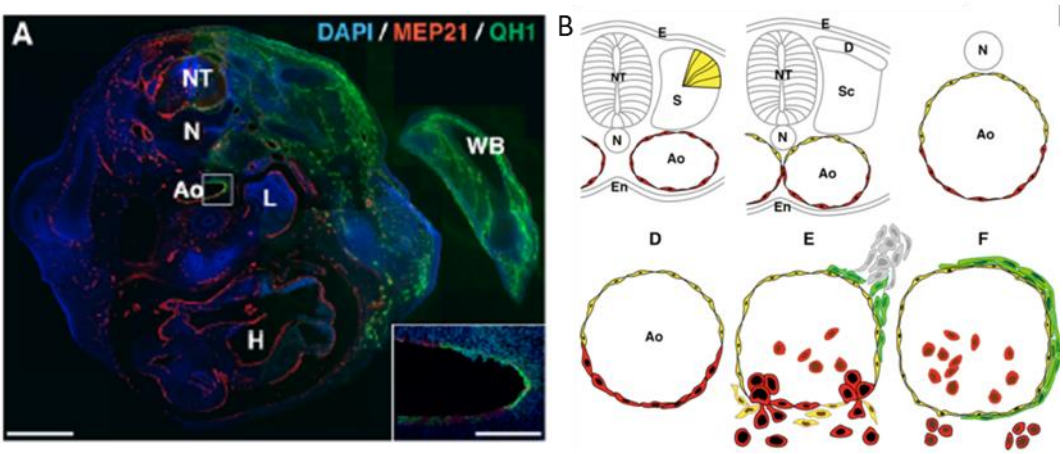


Figure 25: Developmental history of the aorta in relationship with endothelial remodeling
A- *E6.5 chick host, 5 days after the graft. Cross-section at the trunk level. Triple labeling with QH1 (green)/MEP21 (red)/DAPI (blue). The vascular networks of the neural tube, dorsal root ganglia, dermis, mesonephros, body wall and limb bud derive entirely from the graft. Inset: closer view of the right side of the aorta. The aortic endothelium on the grafted side mostly derives from the graft*
B- *Recap of the dynamics of the aorta*
Abbreviations: NT: Neural Tube, N: Notochord, Ao: Aorta, L: Lung, WB: Wing Bud, H: Heart, E: Ectoderm, En: Endoderm, D: Dermomyotom, Sc: Sclerotom.
Illustartions from Pouget et al 2006

D. Other hematopoietic sites: head and early cardiac endothelium

In addition to the aorta, HSCs were suggested to emerge from the mouse head between E10.5 to 11.5 paralleling the AGM. The authors visualized hematopoietic clusters and demonstrated the blood-forming potential from sorted ECs. A recent report confirmed the presence of HSCs and progenitors in the mouse head but failed to detect hematopoietic clusters. Finally, several older reports were in stark contrast with several other studies in the mouse and human embryos, which could not find any HSC activity in the embryonic head (Muller et al., 1994; Kumaravelu et al., 2002; Ivanovs et al., 2011).

Ultimately, a study uncovered a hemogenic activity in the early cardiac endothelium in the mouse. The authors used $Ncx1^{-/-}$ embryos that lack heartbeat and found a hematopoietic-forming potential in a subset of ECs expressing the CD41antigen. These CD41+ cells arise from Nkx2.5+ cells i.e., cardiac cells. This hematopoietic activity was detected in the E8.5-E9.5 embryo anticipating the AGM activity. The hematopoietic production resembles EMP produced in the yolk sac (Nakano et al., 2013).

E. The Fetal liver

The fetal liver FL, of mesodermal and endodermal origin, is the major organ of amplification and differentiation of embryonic HSCs. For the moment, researchers showed that the FL is not capable of producing HSCs de novo and is simply colonized by circulating HCs of extrinsic origin starting E10-E10.5. (Johnson and Moore, 1975; Houssaint, 1981).

The FL is colonized by two successive waves of HCs: the first wave establish a connection between the YS and the embryo via the vitelline vessels and are mainly enriched with primitive and definitive hematopoietic precursors. Since the BM is not yet formed at the time the FL begins to be colonized, the primary hematopoietic role of the FL is to form mature erythrocytes from primitive and definitive YS-derived precursors to ensure the needs of the embryo. The FL contains “erythroblast islands” that also exist in the BM. Those islands consist of a central macrophage that extends

cytoplasmic protrusions to a ring of surrounding erythroblast and are the site of production of all the erythrocytes (Chasis, 2006; Palis, 2004) Figure 26.

The second wave of hematopoiesis in the FL begins at E11.5 and corresponds to populations enriched in hematopoietic progenitors/immature HSCs mainly from the AGM and the Placenta via the umbilical vessels (the second vascular network pathway of HCs). At this stage, FL HCs are capable of forming in vitro BFU-E and CFU-E colonies and are able to proliferate for several days and to generate definitive erythroid cells.

At E12.5, the CFU-S activity of the FL begins to be detected, when the AGM has practically no more activity of this kind. Between E12.5 and E15.5, the total number of definitive hematopoietic precursors/HSCs increases exponentially in the FL, doubling practically every day. The total number of myeloid and lymphoid B precursors sequentially increases till near-birth stage (Barker et al., 1969; Gunji et al., 1991; Paige et al., 1984). Concerning the number of T-lymphoid precursors, they reached a peak at E13 and then decline (Ema et al., 1998). Finally, the number of HSCs increases sharply to E16 (the stage in which the FL contains approximately 1,600) and then decreases gradually (Ema and Nakauchi, 2000; Gekas et al., 2005; Morrison et al., 1995). The results of Kurata et al. show that a small number of hematopoietic precursors subsist for up to five weeks after birth (Kurata et al., 1998), these results are correlated to Taniguchi et al. work which showed that the adult liver contains immature HCs capable of restoring the hematopoietic system of a lethally irradiated mouse (Taniguchi et al., 1996). Several studies have shown that the HSCs present in the FL possess a very strong proliferative capacity compared to the HSCs of the adult BM (Micklem et al., 1972; Rebel et al., 1996a a). However, the frequency of HSCs of FL at E14.5 is comparable to that of the adult BM (Rebel et al., 1996b). Assuming that the AGM is the only real site for the emergence of definitive HSCs and since the AGM at E12.5 is intended to contain only 3 definitive HSCs (Kumaravelu et al., 2002), it was however difficult to explain the large number of HSCs found in the FL at E14.5 only by the speed of cell division. Thus just as the study of the emergence of HSCs in the YS, researchers are looking for the emergence and amplification capacity of the Placenta.

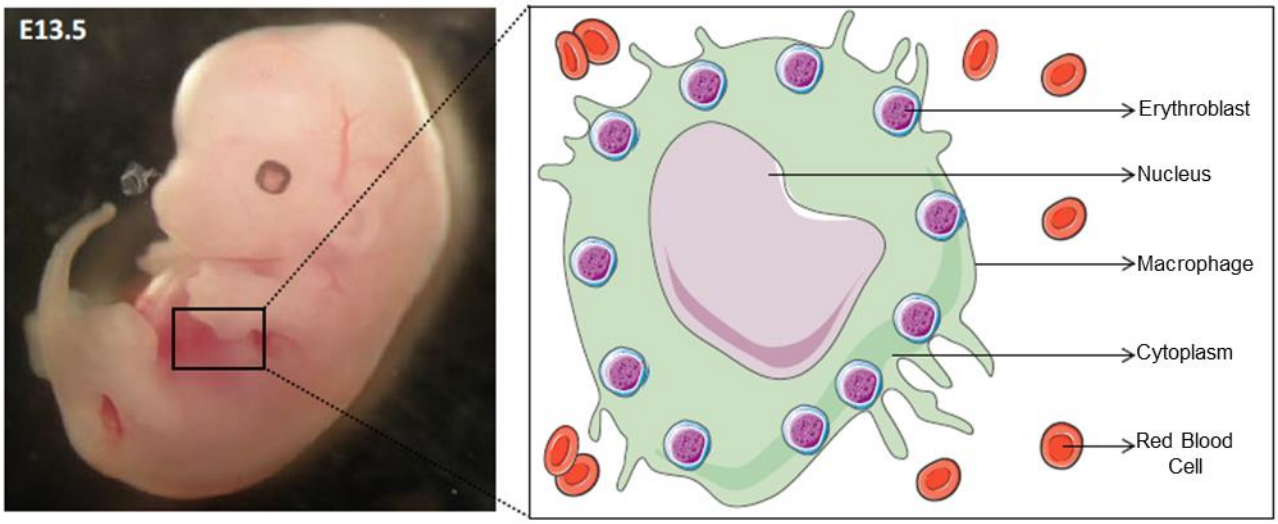


Figure 26: Schematic representation of an erythroblast islands of the FL at E13.5
Red blood cells are produced in the FL hence its red color. The maturation of erythroblasts takes place in erythroblastic islands. Each of them has a macrophage at its centre, in contact with a ring of erythroblasts, immature nucleated red blood cells.
Abbreviations: FL: Fetal Liver.
Adapted from Palis et al 2004 and Chasis et al 2006

F. Definitive hematopoietic organs

Starting E15.5, the frequency of the HSCs present in the FL begins to decrease (Morrison et al., 1995). This decrease is due to the gradual migration of HCs/HSCs from FL to the definitive hematopoietic organs: the thymus, the spleen and finally the BM, where HSCs are maintained during a lifetime. These three organs possess specific microenvironments in which the lymphocyte T differentiation, lymphocyte B differentiation, and the maintenance/differentiation of the HSCs are supported respectively.

a. The thymus

Since HCs capable of colonizing the thymus are present in the blood several days before they enter the thymic rudiment, the onset of stem cell inflow into the thymus is dependent upon maturation of the thymic epithelium, which acquires a specific "receptivity" for lymphocyte precursor cells at a precise stage of development. In early thymic ontogeny, the "receptivity" for hematopoietic cells falls sharply and probably stops completely for a transient period.

Two teams showed that the thymic rudiments are colonized by lymphoid precursors from E10.5 (Fontaine-Perus et al., 1981 ; Owen and Ritter, 1969). During

the following days, the colonization of the thymus accelerated and the first mature lymphocytes appeared at E15. The E12-E16 thymus transplants show the existence of a second wave of thymic colonization. The resulting mature HCs are the only ones that will subsist in the thymus thereafter (Mathieson and Fowlkes, 1984).

b. The spleen

The spleen doesn't have the capacity to generate HSCs. It is colonized by HSCs coming from the FL. The microenvironment of the spleen supports hematopoietic differentiation, but unlike the FL, it is not capable of amplifying HCs.

The first HCs to colonize the spleen are multipotent (lymphomyeloid progenitors) and are detectable starting E12.5-E13 (Godin et al., 1999). The work of Christensen et al. demonstrates, through hematopoietic reconstitution experiments, that LT-HSCs are detectable starting E14.5 (Christensen et al., 2004). The hematopoietic activity of the spleen then increases daily until E17.5 and remains detectable up to 2 weeks after birth. In cooperation with FL and BM, the spleen plays an important role in maintaining the homeostasis of the hematopoietic system (Wolber et al., 2002).

c. The Bone Marrow

The first colonization of the BM occurs starting E15.5 by the progenitors followed by the colonization starting E17.5 by the HSCs (Christensen et al., 2004; Delassus and Cumano, 1996). As we discussed earlier for the thymus, the HSCs able to colonize the BM are present in the blood several days before the colonization, suggesting that the BM microenvironment needs certain maturation before receiving the HSCs. A cascade of chemotactic interactions governs the migration of progenitors / HSCs to the BM. The Stromal Cell-Derived Factor-1 (SDF-1) chemokine, also known as CXCL12, seems to play a particularly important role. Mice deficient for SDF-1 or its receptor CXCR4 (Chemokine Receptor type 4) fail to establish hematopoietic activity in BM although the FL develops normally (Ara et al., 2003; Nagasawa et al., 1996; Zou et al., 1998). In addition, it has been shown that HSCs of the BM migrate selectively in vitro in response to SDF-1 (Wright et al., 2002). In humans, a study has shown that in response to SDF-1, CD34 + BM cells injected into lethally irradiated mice have an increased hematopoietic reconstitution potential (Plett et al., 2002). Thus,

SDF-1 plays an important role both as a chemokine (Imai et al., 1999), as an activator of HSC adhesion molecules (Peled et al., 1999; Peled et al., 2000) and as a molecule involved in the maintenance of embryonic HSCs when they arrived in a hematopoietic niche (Kawabata et al., 1999; Ma et al., 1999).

2. The establishment of human hematopoiesis

Several aspects of hematopoietic development of mice are conserved in humans Figure 13.Human HSCs have been extensively studied in newborns and adults through transplantation into immunodeficient mice (Conneally et al., 1997; Wang et al., 1997), but much less in the embryo. Despite the current difficulties in conducting research on the human embryo, several teams have studied the AGM's HSCs, the FL, and the placenta both in vitro by LTC-IC tests, and in vivo through hematopoietic reconstructions in the NOD-SCID mouse. The study of several animal models is therefore essential to have an overview of the embryonic hematopoietic process during evolution, as well as to better understand these processes in human.

In Human as well as in mouse, hematopoiesis begins with a first wave of primitive HCs appearing prior to the introduction of blood circulation at the level of the blood islands of the YS, visible between E17-E18 and E24 (Luckett, 1978; Tavian et al., 1999). This hematopoietic population is transient since it contains only progenitors with erythro-myeloid potential. However, several types of HCs were detectable in the YS later during development, and between E25 and E50, YS contained immature HCs capable of forming myeloid and erythroid colonies *ex vivo* (Huyhn et al., 1995 ; Dommergues et al., 1992), suggesting that it still possesses hematopoietic activity *in vivo* at this time of embryonic development. The frequency of these precursors remains constant between E35 and E50, but no hematopoietic activity is detectable after E60.

Unlike the YS, the aortic region contains multilineage HCs (lymphoid and myeloid), and these cells were initially characterized *ex vivo*. Starting E28 of development, the embryo contains GMPs capable of forming hematopoietic colonies in clonogenic tests (Coulombel et al., 1995)) and these progenitors are $CD34^+$ (Huyhn et al., 1995). The work of Tavian et al. shows that these intra-embryonic hematopoietic progenitors are adjacent to the ventral wall of the aorta (Tavian et al., 1996) and can produce HCs for 4 to 5 weeks when cocultured with Sys-1 murine stromal cells (Tavian

et al., 1999). This population has a specific phenotypic profile (Lin^- $CD45^+CD34^+CD31^+ckit^+CD38^-$) and molecular profile ($GATA\text{-}2^+GATA\text{-}3^+c\text{-}myb^+Scl^+flk1^+$) characteristic of HSCs (Labastie, 1998; Tavian et al., 1996).

In 2001 the Péault team showed using culture explant followed by coculture on MS5 stromal cells, that hematopoietic progenitors are detected in the aorta region starting E19. Unlike the immature HCs of the YS, which can produce only myeloid cells, the aorta also contains cells capable of producing lymphoid HCs. Starting 2011, Ivanos et al tested the hematopoietic potential of these cells in vivo. Using the reconstitution assays, they showed that the AGM is the only site of definitive HSC emergence between E35 and E44. A single HSCs of the AGM is able to produce at least 300 HSCs on a period of 8 months (Ivanovs et al., 2011).

The hematopoietic potential of the human placenta has been assessed using the flow cytometry analysis and the confocal microscopy (Barcena et al., 2009). Researchers were able to demonstrate the existence, in the placenta, of a population $CD34^+CD45^+$ of fetal origin expressing a combination of markers specific of hematopoietic progenitors and able to produce erythroid, myeloid and lymphoid B populations. Therefore, they suggested that in human as well as in mouse, the placenta is a major hematopoietic site.

The colonization of the definitive hematopoietic organs starts with the colonization of the thymus, between wk 7 and 8.5 of fetal gestation, by the first lymphocyte T precursors (Haynes et al., 1988). Regarding the spleen, the first studies showed that it receives the first HCs between wk 18 and 19 (Vellguth et al., 1985 ; Hann et al., 1983). While new studies showed that the first erythrocytes are present in spleen starting Wk 11 or 12 (Steiniger et al., 2007).

The BM, the principal hematopoietic organ of the adult mammals, starts to be colonized when the hematopoiesis in the YS is finished and the FL one is still present. Using immunohistochemistry, Charbord et al, studied the establishment of hematopoiesis in the medullary cavities which begins during the 11^{th} SG (Charbord et al., 1996). The authors indicate that the $CD15^+$ myeloid HCs are the first to differentiate, followed closely by erythrocytes. According to the *ex vivo* and *in vivo* results of the Péault team, the HSCs of the BM between 18 and 24 wk possess the phenotype thy-1^+ Lin-$CD34^+$ (Baum et al., 1992).

III. The Endothelial to Hematopoietic Transition EHT

As described in the previous chapter, the AGM region is the primordial site of HSCs emergence. The hematopoietic region of the aorta is characterized systematically by the presence of clusters of round cells attached to the endothelial wall budding in the light of the vessel and this is the case of the different studied animal models Figure 20. This process is stereotyped in space and time and researchers aim in understanding the mechanisms governing the onset of the intra-aortic budding.

1. The Hemangioblast concept

Mapping studies have shown that the extraembryonic mesoderm comes from the posterior lateral plate (LP) (Schoenwolf et al., 1992 ; Psychoyos and Stern, 1996; Kinder et al., 1999). It is within this extraembryonic mesoderm that the earliest signs of vertebrate hematopoiesis occur. Based on the pioneering work of Sabin, Maximov and others in birds, Murray coined the term hemangioblast to indicate the thickening of the mesoderm in the chick yolk sac, the mesodermal “masses” located at the sites where later the blood islands emerge (Murray, 1932). It is worth noting that Murray defines the hemangioblast as a population of cells, which, as a whole, gives rise to the blood and endothelium of the blood islands. At the time, he could not say whether extrinsic cues determine the fate of individual hemangioblast cells or whether inherently different precursors for the 2 lineages coexist within his hemangioblast. In birds, the mesoderm from this region colonized the *area opaqua* by extending laterally and anteriorly. The expression profile of the Fetal Liver Kinase FLK-1 (VEGFR2) has been described in detail during this process. FLK-1 is strongly expressed in cells exiting the posterior LP and marks forming blood islands in the yolk sac. When the islands mature, receptor expression becomes restricted to EC (Dieterlen-Lièvre et al., 2006; Eichmann et al., 1993). The earliest experimental proof of the existence of the hemangioblast was brought by the deletion of VEGFR2 by gene targeting. In the absence of VEGFR2, both endothelial and HCs were absent in homozygous null mice (Shalaby et al., 1995). This observation suggested that VEGFR2 could be expressed by the hemangioblast and essential for its further differentiation along both lineages. Moreover, clonal cultures of $VEGFR2^{+}$ cells isolated from the posterior primitive streak of the avian embryo give rise to either a hematopoietic or an endothelial cell colony.

The developmental decision appears to be regulated by the binding of two different VEGFR2 ligands. Thus, endothelial differentiation requires VEGF, whereas hematopoietic differentiation occurs in the absence of VEGF and is significantly reduced by soluble VEGFR2 (Eichmann et al., 1997). At the same time, a clonal mesodermal precursor for blood and endothelium was identified in Embryonic Stem (ES) cell differentiation cultures (Choi et al., 1998; Nishikawa et al., 1998b). Termed blast colony-forming cell (BL-CFC) by G. Keller and co-workers, this cell was thought to be the *ex-vivo* equivalent of the *in vivo* hemangioblast, implying a conceptual change, where "hemangioblast" no longer represents a population of cells, but a clonal bipotent progenitor for blood and endothelium. However, the work by Ema et al. suggests that this precursor may also undertake smooth muscle differentiation (Ema, 2003). The BL-CFC expresses FLK1 (Faloon et al., 2000) and the mesodermal marker Brachyury (Fehling, 2003) but no marker of blood or ECs (with the exception of FLK-1) (Lancrin et al., 2010). In addition to these two markers, Runx1 (North et al., 1999), LIM Only domain 2 or LMO2 (Manaia, 2000) and Scl (Chung et al., 2002) have also been described as expressed by the hemangioblast.

FLK1 and Brachyury allowed Keller's group to finally track the hemangioblast from the primitive streak of the gastrulating E7-E7.5 mouse embryo (Huber et al., 2004) and showed that this cell is present in a very small number (33 per mouse) and for a very short period of time Figure 27. Using different approaches, the presence of hemangioblasts was also shown *in vivo* in the zebrafish (Vogeli et al., 2006), Drosophila (Mandal et al., 2004), mouse (Huber et al., 2004) and avian (Weng et al., 2007) embryos. The percentage of hemangioblasts varies from one species to another and, probably, with the different tracking techniques used. Thus 12.5% of mesoderm cells in zebrafish gives rise to hemangioblasts, while only 7% have this fate in birds. Despite these publications, the *in vivo* existence of the hemangioblast has been challenged by many studies. Fate mapping studies suggested that the endothelial and hematopoietic precursors of the yolk sac are specified at the pregastrula stage (Kinder et al., 1999). ES cells injections into blastocysts followed by thorough clonal analysis failed to formally demonstrate the existence of the hemangioblast (Ueno and Weissman, 2006).

More recently, an elegant *in vivo* clonal analysis conducted in the early developing mouse embryo found that early YS blood and endothelial lineages mostly derive from independent epiblast populations, specified before gastrulation (Padrón-Barthe et al., 2014) in keeping with Kinder et al., (1999). It is worth noting here that the definition of the hemangioblast raised mostly by Keller's group and some others is based on the potential of the cells *ex vivo* and not on their fate. It should also be noted that the last published experimental analysis (Padrón-Barthe et al., 2014) suffers drawbacks. One good example is the use of Tie2-Cre to track the hemangioblast. Tie2 is not expressed by the gastrulating mesoderm but is switched on at the onset of endothelial differentiation. Another one is the fact that Ueno and Weissmann claimed that HCs were derived from $FLK1^-$ cells whereas ECs were generated from $FLK1^+$ cells. This has been refuted by another study using a simpler and more efficient lineage tracing system demonstrating that indeed most, if not all, blood island HCs and adult bone marrow resident HCs are the progeny of FLK-1-expressing cells (Lugus et al., 2009). Taken together the demonstration that hemangioblasts exist *in vivo* is still to be done and needs additional, more appropriate approaches.

The existence of the hemangioblast has also been investigated using human ES cell (hESC) cultures. hES cells form Embryonic Bodies (EBs) and produce hematopoietic progenitors (Thomson et al., 1998; Vodyanik et al., 2005; Zambidis et al., 2005). The existence of a hemangioblast equivalent to that described in the mouse model has been demonstrated later. Placed on proper culture conditions, human EBs indeed produce a progenitor equivalent to murine BL-CFC, expressing FLK1 and which gives rise to cell colonies mixing erythrocytes and ECs. A human hemangioblast dependent of BMP4 and able to clonally generate both endothelial and HCs was first reported in 2007 (Kennedy et al., 2007) Figure 28. Another study using a slightly different approach revealed hemangioblasts that, when injected into an animal with diabetes or vascular injury, were able to contribute to tissue repair (Lu et al., 2007). Two studies used the expression of the angiotensin-converting enzyme (CD133) to enrich in human hemangioblasts. Cultures of CD133+ cells in the presence of VEGF, stem cell growth factor and Flt3 ligand was shown to give rise to cells giving rise to dual endo-hematopoietic colonies in one-quarter of the cells (Loges et al., 2004). CD133, in addition with the lack of CD45 expression and the presence of CD34, was demonstrated as a strong combination to isolate cells capable to give rise to

endothelial and lymphohematopoietic cells from hES cells (Zambidis et al., 2008). Blast cells giving rise to endo-hematopoietic colonies were shown to express low levels of class 1 antigen compared to their non-hemangioblast counterparts.

If the hemangioblast exists *in vivo* and is functional in giving rise to blood and endothelial progenies, it certainly contributes marginally to the formation of the blood system hence asking the question of the mechanisms of blood cell formation.

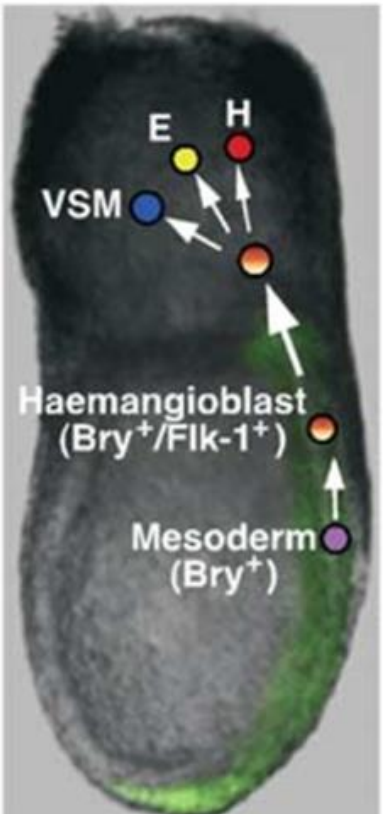


Figure 27: A model of hemangioblast development in the mouse embryo
The hemangioblast is a brachyury+ and FLK-1+ cell that arises in the primitive streak and migrates onto the yolk sac where it differentiates into hematopoietic (H), endothelial (E) and vascular smooth muscle (VSM) progenitor cells.
Figure from Huber et al 2005

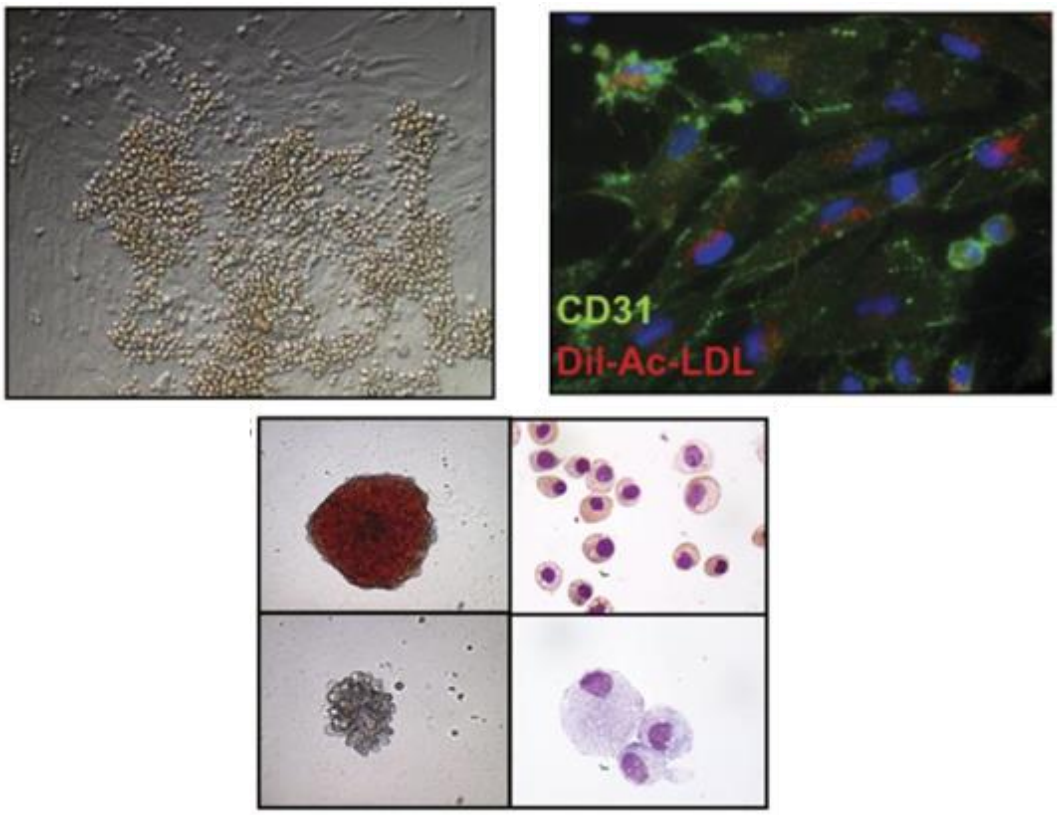


Figure 28: Developmental potential of human blast colonies

A*- Photograph showing the adherent and nonadherent population generated from an individual 6-day-old blast colony plated on a thin layer of matrigel in media containing both hematopoietic and endothelial cytokines*

B*- Immunostaining and DiI-Ac-LDL uptake of adhesive cells generated from a single blast colony. CD31 expression is indicated by green fluorescence and LDL uptake by red fluorescence. Original magnification X400*

C*- EBs were harvested, and the cells were dissociated and plated in methylcellulose supplemented with hematopoietic cytokines. The top row shows a primitive erythroid colony (left) and primitive erythrocytes (right), whereas the bottom row shows a macrophage colony (left) and macrophages (right).*

Photos from Kennedy et al 2007

2. Hemogenic endothelium: a bunch of convincing arguments

The existence of a close link between endothelial and HCs has long been recognized. The formal demonstration that ECs were giving rise to blood was achieved using lineage tracing (Jaffredo et al., 1998; Zovein et al., 2008) and time-lapse imaging (Bertrand et al., 2010; Boisset et al., 2010; Eilken et al., 2009; Kissa and Herbomel, 2010; Lam et al., 2010; Lancrin et al., 2009). These ECs endowed with the potential to give rise to blood were coined “hemogenic endothelium” (Jordan, 1916). This hemogenic endothelium gives rise to blood through an endothelial-to-hematopoietic transition (EHT) not preceded by asymmetric divisions (Kissa and Herbomel, 2010). Hemogenic ECs display an endothelial-specific signature and morphology and is integrated into a vascular tree hence ensuring bona fide endothelial-specific functions. These cells are also capable to give rise to blood cells by turning on a blood-specific program and turning off the endothelial-specific program. Hemogenic ECs have been reported in all the vertebrate species described to date (Dieterlen-Lièvre et al., 2006). Most of the data regarding the existence and the phenotypic and functional characterization of the hemogenic endothelium *in vivo* were collected on the mouse and zebrafish embryos. The hemogenic endothelium has long been associated with the production of HSCs in the aorta however, it also extends to non-HSC production and has been shown in the yolk sac and probably in the placenta.

Since primitive hematopoiesis occurs before the formation of the vasculature, it is unlikely that HCs derived from this primitive wave were generated from hemogenic endothelium. However, experimental evidence indicates that this primitive wave emerges from precursors expressing endothelial cell markers such as VE-Cadherin (Fraser et al., 2002) CD31 and CD34 (Ema et al., 2006). These precursor cells are not organized into endothelial layers but rather grouped into cell masses. Lacaud and co-workers have proposed the term “angiogenic angioblast” for this immature, likely hemogenic ECs. EMPs emerging in the yolk sac at E8.5 arise at the time of yolk sac vascular plexus remodeling (Marcelo et al., 2013; Palis, 1999b). There are also reports indicating the presence of c-Kit$^+$ or CD41$^+$ cells clusters within the yolk sac early during development (Ferkowicz, 2003; Li et al., 2005; Yokomizo et al., 2001).

In 1998, Nishikawa and colleagues demonstrated that E9.5 VE-Cadherin^{+} CD45^{-} yolk sac cells were able to generate HCs in culture (Nishikawa et al., 1998b) hence bringing the first evidence of hemogenic endothelium in the mouse embryo. A recent precise morphological study indicated that EMPs arise from a rare population of c-Kit^{lo} Runx1^{+} cells expressing endothelial markers detected from E8.5 until E10. Interestingly these flat cells were shown to become cuboid and then to round, forming clusters. Although circulation initiates at E8.25, all EMPs remained attached as clusters to the vasculature until E9.5. The study also shows that both arterial identity and blood flow are dispensable for EMP emergence, a major difference regarding the aorta (Frame et al., 2016). In addition, it was shown that yolk sac endothelium gives rise to AA4.1^{+}CD19^{+} B220$^{lo\text{-}neg}$ B progenitor cells that preferentially differentiate into innate type B-1 and marginal zone B cells but not into B-2 cells upon transplantation and can also give rise to T cell progenitors that colonize the thymus (Yoshimoto et al., 2011). Interestingly in the aorta pan endothelial expression of the core binding factor ß (the non-DNA-binding partner of Runx1) in Tek (Tie2) expressing cells was shown sufficient to produce EMPs but not HSCs whereas HSCs are produced from Ly6A-expressing ECs (Chen et al., 2011).

The first hematopoietic site described to date as harboring hemogenic ECs was the aorta and later on its associated arteries i.e., vitelline and umbilical arteries. The existence of the hemogenic endothelium in the aorta was first demonstrated in the avian embryo. Its easy accessibility allowed the use of intracardiac inoculation to tag vascular ECs and trace their progeny. Starting at E.2.5, Intra-aortic Hematopoietic cell Clusters (IAHCs) extended from the branchial arches to the pre-umbilical levels (Yvernogeau and Robin, 2017). They developed as two longitudinal symmetric ridges according to a rostrocaudal gradient located on the floor of the aorta. By injecting 1,10 -dioctadecyl-3,3,30,30-tetramethyl indocarbocyanine perchlorate-labeled human acetylated low-density lipo- protein (AcLDL-DiI) 1 day before the emergence of IAHCs, the whole endothelium of the embryo was readily labeled. Human low-density lipoprotein (LDL) is a large protein complex (~500,000 Da) that binds to a specific receptor on the surface of vertebrate cells and delivers cholesterol via receptor-mediated endocytosis. ECs and macrophages express this specific receptor i.e., the scavenger receptor, and can hence endocytose human acetylated LDLs.

The following day, not only the vasculature but also the newly formed IAHCs identified by expression of CD45 were DiI^{+}, confirming that IAHCs derived from the endothelium (Jaffredo et al., 1998) Figure 29. Similar results were obtained with non-replicative retroviral vectors (Jaffredo et al., 2000) or liposome-delivered eukaryote expression vectors (Bollerot et al., 2005). The use of a long-term tracing method using non-replicative viruses, allowed to follow the fate of the infected ECs (Jaffredo et al., 2000). In addition, the authors showed that the clones obtained are either endothelial or hematopoietic, suggesting that the production of aortic hematopoietic progenitors would pass through a phenotypical switch but not by an asymmetric division of the ECs. Cells of the intra-aortic clusters no longer express FLK1^{+} or VE-Cadherin mRNA but acquired the expression of the pan leukocyte antigen CD45 but also of the transcription factors PU1 and c-MYB (Richard et al., 2013). FLK1/VE-Cadherin and CD45 are mutually exclusive at the level of the aortic floor. No ECs are visible under the budding cells, suggesting that HCs are formed from ECs that have changed their identity to acquire a hematopoietic fate.

The cell tracing approach developed in the bird embryo was adapted to the mouse. Injections of AcLDL-DiI were performed at E10 in the circulation allowing to mark ECs. On sections, DiI staining was observed in the endothelial layer a few hours after AcLDL-DiI inoculation. Flow cytometry analysis showed that the DiI^{+} cells expressed CD31 and CD34 but not CD45. The embryo was cultured *in vitro* for an additional period of 12 hours. Ac-LDL-DiI^{+}, c-Kit^{+} multipotent progenitors as well as erythroid progenitors were revealed using clonogenic tests suggesting that they derived from pre-existing ECs (Sugiyama, 2003). The experimental design does not allow the authors to tell which EC i.e., yolk sac versus embryo, contributed to this production. The endothelial origin of HSCs from the aorta in mice has formally been shown by using a temporally restricted endothelial cell tracing to selectively label the endothelium. Lineage tracing endothelium, via an inducible VE-cadherin Cre line, reveals that the aortic endothelium is capable of HSC emergence. The endothelial progeny migrates to the FL, and later to the BM and are capable of expansion, self-renewal, and multilineage hematopoietic differentiation (Zovein et al., 2008). In addition, this study shows that the sub-aortic tissues are incapable of giving rise to hematopoiesis. Finally, the ultimate proof of hemogenic endothelium has been carried out with the use of time lapse-confocal imaging on living tissues or embryos. Thick

slices made on living Ly-6A-GFP mouse embryos were imaged over a period of 15h and showed flat ly-6A-GFP^{+} cells becoming round and bud into the lumen. Tissue staining with an anti-CD31 antibody showed that the budding cells are hematopoietic. Identical conclusions were drawn using a CD41-YFP mouse strain (Boisset et al., 2010) Figure 30. These conclusions were recently confirmed by (Eich et al., 2018) and extended to the molecular control of EHT by the transcription factor Gata-2.

The zebrafish model also conveyed undisputable evidence for the existence of hemogenic endothelium in the aorta. The use of high-resolution live imaging of ECs of the dorsal aorta in the zebrafish allowed to observe the precise steps leading to the formation of the first HSCs. Using the KDR promoter driving a fluorescent protein, Bertrand et al., (2010) showed that a portion of HCs in the adult was stained indicating that their ancestors had undergone an endothelial step at an earlier stage of their development. Three independent groups used confocal microscopy to look at the aortic region of Kdr-GFP zebrafish embryos during different time points of embryonic development: 24.5–30 hours post fertilization (h.p.f.) (Lam et al., 2010), 30–38 h.p.f. (Bertrand et al., 2010) and 36–52 h.p.f. (Kissa and Herbomel, 2010). In all cases, the emergence of HCs from ventral ECs of the dorsal aorta was observed. This hematopoietic commitment coincided with CD41, Lmo2, c- myb or Runx1 expression (all being HSC markers). The cells bud in the sub-aortic mesenchyme and later access the circulation through the neighboring axial vein. It should be noted that IAHCs do not form in the lumen of the embryonic aorta of zebrafish due to the small size of the aorta. The transition from ECs to HCs does not require cell division but rather cell shape transformations. Indeed, hemogenic ECs bend toward the mesenchyme and round up, leading to detachment from the neighboring ECs. Interestingly, the knock-down of Runx1 leads to the death of the emerging cells (Kissa and Herbomel, 2010). The zebrafish model therefore clearly shows that hematopoietic stem/ progenitor cells emerge from a polarized hemogenic endothelium located in the ventral part of the dorsal aorta.

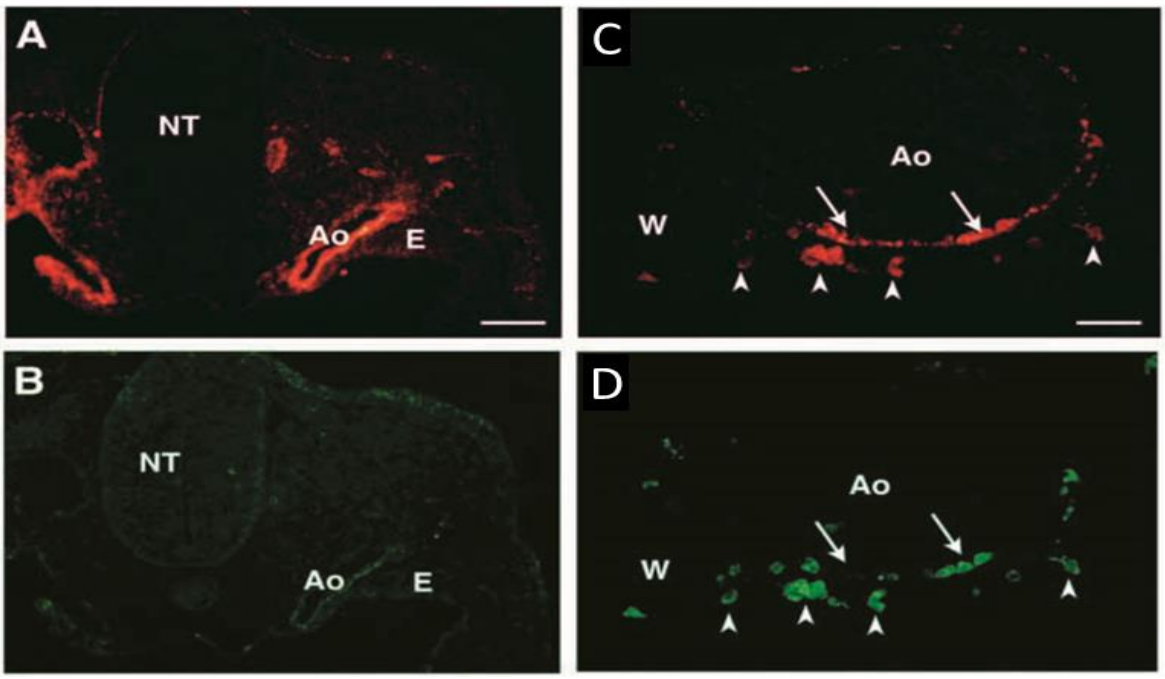


Figure 29: Tracing of aortic endothelium derivatives labelled by Ac-LDL-DIL injection in HH13 chick embryo

(A-B) 3 hours after inoculation. AcLDL-DiI (red) has been endocytosed by endothelial cells) No CD45+ cells in the aorta (B).

(D-E) 20 hours after inoculation. AcLDLDiI has been taken up by CD45− endothelial cells of the aorta and CD45+ budding cells (arrows) as well as CD45+ cells in the dorsal mesentery (arrowheads).

Abbreviations: Ao, Aorta; E, endoderm; NT, Neural tube, W, Wolffian duct

Photos from jaffredo et al 1998

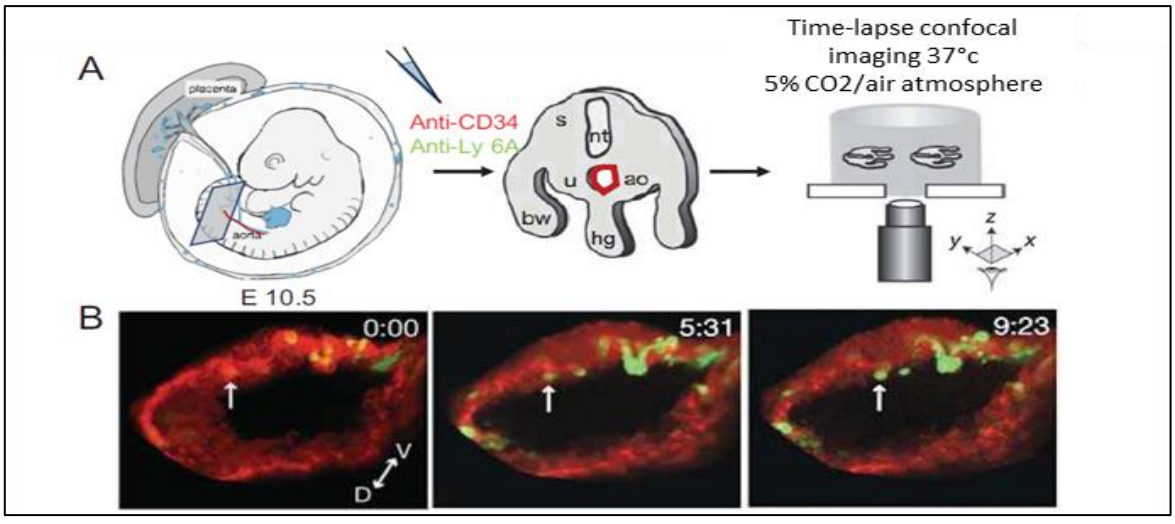


Figure 30: The endothelial origin of HSCs from the aorta in mice

A*- Protocol scheme. Slices were stained with antibodies before or after embryo slicing and imaged by confocal microscopy. Schematic of an embryo slice.*

B*- Emergence of a GFP + cell (arrow). Green, Ly-6A-GFP; red, CD31-PE.*

Abbreviations: s, somites; nt, neural tube; u, urogenital ridge; ao, aorta; bw, body wall; hg,hind gut.

Adapted from Boisset et al 2010, De Bruijn and Dzierzak, 2017

Not all the ECs are endowed with a hemogenic potential. How could we identify them and to what extent do they differ from their non-hemogenic counterparts?

As there are no unique surface marker that allow for the phenotypic identification and isolation of the hemogenic endothelium, people have sought for different cell surface markers. CD31 (PECAM), CD34 and FLK1 are shared by both endothelial and HCs. Positive and negative markers with a more pronounced lineage specificity have been used to enrich in hemogenic ECs. CD144 (VE-Cadherin) was used as a positive marker with CD45 and Ter119 being excluded (Fraser et al., 2002; Nishikawa et al., 1998b). The markers used to isolate the hemogenic ECs have varied in stringency. For instance, in several papers, $CD41^{+}$ progenitor cells were not excluded. Further dissection has included the use of integrin alpha4 (CD49d) (Ogawa et al., 1999). Based on the expression of the transcription factor Runx1 (see chapter IV, 1-D page 76), only a subset of ECs in hematopoietic sites is thought to be hemogenic. However, when considering the aorta, Runx1 is not expressed until the formation of the sub-aortic mesenchyme that acts as a switch to induce Runx1 and to trigger the hematopoietic program (Richard et al., 2013). In this respect, a silencing of Runx1 expression mediated by BMI was recently uncovered elucidating the blockade of Runx1 expression in the angioblasts (Eliades et al., 2016). At the functional level, the hemogenic endothelium is characterized by displaying an endothelial phenotype and morphology and having the capacity to form HCs and endothelial tubules in culture (Nishikawa et al., 1998a b; Nishikawa et al., 1998b a). In an effort to identify hemogenic endothelium in the mouse embryo, De Bruijn and co-workers used the Runx1+23 hematopoietic enhancer they identified (Bee et al., 2010; Nottingham et al., 2007) to generate reporter mice expressing either LacZ or GFP. These mice were used to sort +23-GFP+ hemogenic ECs that were probed for EHT potential and identified for their transcriptional program. They showed that initiation of the hematopoietic program is initiated earlier than expected and is associated with a decreasing endothelial program. Hemogenic ECs were detected from E8.5 until E9.5 the time point at which the hematopoietic program is initiated. At E10.5, no hemogenic EC was detected and all the Runx1+23 cells are hematopoietic (Swiers et al., 2013). Trying to tackle the problem from the hematopoietic side, Medvinsky and co-workers identified two distinct hematopoietic cell types within the embryonic dorsal aorta that they designated type I and type II pre-HSC. They showed that type II pre-HSCs give rise to HSCs. Type I and

II pre-HSCs differ in their expression of several surface markers. Type I and II pre-HSCs have a VE-Cadh^{+}CD41^{lo}CD45^{-} and VE-Cadh^{+}CD41^{lo}CD45^{+} phenotype, respectively, and lack the capacity to form endothelial tubules and sheets in culture (Rybtsov et al., 2011; Taoudi and Medvinsky, 2007). In addition to the in vivo studies, the hemogenic potential of ECs was also probed ex vivo using different culture conditions.

The conditional extinction of the runx1 hematopoietic gene in the cells expressing Ve-Cadherin causes a phenotype similar to that observed in embryos Runx1-/- represented by a lack of hematopoietic and definitive hematopoietic progenitors budding (Chen et al., 2009). The authors thus prove that the cells of the clusters are produced from the aortic endothelium that expresses runx1, and that runx1 is necessary for the transition between the ECs state and HCs state.

In vitro approaches in the human embryo.

Since no manipulation is ethically possible on the human embryo, all the approaches were made using either observation on sections or in vitro/ex vivo culture systems. Cortes et al., (1999) reported the presence of FLK1+CD34- cells likely to develop into endothelial precursors in the mesoderm of the early human embryo at the beginning of the 4th week of gestation. CD34 which marks human ECs at 27 days of gestation has been shown to label IAHCs (Tavian et al., 1999). CD34+CD45- cells purified from the 28-44 gestational days were shown to give rise to myeloid, lymphoid and NK cell populations following co-culture on stromal cells (Oberlin et al., 2002). However, based on the assumption above it is unlikely that these populations were composed of pure ECs.

Since much of the past work has considered the hemangioblast as a source for hematopoietic progenitors and stem cells, several studies have reconciled this view and proposed that the hemogenic endothelium was the progeny of the hemangioblast. The pioneering work by Bhatia and co-workers identified a population of primitive ECs that express CD31, FLK1, and Ve-Cadh that was able to give rise to both ECs and HCs. Despite the fact that the authors designated these cells as hemangioblast, they have all the characteristics of ECs endowed with a hemogenic potential (Wang et al., 2004). Furthermore, using the ES murine cells, several groups demonstrated that BL-

CFC generated blood progenitors through a transient and specialized population of hemogenic ECs (Dieterlen-Lièvre and Jaffredo, 2009; Eilken et al., 2009; Lancrin et al., 2009; Pearson et al., 2010) for review (Dieterlen-Lièvre and Jaffredo, 2009). These ECs form clusters from which blood cells emerge (Eilken et al., 2009; Lancrin et al., 2009) Figure 31. showed that FLK1^{+} cells give rise to VE-Cadh^{+} cells and differentiate into CD41^{+} and finally CD45^{+} HCs. Several studies using human ES cells now confirmed this mouse model (Choi et al., 2012; Kennedy et al., 2007; Kennedy et al., 2012; Wang et al., 2004). Examining the hematopoietic commitment using VE-Cadh and CD41 reporters, it was shown that there are several subsets of hemogenic ES cells in keeping with in vivo studies and that the quality of the hematopoiesis produced varies with time un culture (Rafii et al., 2013). In an effort to produce engraftable HSCs, several groups have refined their differentiation protocol of either murine or human EC or iPS cells and in particular have avoided the use of serum in culture. Lacaud and Kousskof's groups produced hemogenic endothelium from mouse ES cells that were capable of giving rise to myeloid, erythroid and lymphoid lineages upon engraftment into immunocompromised recipients. Primary transplantation was successful up to 22 weeks, but secondary transplantation failed. In addition, they showed that the hemogenic ECs were extremely transient in culture (Pearson et al., 2015). Another strategy to generate hemogenic endothelium is to reprogram ES cell-derived cell types of differentiated cells. Overexpression of SOX17, a transcription factor involved in EC differentiation, in CD34^{+} CD43^{-} cells generated hemogenic ECs with hematopoietic potential. Sox17 was shown to play a key role in priming hemogenic potential in ECs (Nakajima-Takagi et al., 2013). Another strategy consisted of screening an ensemble of 27 transcription factors for their capacity to induce hematopoiesis from human pluripotent stem cells. The authors identified two optimal combinations of TFs capable of inducing distinct, robust hematopoietic programs from pluripotent stem cells: pan-myeloid (ETV2 and GATA2) and erythro-megakaryocytic (GATA2 and TAL1). Interestingly, both transcription factor combinations directly induced hemogenic ECs, which subsequently transformed into blood progenitors with a distinct spectrum of hematopoietic differentiation (Elcheva et al., 2014).

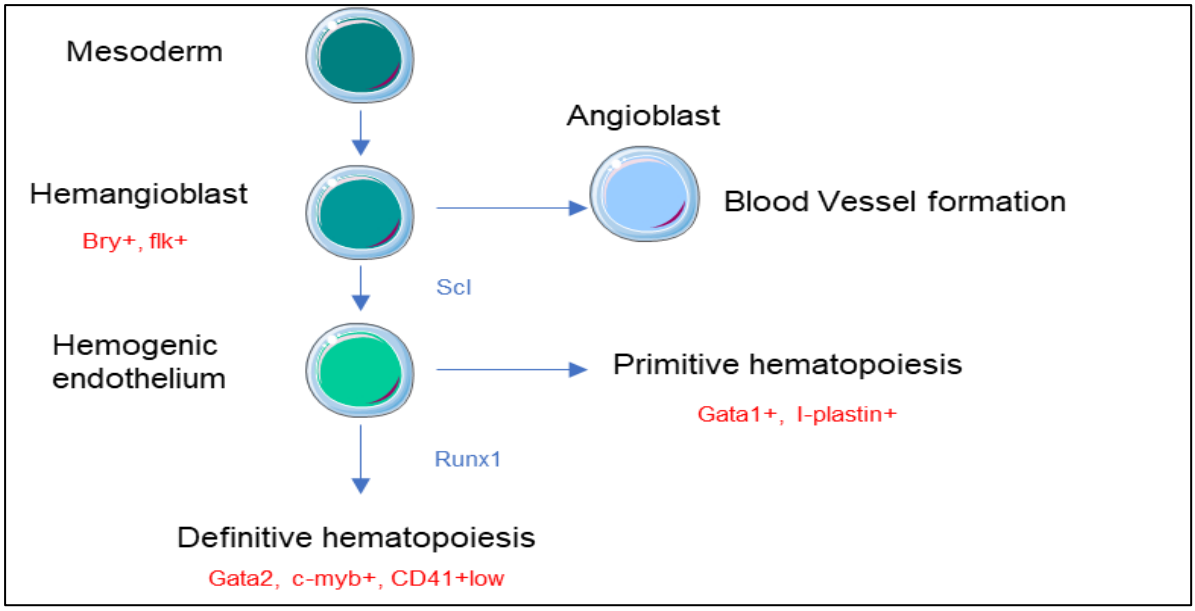


Figure 31: The endothelial origin of HSCs from the aorta in mice
Link between HSCs, hemangioblast and hemogenic endothelium
Genes in red are cellular markers of the different cell types, genes in blue are differentiation factors.
Adapted from Eilken et al 2009, Lancrin et al 2009

Another aspect consists in generating hemogenic ECs directly from differentiated cells through direct reprogramming. The team of K. Moore showed that the overexpression of four transcription factors, Gata2, Gfi1b, c-Fos, and Etv6 within fibroblasts induced a hemogenic endothelial cell fate in murine fibroblasts that subsequently produced hematopoietic precursor cells that express CD34, SCA1 and PROMININ1 within a global endothelial program (Pereira et al., 2013). However, these cells displayed poor multilineage potential in vitro. This hemogenic phenotype completed by the absence of CD45 expression was later shown to be present in ECs of the placenta. When grafted into irradiated mice following co-culture with stromal cells, these hemogenic ECs engraft into recipient mice and give rise to primary and secondary recipients (Pereira et al., 2016). Lacaud and co-workers showed that the ectopic expression of five transcription factors (Erg, Gata2, Lmo2, Runx1c, and Scl) reprogrammed fibroblasts to hematopoietic progenitors with erythrocyte, granulocyte, macrophage, and megakaryocyte potential. This reprogramming passes through a hemogenic endothelium state, that in contrast to Peirera et al. gives rise to robust clonogenic potential (Batta et al., 2014). In recent reports, Daley and co-workers showed that seven transcription factors (*ERG*, *HOXA5*, *HOXA9*, *HOXA10*, *LCOR*, *RUNX1,* and *SPI1*) that are sufficient to convert hemogenic endothelium into hematopoietic stem and progenitor cells that engraft myeloid, B and T cells in primary

and secondary mouse recipients. These factors are transferred into so-called $CD34^+$ hemogenic ECs and are likely to mimic molecular events occurring within the dorsal embryonic aorta (Sugimura et al., 2017). In another study published in the same issue of Nature, Rafii and colleagues reported constitutive expression of transcription factors *FOSB*, *GFI1*, *RUNX1*, and *SPI1* within adult ECs and cocultivation of these cells with vascular niche induce a hemogenic state characterized by the expression of RUNX1 and the subsequent expression of CD45. When transplanted into immunocompromised recipients, these cells were capable of primary and secondary engraftment (Lis et al., 2017). Interestingly, this protocol is a refinement of a previous study published in 2014 identifying the transcription factors and the requirement for a vascular niche (Sandler et al., 2014).

3. SubAortic Patches

At E4 in the bird embryo, IAHCs disappeared from the aortic floor. Instead of being produced into the lumen, some cells accumulate in the inner side of the aorta being also the progeny of the aortic endothelium (Jaffredo et al., 2000). Those cells emitted toward the inner side of the aorta would multiply in the mesentery giving rise to the para-aortic foci (Dieterlen-Livre and Martin, 1981; Jaffredo et al., 2000) Figure 32. These para-aortic foci contain cells expressing the surface molecules $\alpha_2\beta_1$ integrin, c-kit, thrombomucin/MEP21, HEMCAM, and chL12. Importantly, the presence of these antigens was found to correlate with the recolonization of the recipient thymus following intrathymic cell transfers. These intra-embryonic cells were also found to express the Ikaros transcription factor, the molecular function of which is considered to be prerequisite for embryonic lymphoid development (Lampisuo et al., 1999). In the mouse embryo, Godin and co-workers described the presence of cells localized in the sub-aortic space that express c-kit, AA4.1 (Petrenko et al., 1999), CD41 and Gata3 (Manaia, 2000)Figure 33 but lack the expression of FLK1 and CD45. They designated these cells as sub-aortic patches and proposed they were at the origin of IAHCs (Bertrand et al., 2005). However, in accordance with the results of Bertrand et al. (2005) Rybtsov et al. showed that the population enriched with HSCs $VeCad^+$ $CD45^+$ present in aortic clusters at E11.5 originates from a $VeCad^+$ $CD45^-$ $CD41^+$ E10.5 population localized not only in the aortic endothelium but also in the subaortic mesenchyme, relaying again the debate on a potential mesenchymal origin

of HSCs (Rybtsov et al., 2011). Then, one study has shown that the ubiquitin ligase mind bomb-1, a regulator of the Notch pathway, is implicated in the hematopoiesis of SAPs (Yoon et al., 2008). Although not totally solved these Gata3-positive patches were shown to be part of the AGM microenvironment and belonging to the sympathetic nervous system and catecholamine-producing cells (Fitch et al., 2012).

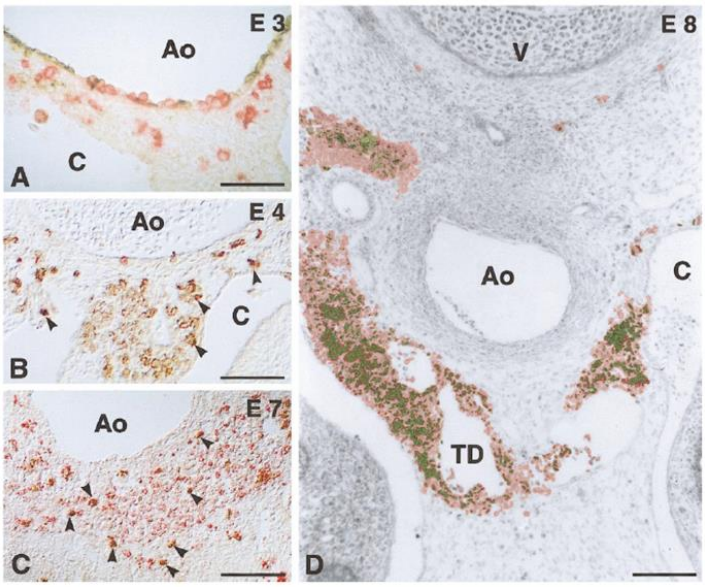


Figure 32: Labeling of hematopoietic cells in the aortic region

Cross-sections of the aortic region of E3, E4, E7 and E8 chicken embryos, first showing the hematopoietic cells inserted into the aortic endothelium and then the para-aortic foci that form in the mesenchyme underlying the aorta.

A. In green: Endothelial cells FLK1, in pink the hematopoietic cells CD45.

B, C. CD45 cells that have ingressed in mesenchyme. Arrowheads: macrophages.

D. Para-aortic foci (pink and green) are basophile cells.

Photos from jaffredo et al 2000

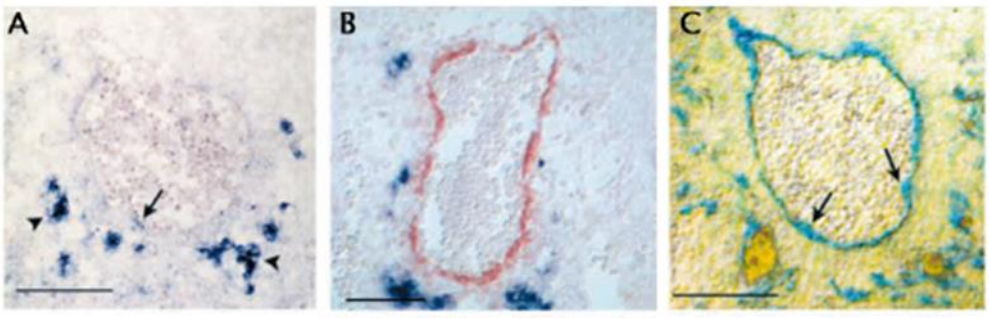


Figure 33: Existence of para aortic foci in mice

The analysis of Gata3 showed the existence of cells underneath the aorta at E11. the authors called this group of cells "para-aortic foci".

A, B. In blue: Gata3

B. Gata3 positive cells do not overlap with Smooth muscle actin positive cells in red

C. AA4-positive cells in blue encircle GATA-3-positive patches in yellow;

Photos from Manaia et al, 2000

IV. Molecular regulators of developmental hematopoietic cell generation

The identification of specific markers related to the emergence and amplification of HSCs is an important step in understanding the maturation stages of definitive HSCs during embryonic development. In this part, I will aim to establish a non-exhaustive list of the main transcription factors and surface markers involved in the emergence and differentiation of HSCs during hematopoietic development.

1. Hemogenic endothelium and EHT induction factors

A. FLK1: Fetal liver kinase 1 or Vascular Endothelial Growth Factor Receptor 2 (VEGFR2)

FLK1, a receptor with tyrosine kinase activity, is activated by VEGF; It is also called VEGF-R2 or KDR (Millauer, 1993). Flk1 is expressed by cells dedicated to giving the cardiac tissue at the level of the primitive streak and is described as a marker of endothelial precursors (Yamaguchi et al., 1993). Flk1 knock out shows that this receptor is essential for both the development of HCs and vasculogenesis. Embryos invalidated for the Flk1 gene present a total absence of ECs and HCs and die in utero between E 8.5 and E 9.5. In this respect, the Flk1 knockout phenotype has been considered as the first functional proof of the existence of the hemangioblast (Eichmann et al., 1997; Shalaby et al., 1995; Shalaby et al., 1997). Later in development, Flk1 is expressed at the base of IAHCs and in the adjacent ECs at the protein level for the chicken (Jaffredo et al., 1998) and mouse (Yokomizo and Dzierzak, 2010), and at the mRNA level for the human embryo (Labastie, 1998). The expression profiles of Flk1 and CD45 are therefore mutually exclusive in the aorta as previously shown by Nishikawa et al. with ES cells (Nishikawa et al., 1998b a). Interestingly, Runx1 represses the expression of Flk1 contributing to orchestrate EHT. In view of all these results, the Flk1 expression was attached to the hemangioblast. The work of Fehling et al. on ES cells confirm this hypothesis (Fehling, 2003).

B. Stem Cell Leukemia Scl, Lmo2 and Lyl1

The SCL gene encoding for a bHLH domain transcription factor was initially identified by its involvement in a number of chromosomal translocations observed in leukemic T cells (Begley and Green, 1999; Begley et al., 1989; Chen et al., 1990). This gene is expressed early in the blood islands and when differentiated, remains expressed in the hematopoietic cells and is detected very weakly in ECs (Drake and Fleming, 2000; Elefanty et al., 1999; Minko et al., 2003). In the aortic region, SCL is expressed in the aortic endothelium prior to the onset of intra-aortic budding, and its expression appears to be strengthened at the level of the ventral endothelium. When the hematopoietic budding appears, SCL is strongly expressed and its expression in endothelial cells is low (Jaffredo et al., 2005). Functional experiments in mice and zebrafish have helped to highlight the role of SCL during hematopoiesis. When the gene is invalidated in mice, the embryos die at E9,5 with a complete absence of hematopoiesis in the yolk sac (Robb et al., 1995). Later on, the authors showed that SCL is indispensable for the development of all the hematopoietic lineages, primitive and definitive (Porcher et al., 1996; Robb et al., 1996). Although it was proposed that SCL would be secondarily required for remodeling of the primitive vascular network (D'Souza et al., 2005; Elefanty et al., 1999; Visvader et al., 1998), several reports indicate a crucial role for the specification of the hemogenic endothelium. In the absence of SCL in BL-CFC, all the hematopoietic lineages are, as expected, compromised but all ECs differentiation including hemogenic ECs is abrogated (Lancrin et al., 2010). This is consistent with a report showing that SCL was required for the formation of VE-Cadherin+ cells from $Flk1^{+}$ cells (Endoh et al., 2002). Several years after, Mikkola and co-workers showed that SCL activated a network of transcription factors necessary for the specification of the hemogenic endothelium and showed a repressive role of SCL to limit the commitment of ectopic cardiomyocytes (Van Handel et al., 2012). On the other hand, it has been shown that the ectopic expression of SCL can restore the differentiation of all hematopoietic lineages in $SCL^{-/-}$ mouse embryos, suggesting a role in the differentiation of hemogenic progenitors (Visvader et al., 1998). In addition, in zebrafish, morpholinos targeting of SCL blocks hematopoiesis and GATA1 expression without altering the expression of other hematopoietic markers such as GATA2 or Lmo2. The absence of hematopoiesis would not be due to the absence of progenitors but SCL would play a key role in

hematopoietic differentiation (Dooley et al., 2005; Patterson et al., 2005). Interestingly, SCL becomes dispensable at the end of hematopoietic development due to redundancy with another bHLH protein, Lyl1.

Lmo2 is also a transcription factor involved in lymphoid leukemia induced by chromosomal translocations (Royer-Pokora et al., 1991). Its expression in hematopoietic sites is similar to that of SCL although a little more extensive. Invalidation of the gene in mice leads to the same phenotype as the one described for SCL (Yamada et al., 2000). *In vitro* functional interaction experiments have shown that SCL and Lmo2 interact within of multimeric transcriptional complexes: through its LIM domain, Lmo2 plays a role in between SCL and GATA1 (Wadman et al., 1997).

Studies to determine whether there was a synergy of action between the two transcription factors were performed in zebrafish: the combined overexpression of SCL and lmo2 have indeed shown that the presence of both factors is necessary to respecify a mesodermal territory of origin inducing the expression of endothelial and hematopoietic genes.

The lyl-1 gene also codes for a bHLH domain transcription factor and has been identified, along with SCL, as being implicated in some lymphoid leukemia in humans. Although gene deletion in mice is not lethal, it results in a decrease in the number of B lymphocytes as well as immature progenitors, reducing their potential for reconstitution (Capron et al., 2006). Analysis of the expression patterns at different stages shows that the SCL and lyl-1 expression overlap widely in hematopoietic sites in mice, particularly in the aortic endothelium and the fetal liver, suggesting possible cooperation with SCL in the regulation of hematopoiesis (Giroux et al., 2007). Analysis of the regulation of the expression of lyl-1 as well as functional studies through the obtaining of $SCL^{-/-}$and $Lyl\text{-}1^{-/-}$ embryos yielded a number of responses to the respective roles of Lyl-1 and SCL in hematopoiesis: If both genes are regulated by the same molecular combinatorial, the SCL expression is initiated before that of lyl-1. In addition, Lyl-1 is unable to save the hematopoietic phenotype of ES SCL-/-cells, suggesting that these two transcription factors play a different role during hematopoiesis.

C. Gata transcription factors

The Gata protein family includes six highly conserved zinc finger transcription factors, of which three members (GATA-1, 2 and 3) are involved in hematopoiesis at different levels. GATA-1 and 2 are required very early for primitive hematopoiesis (Fujiwara et al., 2004). The expression patterns established in different species have the same characteristics: the expression of GATA-1 and 2 is detected very early in the mesodermal masses in the blood islands. During their formation, the expression of GATA-1 persists while that of GATA-2 gradually disappears from the most mature islands (Minko et al., 2003 ; Silver and Palis, 1997; Walmsley et al., 1994; Zon et al., 1991). These observations suggest an early role for both genes in definitive erythropoiesis, which could be subsequently confirmed through GATA-1 invalidation experiments: erythroid cells are present in the blood islands, but their maturation is blocked, and they enter into apoptosis (Fujiwara et al., 1996; Fujiwara et al., 2004). The *in vivo* deletion of GATA-2 leads to a reduced number of hematopoietic progenitors that normally develop (Tsai et al., 1994). Thus, the Gata-2 factor would be required for the commitment to the erythropoietic pathway while GATA-1 would allow the maturation of erythroid progenitors. However, double mutant embryos have a lack of primitive erythrocytes in blood islands, suggesting a partially redundant role of two transcription factors (Fujiwara et al., 2004; Takahashi et al., 2000). In the aortic endothelium, GATA-2 appears in the ventral endothelium of the aorta at the same time as SCL and lmo2, at the time of fusion of the two pair aorta. The expression of GATA-2 is then strengthened in the budding cells the expression of Gata-3 is also initiated but in a lower manner (Jaffredo et al., 2005). In 2004, a double role of Gata2 dose-dependent could be shown in mice: it would be involved in the production and expansion of HSCs in the AGM as well as the proliferation stage of HSCs in the adult bone marrow (Ling et al., 2004). Notch1 specifically associated with the GATA2 promoter suggesting a role in the formation of HSPCs by the hemogenic ECs (Robert-Moreno et al., 2005) but multiple roles for GATA2 at different levels has been demonstrated in culture using conditional GATA2 induction in BL-CFC. Conditional GATA2 deletion where CRE is under the control of the vascular-specific GATA2 enhancer leads to defect of HSCs in the fetal liver and in the bone marrow. Interestingly, enforced expression of GATA2 in fibroblasts together with Gfi1b, Fos and ETV6 induces a hemogenic endothelium program in fibroblasts with the

subsequent appearance of blood cells (Pereira et al., 2013). Another study using Erg, Lmo2, Runx1c SCl and GATA2 reached to the same conclusions (Batta et al., 2014) Similar to Runx1 (see the Runx1 chapter), the requirement for GATA2 has been studied both in the context of ECs or HSPCs. Deletion of GATA2 in an endothelial i.e., VE-Cadherin-Cre, context leads to a complete loss of HSCs in the AGM whereas deletion in hematopoietic cells i.e., Vav-Cre impacts HSC survival (de Pater et al., 2013). In this respect, abrogation of the +9.5 GATA2 regulatory element completely abrogates IAHCs formation and HSC emergence from the AGM. Transcriptome analysis indicates that the +9.5 element regulates an ensemble of genes involved in hemogenic endothelium and HSCs (Gao et al., 2013). Interestingly, the existence of two GATA2 orthologues in the zebrafish reveals a tight control of GATA2b, under the control of Notch, over the hemogenic endothelium and Runx1 expression and a broader role of GATA2a on the vascular endothelium (Butko et al., 2015). Conditional expression of GATA2 in GATA2 knockout human embryonic stem cells under hematopoietic differentiation reveals that GATA2 is not required for the generation of EC or hemogenic ECs but is required for hemogenic ECs to proceed EHT (Kang et al., 2018). Recently, time-lapse imaging of GATA2 expression within the aorta has uncovered the existence of a pulsatile expression in cells undergoing EHT. GATA2 haploinsufficiency alters the pulsatile pattern, reduces the number of EHT and reduces hematopoiesis (Eich et al., 2018). Concerning Gata3, it is expressed in adult and embryonic HSCs (Bertrand et al., 2005; Frelin et al., 2013) It is also involved in the development of T lymphocytes (Ting et al., 1996) and NK cells (Vosshenrich et al., 2006). KO embryos for Gata3 die around E11.5 and exhibit abnormal fetal hematopoiesis (Pandolfi et al., 1995). Nevertheless, the deletion of Gata3 under the control of pan-hematopoietic marker Vav1 does not affect the activity of HSC (Buza-Vidas et al., 2011). However, another analysis showed that Gata3 was required for HSC maintenance by regulating their entry into the cell cycle.

D. The Runx1 transcription factor

Runt-related transcription factor 1 (RUNX1) also known as acute myeloid leukemia 1 protein (AML1) or core-binding factor subunit alpha-2 (CBFA2) is a protein encoded by the RUNX1 gene. Three major isoforms are transcribed by use of two promoters and alternative splicing. Isoforms 1A and 1B are transcribed from the proximal promoter P2 and differ at their carboxyl termini. Isoform 1C is transcribed from the distal promoter P1 and is identical to isoform 1B except for 32 amino acids encoded by alternative exons at its amino terminus. The three isoforms are expressed in a temporal and tissue-specific manner. Isoform 1C is expressed at the time of emergence of definitive hematopoietic stem cells, whereas isoforms 1A and 1B are expressed throughout hematopoietic differentiation (Sroczynska et al., 2009) it was recently demonstrated that Runx1b levels need to be kept low for initiation and successful completion of EHT, since increased levels of Runx1b accelerate the onset of EHT, but are unable to drive hematopoietic maturation (Lie-A-Ling et al., 2018).

Runx1 needs a heterodimeric partner to bind DNA. CBFß has no DNA binding activity by itself but potentiates Runx1 binding to DNA and protects it from degradation by the ubiquitin-proteasome (Wang et al., 1996).

Runx1 knockout mice show a drastic decrease in the number of erythrocytes in fetal liver, and premature death at E12.5-E13.5 accompanied by large brain hemorrhages (Wang et al., 1996). If primitive erythropoiesis is marginally affected (Yokomizo et al., 20010), no definitive hematopoietic progenitor and HSCs is detected in the Runx1-/- embryo and are severely reduced in CBFß-/- embryos (Wang et al., 1996). Of the primitive lineages affected, the yolk sac -derived macrophages are absent from Runx1-/- embryos and from ES cell cultures (Lacaud et al., 2002). These observations therefore suggest that Runx1 would play a key role in the induction of definitive hematopoietic progenitors at the level of the aortic endothelium.

In the embryo, Runx1 expression is found in all sites of hematopoiesis. Strikingly, Runx1 expression is found in a subset of ECs in the vitelline and umbilical arteries, in the yolk sac, the placenta and the ventral aspect of the dorsal aorta in the mouse embryo. Runx1 is also found expressed in some peri-aortic tissue underneath the aorta. A polarized expression of Runx1 is also found conserved in all the vertebrate species described (Bollerot et al., 2005; Ciau-Uitz et al., 2000; Kalev-Zylinska, 2002;

North et al., 1999) giving a molecular credit to the existence of hemogenic endothelium. In the aorta, Runx1 knock-down abrogates the formation of the IAHCs (North et al., 1999; Yokomizo et al., 2001). In the Zebrafish, morpholino knock-down impedes IAHCs formation without affecting the aortic endothelium (Kalev-Zylinska, 2002 ; Burns et al., 2005; Gering and Patient, 2005). In contrast, in the adult, Runx1 deletion causes rather an expansion of HSCs and progenitors in the bone marrow (Chen et al., 2009 ; Growney et al., 2005) indicating a change in HSPC requirements regarding Runx1. Interestingly, conditional deletion of Runx1 made either in an endothelial context where the Cre is driven by the VE-Cadherin regulatory sequences or in a hematopoietic context where the Cre is driven by the hematopoietic-specific Vav gene regulatory sequences demonstrates that Runx1 is required for the EHT but is no longer required after hematopoietic cells were formed (Chen et al., 2009). In line, retroviral transduction of Runx1 into cells from E11.5 AGM did not rescue hematopoiesis wile the same approach made at E9.5 did so (Goyama et al., 2004; Mukouyama et al., 2000). Restauration of Runx1 expression in ECs, using a Tek-driven construct showed that Runx1 expression is sufficient to rescue lymphoid, myeloid and HSCs and to prolong the life of the embryos until birth (Liakhovitskaia et al., 2009), post-natal lethality resulting from the loss of Runx1 expression in non-hematopoietic tissues. Conditional, timed deletion shows that Runx1 deletion between E7.5 and E10.5 impairs EMP formation but not HSCs which are eliminated at E11.5. hence revealing a temporal requirement for Runx1 function (Tober et al., 2013). Conditional expression of Runx1 in endothelial cells at different time points during development indicates a temporal window for Runx1 to generate hemogenic endothelium. Runx1 expression between E7.5 and 8.5 in non-hemogenic ECs promotes the formation of erythro-myeloid progenitors in the yolk sac, the AGM and the heart. This formation is abrogated when Runx1 is expressed at E8.5-9.5 (Yzaguirre et al., 2018).

Runx1 thus appears as an orchestrator of EHT rather than a master gene to promote the hemogenic endothelial fate. Runx1 was first shown to regulate the expression of a number of hematopoietic-specific genes such as Spi1, IL3, GM-CSF but also to interact with either activation cofactors, such as p300 and CBP (Creb Binding Protein), histone acetyltransferases, or repression cofactors such as SUV39H1, a histone methyltransferase. Regarding the EHT, two direct targets of

Runx1, Gfi1 and Gfi1b, that encode nuclear zinc finger transcriptional repressors have been identified and shown to inhibit the expression of endothelial-specific genes (Lancrin et al., 2012). Embryos deficient for Gfi1/Gfi1B display a block in EHT wherein the cells remain attached to the basal layer in the yolk sac. However, when cells are set free, they are able to form clonogenic colonies (Lancrin et al., 2012). However, expression of Gfi1/Gfi1b in Runx1 deficient mouse embryos shows that cells can undergo EHT but don't form hematopoietic colonies (Lancrin et al., 2012). A more recent study indicates that IAHCS did not form in Gfi1/Gfi1b-deficient embryos but cells remain attached to the endothelial layer. Contrary to the yolk sac, cells are not able to give rise to hematopoietic colonies when dissociated and placed in clonogenic conditions (Thambyrajah et al., 2016). Despite advances, more significant information is needed to understand how EHT is controlled. In order to broaden the search, Lacaud and co-workers implemented a highly sensitive DNA adenine methyltransferase (Dam) identification-based methodology, to map early RUNX1 transcriptional targets in hemogenic ECs differentiated from mouse ES cells deficient for Runx1 and containing doxycycline inducible Runx1-Dam. Comparison of the Runx1-DamID methylation and RNA-Seq datasets led to the identification of 235 genes that were both bound by Runx1 and differential expressed in hemogenic ECs generated from wild type and $Runx1^{-/-}$ ES cells. The expression of 80 of the genes was negatively correlated with Runx1 occupancy and 155 genes were positively correlated (Lie-A-Ling et al., 2018), consistent with Runx1's ability to function as a transcriptional repressor or activator in the same cell type (Canon and Banerjee, 2003). Interestingly, target genes that were positively correlated with Runx1 expression were associated with cell adhesion, integrin signaling, cellular movement and interaction with the extracellular matrix but few hematopoietic-specific genes were identified (Lie-A-Ling et al., 2018).

In addition to the activation or repression of target genes, Runx1 was also shown to have an epigenetic role in chromatin remodeling. It was shown that the transient binding of Runx1 to PU1 cis-regulatory elements is required for further DNA binding-dependent transactivation of PU1(Hoogenkamp et al., 2009). More widely, Runx1 was shown to be able to associate with numerous de novo DNA sites upon Runx1 expression hence initiating a global reorganization of transcription factors assemblies (Lichtinger et al., 2012). A genome-wide analysis of Runx1-dependent

binding of factors involved in transcription elongation shows that Runx1 induction moves several transcription factors from their previous sites and relocates them to other ones and recruits the basal transcription factors CDK9, BRD4, the Mediator complex, and the looping factor LDB1 thus being able to shortly activates genes essential for EHT (Gilmour et al., 2018). Runx1 also contributes to site-specific demethylation via the recruitment of Tet2, a demethylation enzyme (Suzuki et al., 2017).

E. Gfi1 and Gfi1b

Gfi1 and *Gfi1b* genes encode 2 highly homologous nuclear zinc finger proteins that function as transcriptional repressors (Grimes et al., 1996; Tong et al., 1998). They both harbor a DNA binding domain localized in the C-terminal domain of the molecules and embedded within 6 zinc finger motifs while the repressor activity is localized within the N-terminal domain (Grimes et al., 1996; Zweidler-Mckay et al., 1996). Gfi1 is expressed in many hematopoietic cells including HSCs, lymphoid and myeloid progenitors and mature cells and Gfi1b is mostly expressed in erythroid and megakaryocytic lineages. Invalidation of Gfi1 leads to neutropenia and reduction of HSCs whereas invalidation of Gfi1b leads to embryonic lethality at E14.5 due to an impairment in erythropoiesis and megakaryocytopoiesis.

Both were identified as direct targets of Runx1 and critical regulators of EHT based on pluripotent-based cell culture approaches. In the absence of Runx1, Gfi1 and Gfi1b are able to induce the loss of EC identity required for EHT and induce the formation of round cells that are not endowed with the full hematopoietic competence. Conversely, blood cells generated in the absence of Gfi1 and Gfi1b maintain the expression of EC markers and cannot be released into the circulation (Lancrin et al., 2012). Examination of Gfi1 and Gfi1b expression patterns indicates that during the period of hematopoietic production Gfi1 expression is associated with ECs and IAHCs whereas Gfi1b is more associated with IAHCs. Transplantation of Gfi1/1b $^{+}$ cells from the AGM into irradiated mice indicates that the reconstituting fraction is associated with the Gfi fraction. Double Gfi1/Gfi1b inactivation results in earlier lethality than either single invalidation. IAHCs formation is totally impaired in the AGM but cells expressing c-Kit or CD31 markers were found embedded within the endothelial layer. When isolated and dissociated, these cells were able to form clonogenic colonies. To gain

insights into the molecular regulation of EHT by Gfi1 molecules, Lacaud and co-workers examined the role of LSD1, belonging to the CoREST complex, known to interact with Gfi1 proteins. Pharmacological or genetic inactivation of LSD1 leads to an impairment of EHT in culture. The analysis of global gene expression profiles upon LSD1 inhibition reveals an upregulation of cardiac-specific genes at the expense of hematopoietic-specific genes. Comparison of transcriptome in LSD1 inactivation and Gfi1 and Gfi1b binding sites obtained using Dam-ID reveals genes associated with cell adhesion, integrin signaling, cellular movement and interaction with the extracellular matrix (Thambyrajah et al., 2016). In addition, Gfi1 or Gfi1b also belong to the small set of genes used to reprogram cells into HSCs (Pereira et al., 2013; Sandler et al., 2014; Tsukada et al., 2017).

F. Sox 7, 17 and 18

Sox genes encompasses a family of SRY (Sex Determining Region Y)-related HMG box of DNA binding proteins, referred to as SOX transcription factors, which consists of more than 20 *Sox* genes characterized by the evolutionarily conserved high mobility group (HMG) box, a 79-amino-acid DNA-binding motif that binds to a common consensus site with variable efficiency (She and Yang, 2015). The Sox transcription factors associated with the vasculature are Sox7, Sox17 and Sox18. They are expressed at the onset of endothelial differentiation as early as E7.5 in a subset of $Etv2^{+}$ $Flk1^{+}$ $CD41^{-}$ cells. Sox7 and Sox 18 continue to be expressed in the aorta, cardinal vein and intersomitic vessels at later stages whereas Sox17 becomes restricted to arterial ECs from E10.5. This pattern continues after birth. Besides roles in cardiovascular development (reviewed in (Lilly et al., 2017)), Sox7, 17 and 18 were shown to play a growing role in the hemogenic endothelium and EHT. Sox 17 was shown to be expressed in hemogenic ECs during ES cell differentiation and to bind and activate the expression of VE-Cadherin (Costa et al., 2012). Another study demonstrated that Sox17 is expressed by hemogenic ECs *in vivo* and that it is required for the expansion of hemogenic endothelium through the Notch signalling pathway (Clarke et al., 2013). Sox17 was shown to actively repress hematopoietic-specific genes and thus contribute to maintain and modulate the endothelial identity and the EHT (Lizama et al., 2015) through an interplay with Runx1 (Bos et al., 2015).

G. HoxA3

Another gene which is shown to control Runx1 expression and EHT is HoxA3. HoxA3 is a member of the paralogue group 3 that was shown to play a role in the specification of pharyngeal tissues. One of these paralogue HoxA3 is also shown to play a role in vascular development. Mice lacking HoxA3 displays cardiovascular anomalies (Chisaka and Capecchi, 1991) and vascular expression of group 3 paralogues is associated with revascularization and wound repair (Boudreau et al., 1997; Myers et al., 2000). In the embryo, HoxA3 is expressed in vascular ECs of the embryo as early as 8.25 and is found expressed in aortic ECs but is absent from ECs of the yolk sac. As EHT takes place, HoxA3 and Runx1 expression are mutually exclusive. Forced expression of HoxA3 in ES cells under endo-hematopoietic differentiation results in the repression of CD41 and CD45 hematopoietic-specific genes and the enforced expression of VE-Cadherin and Flk1. This correlates with the loss of hematopoietic-forming potential in a clonogenic assay. A similar effect has been observed when HoxA3 is transduced in AGM cells. Transcriptome analysis and chromatin immunoprecipitation indicates that HoxA3 targets the key hematopoietic transcription factors Runx1, Gata1, Gfi1b, Ikaros and Pu1 and down regulates them to maintain an endothelial phenotype (Iacovino et al., 2011, 3). A recent report indicates that this repressive mechanism is also passing through the Notch pathway in upregulating Jag1 in ECs thus reinforcing the endothelial phenotype (Sanghez et al., 2017).

H. NOTCH pathway

The involvement of the Notch pathway in developmental processes has been described in multicellular organisms, from nematodes to humans. It plays an essential role in many cell lineage selections and can regulate processes such as proliferation, survival or cell differentiation, both during development and in adults. The Notch gene was characterized for the first time in Drosophila in 1917: its mutation causes in animals the appearance of a notch on the wings. It was not until 1985 that this gene was cloned in Drosophila (Kidd et al., 1986) and then in 1990 in Vertebrates (Coffman et al., 1990). In mammals, four Notch genes were identified and cloned, in the Bird, several sequences with homologies to the mouse genes were identified in the avian genome, but only two genes, notch1 and notch2, were cloned. Notch is a transmembrane protein and has a stereotyped structure, whatever the isoform. The extracellular part of Notch includes repetitions of a particular motif, the EGF-like motif, which allows it to link with other proteins: in fact, Notch acts as a receptor on the surface of the target cell and its extracellular domain allows him to interact with the ligands of the signaling pathway. Notch ligands are numerous; they are also transmembrane proteins which, in addition to EGF-like repeats, have a DSL (Delta Serrate Lag2) domain in their extracellular domain. Notch ligands include Delta proteins as well as Serrate / Jagged proteins.

Notch signaling is pleiotropic but has been shown to play a major role in both embryonic and adult hematopoiesis. Indeed, Notch 1, 2 and 3 are expressed in mature and immature progenitors, as well as in blood cells and lymphocytes. Notch ligands are predominantly expressed in the stroma of adult hematopoietic sites as well as on antigen-presenting cells. The Notch pathway is implicated, for example, in the induction of the T lymphoid lineage (Radtke et al., 1999) at the expense of lymphocyte B differentiation (Pui et al., 1999). In addition, in mice a GFP reporter study of Notch activity showed that in the cKit+ HSCs in the bone marrow, Notch pathway is activated. The authors suggest that it plays a role in maintaining the undifferentiated state of HSCs. Bone marrow stromal cells express Notch ligands such as Jagged-1, Delta-1 or Delta-4. Activation of the Notch pathway in HSCs is probably involved in the differentiation of hematopoietic cells such as dendritic cells (Cheng and Gabrilovich, 2008). In vitro, HSC culture experiments in which a constitutively active form of Notch

is transfected showed that the Notch pathway inhibits the differentiation of HSCs and allows expansion of the number of progenitors.

During development, NOTCH1 a member of the notch pathway has been defined as a key gene in EHT. Evidence for the involvement of the notch pathway in embryonic aortic hematopoiesis was provided by the study of mutant mice for Notch1 or Notch2 (Kumano et al., 2003). The authors showed that the hematopoietic activity of the endothelium in Notch1-/- embryos was abolished, whereas Notch2-/- embryos exhibited normal hematopoietic activity, showing that, unlike Notch2, Notch1 was essential for normal aortic hematopoiesis. These results have been confirmed by the analysis of RBP-JK-/- mutants, which are therefore unable to transduce Notch signaling via the classical pathway. These mutants exhibit an absence of hematopoietic progenitors, accompanied by a lack of expression of the transcription factors Runx1, SCL and GATA-2 (Robert-Moreno et al., 2005). At the aortic endothelium level, Notch1 and Notch4 were conserved in the mutant but Jagged-1, and Jagged-2, expressions were lost in the mutant RBP-JK-/-. In addition, Notch 1 and Jagged-1 expressions are regionalized to the ventral endothelium of the aorta, suggesting a specific role in aortic hematopoiesis. Finally, careful examination of the Notch1 and Gata-2 patterns shows that these two genes are co-expressed at E9, 5 in the ventral endothelium of the aorta. A direct interaction between Notch1 and the Gata-2 promoter has also been highlighted, reinforcing the idea that the Notch pathway plays a fundamental role in the molecular network controlling the hematopoietic process in the aortic region (Robert-Moreno et al., 2005). A relationship between the Notch pathway and the regulation of hematopoiesis by Runx1 was also shown in the zebrafish. The transient activation of the notch pathway in mutant embryos deficient for Notch induces expression of Runx1 in the aortic region as well as the production of HSC (Burns, 2005). In addition, the inhibition of the notch pathway in zebrafish inhibits definitive hematopoiesis, while primitive hematopoiesis occurs normally (Gering and patient, 2005). Finally, it has been shown that the rescue of Notch1-/- by in vitro transfection of Runx1 was capable of restoring the hematopoietic potential, suggesting an effect of Runx1 downstream from the notch pathway (Nakagawa et al., 2006). Even though Notch1 signaling is required to initiate expression of key hematopoietic regulators such as Gata2, its activity also needs to be restricted to allow the completion of the EHT, as high Notch activity promotes an arterial program in

endothelial cells, while suppressing the hematopoietic program (Lizama et al., 2015). In fact, a reporter system that differentially labels cells with high or low Notch activity has identified two populations in the dorsal aorta, an arterial-restricted population with high Notch activity and a hematopoietic competent population with low Notch activity (Gama-Norton et al., 2015). That same study further revealed the involvement of two different Notch ligands, with Dll4 inducing high Notch activity and thus implementing the arterial program, while Jag1 facilitated the hematopoietic program by restricting Notch activity. A down-regulation of Notch signaling was also observed during the maturation of pre-HSCs into fully competent HSCs.

Furthermore, the notch ligands regulator mind bomb-1 was shown to be crucial for intra-embryonic hematopoietic progenitor formation, mainly when expressed in the non-endothelial compartment. This suggests a paracrine action of the notch ligands on the endothelial cells (Yoon et al., 2008). c-Mpl, the receptor for thrombopoietin has been shown to be expressed by IAHC in mice, and its absence led to diminished HSC activity in E11 AGMs (Petit-Cocault et al., 2007).

V. Mechanisms that allow a cell to leave a tissue during development

1. Epithelial-mesenchymal transition (EMT)

Epithelial-mesenchymal transition (EMT) is a morphogenetic process of cells that adopt an epithelial organization in their developmental ontogeny or homeostatic maintenance. During EMT, epithelial cells lose their cell polarity and cell-cell adhesion and gain migratory and invasive properties to become mesenchymal stem cells; these are multipotent stromal cells that can differentiate into a variety of cell types. EMT is essential for numerous developmental processes of many tissues and organs in the developing embryo, and numerous embryonic events such as gastrulation, neural crest formation (Shook and Keller, 2003), heart valve formation, secondary palate development, and myogenesis. In multicellular organisms, epithelial and mesenchymal cells differ in phenotype as well as function, though both share inherent plasticity. Epithelial cells organize themselves with a uniform stable apicobasal polarity and are often associated with barrier junctions and a specialized layer of the extracellular matrix, whereas mesenchymal cells lack this polarization, have a spindle-shaped morphology and interact with each other only through focal points (Hamidi and Sheng, 2018). It is well known that epithelial cells express high levels of E-cadherin, whereas mesenchymal cells express those of N-cadherin, fibronectin, and vimentin. Hence, EMT entails profound morphological and phenotypic changes to a cell Figure 34.

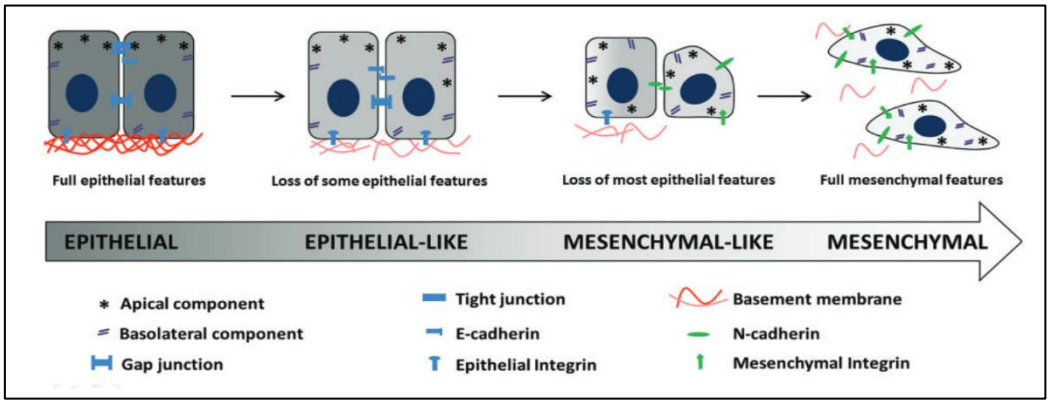


Figure 34: EMT entails profound morphological and phenotypic changes to a cell
EMT involves progressive loss of epithelial markers and acquisition of mesenchymal ones
Co-expression of both epithelial and mesenchymal markers defines intermediate states of EMT. The relative ratio between epithelial and mesenchymal markers is used to define either an epithelial- or mesenchymal-like intermediate state.
Figure from hamidi and Sheng 2018

EMT and its reverse MET (mesenchymal-epithelial transition) regulation involves a conserved set of master transcription factors which are able to directly or indirectly repress E-cadherin i.e. ZEB, SNAIL, TWIST, TCF, KLF … as well as several pathways i.e. TGF-β, FGF, EGF, HGF, Wnt/beta-catenin, and Notch… and EMT promoting factors, with this time an impact on cellular physiology i.e. hypoxia, E3 ubiquitin ligase, inflammation (Zeisberg and Neilson, 2009).

EMT is defined as a metastable state which describes a transiently stable state that is different from the more stable end state and often expresses both epithelial and mesenchymal markers. This definition allowed the description of three separable elements in the EMT: two end stable states with epithelial or mesenchymal identity and a third transitional one which corresponds to the overlapping of transitional states named intermediate or metastable states. It is likely that all three subtypes of EMT will share part of a signature, but some aspects of the signature may be unique to a particular subtype. The most visible protein signatures include at least one changing cytoskeletal protein, transcription factor, adherence protein, and perhaps the changing levels of discrete microRNAs.

2. The parallel between EHT and EMT

During development, Endothelial to Hematopoietic Transition EHT is described as a partial EMT process in which the hemogenic endothelium, like most of the endothelial structures, contains stereotypic epithelial characteristics such as apicobasal polarity, tight junctions, adherens junctions, and basement membrane. It also involves similar cellular processes found in canonical EMTs.

A parallel can be drawn between EHT and EMT. As described in the previous chapter (genes and pathways) it has been shown that NOTCH1/Jagged1 as a main EHT driver (Robert-Moreno et al., 2008). When this pathway is activated in non-hemogenic ECs, it results in an EMT like process called endothelial to mesenchymal transition EndMT and ECs lose their endothelial markers (PECAM1, Ve-Cad) but this time instead of acquiring hematopoietic fate they acquire mesenchymal features and expresses FN1 and a-SMA. In addition to this pathway Runx1 as described earlier is a critical factor for the EHT, it has been shown to be implicated in the initiation of the EMT process through TGFB dependent pathway, resulting in decreasing epithelial markers or through TGFB independent process, resulting in increasing ZEB2

expression which is expressed in the AGM of E10.5 mouse. HSPCS need ZEB2 mediated transition towards a more mesenchymal intermediate EMT state in order to acquire migratory properties necessary for the invasion and colonization of primary and secondary hematopoietic organs Figure 35.

In the bone marrow, several lines of evidence suggested that the interaction of HSCs with the different components of the niche is essential for their maintenance, quiescence and self-renewal balance. At the molecular level, in addition to E-cadherin and Notch pathway members, several groups identified other cell-cell interaction proteins i.e Tie2, JamC, Esam and EPCR as playing a role in HSC niche interactions. (Roch et al., 2017). Together all these findings suggest that LT-HSCs are strongly associated with an appropriate metastable state of EMT regulated by the niche architecture.

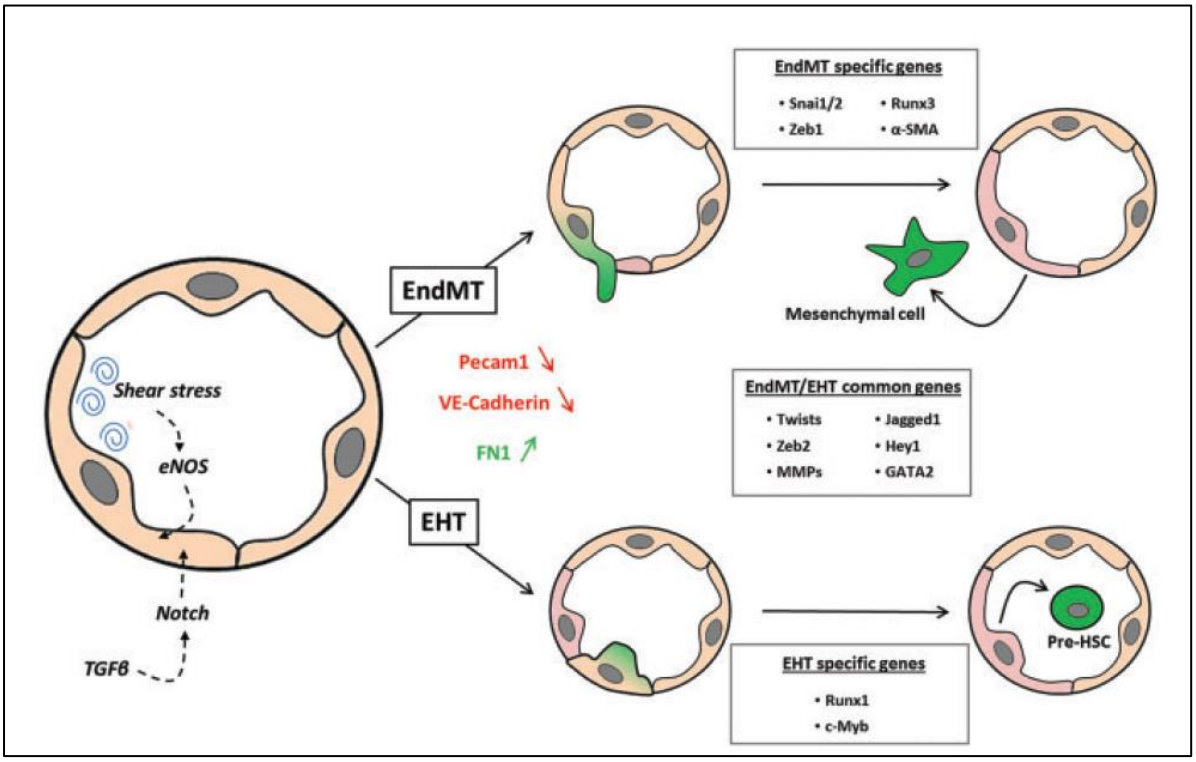


Figure 35: Comparative view of molecular determinants involved in EndMT and EHT processes.
Upon TGF-b and shear stress stimulation, some endothelial cells undergo an EMT-like process characterized by downregulation of epithelial markers and upregulation of mesenchymal ones. EndMT and EHT processes share a set of EMT regulators, but fundamental differences can also be observed (e.g. Runx3 for the EndMT and Runx1 for the EHT).
Figure from hamidi and Sheng 2018

3. POFUT2 in EMT and EHT

GDP-fucose protein O-fucosyltransferase 2 (POFUT2) is an enzyme responsible for adding fucose sugars in O linkage to serine or threonine residues in Thrombospondin Repeats TSRs while GDP-fucose protein O-fucosyltransferase 1 (POFUT1) is known to O fucosylate EGF repeats. POFUT1 and POFUT2 are highly selective, modifying only properly folded repeats containing the appropriate consensus sequences. Therefore, POFUT2 modifies TSRs with a specific consensus sequence: C-X2–3-(S/T)-C-X2-G and this O-fucosylation of TSR containing proteins is critical for their function. Deficiency of O-fucosylation leads to malignancy and developmental abnormalities like Peter's Plus syndrome.

Almost all glycosyltransferases reside in the Golgi apparatus. However, POFUT2, as well as the related enzyme POFUT1, have recently been shown to reside in the endoplasmic reticulum.

In order to confirm that POFUT2 has TSR specific O-fucosyltransferase activity in mouse, Du et al used purified POFUT2 (POFUT2-myc-His) and assayed its ability to O-fucosylate the third TSR (TSR3) of human thrombospondin 1 (TSP1). O-Fucosylation of TSP1-TSR3 increased linearly with an increasing amount of POFUT2-myc-His. In addition, they showed that ERE motif is essential for the protein O-fucosyltransferase activity of mouse POFUT2 (Du et al., 2010).

POFUT2 seems to have 3 isoforms produced by alternative splicing and it is widely expressed in the embryo. In normal development, isoform A is expressed in fetal liver and peripheral blood lymphocytes. Isoform B is expressed in spleen, lung, testis, bone marrow, thymus, pancreas, prostate, fetal brain, fetal liver, and fetal kidney. Isoform C is expressed in brain, heart, spleen, liver, lung, stomach, testis, placenta, skin, thymus, pancreas, mammary gland, prostate, fetal brain, fetal liver and fetal heart (Menzel et al., 2004). Using SMART database, Du et al looked for the TSR proteins containing the POFUT2 consensus recognition sequence, they identified 51 putative targets in mice. 10 of these targets are transmembrane proteins and 41 are secreted proteins Table 2.

Table 2: Mouse Proteins with putative O-fucosylation site within TSRs

TSR class	Mouse Protein	# Consensus Sequences	Family
Group 1	ADAMTS1-4, ADAMTS5 [95], ADAMTS6-10, ADAMTS12, ADAMTS13 [16], ADAMTS14-20	1–12	a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS)
	ADAMTSL1 [18,95], ADAMTSL2-6	1–10	ADAMTS-like
	BAI1-3	4	brain-specific angiogenesis inhibitor
	Cfp (Properdin) [15]	4	complement component
	CILP2	1	cartilage intermediate layer protein/nucleotide pyrophosphohydrolase
	HMCN1	6	hemicentin
	PAPLN (Papilin)	4	
	SEMA5A, SEMA5b	2	semaphorin
	THBS1 (TSP-1) [1,15], THBS2(TSP-2)	3	thrombospondin
	ISM1, ISM2, XM_283765	1	isthmin
	UNC5a	1	UNC5
Group1 and Group 2	C6	1	complement component
	SSPO (Sco-spondin)	12	spondin
	THSD7a, THSD7b	4–5	thrombospondin, type I, domain-containing
Group 2	CCN1 (CYR61), CCN2 (CTGF), CCN3 (NOV), CCN4(ELM1/WISP1), CCN5(COP1/ WISP2), CCN6 (WISP3)	1	CCN (Cyr61, Ctgf, Nov)
	SPON1 (f-spondin) [15]	4	spondin

The majority of POFUT2 targets are constitutive component of ECM or secreted matrix-associated proteins able of influencing several processes as cell adhesion/migration, ECM synthesis, and remodeling and finally modulating growth factor synthesis. Consequently, the loss of POFUT2 disrupts embryonic processes.

To investigate the biological role of POFUT2 in mice, they characterized two POFUT2 gene trap insertions: Pofut2$^{Gt(RST434)Byg}$ and Pofut2 $^{Gt(neo)699Lex}$ inactivating one copy of POFUT2 at a time. They could demonstrate that POFUT2 was a key modulator of gastrulation (cell differentiation, tissue morphology, and tissue boundaries) and that the disruption of POFUT2 gene resulted in embryonic lethality by E10.5. POFUT2 mutant showed from one hand an abundance of mesenchymal cells, markedly expanded the expression of SNAIL and a reduction in the expression of pluripotency markers in the epiblast suggesting that O-fucosylation of TSR protein by POFUT2 is essential to restrict EMT. From the other hand, it shows an abundance of FLK1 and PECAM1 positive cells and absence of differentiated primitive erythrocytes, demonstrating that O-fucosylation of TSR by POFUT2 was important for the distinction between the hematopoietic and vascular endothelial lineage.

Together these results showed that the loss of POFUT2 in mice resulted in unrestricted EMT and biased differentiation of vascular endothelial cells.

VI. The Extracellular Matrix and cytoskeleton and their implication in hematopoiesis

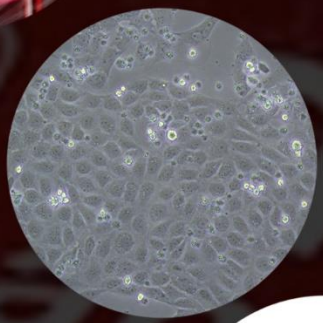
The extracellular matrix (ECM) is a three-dimensional network of extracellular macromolecules, such as collagen, enzymes, and glycoproteins, that provide structural and biochemical support of surrounding cells. It is well described that ECM provides critical support for vascular endothelium through adhesive interaction with integrin on the ECs surface. Therefore, ECM provides a scaffold, important for maintaining the organization of vascular ECs into blood cells. As angiogenesis proceeds, the basement membrane matrix is degraded by proteinases among which membrane-type matrix metalloproteinases (MT-MMPs) are particularly significant. In this case ECM serves essential function in supporting key signaling events involved in regulating EC migration, invasion, proliferation and survival. Moreover, through specific integrin-dependent signal transduction pathways, ECM controls the EC cytoskeleton (a complex, dynamic network of interlinking protein filaments that extends from the cell nucleus to the cell membrane) to orchestrate the complex process of vascular morphogenesis (Davis and Senger, 2005).

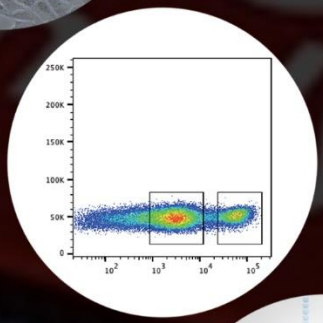
Concerning hematopoiesis, ECM was described as an important component of the BM hematopoietic microenvironment regulating some cytokines and their downstream signaling. Earlier during development, in the pre-umbilical AGM region, directly underlying the hematopoietic cluster, this region of mesenchyme expresses tenascin-C, an extracellular matrix glycoprotein known to facilitate cell-cell interactions and migration. This region of cells may, therefore, provide the microenvironmental support for the intraembryonic development of definitive hematopoietic stem cells, a process in which tenascin-C may play a pivotal role (Marshall et al., 1999). However, the cellular mechanism(s) allowing HSPC egress and migration to secondary niches are incompletely understood. Using zebrafish,(Theodore et al., 2017) showed that Mmp2 and Mmp9 are induced in response to inflammatory signals to perform necessary, but spatially and temporally distinct, functions during definitive hematopoiesis. The inhibition of MMP2 function caused accumulation of fibronectin-rich ECM, retention of runx1/ cmyb+ HSPCs in the VDA, and delayed caudal hematopoietic tissue (CHT) colonization; these defects were absent in fibronectin mutants, indicating that Mmp2 facilitates endothelial-to-hematopoietic transition via

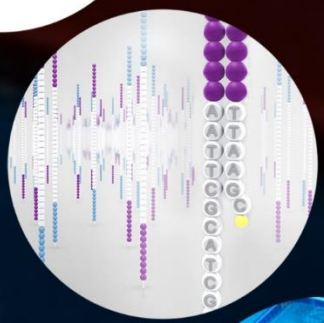
ECM remodeling. Whereas, Mmp9 was dispensable for HSPC budding, being instead required for proper colonization of secondary niches. They demonstrated that MMP2 and MMP9, which represent the gelatinase family, have been implicated in angiogenic ECM remodeling downstream of Notch and VEGF signaling.

Further understanding of the intersections between inflammatory signaling, ECM remodeling, cytoskeleton, and hematopoiesis could be exploited to improve in vitro and in vivo HSPC expansion efforts, as well as HSC mobilization and transplantation.

EMILIN 1, an extracellular matrix could potentially play a role in hematopoiesis. It is an ECM glycoprotein associated with elastic fibers and composed of an N-terminal cysteine-rich domain and the EMI domain, followed by a coiled-coil structure, a short collagenous stalk, and a C-terminal gC1q domain. EMILIN1 is particularly abundant in the walls of large blood vessels, such as the aorta, and has been implicated in multiple functions. EMILIN1 is involved in elastogenesis and in the maintenance of blood vascular cell morphology. It interacts with the α4β1 integrin through the gC1q1 domain and has strong adhesive and migratory properties for different cell types (Zanetti et al., 2004). Concerning EMILIN 2 (multimerin) a closely related molecule to EMILIN1, it is a protein shown to be secreted by endothelial cells and platelets (Hayward et al., 1995). Therefore, EMILIN1 and 2 could play a role in the endothelial to hematopoietic transition.







MATERIALS & METHODS

I. Pre-somitic mesoderm culture system

1. PSM harvesting in the quail embryo

We used quail (*Coturnix coturnix japonica*) embryo pre-somitic mesoderm (PSM) according to Yvernogeau et al, (2016). Eggs were incubated for 36-45 hours at 37±1°C in a humidified atmosphere to reach 10-18 somite pairs. Microsurgery was performed as previously described Pardanaud, 1996 #2110]. The ectoderm layer was gently cut following a line between the neural tube and the PSM. Dispase (1U/ml) was used to facilitate tissue separation. The PSM was harvested over a length of approximately 10 somites (from the most posterior formed somite to the rhomboidal sinus) on both sides of the embryo (Figure M1 B, C). Then they were sectioned into 5-6 pieces each to increase the surface exposed to the culture media and allow a better spreading during the first phases of cell attachment. Five embryos (i.e. 10 PSM) were used per culture dish. Each PSM was cut into 4 to 5 equal pieces, rinsed in Opti-MEM plus GlutaMAX™ I (Gibco Life Technologies) +5% Fetal Calf Serum (FCS, Gibco Life Technologies) +100 Units per ml penicillin/streptomycin (Gibco Life Technologies) + 1% Chicken Serum (Gibco Life Technologies) before culture.

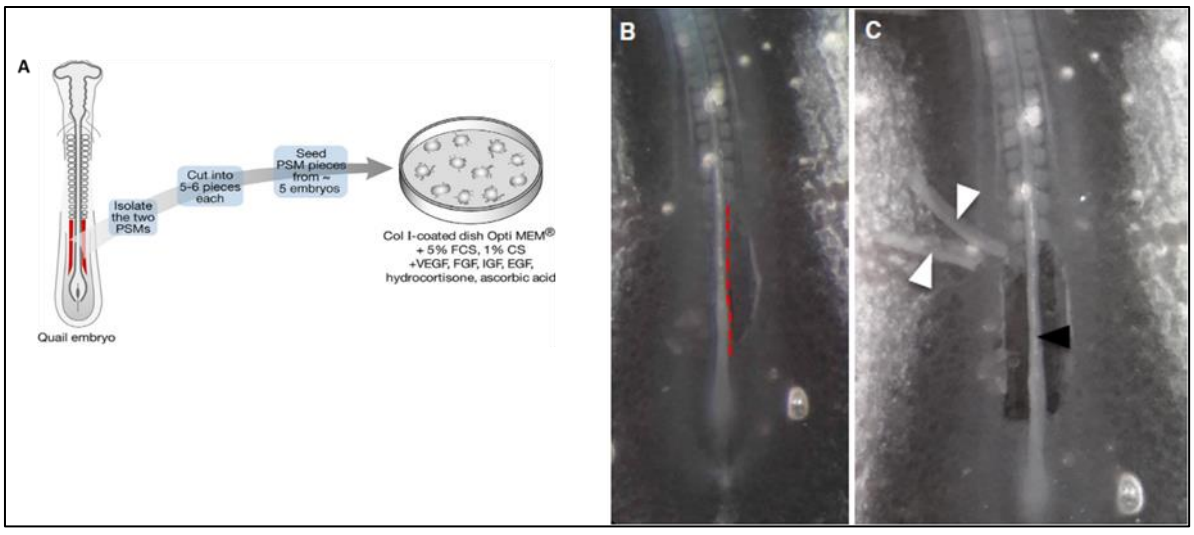


Figure M1: PSM harvesting in the quail embryo

A. Culture system (PSM in red) (Yvernogeau et al 2016)

B-C. Photos of the quail embryo dissection, Leica Binocular Loupe, magnification 40X. The ectoderm is cut following the red line (B). PSM (white arrows) are detached from the embryo on both sides of the neural tube (black arrow) (C).

2. Cell culture

PSM were cultured under the following conditions Opti-MEM® with GlutaMAX™I supplemented with 5% FCS, 1% Chicken Serum, 100 Units per ml penicillin/streptomycin and the following growth factors (PromoCell/PromoKine): hVEGF (≠C64410); 2ng/ml), hFGF (≠C60240) ; 4ng/ml), hIGF (≠C60840) ; 3ng/ml), hEGF (≠C60170) 10ng/ml), hydrocortisone (200ng/ml SIGMA ≠H6909), ascorbic acid (75ug/ml SIGMA ≠A4544). PSM were cultured in 35mm collagen I- or IV -coated dishes from Corning BioCoat (Corning - Dutscher France). Medium was changed every two days unless specified (Figure M1A).

3. Functional tests with small molecules

We selected 3 small molecules to test. PKF 118-310 (K4394 Sigma-Aldrich) an inhibitor of the Wnt pathway, SB216763 (S3442 Sigma-Aldrich) an activator of the Wnt pathway and DAPT (D5942 Sigma-Aldrich) an inhibitor of the Notch pathway. For the initial experiments, we started from reported concentrations found on the literature.

PKF118-310: 0,66µM/mL, 1µM/mL, 1,5µM/mL

DAPT: 25µM/mL and 50µM/mL

SB216763: 10µM/mL, 20µM/mL and 30µM/mL

DAPT and SB216763 were diluted in DMSO, and PKF118-310 was diluted in H20. To ensure the diluent (H20 or DMSO) did not influence the culture, we also cultured cells with media supplemented with either DMSO or water (the same volume as for the solutions of small molecules) as positive controls.

4. siRNA design and lipofection-mediated transfer

We used a knock-down strategy utilizing siRNA to inhibit our candidate genes. We designed three different siRNAs against POFUT2 and three others against TESTIN (Table M1). Lipofectamine RNAiMax was used to transfect cells at day 1 or 2 with specific siRNA or Scramble si control following the manufacturer’s protocols. The transfection efficiency was evaluated by analysing the mRNA expression of the targeted gene by qPCR.

Table M1: Sequences of the siRNA POFUT2, TESTIN and Scramble

siRNA	Sequence
Scr sense	GGUUUCGUCUAUAGAUUGUTT
Scr antisense	ACAAUCUAUAGACGAAACCTT
POFUT2 si1 sense	GCUUCACUUAUGAAGACUUTT
POFUT2 si1 antisense	AAGUCUUCAUAAGUGAAGCTT
POFUT2 si2 sense	GCAGAUUUAUGUCCUGCAATT
POFUT2 si2 antisense	UUGCAGGACAUAAAUCUGCTT
POFUT2 si3 sense	AAAUACCAUACUACGACGGTT
POFUT2 si3 antisense	CCGUCGUAGUAUGGUAUUUTT
TESTIN si1 sense	GGAUAAGUGUGAAGGAUUUTT
TESTIN si1 antisense	AAAUCCUUCACACUUAUCCTT
TESTIN si2 sense	CCAGAAGGCAUCUCAGUAUTT
TESTIN si2 antisense	AUACUGAGAUGCCUUCUGGTT
TESTIN si3 sense	GCAUGCUGCGAUUUGCCAATT
TESTIN si3 antisense	UUGGCAAAUCGCAGCAUGCTT

5. Cell preparation and isolation for RNA extraction

Non-adherent cells were collected by thoroughly pipetting and rinsing the culture dishes. Adherent cells were first trypsinized, washed in PBS/FCS (10%), centrifuged and collected. All RNA extractions were performed using the RNeasy™ kit from Qiagen SAS France. Freshly isolated PSM (10) were re-suspended into the RNeasy™ buffer solution (RLT). For RNA extraction on cultured cells, the supernatant cells are first centrifuged to remove the medium and re-suspended into the RNeasy™ buffer solution. Adherent cells after trypsination are re-suspended into the RNeasy™ buffer solution. Quality/quantity of the extracted RNA was evaluated using a Nanodrop.

6. Staining of in vitro culture, FACS analysis and cell sorting

Endothelial cells were incubated with Acetylated LDL (10µL/mL) (acetylated low-density lipoprotein from human plasma coupled to Alexa Fluor 488- AcLDL-A488; 1 mg/ml; (Life Technologies) diluted in culture medium for two hours at 37°C, 5% CO2, the dishes were rinsed in PBS/SVF (PBS 1X, 10% SVF) solution. Cells were dissociated using 0,25% Trypsin-EDTA 1X (Gibco Life Technologies) for 1min. After 2 washes with PBS/SVF to inhibit the trypsin, the supernatant was removed, the bottom was resuspended in remaining solution and the obtained samples were plated

in round bottom, 96 wells cytometry plate. Plates were centrifuged, and a DAPI solution added (200µL of DAPI/BSA solution at 0,2% DAPI). Flow cytometry was performed with MACSQuant VYB, and data were analysed with Flowjo. Cell sorting was performed with BD FACSAria Cell sorter.

II. Transcriptomic analysis

1. Next Generation Sequencing

RNA libraries were prepared by Fasteris Life Science (Switzerland). In brief, total RNA was submitted to a poly-A mRNA purification using oligodT magnetic beads. Supernatants were kept for small RNA library preparation, and poly-A–RNAs after elution from the beads were prepared according to the manufacturer's protocol (TruSeq RNA Sample Prep- Kit V2; Illumina). Transcripts were broken at 95°C in presence of zinc, and first-strand cDNA syntheses were performed using random primers. A second-strand cDNA synthesis was performed in the presence of deoxyuridine triphosphate, and after a 3′ A addition step, adapters were ligated, and an amplification by PCR was per- formed to generate the DNA colony template libraries. Small RNA libraries were performed according to the manufacturer's protocol (TruSeq Small RNA Library Prep kit; Illumina). After acrylamide gel purification of small RNA between 18–30 nt, single-stranded ligation of 3′ adapter and the 5′ adapter were performed before reverse transcription and PCR amplification to generate the DNA colonies template. All the samples were sequenced using 1 × 125–bp single reads high-throughput sequencing (RNA-Seq) in single lane on a HiSeq 2000 sequencing system (Illumina).

2. Single-cell RNA-seq from D0, D4 and D6 cell culture using 10x Genomics Chromium

Samples were prepared as outlined by the 10x Genomics Single Cell 3′ v2 Reagent Kit user guide. Briefly, the samples were washed twice in PBS (Life Technologies) + 0.04% BSA (Sigma) and re-suspended in PBS + 0.04% BSA. Sample viability was assessed via Trypan Blue (Thermo Fisher). Following counting, the appropriate volume for each sample was calculated for a target capture of 6000 cells. Samples below the required cell concentration as defined by the user guide (i.e., <400 cells/µl) were pelleted and re-suspended in a reduced volume and counted again prior

to loading onto the 10x Genomics single-cell-A chip. After the reverse transcription, cDNA was recovered using Recovery Agent provided by 10x followed by a Silane DynaBead clean-up (Thermo Fisher) as outlined in the user guide. Purified cDNA was amplified for 12 cycles before being cleaned up using SPRIselect beads (Beckman). Samples were diluted 4:1 (elution buffer (Qiagen): cDNA) and run on a Bioanalyzer (Agilent Technologies) to determine cDNA concentration. cDNA libraries were prepared as outlined by the Single Cell 3′ Reagent Kits v2 user guide with appropriate modifications to the PCR cycles based on the calculated cDNA concentration (as recommended by 10X Genomics). Analysis were performed using partek flow single cell workflow.

3. Bioinformatic analysis

A. RNA-Seq data analysis

RNA-Seq analysis was performed on Partek Flow™. For mRNA libraries, sequence reads in fastq format were aligned to the quail reference genome assembly RefSeq GCF_001577835. using the STAR aligner. The number of reads for all the features were then counted using the quantification to a model E/M and normalized for each library, and then ANOVA two ways were performed to compare read values between different libraries (Figure M2).

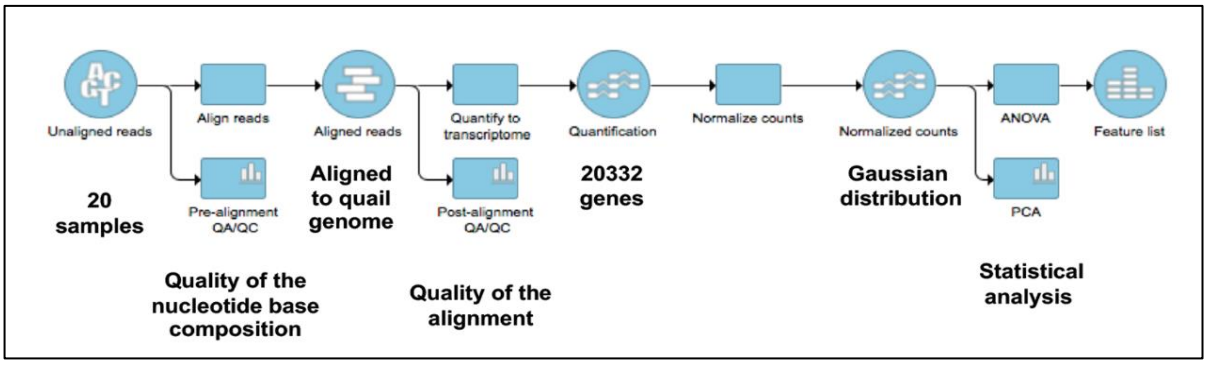


Figure M2: Workflow of RNAseq analysis

The unaligned reads of the 20 samples were analyzed for the quality of the nucleotide base composition and then aligned to the quail genome. The quality of the alignment was assessed by the QA/QC post-alignment quality. We quantified 20332 genes and then normalized to be able to perform the statistical analysis.

The lists of differentially expressed genes were generated using ANOVA and Venn Diagram. GO was performed using DAVID (http://david.abcc.ncifcrf.gov). To find out the network structure of the gene set HEC, we used WGCNA (Langfelder and Horvath, 2008). Adjacencies were given according to signed Pearson correlation. The soft threshold power b that resulted in approximate scale-free topology was 36. Modules were constructed with average linkage hierarchical clustering.

B. Single cell data analysis

10x data were processed using Cell Ranger 2.1.0 with default parameters. Using Partek Flow reads were aligned to the coturnix japonica reference GCF_001577835. Partek Flow single cell workflow was used to perform graph-based clustering of the cells. Briefly, cells with <200 detected genes and with a cell counts > 9000 were filtered out. In addition, cells that showed lower number of genes than expected from the number of UMIs detected were filtered out as low quality cells. We then excluded features where value equal to zero. The data were log normalized (Counts per million, log2+1). Principle Component analysis PCA was performed using variable genes. Eigenvalues was used to determine the statistically Significant PCs to be used for graph-based clustering. t-SNE was used to visualize the clusters. Cell types were classified according to the expression of previously reported marker gene and GO categories.

4. Statistical analysis

Data were processed using GraphPad Prism v.5. All analyses were performed using Student's t tests, except where stated otherwise. Graphs and error bars reflect means ± s.e.m. In all corresponding figures, $*$ represents $p < 0.05$. $**$ represents $p < 0.01$. $***$ represents $p < 0.001$. ns represents $p > 0.05$. Replicate information is indicated in the figures.

RESULTS

CHAPTER 2

I. LDL^{int} cells represent an EC population upregulating RUNX1 and down-regulating CD31 and CD144.

One advantage of the culture system (Yvernogeau et al., 2016) is the stepwise progression of the mesoderm towards the hematopoietic fate that allows isolating discrete cellular steps over the culture period of 12 days. In addition, by removing serum or VEGF from the culture, it is possible to prevent hemogenic EC commitment and to orient the cells towards a more advanced vascular endothelium state. We previously showed that ECs emerging in culture uptake LDL in keeping with their endothelial phenotype including those expressing RUNX1. We reasoned to use the levels of LDL uptake to separate ECs from hemogenic ECs. This is based on the observation that RUNX1 initiated expression is initiated at D3 in only a few ECs but is found expressed in almost 100% of the cells at D5. RUNX1 expression is shortly followed by that of the earliest hematopoietic-specific transcription factors, in particular, PU1, a direct target of RUNX1. In this respect, hemogenic ECs i.e., RUNX1-expressing cells, should undergo a decrease of LDL uptake in keeping with the progressive loss of the endothelial program. Based on this assumption, D2 and D4 adherent cells were incubated with Ac-LDL coupled to Alexa 488 (hereafter referred as LDL) over a period of two hours, trypsinized and Fluorescence-Activated Cell Sorted (FACS) based on the LDL level of expression. At D2, FACS analysis showed cells harboring a wide, albeit continuous, range of fluorescence from nil to high (Figure R1A). At D4, we separated two populations: one with a high level of expression of LDL (LDL^{hi}) and another with an intermediate level of expression (LDL^{int}) (Figure R1B). In the absence of serum, most of the population is within the LDL^{hi} (Figure R1C).

The different LDL populations displayed on Figures R1A-C were sorted and analyzed for the expression of two endothelial markers i.e., CD31 (PECAM) and CD144 (VE-Cadherin) and the canonical hemogenic endothelial marker RUNX1. Gene expression on the D0 pre-somitic mesoderm served as a baseline. Both CD31 mRNA was highly expressed in LDL^{hi} vs LDL^{int} cells with a mean fold change of two (Figure R1D) CD144 mRNA expression was high at D2 and D4 with no clear change on the level of expression between D2 and D4 and between D4 and D4 without SVF but was at least 8-fold decreased in D4 LDL^{int} cells (Figure R1E). However, RUNX1

mRNA expression was strongly increased in LDL^{int} compared with the other LDL-positive populations (Figure R1F).

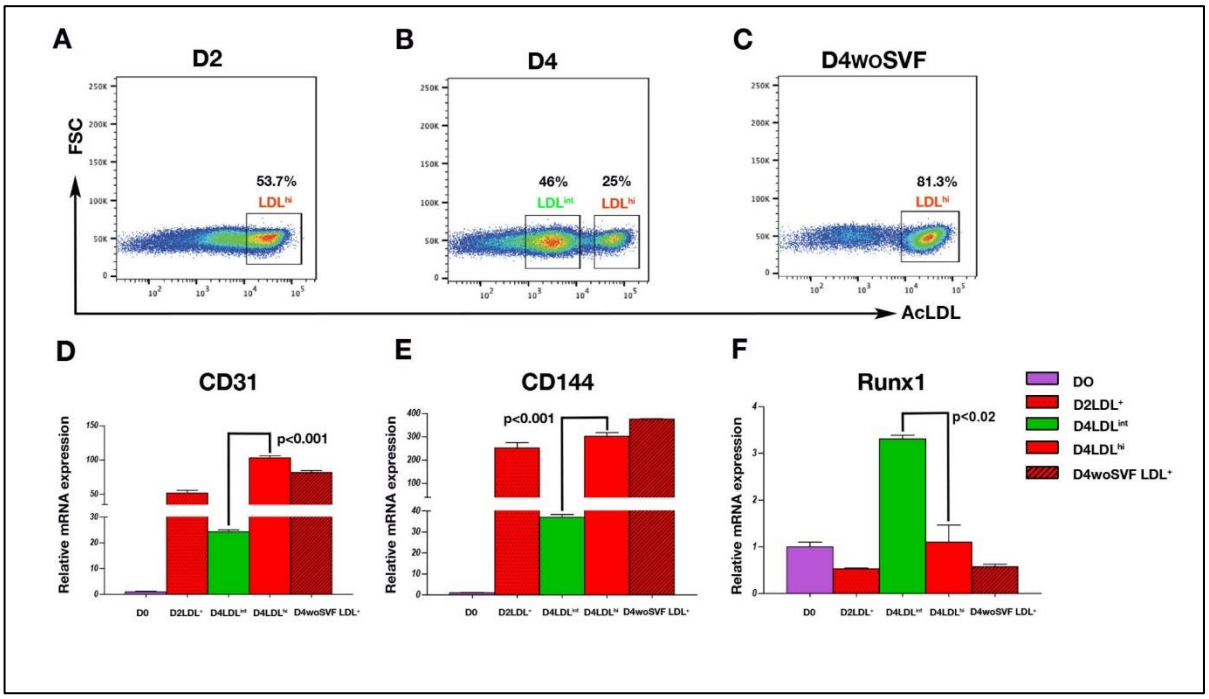


Figure R1: The level of LDL uptake separates endothelium from hemogenic endothelium

A-C: Flow cytometry analysis showing the expression of LDL uptake in the culture. A: analysis at D2. The cells display a continuum of LDL uptake from nil to high. Frame shows the sorted population. B: analysis at D4 in normal culture conditions. Two distinct populations with contrasted LDL uptake are seen. The left frame isolates the LDL^{int} population and the right frame, the LDL^{hi} population. C: analysis at D4 in the absence of fetal calf serum known to orientate the cells towards the vascular endothelium lineage (Yvernogeau et al., 2016). A large percentage of the population is contained within the LDL^{hi} population in keeping with their strong endothelial commitment. D-F: quantitative PCR on the different cell fractions for CD31 (D), CD144 (E) and RUNX1 (F). CD31 and CD144 are decreased in the LDL^{int} population compared with the other endothelial populations. RUNX1 is strongly increased in the LDL^{int} population compared with the other endothelial fractions (n=2). Significance is given compared with the D4 LDL^{hi} fraction. Data are shown as means ± SEM.

II. Sample isolation strategy

Based on this result, we isolated the different cell fractions to perform NGS. We used the level of LDL uptake to isolate ECs (LDL^{high} or $^{+)}$ from hemogenic ECs (LDL^{int}). Presomitic mesoderm was directly isolated from the embryo whereas hematopoietic cells were collected as the floating fraction, in keeping with Yvernogeau et al. (2016) (Figure R2A). We used high-throughput sequencing to compare the molecular signatures of the different cell fractions. We isolated 20 samples corresponding to the different cell fractions. mRNAs were isolated and the libraries were sequenced using next-generation sequencing Illumina technology. According to the quality control, we detected 20 332 genes. The review of the following analysis is shown in Figure R2B.

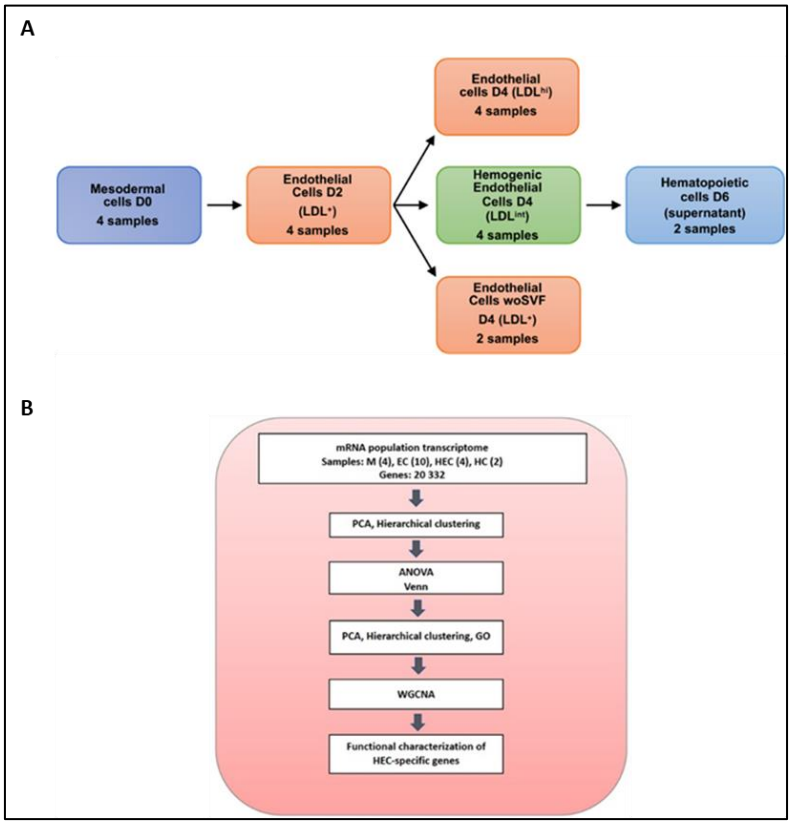


Figure R2: Sample isolation strategy

A- In total we isolated 20 samples. 4 samples for each of the populations: PSM, EC D2, EC D4, HEC D4 and 2 samples for each of HC and EC woSVF populations. PSM were directly isolated from embryos, HC were collected as floating cells whereas EC and HEC were cell sorted based on the level of LDL uptake.

B- Roadmap of the used strategy of the analysis. Starting from the sequenced data going through the bioinformatics and statistical analysis and finishing with the functional characterization of HEC specific genes.

III. Bioinformatic analysis

PCA on the entire set of mRNAs and samples (20 samples; 20 332 genes) indicated differences between the different cell fractions. PC1 (30.2% of the variance) and PC3 (9.85% of the variance) allowed to discriminate ECs from pre-somitic mesoderm (M) and hemogenic ECs (HEC) from hematopoietic cells (HC), respectively (Figure R3A). A first hierarchical clustering on the entire set of mRNAs and samples showed some discrepancies in the ordering with the M fraction found between the HEC and HC fractions (Figure R3B).

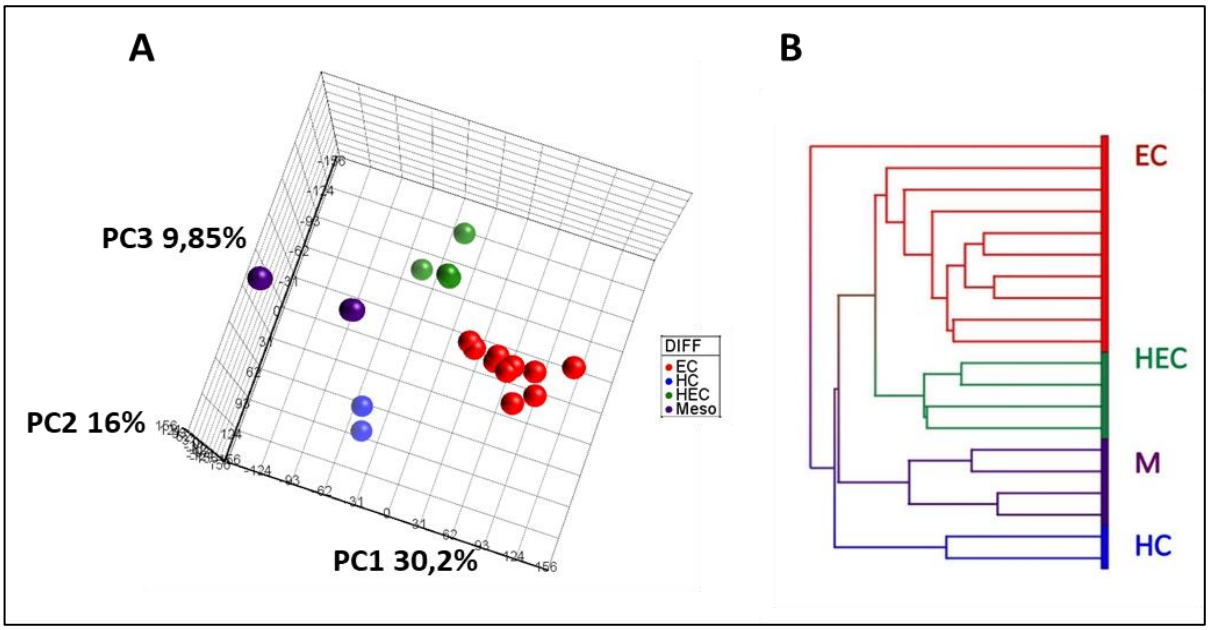


Figure R3: Analysis of the entire set of mRNA and samples (20 samples, 20332 genes)

A- 3D PCA: PC1 (30.2% of the variance) and PC3 (9.85% of the variance) allowed to discriminate ECs from pre-somitic mesoderm (M) and hemogenic ECs (HEC) from hematopoietic cells (HC), respectively.

B- Hierarchical clustering: showed some discrepancies in the ordering with the M fraction found between the HEC and HC fractions.

To identify the genes that were specifically up- or down-regulated in the different cell populations we followed a two-step procedure. First, we compared pairwise using ANOVA (with p-values <0.06 and fold changes ≥1.45 or ≤−1.45) the M, EC, HC and HEC cell populations. Second, we considered using Venn diagrams the gene sets that were up-regulated (or down-regulated) in a given population according to the 4 comparisons (Figure R4A). This procedure led to the identification of 2462 up-regulated and 2087 down-regulated genes (upregulated DEGs in the 4 populations Table S1; Table S2). We then performed a novel PCA that provided a clear-cut discrimination between the 4 populations, according to PC1 (43% of the variance) corresponding to the contrast between M and EC and to PC3 (17% of the variance) corresponding to that contrast between HC and HEC (Figure R4B). And this time hierarchical clustering showed a perfect ordering going from M through EC and HEC fractions until HC population (Figure R4C).

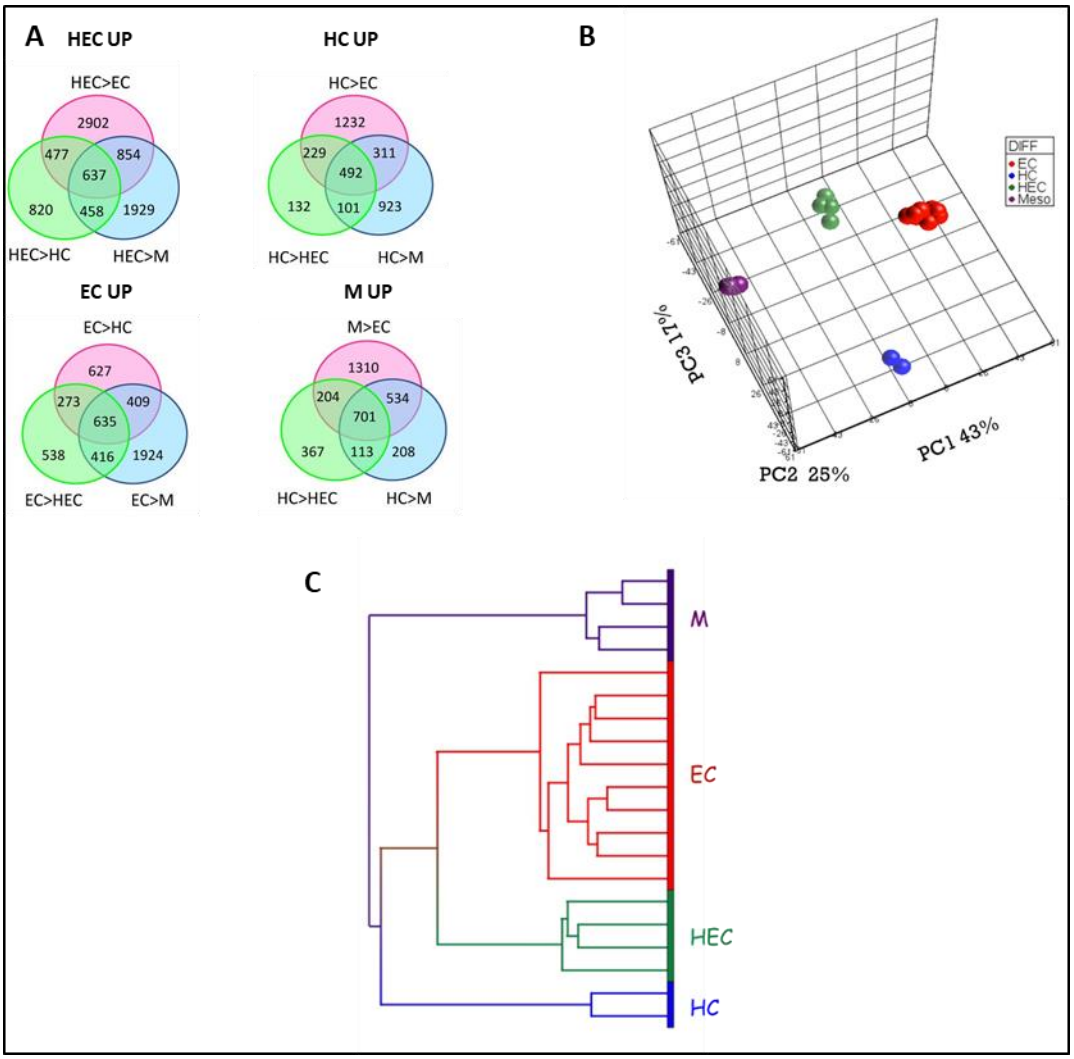


Figure R4: Identification of the specifically up-regulated genes in the different cell populations

A- we used Venn diagrams to isolate the gene sets that were upregulated in a given population according to the 3 comparisons. This led to 4 Venn diagrams one for each population.

B- 3D PCA provided a clear-cut discrimination between the 4 populations, according to PC1 (43% of the variance) corresponding to the contrast between M and EC and to PC3 (17% of the variance) corresponding to that contrast between HC and HEC.

C- Hierarchical clustering showed a perfect ordering going from M through EC and HEC fractions until HC population.

Analysis of the GO categories clearly confirmed the phenotypic differences between the 4 distinct populations. The M cell population was enriched in genes implicated in spliceosome, Hippo and Hedgehog signaling pathways and shared the category adherens junctions with the EC population. The EC population showed enrichment in lysosome, FoxO, Wnt and Rap1 signaling pathways and glycosaminoglycan degradation. The HEC population was enriched in genes implicated in focal adhesion, ECM, regulation of actin cytoskeleton, MAPK and PI3K-Akt pathways, glycosaminoglycan biosynthesis and axon guidance. Finally, the HC population was enriched in genes implicated in metabolic pathways, oxidative phosphorylation, Jak-STAT and B and T cell receptor signaling pathways (Figure R5).

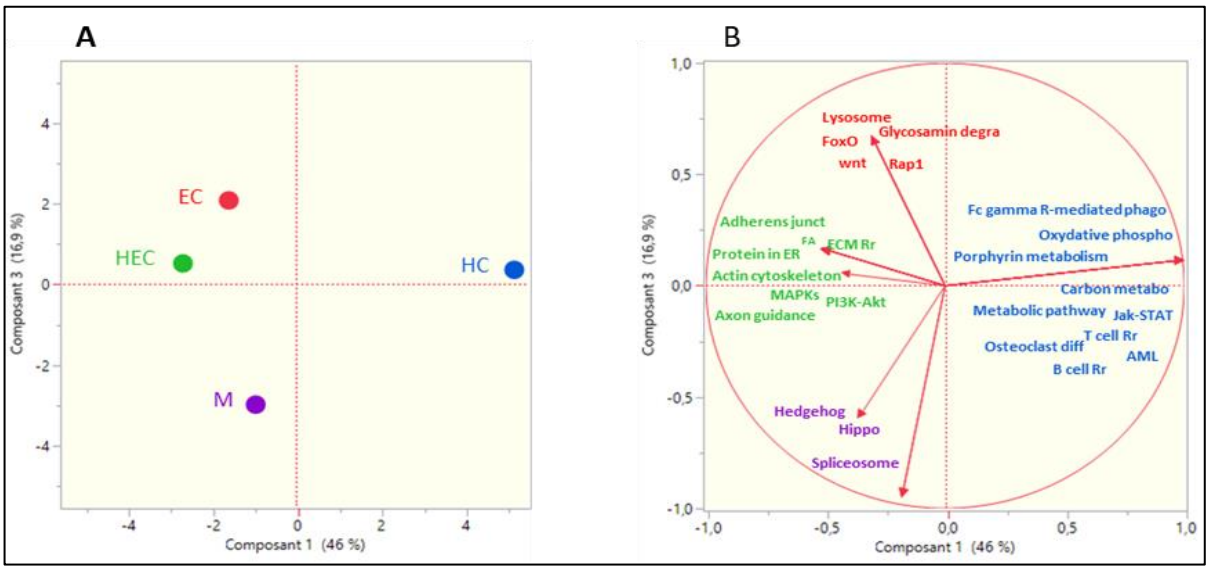
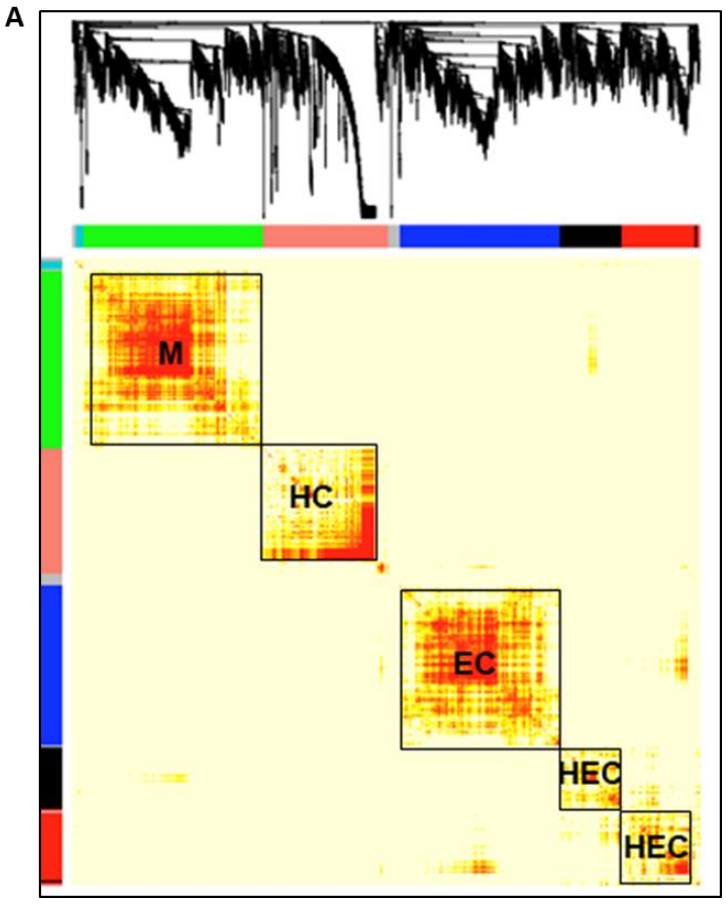


Figure R5: Analysis of the gene ontology categories

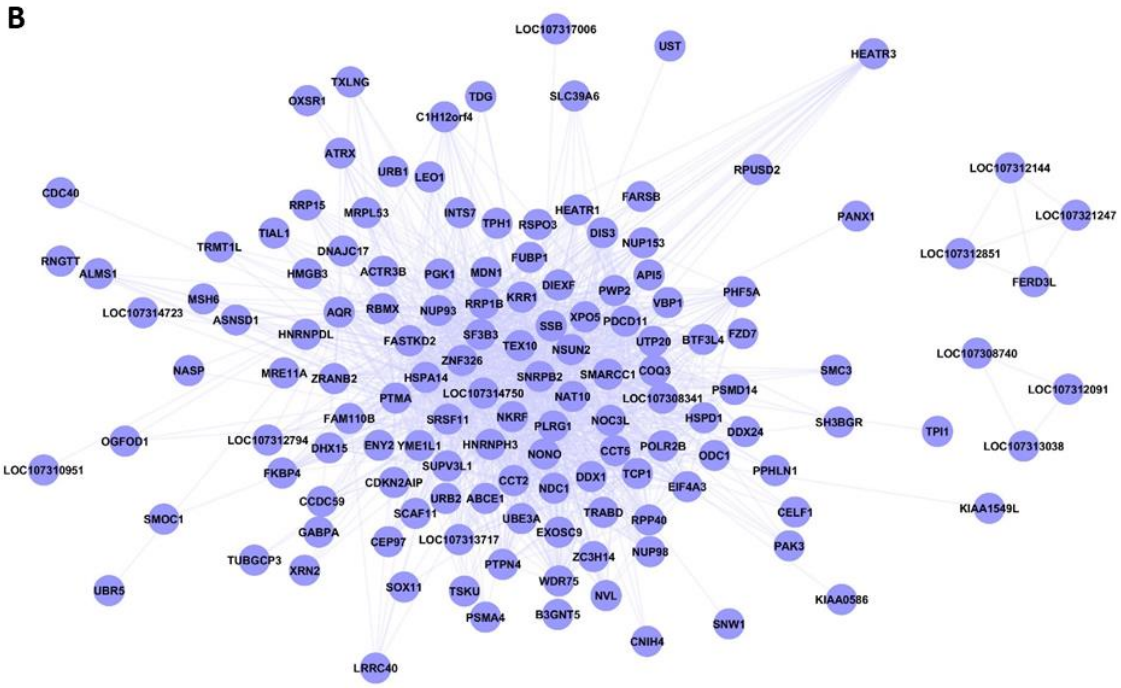
A- Score plot (PC1 v/s PC3) on 28 GO categories and 4 populations allowed the discrimination between HC and HEC (PC1 46%), EC and M (PC3 16.9%).

B- Loading plot (variables as red arrows) defined the specific categories of GO for each population.

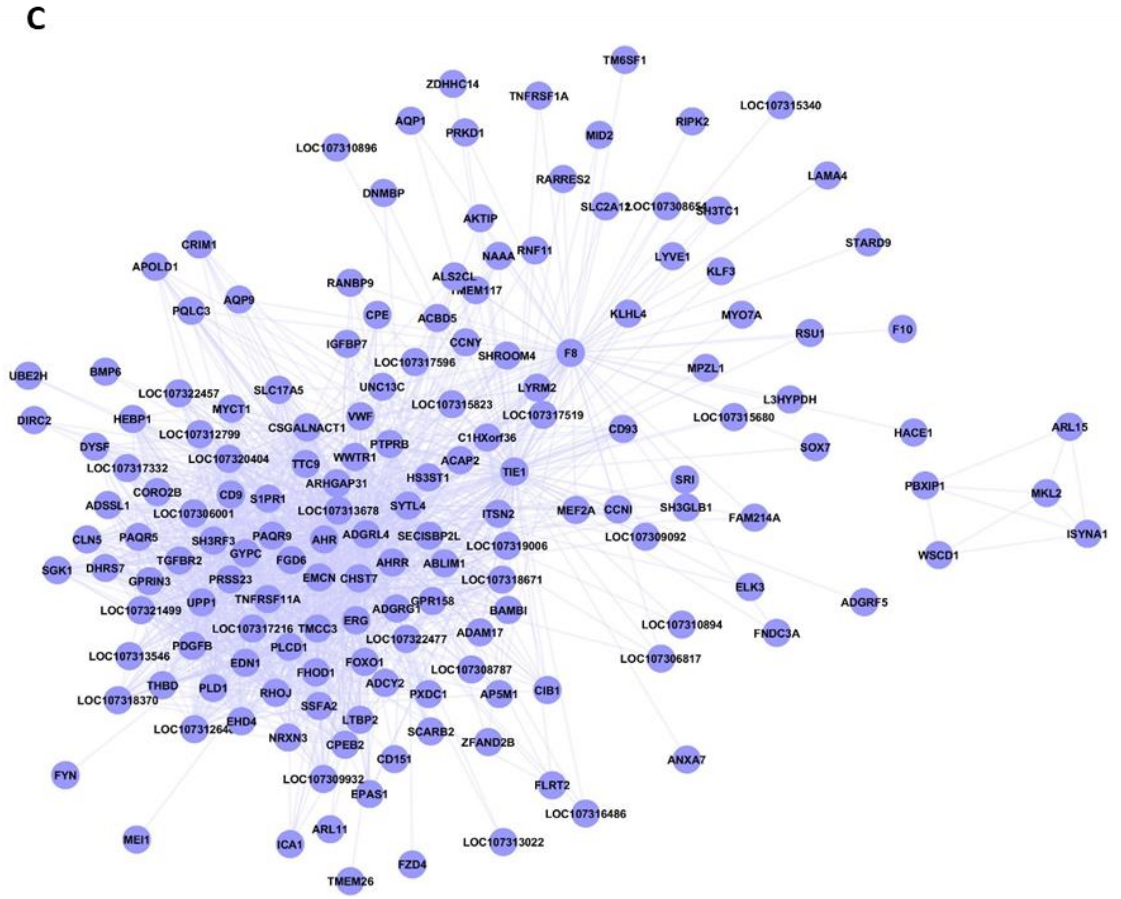
To find the structure of the gene network indicating how genes from our dataset are interconnected, we used weighted gene correlation network analysis (WGCNA). WGCNA makes use of correlation between genes to identify coordinately expressed genes (Langfelder and Horvath, 2008). Moreover, this method indicates how genes are correlated to an external genetic trait, corresponding in this work to the factor support quantified for each line by PCA. This analysis allowed network construction corresponding to the set of 2452 up regulated genes without relying upon literature mining. WGCNA revealed the existence of 5 modules, one for EC, one for M, one for HC and two for HEC (Figure R6A). Each module corresponds to a network, in order to visualize these networks, we used Cytoscape an open source software platform for visualizing complex networks. Each node represents a gene and edges represent the connection between genes (Figure R6 B, C, D) and (Figure R7 A, B).



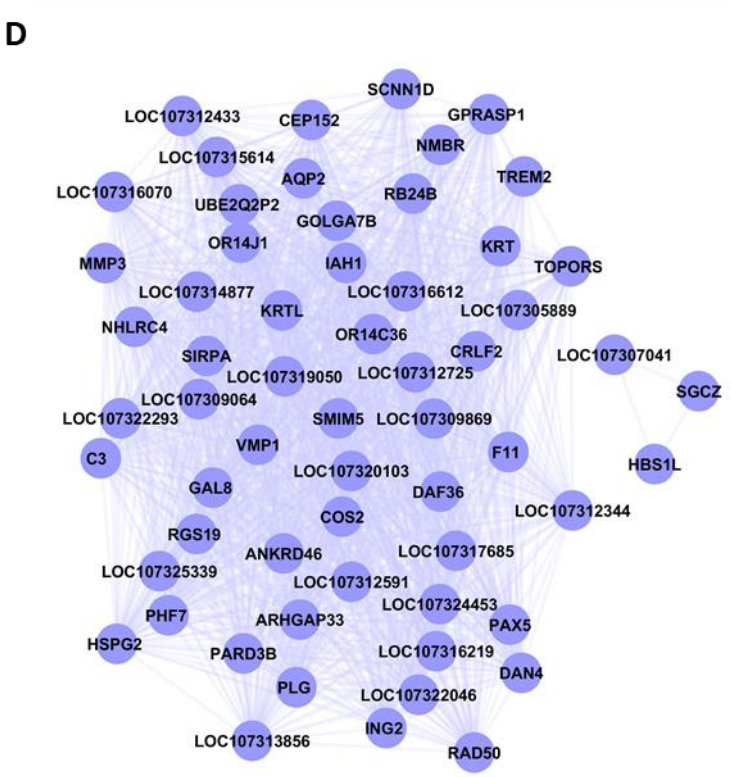
A- Dendrogram of modules identified by WGCNA



B- Green Mesoderm network



C- Blue EC network

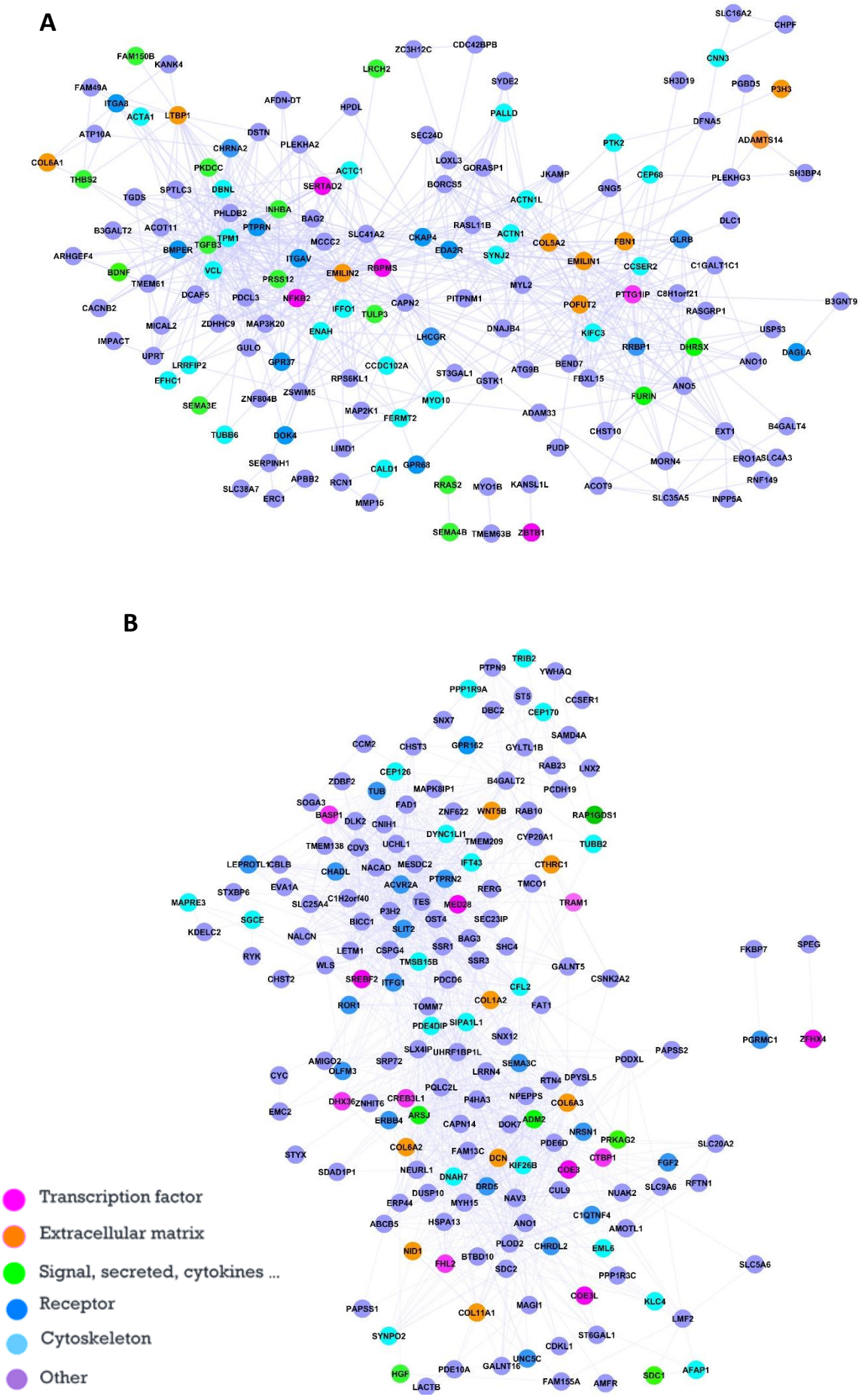


D- Salmon HC network

Figure R6: Identification of five cell specific modules using WGCNA and their attributed networks.

A- Dendrogram of modules identified by WGCNA: Green module was attributed to Meso population, Salmon module to HC population, Blue module to EC population, Black module to HEC as well as Red module.
B- Cytoscape visualization of the Green M network
C- Cytoscape visualization of the Blue EC network
D- Cytoscape visualization of the Salmon HC network

From this point, we decided to focus on the two HEC modules of up-regulated genes (Figure R7 A, B). The designated red and red black (corresponding to 192 and 253 genes respectively significantly upregulated in HEC) had a very particular extended topology consisting in relatively dense cores connected by a few genes that appear as very prominent hubs: the ECM components POFUT2, EMILIN1, EMILIN2, COL5A1, FBN1 and COL1A2, COL6A2, COL6A3, CSPG4 in the red and black modules, respectively, the ECM enzyme PLOD2, the migration cue SLIT2, the transcription factor CREB3L1 and the protein TES in the black module, and the transcription factor NFKB2 and the cytokines TGFB3 and INHBA in the red module. The gene connectivity of the most connected genes in each module is displayed in Figure R7 C,D.


A
B
Transcription factor
Extracellular matrix
Signal, secreted, cytokines ...
Receptor
Cytoskeleton
Other

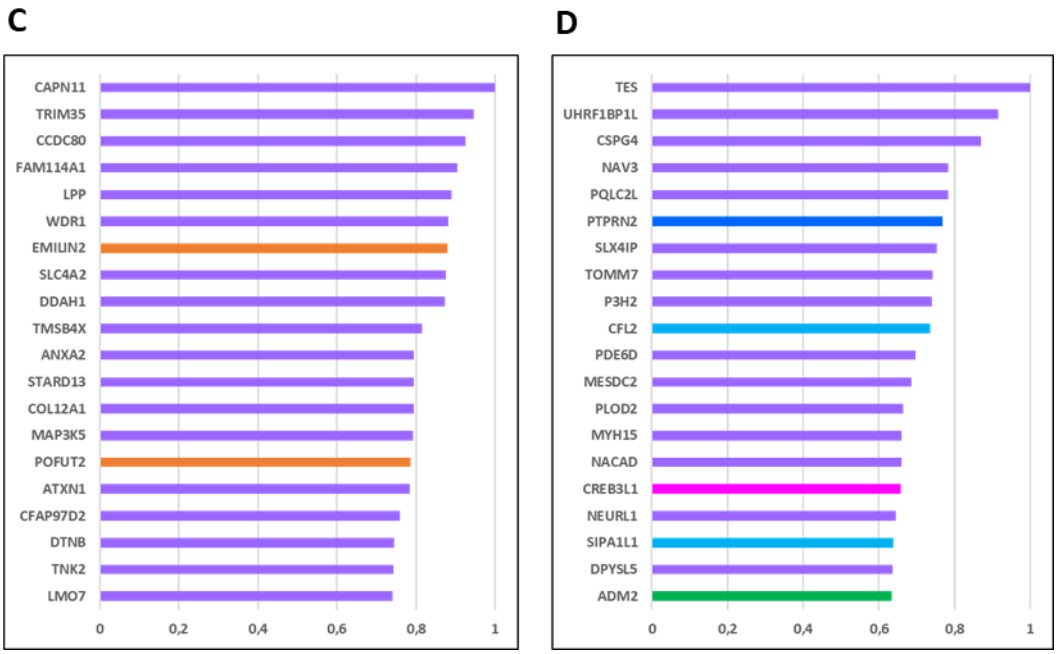


Figure R7: HEC modules and gene connectivity

A- Cytoscape visualization of the Black HEC network
B- Cytoscape visualization of the Red HEC network
C- The gene connectivity chart of the most connected genes in red module
D- The gene connectivity chart of the most connected genes in black module

Analysis of the GO categories revealed shared and distinctive patterns between the two HSCs networks. Both networks showed an enrichment in molecules implicated in the ECM, with Pleckstrin homology domain, SH3 domain, components of the Golgi apparatus, cell adherent junction, membrane and cell signaling. On the contrary, genes implicated in Focal adhesion were enriched in the red module only, while the black module showed enrichment in genes implicated in microtubule, migration and phosphatase activity (Figure R8).

In order to confirm the predictive value of our culture system and in order to validate candidate genes that potentially play a crucial role in the EHT, we selected two genes i.e., POFUT2 and TESTIN corresponding to two of the hubs of the HEC networks. To fully validate the system, we also selected the Wnt and Notch pathways already demonstrated as playing a role in the EHT. In order to do so, we followed a lipotransfection strategy with siRNA for POFUT2 and TESTIN and small molecules compounds for Notch and Wnt.

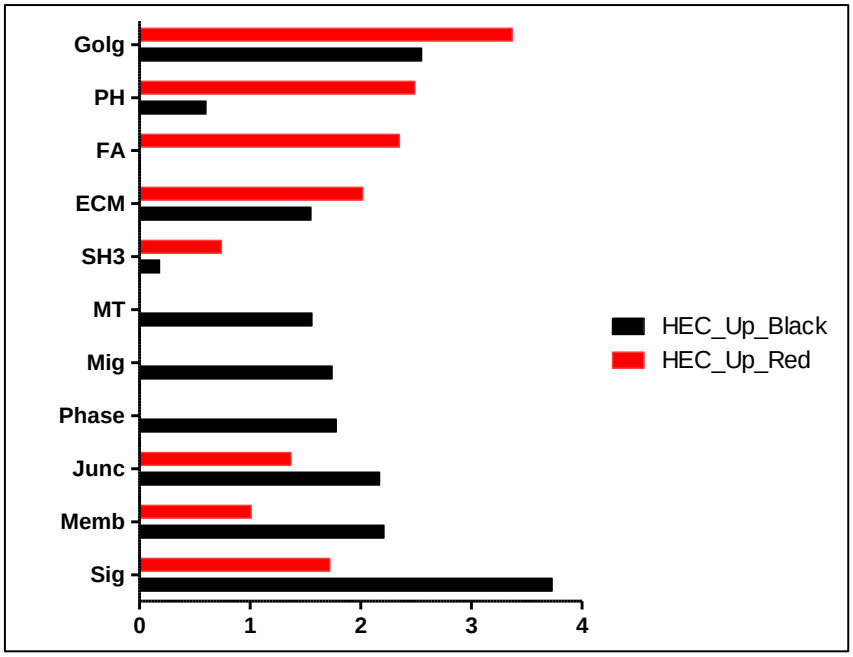


Figure R8: Analysis of the gene ontology categories of HEC networks

Both networks showed an enrichment in molecules implicated in the extracellular matrix ECM, with Pleckstrin homology domain PH, SH3 domain, components of the Golgi apparatus Golg, cell adherent junction Junc, membrane Memb and cell signaling sig. On the contrary, genes implicated in Focal adhesion FA were enriched in the red module only, while the black module showed enrichment in genes implicated in microtubule MT, migration Mig and phosphatase activity Phase.

IV. Modulation of the Wnt pathway contributes to the endo-hematopoietic balance

To investigate whether ß-catenin activity influences the generation of HEC and the EHT, we tested the effect of inhibitors or activators of the pathway on culture. As an inhibitor, we used the small molecule PKF118-310 (hereafter referred as PKF118), an inhibitor of the TCF4 ß-CATENIN protein interaction (Hallet et al., 2012). We tested 3 different concentrations of PKF118 (0.66µM, 1µM and 1.5µM). The last concentration provoked toxic effects whereas the first one did not produce significant results (not shown). The concentration of 1µM caused drastic changes to the culture. Treated cells were flatter and had fewer round cells than the non-treated samples (Figure R9A; Figure R9B; Figure R9D; Figure R9E). Flow cytometry analysis confirmed the increase of LDL^{hi} cells at the expense of LDL^{int} cells (Figure R9C; Figure R9F; Figure R9G; Figure R9H) and the strong decrease of the floating cells (Figure R9I) (n=5). QRT-PCR on adherent cells revealed the reinforcement of CD144 mRNA expression (Figure R9J) in keeping with the increase of the endothelial phenotype and the decrease of RUNX1 expression (Figure R9K) consistent with the lower number of LDL^{int} cells.

We also tested the Wnt activator SB216763 hereafter referred as SB, that inhibits GSK3 alpha and GSK3 beta. We used two different concentrations i.e., 10µM and 20µM with 20µM displaying the strongest effect. The result is an increase in the number of round hematopoietic cells produced compared with the mock- (DMSO) treated culture (Compare Figure R10A to Figure R10D and Figure R10B to Figure R10E). Flow cytometry analysis reveals a 1.6-fold change in the number of LDL^{int} cells at the expense of LDL^{hi} cells (Figure R10C; Figure R10F and Figure R10G; Figure R10H). Taken together the Wnt pathway plays a role in EHT. These results are completely in keeping with results published by Ruiz-Herguido et al., 2012 on the mouse AGM.

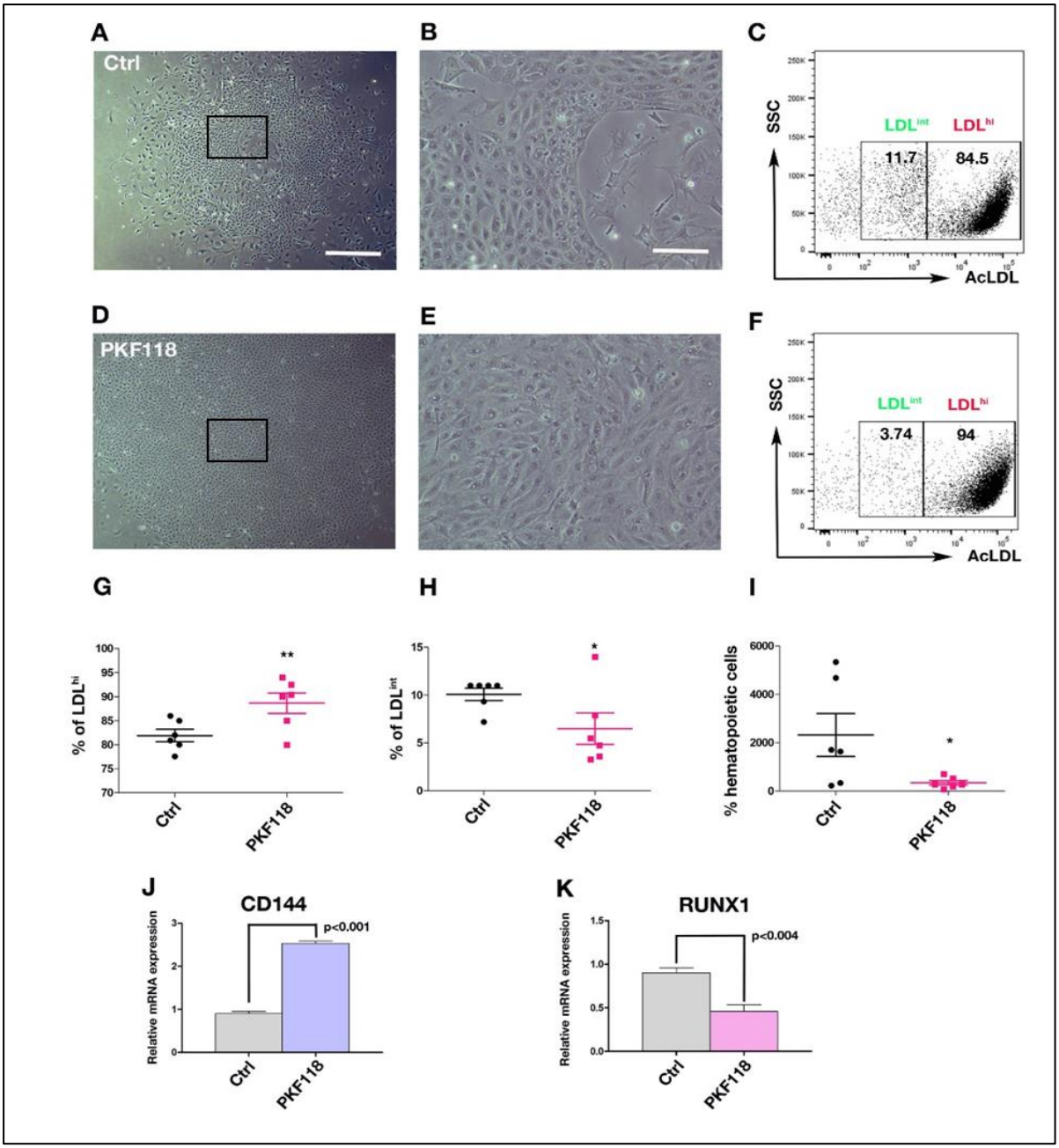


Figure R9: The Wnt antagonist, PKF118, inhibits EHT by reinforcing the endothelial phenotype

A: PSM culture mock treated at D1 and observed at D4. B: enlargement of the frame in A showing the flat endothelial layer and the few round cells resulting from the initiation of EHT at this stage. C: Representative FACS analysis showing the percentages of LDL^{hi} et LDL^{int} in the mock- (water) treated culture. D: PSM culture treated with PKF118 1µM at D1 and observed at D4. Note the flattening of the cell layer treated with PKF118 compared with the mock-treated culture. E: Enlargement of the frame in D showing the flattening of the cells and the presence of fewer round cells. F: representative FACS analysis showing the percentages of LDL^{hi} et LDL^{int} in the PKF118-treated culture (G). G: Percentage of the LDL^{hi} cells in mock- vs PKF118-treated cells (n=5, p<0.02). H: Percentage of the LDL^{int} cells in mock- vs PKF118-treated cells (n=5, p<0.05). I: Percentage of floating cells in mock- vs PKF118-treated cells (n=5, p<0.05). J: qRT-PCR for CD144 on mock- and PKF118-treated cells. A significant increase of CD144 mRNA expression is detected in the PKF118-treated cultures. K: qRT-PCR for RUNX1 on mock- and PKF118-treated cells. A significant decrease of RUNX1 mRNA expression is detected in the PKF118-treated cultures. Means are ± SEM. A, D, bar= 400µm; B, E, bar= 20µm.

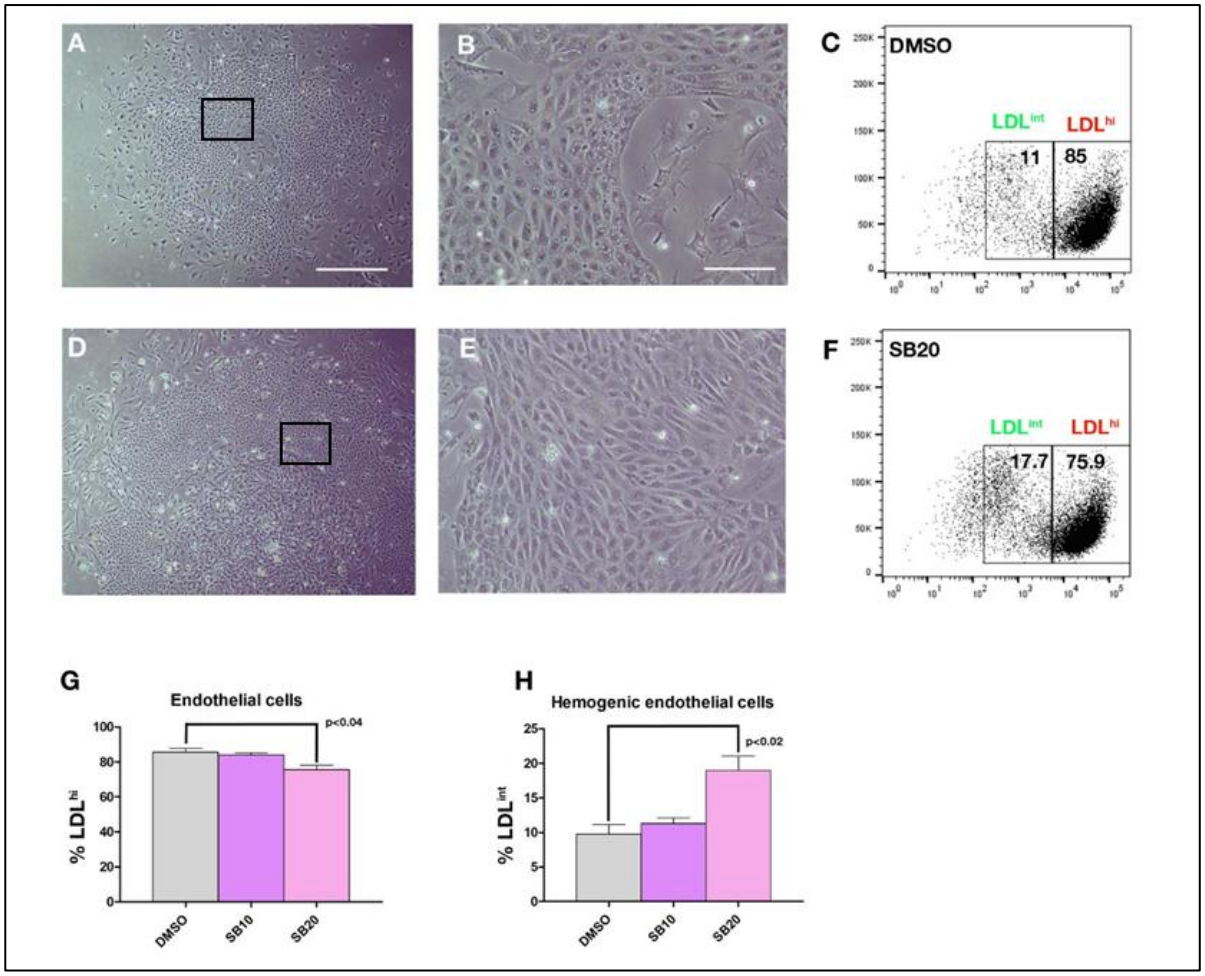


Figure R10: The Wnt agonist SB enhances EHT

A: PSM culture mock treated at D1 and observed at D4. B: enlargement of the frame in A showing the flat endothelial layer and the few round cells resulting from the initiation of EHT at this stage. C: Representative FACS analysis showing the percentages of LDLhi and LDLint in the mock- (DMSO) treated culture. D: PSM culture treated with SB216763 20µM at D1 and observed at D4. Note the presence of more round hematopoietic cells in the sample treated with SB216763 compared with the mock-treated culture. E: Enlargement of the frame in D showing the flattening of the cells and the presence of fewer round cells. F: representative FACS analysis showing the percentages of LDLhi et LDLint in the SB216763-treated culture. G: Percentage of the LDLhi cells in mock- vs SB216763-treated cells (n=5, p<0.04). H: Percentage of the LDLint cells in mock- vs SB216763-treated cells (n=5, p<0.02). Means are ± SEM. A, D bar= 400µm; B, E, bar= 20µm.

V. The Gamma secretase inhibitor DAPT stimulates the EHT

We used gamma secretase as an inhibitor of Notch to inquire into the role of this signaling pathway in hemogenic endothelium and EHT. DAPT 25µM was applied to the culture at D2. Analysis was performed at D4. DAPT readily produced a dramatic change in culture. Within the two days of DAPT contact, many cells rounded up, detached from the adherent layer and were found in the supernatant (Compare Figure R11A; Figure R11B). Cells in the supernatant were round and very bright suggestive of a hematopoietic production (Figure R11B) and Yvernogeau et al., 2016). FACS analysis showed a sharp increase in the number of LDL^{int} cells at the expense of LDL^{hi} cells (Figure R11C; Figure R11D and Figure R11G; Figure R11H) accompanied by a parallel increase of hematopoietic cells (Figure R11E; Figure R11 F and Figure R11I). QRT-PCR analysis showed a highly significant increase of CD45 expression characteristic of hematopoietic differentiation (Figure R11J).

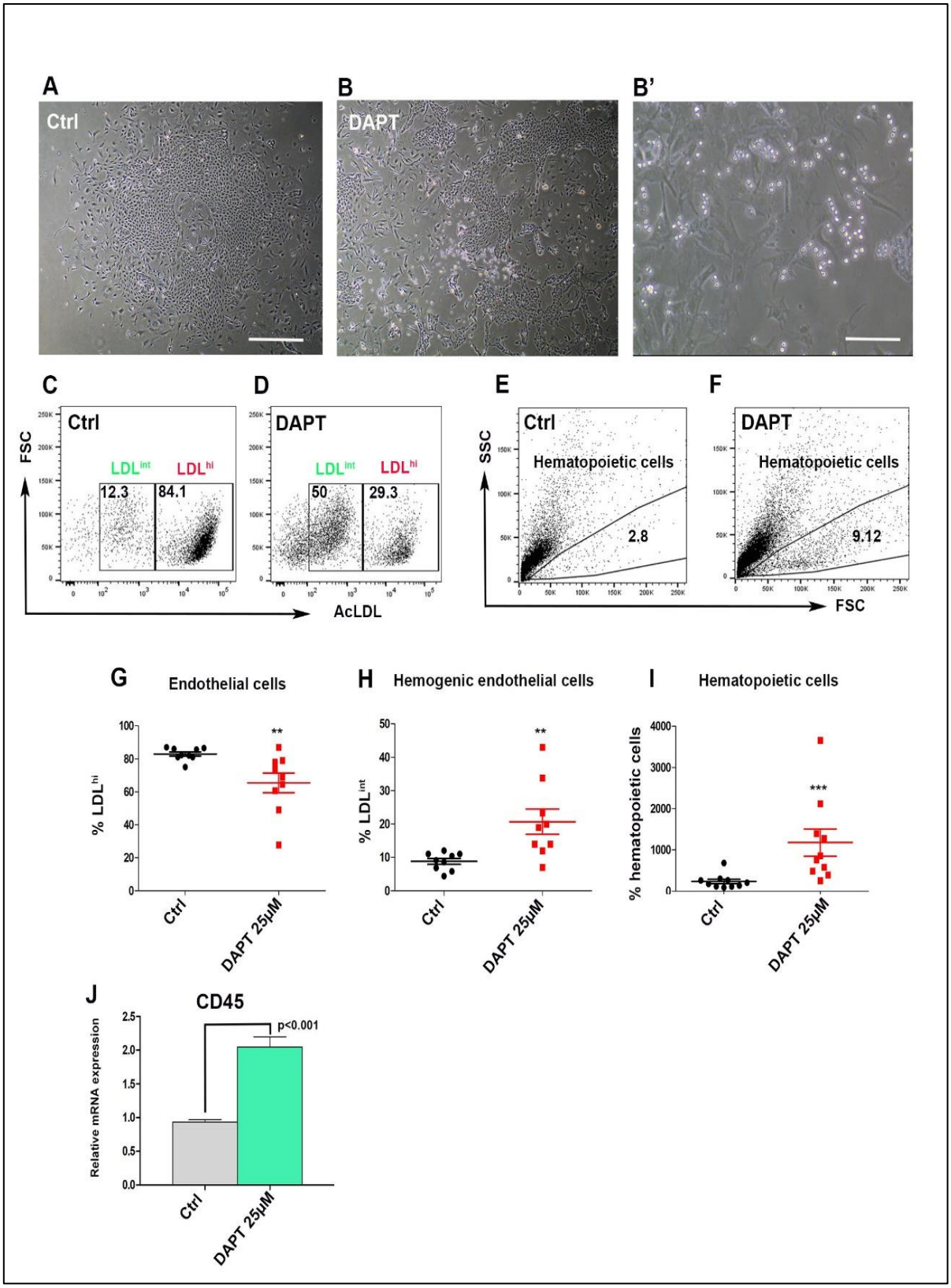

A
Ctrl
B
DAPT
B'
C
Ctrl
LDL^int
LDL^hi
12.3
84.1
FSC
D
DAPT
LDL^int
LDL^hi
50
29.3
AcLDL
E
Ctrl
Hematopoietic cells
2.8
SSC
F
DAPT
Hematopoietic cells
9.12
FSC
G
Endothelial cells
% LDL^hi
**
Ctrl
DAPT 25μM
H
Hemogenic endothelial cells
% LDL^int
**
Ctrl
DAPT 25μM
I
Hematopoietic cells
% hematopoietic cells
***
Ctrl
DAPT 25μM
J
CD45
Relative mRNA expression
p<0.001
Ctrl
DAPT 25μM

Figure R11: The gamma secretase inhibitor DAPT enhances EHT

A: Aspect of a control culture at D4. B: aspect of the DAPT-treated culture (25µM) at D4. Note the presence of numerous round cells. B': Enlargement of a significant area of a DAPT-treated culture. Numerous rounds, bright, floating cells are visible. C: Representative FACS analysis showing the percentages of LDLhi and LDLint in the mock- (DMSO) treated culture.

D: Representative FACS analysis showing the percentages of LDLhi and LDLint in the DAPT-treated culture. E: Representative FACS analysis showing the percentages of floating hematopoietic cells in the mock- (DMSO) treated culture. F: Representative FACS analysis showing the percentages of floating hematopoietic cells in the DAPT- treated culture. A 3.2-fold increase is obtained compared to the DMSO-treated culture. G: LDLhi cells in DMSO- vs DAPT-treated cultures. H: LDLint cells in DMSO- vs DAPT-treated cultures. I: Hematopoietic cells in DMSO- vs DAPT-treated cultures ($p<0.001$). J: q-RT-PCR of CD45 expression in DMSO- vs DAPT-treated cultures showing a more than two-fold increase in CD45 expression. A, B bar= 400µm; B' bar= 20µm.

VI. POFUT2 loss of function enhances the EHT

Based on the list of genes positively correlated to the HEC state obtained following WGCNA analysis and the measure of gene connectivity, we selected POFUT2 for potentially playing a strong role in HEC. POFUT2 is upregulated in HEC compared to EC or HC. POFUT2 (Protein O-Fucosyltransferase 2) catalyzes the reaction that attaches fucose through an O-glycosidic linkage to a conserved serine or threonine residue in the consensus sequence C1-X(2,3)-S/T-C2-X(2)-G of thrombospondin type 1 repeats where C1 and C2 are the first and second cysteines, respectively. It fucosylates members of the ADAMTS family, the Trombosporin and Spondin families. This fucosylation is required to regulate EMT during formation of the mesoderm and definitive endoderm. We used a knock-down strategy utilizing siRNA. We designed three different siRNAs against POFUT2 (Table M1), that were transfected at the culture incipience at Day 1 or 2. The analysis was performed at day 4 (Figure R9A) and was compared to Lipofectamine alone or a scrambled POFUT2 siRNA.

We tested 3 concentrations of siRNA against POFUT2 i.e. 30, 60 and 120nM (n=5) and compared their knock-down efficiency by measuring POFUT2 mRNA expression by q-PCR on transfected cells. The lowest concentration decreased POFUT2 mRNA expression by 80% whereas the highest displayed signs of toxicity and was not retained (not shown). The 60nM siRNA concentration decreased POFUT2 by 90% on transfected cells and was selected for further use (Figure R12B, R14B). POFUT2 mRNA knock-down at D1 resulted in the appearance of more round hematopoietic cells in the culture compared with Lipofectamine or scramble si controls (Figure R12C; Figure R12E; Figure R12G). Flow cytometry analysis revealed an increase in the LDL^{int} population at the expense of the LDL^{hi} population (Figure R12 D, F, H). Taken together POFUT2 knock-down resulted in a decrease of the number of ECs ($p<0.001$) and an increase in the number of HEC ($p<0.05$) (Figure R12I; Figure R12J).

To further validate the gene network and its predictive value, we selected the first neighbors of POFUT2 (Figure R13) and examined if a decrease of POFUT2 mRNA expression was also associated with a modulation of several of its first neighbors. Based on gene connectivity, we selected 3 genes among the most

connected to POFUT2. COL5A2 encodes an alpha chain for one of the low abundances fibrillar collagens; DNAJB4, is a molecular chaperone, tumor suppressor, and member of the heat shock protein-40 family that binds and regulates E-Cadherin and EMILIN1A (Elastin Microfibril Interfacer 1) is thought to regulate smooth muscle cell adhesion and vessel assembly. COL5A2, the most connected gene to POFUT2 showed a significant decrease upon POFUT2 siRNA mediated knock-down when compared to the Lipofectamine control but not to the scramble ($p<0.001$and 0.3 respectively) (Figure R12K). DNAJB4 and EMILIN1A both showed a significant decrease upon POFUT2 siRNA knock-down whatever the control used (Figure R12L; Figure R12M). Taken together, this indicates that POFUT2 and, at least, 3 of their first neighbors are acting coordinately to regulate the HEC fate and the hematopoietic production.

We also tested POFUT2 mRNA knock-down at D2. The 60nM siRNA concentration decreased POFUT2 by almost 90% on transfected cells (Figure R14B). This transfection did not show any significant difference in the culture compared with Lipofectamine or scramble si controls (Figure R14 C-E-G). Flow cytometry analysis revealed no changes in LDLint population nor the LDLhi population (Figure R14 D-F-H). Taken together POFUT2 knock-down at D2 resulted in no changes in the number of ECs and HEC (Figure R14 I-J). When looking at the RPKM value of the expression of POFUT2 in our transcriptome, we noticed that POFUT2 mRNA starts to be expressed at D1, meaning that POFUT2 protein is already present, even at low concentrations, at D2.

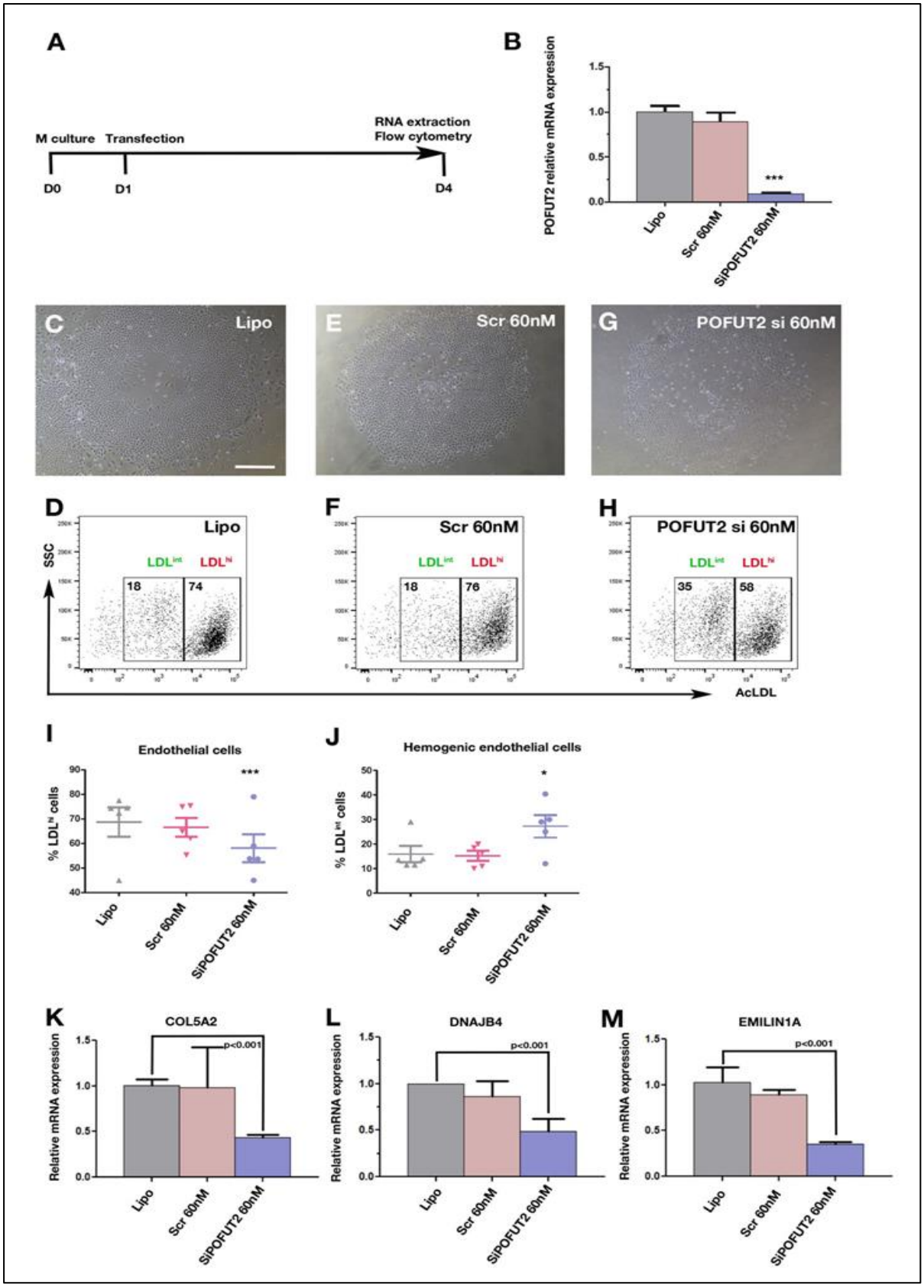

A
M culture
Transfection
RNA extraction
Flow cytometry
D0
D1
D4
B
POFUT2 relative mRNA expression
1.5
1.0
0.5
0.0
***
Lipo
Scr 60nM
SiPOFUT2 60nM
C
Lipo
E
Scr 60nM
G
POFUT2 si 60nM
D
Lipo
SSC
LDLint
LDLN
18
74
F
Scr 60nM
LDLint
LDLN
18
76
H
POFUT2 si 60nM
LDLint
LDLN
35
58
AcLDL
I
Endothelial cells
% LDLN cells
90
80
70
60
50
40
***
Lipo
Scr 60nM
SiPOFUT2 60nM
J
Hemogenic endothelial cells
% LDLint cells
50
40
30
20
10
0
*
Lipo
Scr 60nM
SiPOFUT2 60nM
K
COL5A2
Relative mRNA expression
1.5
1.0
0.5
0.0
p<0.001
Lipo
Scr 60nM
SiPOFUT2 60nM
L
DNAJB4
Relative mRNA expression
1.5
1.0
0.5
0.0
p<0.001
Lipo
Scr 60nM
SiPOFUT2 60nM
M
EMILIN1A
Relative mRNA expression
1.5
1.0
0.5
0.0
p<0.001
Lipo
Scr 60nM
SiPOFUT2 60nM

Figure R12: POFUT2 knock down enhances EHT.

A: experimental design. PSM cells were put in culture at D0, transfected at D1 and analyzed at D4 for LDL uptake by flow cytometry and q-PCR for the candidate gene and some of its immediate neighbors in the gene network. B: Effect of Lipofectamine (Lipo), scramble (Scr) and POFUT2 siRNA transduction on the expression of endogenous POFUT2. At 60nM, POFUT2 mRNA expression was decreased by more than 90% compared with the liposome and scramble siRNA controls.
C, E, G: culture at D4 that received Lipofectamine (C), scramble si (E) and POFUT2 si (F) at D1. Note that the number of round cells is more important in the siPOFUT2 condition than with the two other ones. D, F, H: representative flow cytometry experiment showing the balance between LDL^{int} and LDL^{hi} cells when cells received Lipofectamine (D), scramble si (F) and POFUT2 si (H). Note that the number of LDLint cells has increased at the expense of LDL^{hi} cells in the POFUT2 si condition. I: percentage of LDL^{hi} cells in the three conditions. A significant (p<0,001) decrease of LDL^{hi} cells is visible in the POFUT2 si conditions (n=5).
J: percentage of LDL^{int} cells in the three conditions. A low but significant (p<0.05) increase of LDL^{int} cells is observed in the POFUT2 si condition compared with the two others. K: mRNA expression of COL5A2 is significantly decreased when POFUT2 is knock down by siRNA. L: DNAJB4 mRNA expression is decreased in POFUT2 knock down condition. M: EMILIN1A mRNA expression is significantly decreased in POFUT2 knock down condition. C, E, G, bar= 400µm.

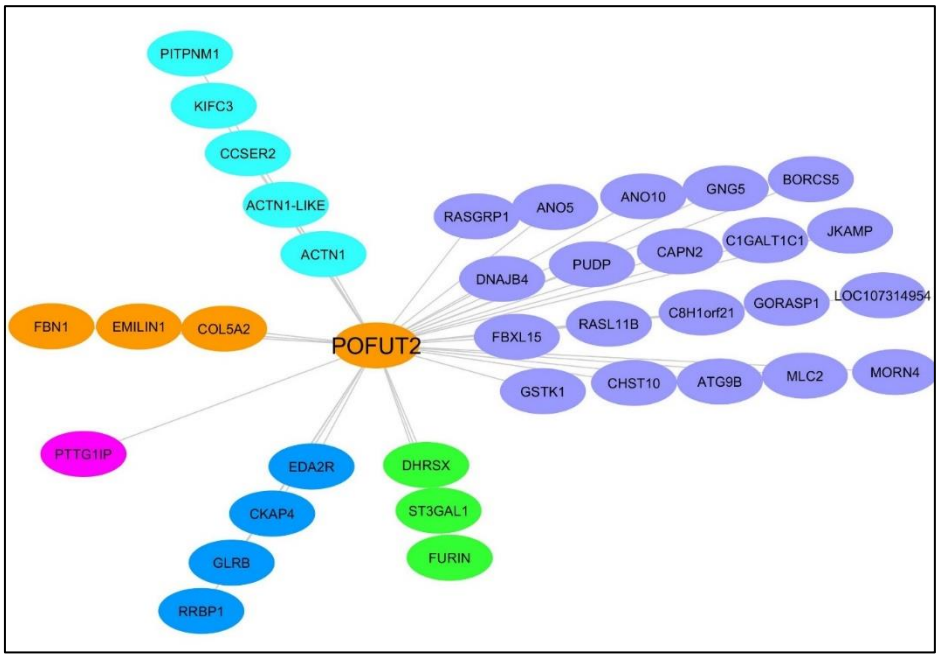


Figure R13: POFUT2 first neighbors

The color code is the same as Figure R7

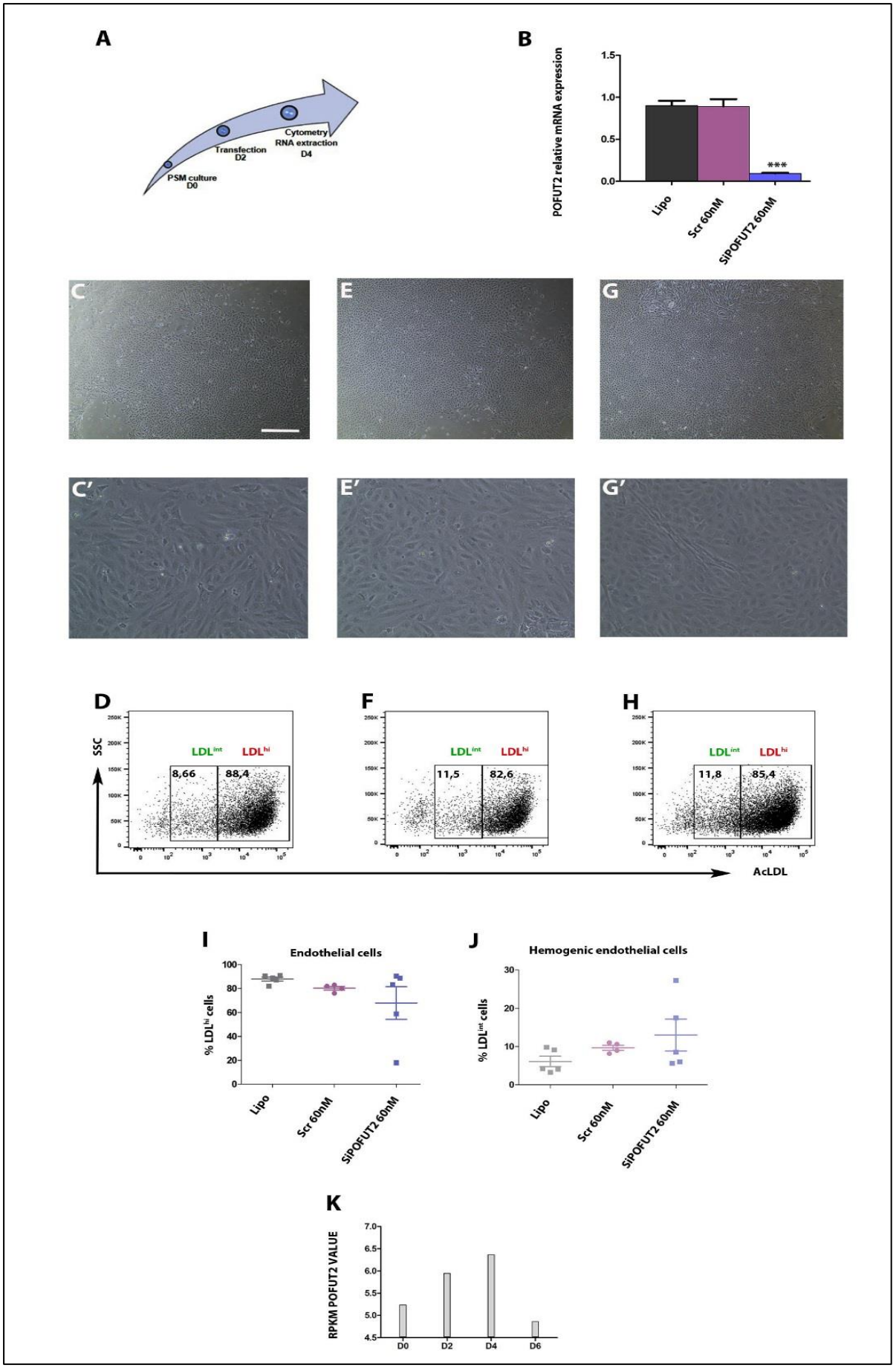

A
PSM culture D0
Transfection D2
Cytometry RNA extraction D4
B
POFUT2 relative mRNA expression
1.5
1.0
0.5
0.0
***
Lipo
Scr 60nM
SiPOFUT2 60nM
C
E
G
C'
E'
G'
D
F
H
SSC
LDLint
LDLhi
8,66
88,4
11,5
82,6
11,8
85,4
AcLDL
I
Endothelial cells
% LDLhi cells
J
Hemogenic endothelial cells
% LDLint cells
K
RPKM POFUT2 VALUE
D0
D2
D4
D6

Figure R14: POFUT2 knock down at day 2 doesn't affect EHT.

A: experimental design. PSM cells were put in culture at D0, transfected at D2 and analyzed at D4 for LDL uptake by flow cytometry. B: Effect of Lipofectamine (Lipo), scramble (Scr) and POFUT2 siRNA transduction on the expression of endogenous POFUT2. At 60nM, POFUT2 mRNA expression was decreased by more than 90% compared with the liposome and scramble siRNA controls.

C, E, G: culture at D4 that received Lipofectamine (C,C'), scramble si (E,E') and POFUT2 si (G,G') at D2. Note that no changes in the number of round cells in the siPOFUT2 condition compared to the two other ones. D, F, H: representative flow cytometry experiment showing the balance between LDL^{int} and LDLhi cells when cells received Lipofectamine (D), scramble si (F) and POFUT2 si (H). I,J: percentage of LDLhi and LDLint cells in the three conditions. Note that no significant changes in the % of LDL^{hi} nor the LDL^{int} fraction with POFUT2 compared to Scr and Lipo. K: RPKM value of the expression of POFUT2 in our transcriptome.

VII. TESTIN a hub of the HEC network

Based on the list of genes positively correlated to the HEC state obtained following WGCNA analysis and the measure of gene connectivity, we selected TESTIN for potentially playing a role in HEC. TESTIN is upregulated in HEC compared to EC or HC. It is a protein expressed in almost all normal human tissues. It locates in the cytoplasm along stress fibers being recruited to focal adhesions. it forms complexes with various cytoskeleton proteins such as Actin, Talin and Paxilin playing a significant role in cell motility and adhesion. Decreased Testin expression associate with loss of epithelial morphology and gain of migratory and invasive properties of mesenchymal cells. Testin plays a significant role in cell adhesion, cell spreading and the reorganization of actin cytoskeleton.

We designed three different siRNAs against TESTIN (Figure R15A) that were transfected at day 1 of culture incipience. The analysis was performed at day 4 (Figure R15B) and was compared to Lipofectamine alone or a scrambled TESTIN siRNA. We tested 3 concentrations of siRNA against TESTIN i.e. 30, 60 and 120nM (n=3) and compared their knockdown efficiency by measuring TESTIN mRNA expression by q-PCR on transfected cells. The lowest concentration decreased TESTIN mRNA expression by 70% whereas the highest displayed signs of toxicity and was not retained (not shown). The 60nM concentration decreased TESTIN by 80% on transfected cells and was selected for further use (Figure R15C). TESTIN mRNA knockdown at D1 did not show any significant difference in the culture compared with Lipofectamine or scramble si controls (Figure R15D-E'). Flow cytometry analysis revealed no changes in LDL^{int} population nor the LDL^{hi} population (Figure R15F-H). Taken together TESTIN knock-down resulted in no changes in the number of ECs and HEC (Figure R15I-J).

To further validate if the inhibition of TESTIN even affects the gene network, we selected the TESTIN's first neighbors and examined if a decrease of TESTIN mRNA expression was also associated with a modulation of several of its first neighbors. Based on gene connectivity (Figure R16) we selected 3 genes among the most connected to TESTIN. BICC1 is a putative RNA-binding protein that acts as a negative regulator of Wnt signaling, CREB3L1 a protein with a DNA-binding transcription factor activity and chromatin binding required for the assembly of collagen extracellular

matrix and CSPG4 which is a Proteoglycan playing a role in cell proliferation and migration which stimulates endothelial cells motility during microvascular morphogenesis. BICC1, CREB3L1 and CSPG4 showed no significant changes upon TESTIN siRNA mediated knockdown when compared to the Lipofectamine and scramble controls (Figure R15L-N). Taken together, this indicates that the loss of TESTIN may be compensated by other cytoskeleton and ECM proteins and TESTIN does not play a crucial role in the regulation of HEC fate and the hematopoietic production.

A

Couple	Sequence	
1	Sens	GGAUAAGUGUGAAGGAUUUTT
	Antisens	AAAUCCUUCACACUUAUCCTT
2	Sens	CCAGAAGGCAUCUCAGUAUTT
	Antisens	AUACUGAGAUGCCUUCUGGTT
3	Sens	GCAUGCUGCGAUUUGCCAATT
	Antisens	UUGGCAAAUCGCAGCAUGCTT

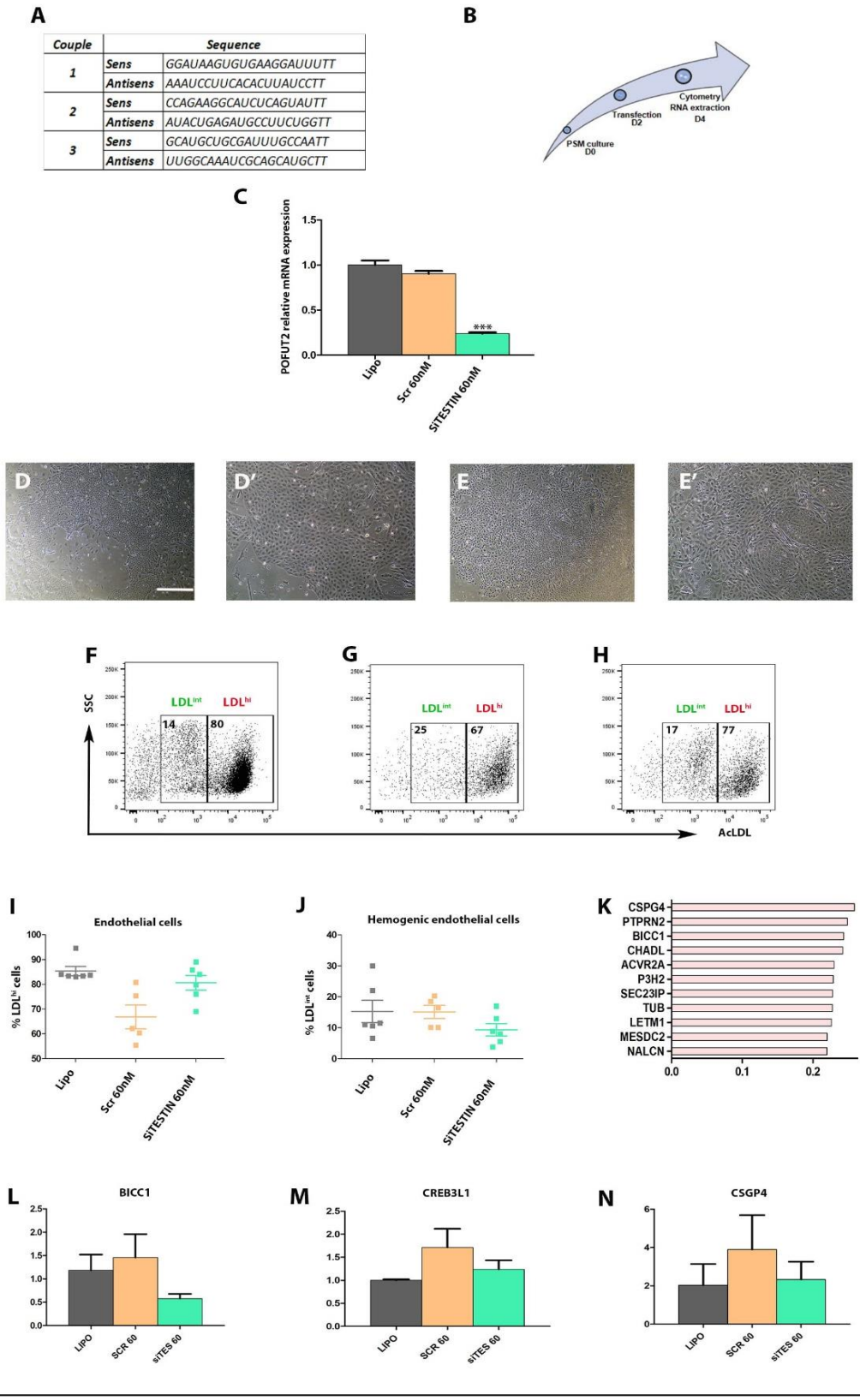


Figure R15: TESTIN knock down doesn't affect EHT.

A: Sequences of the siRNA and scramble designed for Testin knock down.
B: experimental design. PSM cells were put in culture at D0, transfected at D1 and analyzed at D4 for LDL uptake by flow cytometry. C: Effect of Lipofectamine (Lipo), scramble (Scr) and POFUT2 siRNA transduction on the expression of endogenous TESTIN. At 60nM, TESTIN mRNA expression was decreased by more than 90% compared with the liposome and scramble siRNA controls.
D, E: culture at D4 that received Lipofectamine (D,D'), TESTIN si (E,E') at D1. Note that no changes in the number of round cells in the siTESTIN condition compared to the two other ones. F, G,H: representative flow cytometry experiment showing the balance between LDL^{int} and LDL^{hi} cells when cells received Lipofectamine (F), scramble si (G) and TESTIN si (H). I, J: percentage of LDL^{hi} and LDL^{int} cells in the three conditions. Note that no significant changes in the % of LDL^{hi} nor the LDL^{int} fraction with TESTIN si compared to Scr and Lipo. K: The 11 first highly connected first neighbors of TESTIN. L: mRNA expression of BICC1 is decreased when TESTIN is knock down by siRNA. M, N: CREB3L1 mRNA expression and CSPG4 showed no differences in TESTIN knock down condition.

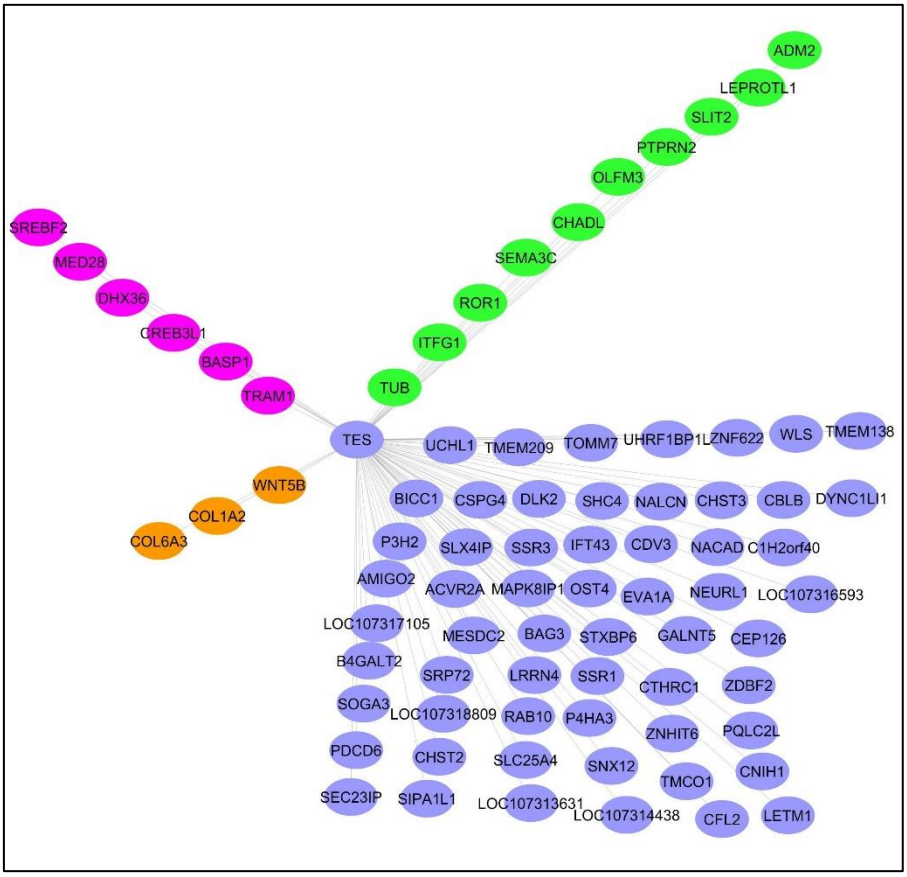


Figure R16: TESTIN first neighbors

The color code is the same as Figure R7

VIII. Single-cell profiling identifies several well-defined clusters

In order to reveal specific signatures of cells undergoing EHT and to decipher the changes that occur between endothelial and hemogenic endothelial fates, we performed single-cell RNA-seq (scRNAseq) using 8326 cells from 4 independent samples (1 D0, 2 D4, 1 D6). Single-cell t-Distributed Stochastic Neighbor Embedding (tSNE) analysis revealed 12 well defined clusters (Figure R17).

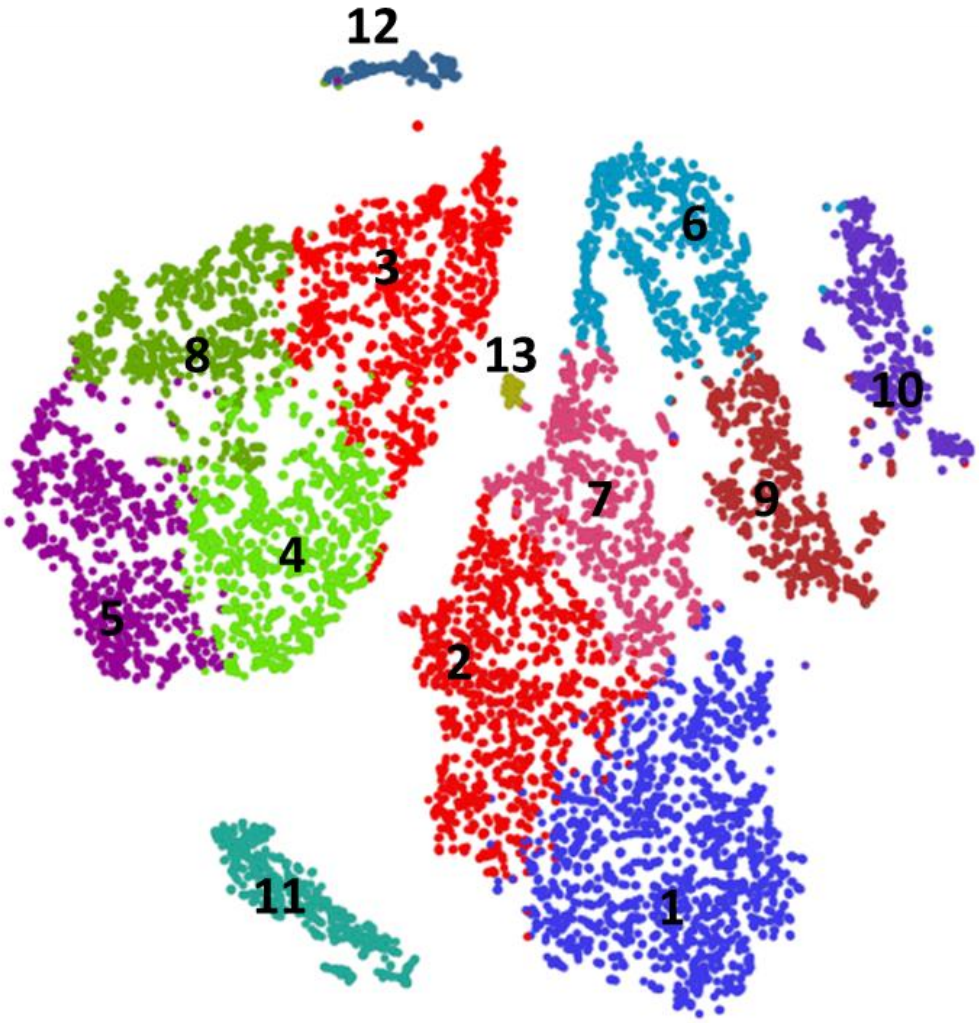


Figure R17: t-SNE analysis revealed 12 well defined clusters

The next step was to attribute each of these clusters to a well-defined population according to the expression of previously reported marker genes. Clusters 3, 4, 5, 8 and 12 showed enrichment in mesodermal markers: FGF8, MEOX1, TBXT, HOXA1, ADD3 (Figure R18). Clusters 1, 2, 7, 13 showed enrichment in endothelial markers: CDH5, ESAM, KDR, CD34 (Figure R19). Whereas, cluster 6 showed enrichment in hemogenic endothelial markers: RUNX1, CD44 (Figure R20); The cluster 9 and 10 did not show a specific enrichment with defined markers apart from some cells expressing PTPRC and RUNX1. No evidence of endothelial or mesodermal expression were detected in these two populations suggesting that these clusters might be of hematopietic nature. In summary, the single cell approach recapitulates the presence of 4 defined populations: mesodermal, endothelial cells, hemogenic endothelial cells and hematopoietic cells.

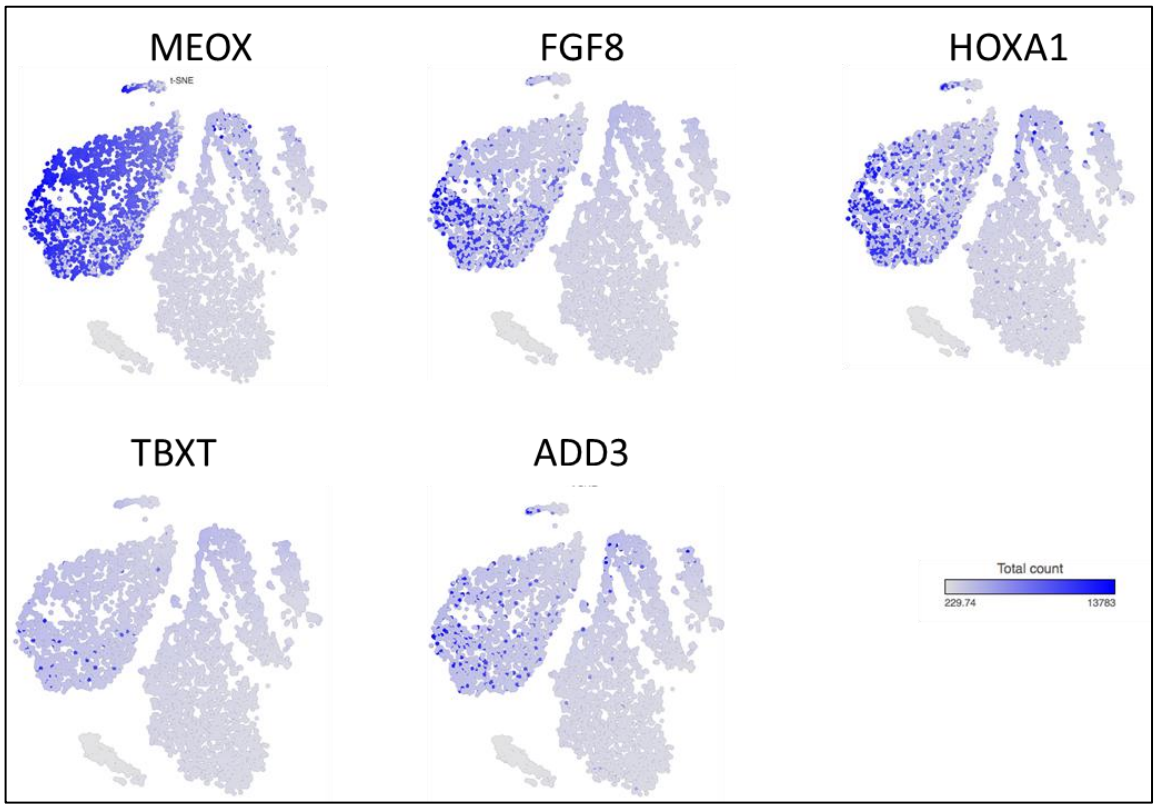


Figure R18: Mesodermal markers

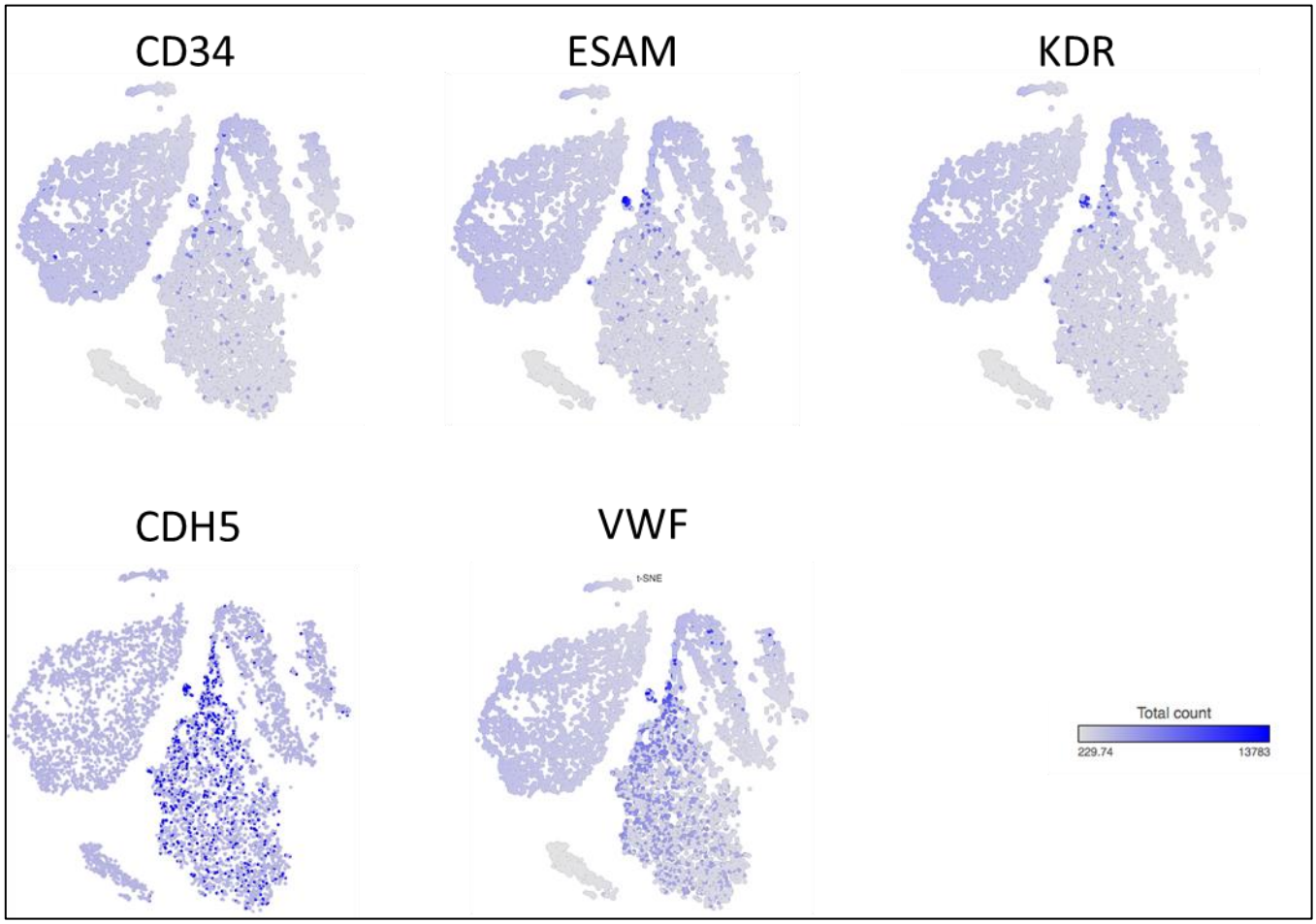


Figure R19: Endothelial markers

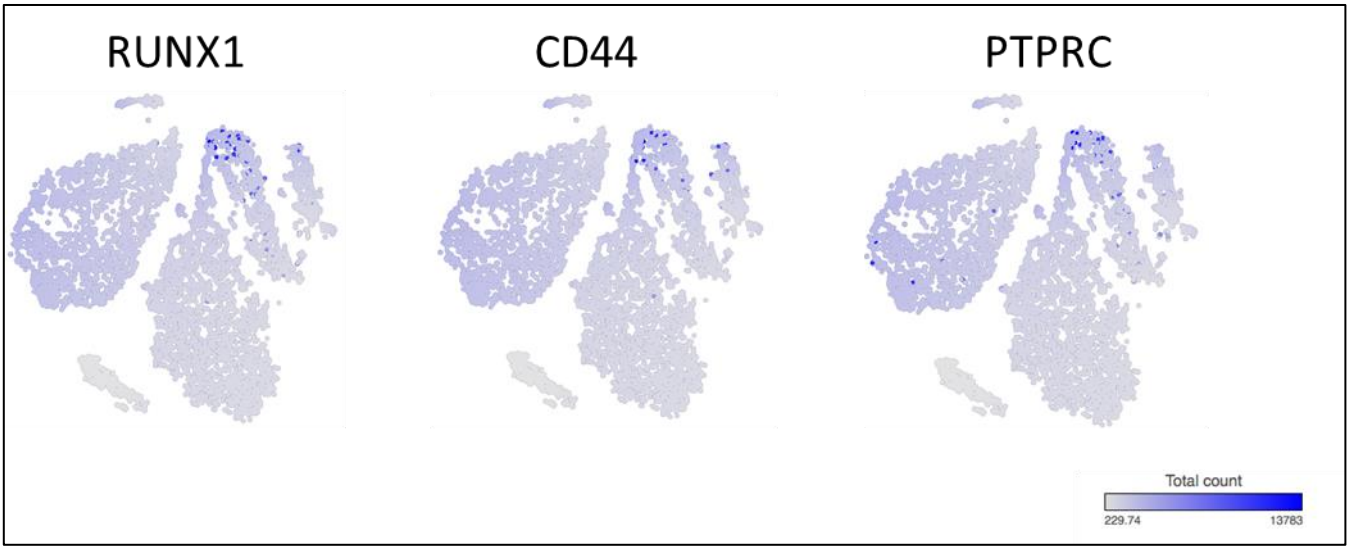


Figure R20: Hemogenic endothelial and hematopoietic markers

In order to validate the attribution of these clusters to a defined cell category, we performed Gene Ontology analysis and identified the major KEGG pathways. The Mesodermal population (Cluster 3, 4, 5, 8 and 12) were enriched in genes implicated in Spliceosome, Cell cycle, DNA replication, Ribosome, RNA transport, DNA replication… The Endothelial population (cluster 1, 2, 7, 10 and 13) showed an enrichment in Lysosome, Metabolic pathways, adherens junctions, PI3K-Akt signaling pathways, endocytosis, MAPK… The cluster 11 showed a specific enrichment in oxidative phosphorylation, Parkinson's and Alzheimer's diseases corresponding to a mitochondrial cluster. Whereas the cluster 6 corresponding to the hemogenic endothelium showed an enrichment in Focal adhesion, actin cytoskeleton, ECM receptor interaction, Ras signaling pathway, adherens junction… Finally, the cluster 9 showed an enrichment in carbon metabolism, apoptosis and hemopoiesis, which could correspond to hematopoietic cluster (Figure R21).

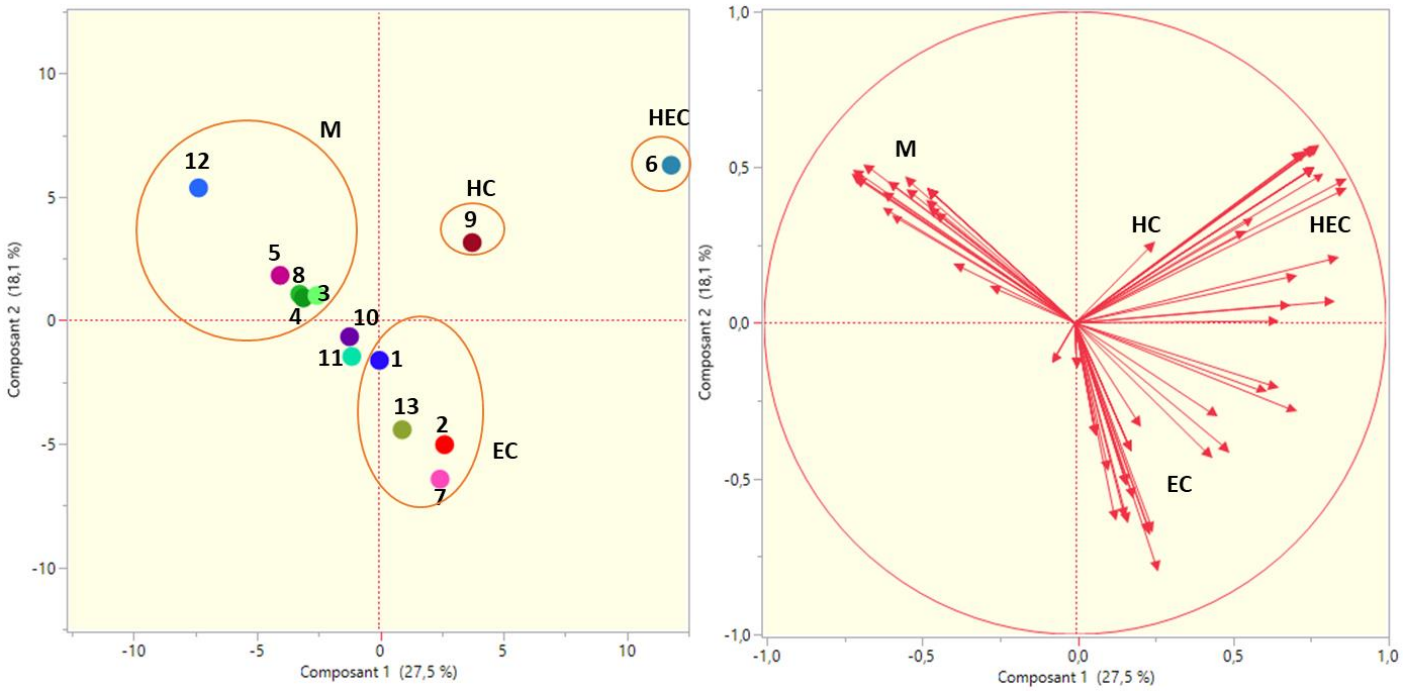


Figure R21: Analysis of the KEGG pathways categories by PCA

- Left: the score plot (PC1 v/s PC2) on 82 KEGG pathways categories and 13 clusters allowed the discrimination between 4 major populations HEC, HC, M and EC
- Right: Loading plot (variables as red arrows) defined the specific categories of KEGG pathways for each population.

This analysis allowed us to confirm the presence of the 4 specific populations in our culture system going from Mesoderm to Endothelium to Hemogenic endothelium and finally the Hematopoietic population. This part of the work is still preliminary, and this data set will be further investigated in more refined way.

IX. Complementary results

1. LOC assignment workflow

One of the issues with our gene list (after alignment to the quail genome) is that it included around 8000 non-assigned sequences, named LOC. Those LOC correspond to sequences not yet linked with any orthologs. It was vital for our future investigation that these sequences were matched with known genes, so we had to identify most of them manually. We used the databases of NCBI Genes, NBCI Nucleotides and Uniprot, as well as the alignment search program BLAST (blastn algorithm) to assign our ARN sequences to known genes (Figure CR1). We could almost identify 4000 of these LOCs (Annex table 1).

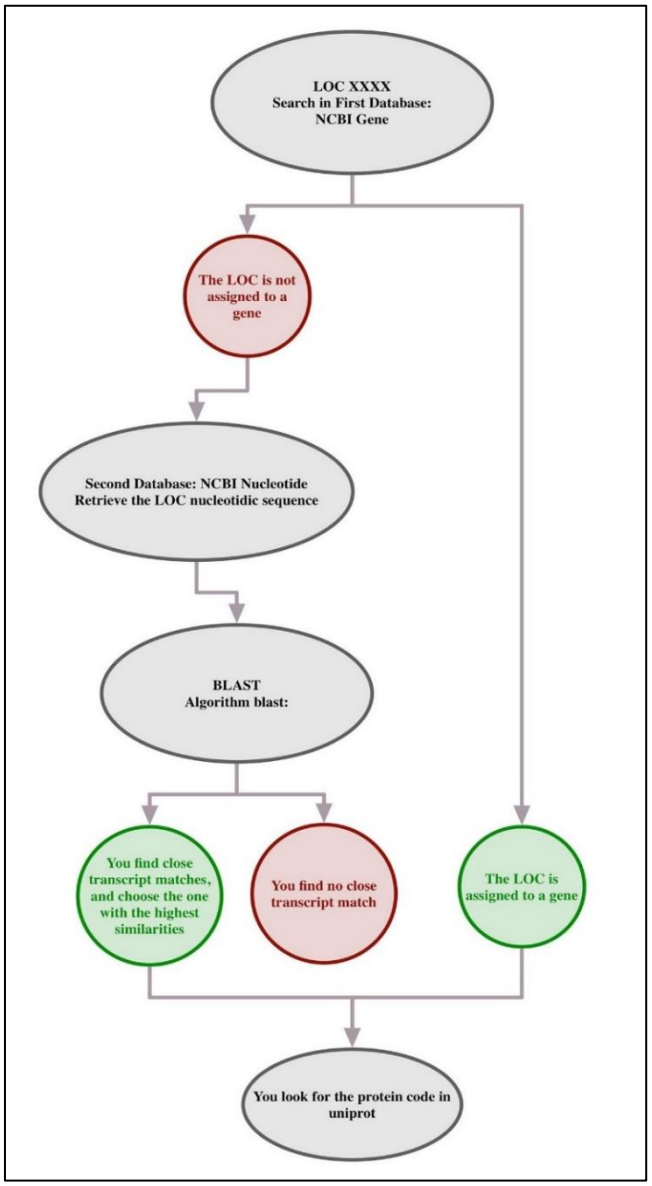


Figure CR1: LOC Assignment workflow

Table S1: upregulated DEGs in HC and EC

HC UP			EC UP			
ABCB10	LOC107309112	PPM1H	A4GNT	HEG1	METTL21A	TSPAN12
ABCB7	LOC107309121	PPOX	AASS	HEXA	MFGE8	TSPAN13
ACOT13	LOC107309332	PRAM1	ABCA13	HGSNAT	MFSD1	TSPAN6
ACSS2	LOC107309555	PRKCB	ABCD4	HIPK2	MFSD10	TSPAN9
ADA	LOC107309690	PSTPIP1	ABLIM1	HIPK3	MFSD4B	TTC38
ADAP2	LOC107309869	PTAFR	ABLIM2	HIVEP1	MFSD6	TTC9
ADIPOR2	LOC107310095	PTPN6	ACAP2	HMGCLL1	MGAT5	TTLL2
ADRBK2	LOC107310166	PTPRC	ACBD5	HRASLS	MICAL3	TUSC3
AFTPH	LOC107310573	RAB11FIP5	ACER3	HS1BP3	MID2	TXNDC16
AGL	LOC107311215	RAB27A	ADAM17	HS3ST1	MKL2	TXNDC5
AGTPBP1	LOC107311272	RAB32	ADAMTS7	HTATIP2	MMRN1	UACA
AK2	LOC107311500	RAB44	ADCK2	HYI	MMRN2	UBE2H
ALDOB	LOC107311523	RASSF2	ADCY2	ICA1	MOB1B	UEVLD
ALOX5	LOC107311560	RASSF5	ADGRA1	IDS	MPPE1	UNC13C
ANK1	LOC107311590	RBPJ	ADGRF4	IFNGR1	MPZL1	UPP1
ANKRD12	LOC107311827	RGR	ADGRF5	IGDCC4	MTMR10	UPP2
ANKRD13A	LOC107312292	RHCE	ADGRG1	IGFBP3	MTRF1	USP12
AP1AR	LOC107312344	RHOBTB3	ADGRL4	IGFBP7	MYCT1	USP38
AQP2	LOC107312365	RHOF	ADSSL1	IMPA1	MYLK	USP6NL
ARFGEF2	LOC107312428	RIMS4	AGA	ING4	MYO6	UTRN
ARL8B	LOC107312433	RNF145	AGT	INPP4B	MYO7A	VANGL1
ARSB	LOC107312463	RNF212	AHR	INTS10	MYRIP	VAT1
ASB6	LOC107312591	RRP36	AHRR	ISCA2	NAAA	VPS33B
ASB8	LOC107312649	S1PR4	AK4	ISPD	NAGA	VWC2
ASTL	LOC107312725	SAMSN1	AKT3	ISYNA1	NCKAP1	VWF
ASXL2	LOC107313061	SASH3	AKTIP	ITM2A	NDFIP2	WDFY2
ATP6	LOC107313680	SBNO1	ALS2CL	ITM2B	NECTIN3	WSCD1
ATP6V0B	LOC107313817	SCIN	ANGEL1	ITSN2	NEMP2	WWTR1
ATP6V1A	LOC107313856	SCNN1D	ANKRD50	IVNS1ABP	NFIA	XYLB
ATP6V1B2	LOC107313879	SESN1	ANPEP	JAG2	NHLRC1	ZC3H7B
ATP6V1C1	LOC107314004	SGCZ	ANXA7	KCNB2	NOS1AP	ZCCHC24
ATP6V1C2	LOC107314262	SLC15A5	AP5M1	KCNE3	NOVA1	ZDHHC14
ATP8	LOC107314505	SLC1A1	APOLD1	KCNMB2	NPL	ZDHHC23
B4GALNT1	LOC107314560	SLC25A37	APP	KCTD18	NR0B1	ZFAND2B
BCL11A	LOC107314689	SLC35D1	AQP1	KDR	NR3C2	ZFPM2
BHLHA15	LOC107314794	SLC40A1	AQP9	KIAA1210	NR5A2	ZMYND11
BPI	LOC107314877	SLC43A3	ARFGEF3	KIAA1462	NRBP2	ZNF521
BRI3	LOC107315482	SLC4A1	ARHGAP31	KIF13A	NRXN3	
BRIP1	LOC107315614	SLC9A3R1	ARHGEF10	KIF13B	NTNG1	
BST1	LOC107316070	SMIM5	ARHGEF26	KLF3	NTRK3	
BTK	LOC107316219	SPATA2L	ARHGEF9	KLHL24	NUDT14	
C17H9orf91	LOC107316229	SPG21	ARID4B	KLHL3	OSBPL10	
C18H17orf62	LOC107316493	SPNS3	ARL11	KLHL4	OSBPL8	
C18H17orf89	LOC107316612	STARD5	ARL15	KLHL5	OSBPL9	

C21H1orf158	LOC107316955	STK10	ARPP19	L3HYPDH	PAOX	
C21H1orf167	LOC107317373	STK4	ARSA	LAMA4	PAQR5	
C26H6orf106	LOC107317685	STX11	AS3MT	LAMB1	PAQR9	
C6H10orf54	LOC107318005	SULT1B1	ASAP1	LAMP2	PBXIP1	
C8H1orf228	LOC107318300	SUOX	ASB5	LDB2	PCNX4	
CADM3	LOC107318468	SUSD3	ATG13	LEMD3	PCSK2	
CARD9	LOC107318673	SYNGR3	ATG4A	LGMN	PDE8A	
CC2D1B	LOC107318831	TALDO1	ATOH8	LGR5	PDGFB	
CCDC171	LOC107318879	TAS1R1	ATP8B2	LIN7C	PEX26	
CCND3	LOC107319031	TDP2	ATP9B	LNX2	PEX5	
CCNDBP1	LOC107319050	TECTA	AXDND1	LOC107306001	PHACTR2	
CD59	LOC107319126	TFRC	B3GLCT	LOC107306678	PHKA1	
CEBPA	LOC107319205	TGM4	BACH2	LOC107306761	PIK3C2G	
CEL	LOC107319240	TMBIM4	BAMBI	LOC107306762	PIK3CA	
CHST12	LOC107319258	TMED3	BBX	LOC107306776	PKP4	
CLEC3A	LOC107319261	TMEM156	BMF	LOC107306817	PLA2G12A	
CLK4	LOC107319266	TMEM173	BMP6	LOC107306852	PLA2G7	
CLPB	LOC107320068	TMEM40	BMPR2	LOC107307576	PLCB1	
CNIH3	LOC107320098	TMEM52	BTD	LOC107308050	PLCD1	
COX1	LOC107320103	TMEM56	BVES	LOC107308084	PLD1	
COX2	LOC107320131	TMEM63A	C1HXorf36	LOC107308152	PLEKHB2	
COX3	LOC107320199	TMEM71	C2CD2	LOC107308164	PPFIBP2	
CPN2	LOC107320215	TMOD4	C2H18orf8	LOC107308181	PPM1K	
CPOX	LOC107320470	TNFAIP8	C4H4orf33	LOC107308460	PPP3CA	
CREG1	LOC107320472	TNFRSF18	C4H8orf48	LOC107308462	PQLC3	
CRLF2	LOC107320548	TRADD	C5H11orf74	LOC107308654	PRDM11	
CSF1R	LOC107320703	TRPV2	C8H1orf210	LOC107308722	PRDM4	
CTAGE5	LOC107320986	TTC39C	C8H1orf226	LOC107308787	PREX2	
CTSC	LOC107321029	TTC7A	CABLES1	LOC107308858	PRKD1	
CTSD	LOC107321102	TTPAL	CADPS2	LOC107308953	PROCR	
CYTB	LOC107321358	TUBB1	CALCRL	LOC107309030	PROX1	
CZH5orf51	LOC107321408	TUBD1	CAPN5	LOC107309087	PRR5L	
DECR1	LOC107321519	TUSC1	CAT	LOC107309092	PRRG3	
DNAAF1	LOC107321546	TYR	CCDC85A	LOC107309179	PRSS23	
DNAH12	LOC107321550	TYRP1	CCNI	LOC107309559	PTCHD1	
DNAH14	LOC107321620	UIMC1	CCNY	LOC107309563	PTP4A1	
DNAH17	LOC107321658	UROD	CCNYL1	LOC107309600	PTPN14	
DNAH3	LOC107321681	UROS	CD151	LOC107309698	PTPRB	
DOCK11	LOC107321684	USP45	CD2AP	LOC107309880	PTPRM	
DOCK5	LOC107322046	VPS11	CD81	LOC107309932	PXDC1	
DOCK8	LOC107322132	VRK3	CD9	LOC107310100	PXDN	
DOPEY2	LOC107322293	WAS	CD93	LOC107310202	RAB12	
DSCR3	LOC107322314	WDR41	CDC42EP1	LOC107310334	RAI2	
EFHD2	LOC107322386	WDTC1	CDK17	LOC107310341	RANBP9	
EMB	LOC107322421	WIPF1	CDKL2	LOC107310343	RAP2B	
ENO3	LOC107322440	ZBTB8OS	CEP85L	LOC107310620	RAPGEF5	

ENTPD5	LOC107322461	ZNF831	CGNL1	LOC107310625	RARRES2	
EPN3	LOC107322623	ZNFX1	CHMP4C	LOC107310736	RASAL2	
ERAL1	LOC107322695	rna47256	CHRM3	LOC107310737	RASGEF1A	
EVI2A	LOC107322790	rna47258	CHST7	LOC107310746	RASGRP3	
EVI2B	LOC107322824		CIB1	LOC107310815	RASSF8	
F13A1	LOC107322904		CLCN3	LOC107310833	RECK	
FABP3	LOC107322910		CLN5	LOC107310894	RGL1	
FAHD1	LOC107322928		CNDP2	LOC107310896	RGS2	
FAM117A	LOC107323028		CNTNAP2	LOC107310913	RHOB	
FAM129A	LOC107323038		COBLL1	LOC107311894	RHOJ	
FAM134C	LOC107323081		COLGALT2	LOC107312011	RIPK2	
FAM13B	LOC107323147		CORO2B	LOC107312057	RMDN2	
FAM210B	LOC107323592		CPE	LOC107312129	RNF11	
FBXO18	LOC107323637		CPEB2	LOC107312133	RNF144A	
FBXO9	LOC107323914		CPN1	LOC107312528	RNF150	
FCHO1	LOC107323924		CRIM1	LOC107312575	RNF19A	
FECH	LOC107324314		CRY1	LOC107312640	RSU1	
FETUB	LOC107324323		CSGALNACT1	LOC107312798	RUFY3	
FGFR1OP2	LOC107324364		CTDSP1	LOC107312799	RUNX1T1	
FLVCR2	LOC107324389		CTDSPL	LOC107312816	S1PR1	
FMNL1	LOC107324445		CTNNB1	LOC107312822	SAMD12	
FNIP1	LOC107324452		CUL4A	LOC107312840	SASH1	
FRRS1	LOC107324453		CYYR1	LOC107312977	SCARB2	
FUCA2	LOC107324456		DCAF6	LOC107313022	SCP2	
FUT7	LOC107324465		DEPTOR	LOC107313134	SDCBP	
FUT8	LOC107324551		DESI1	LOC107313185	SDPR	
FYB	LOC107324681		DGKZ	LOC107313228	SECISBP2L	
GCNT4	LOC107324776		DHRS12	LOC107313257	SERF2	
GDA	LOC107324840		DHRS7	LOC107313495	SERHL2	
GDPGP1	LOC107324841		DIRC2	LOC107313546	SETD7	
GHITM	LOC107324873		DISC1	LOC107313549	SFRP4	
GMIP	LOC107324930		DLEC1	LOC107313570	SFT2D1	
GNA13	LOC107324939		DNM3	LOC107313670	SGCB	
GNAT2	LOC107325029		DNMBP	LOC107313674	SGK1	
GOLGA7B	LOC107325190		DOCK1	LOC107313678	SH3GLB1	
GPNMB	LOC107325339		DOCK9	LOC107314087	SH3RF3	
GPR157	LOC107325340		DPYD	LOC107314285	SH3TC1	
GPR174	LOC107325522		DYNC1I1	LOC107314378	SHANK3	
GPRC6A	LOC107325555		DYSF	LOC107314473	SHROOM4	
GPT2	LOC107325677		DZIP1	LOC107314525	SLC12A7	
GPX1	LOC107325871		EDN1	LOC107314526	SLC12A9	
GRB2	LRCH3		EDNRB	LOC107314670	SLC17A5	
GUCY2F	LRRC25		EEPD1	LOC107314707	SLC22A23	
HAX1	LY86		EGLN1	LOC107314714	SLC2A12	
HCK	LYRM9		EHD4	LOC107314972	SLC2A13	
HDHD3	M6PR		ELF2	LOC107315000	SLC30A1	

HFM1	MADD		ELK3	LOC107315001	SLC30A4	
HIPK1	MAP3K6		ELMO1	LOC107315340	SLC35C1	
HK3	MAP3K8		EMCN	LOC107315364	SLC7A6	
HMBS	MAPT		ENPP2	LOC107315680	SLCO2A1	
HMGCL	MAT1A		EPAS1	LOC107315726	SOAT1	
HPCAL1	MATN4		ERBIN	LOC107315820	SOCS6	
HSPB9	MBNL1		ERG	LOC107315823	SORBS2	
HYAL3	MGST1		ETFBKMT	LOC107315983	SOS2	
HYLS1	MITF		EVC	LOC107316143	SOX7	
IFITM10	MKNK1		EVI5	LOC107316352	SPG20	
IFNAR2	MOB1A		EXOC3L1	LOC107316486	SPPL2A	
IGF2R	MPHOSPH9		F10	LOC107316541	SPTBN1	
IGLL1	MPP1		F8	LOC107316772	SRI	
IL22	MS4A15		FAM120B	LOC107317020	SRPK2	
IL2RG	MSL1		FAM124B	LOC107317216	SSFA2	
IL5RA	MTMR3		FAM13A	LOC107317314	ST3GAL2	
ING5	MXD1		FAM160A1	LOC107317332	ST3GAL6	
IRAK4	MXD3		FAM160B1	LOC107317338	STAB2	
IRF5	MYD88		FAM171A1	LOC107317376	STARD3NL	
IRF8	MYO1F		FAM185A	LOC107317519	STARD9	
IRG1	MYO7B		FAM214A	LOC107317595	STK32C	
KATNA1	NADK		FAM21C	LOC107317596	STON1	
KCNG2	NARF		FAM234B	LOC107317652	STS	
KCTD12	NCEH1		FAM43A	LOC107318030	STXBP3	
KIF1BP	NCF2		FAM89A	LOC107318190	STYK1	
KIT	ND1		FAT4	LOC107318199	SYNJ2BP	
KMT5A	ND2		FBXL2	LOC107318369	SYT16	
LARP1B	ND4		FBXO8	LOC107318370	SYTL4	
LDB1	ND4L		FCHSD2	LOC107318671	SYTL5	
LECT2	ND5		FGD6	LOC107318756	TANC1	
LOC107305709	NEU2		FHOD1	LOC107318851	TBC1D1	
LOC107305764	NFAM1		FJX1	LOC107319006	TBC1D22A	
LOC107305832	NFE2L2		FLRT2	LOC107319130	TBC1D30	
LOC107305842	NHLRC4		FNDC3A	LOC107320404	TBC1D8	
LOC107305889	NIT1		FOXO1	LOC107321499	TBCK	
LOC107305946	NKX3-1		FOXO4	LOC107322457	TDRD5	
LOC107306192	NKX6-3		FPGT	LOC107322477	TGFBR2	
LOC107306243	NLRC5		FRYL	LOC107324276	TGFBR3	
LOC107306265	NMBR		FYCO1	LPAR4	THBD	
LOC107306618	NOS2		FYN	LPAR6	THPO	
LOC107306830	NRROS		FZD4	LRCH1	TIE1	
LOC107306906	NT5DC4		FZD6	LRP1B	TJP1	
LOC107306909	NUDT7		GAB1	LRP6	TM2D1	
LOC107307041	NUP210		GALK2	LRRC57	TM6SF1	
LOC107307042	OAZ1		GALNT4	LRRC58	TMCC3	
LOC107307139	OMG		GDF7	LRRK2	TMCO3	

LOC107307169	OXGR1		GLB1	LSAMP	TMEM106B	
LOC107307292	PAPD4		GLB1L	LTBP2	TMEM108	
LOC107307293	PAX5		GLTSCR1L	LYRM2	TMEM117	
LOC107307355	PDE3B		GNAL	LYVE1	TMEM135	
LOC107307435	PGD		GNB4	MAB21L3	TMEM140	
LOC107307508	PGP		GNB5	MAN2A2	TMEM181	
LOC107307602	PI16		GPHN	MANSC1	TMEM241	
LOC107307701	PI4KA		GPR158	MAP3K9	TMEM26	
LOC107307710	PIGV		GPR50	MAP4K4	TMEM30A	
LOC107307799	PIK3AP1		GPR75	MAST4	TMOD3	
LOC107307838	PIK3CD		GPR89B	MBTPS1	TMPPE	
LOC107307845	PIK3CG		GPRIN3	MC5R	TNFRSF11A	
LOC107308175	PLA2G15		GYPC	MCEE	TNFRSF1A	
LOC107308400	PLCB2		HACE1	MCTP2	TRAF3IP2	
LOC107308694	PLCL1		HAPLN3	MEF2A	TRIM45	
LOC107308713	PLEKHA2		HDLBP	MEI1	TRIM67	
LOC107309064	PNPLA2		HEBP1	MERTK	TSC22D1	

Table S2: upregulated DEGs in M and HEC

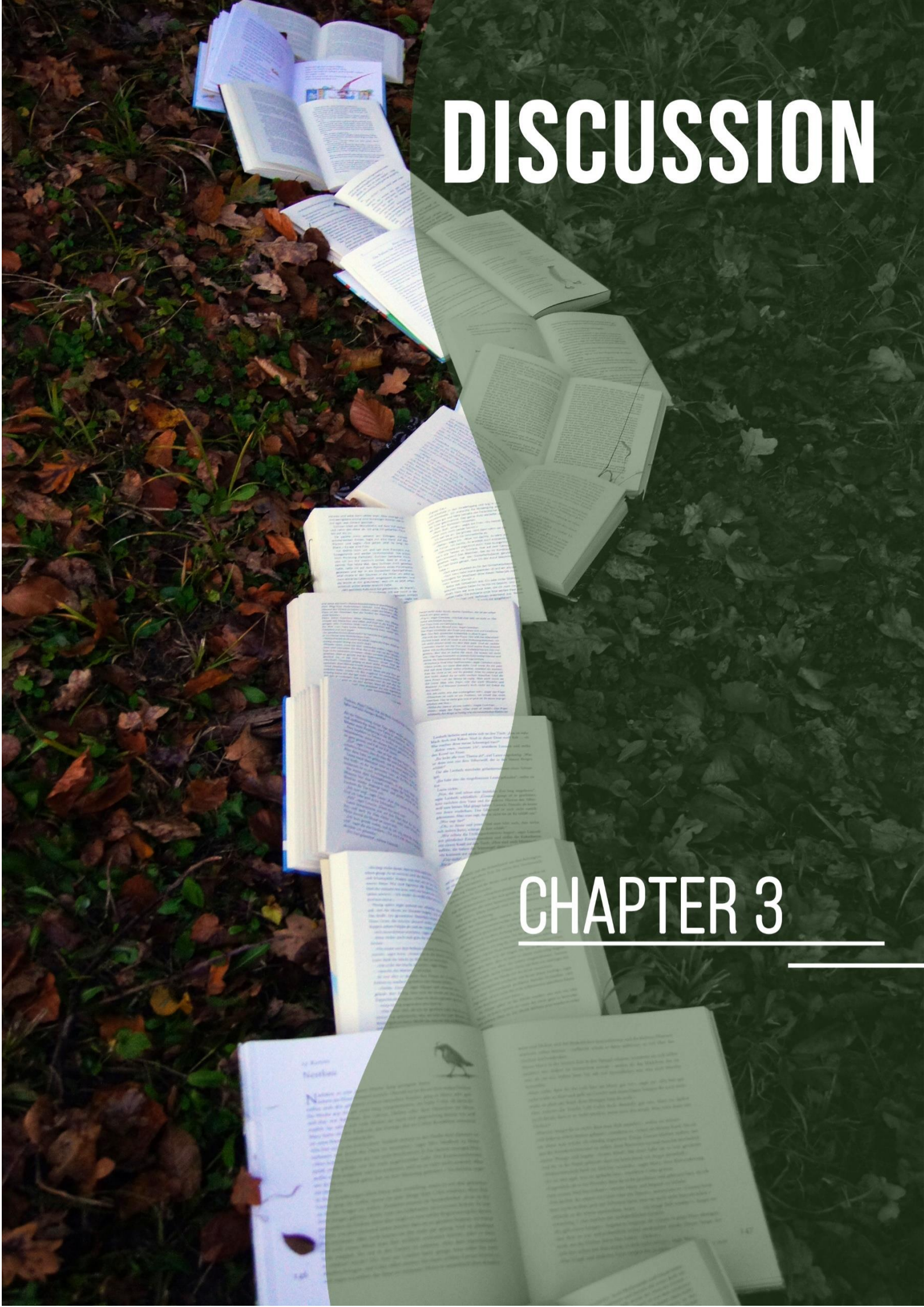
M UP				HEC UP			
AAGAB	GTF3C3	NAT10	TFAM	ABCB5	GALNT16	NTF3	TUBB6
AASDH	GUK1	NBAS	TFDP2	ACOT11	GALNT5	OLFM3	TULP3
ABCA4	HACD2	NCOA1	TGFBR1	ACOT9	GALNT7	OPHN1	TWSG1
ABCE1	HACL1	NDC1	TGS1	ACTA1	GCK	OPN4	UCHL1
ABCG5	HAS2	NDP	THOC2	ACTA2	GIPC2	OST4	UCK2
ABI2	HDAC2	NEBL	THUMPD2	ACTC1	GJC2	OTUD7B	UHRF1BP1L
ACBD7	HDAC4	NEDD1	TIAL1	ACTR1A	GLRB	OTULIN	UNC45A
ACTL6A	HDX	NEO1	TM7SF3	ACTR3	GLUD1	P3H2	UNC5C
ACTR3B	HEATR1	NFXL1	TMEFF1	ACVR2A	GNA11	P3H3	UPRT
ACVR2B	HEATR3	NGB	TMEM151B	ADAM33	GNAS	P4HA1	USP46
ADAMTS20	HMGB3	NKAIN3	TMEM163	ADAM8	GNG12	P4HA3	USP53
ADCK3	HMGN1	NKD1	TMEM168	ADAMTS14	GNG5	PAFAH1B2	USP8
ADGRA3	HNRNPDL	NKRF	TMEM169	ADGRG5	GORASP1	PALLD	USP9X
ADNP2	HNRNPH3	NKTR	TMEM192	ADM2	GPR162	PAMR1	VCL
ADSL	HOMER2	NOC3L	TMEM237	AFAP1	GPR37	PAPD5	VIM
AGPAT3	HOXA1	NONO	TMEM266	AHDC1	GPR68	PAPSS1	VOPP1
AKAP12	HOXA2	NRDE2	TMEM33	AKAP6	GRAMD4	PAPSS2	VTCN1
AKT1	HOXA3	NRG4	TMEM55A	ALG5	GRIA3	PARVA	WDR1
ALMS1	HOXA4	NSUN2	TMEM64	AMFR	GSTK1	PCDH11X	WLS
ALS2	HS6ST2	NSUN3	TMPRSS3	AMIGO2	GYLTL1B	PCDH18	WNT5B
AMER1	HSF2	NUDCD1	TMTC4	AMOTL1	HGF	PCDH19	WNT9A
ANK3	HSPA14	NUFIP1	TNFRSF19	ANGPT1	HHAT	PCGF2	YBX3
ANOS1	HSPD1	NUP153	TOMM70	ANKMY2	HIBADH	PCOLCE2	YWHAQ
AP3B2	HUNK	NUP93	TP53BP1	ANKRD22	HPDL	PDCD6	ZBED1
API5	ICE1	NUP98	TPH1	ANKRD29	HPS6	PDCL3	ZBTB1
AQR	ICE2	NVL	TPI1	ANKRD34C	HSPA13	PDE10A	ZC2HC1A
ARGLU1	IFT27	ODC1	TPP2	ANKRD52	HSPA2	PDE5A	ZC3H12C
ARHGAP28	IGF2BP3	OGFOD1	TPR	ANO1	HTR2B	PDE6D	ZDBF2
ARHGAP8	ILDR2	ORC2	TRA2A	ANO10	HTRA3	PDGFC	ZDHHC5
ARID1B	ING3	ORC3	TRA2B	ANO3	IFFO1	PGAM1	ZDHHC9
ARID3B	INTS7	OTUD4	TRABD	ANO5	IFNGR2	PGBD5	ZFHX4
ARL13B	INTS8	OXSR1	TRHR	ANTXR2	IFT43	PGM3	ZNF362
ARL3	INTS9	PAK3	TRMT1L	ANXA2	IGF2BP2	PGRMC1	ZNF425
ARL5A	IPO13	PANK1	TSHZ1	ANXA5	IGHMBP2	PHLDA1	ZNF622
ARL5B	IPO5	PANX1	TSKU	APBB2	IGSF3	PHLDB2	ZNF804B
ASB7	ITPR2	PAQR3	TSN	ARFIP1	IKBIP	PICK1	ZNHIT6
ASNSD1	JARID2	PARD3	TSSC1	ARFIP2	IL34	PITPNM1	ZSWIM5
ATG4C	JMJD4	PARD3B	TTC14	ARHGAP42	ILK	PKDCC	ZYX
ATRX	JMJD7	PARG	TTC21B	ARHGEF11	IMPACT	PKIA	
ATXN10	KCNAB1	PARP1	TTC26	ARHGEF17	INHBA	PLCE1	
B3GALNT1	KCNMB4	PAX2	TTLL12	ARHGEF4	INPP5A	PLEKHG3	
B3GNT5	KCNS2	PCCA	TUBGCP2	ARID5B	INPP5F	PLEKHH2	
BAZ2B	KCNS3	PCID2	TUBGCP3	ARL1	IQGAP1	PLOD2	
BBS5	KDM4A	PCM1	TULP4	ARNTL	IRX2	PLPP6	

BCHE	KDM5A	PCMTD1	TXLNG	ARPC2	ITFG1	PLXDC2	
BCLAF1	KIAA0586	PCNP	U2AF1	ARSJ	ITGA11	PODXL	
BEND3	KIAA1549	PDCD11	U2SURP	ASB2	ITGA8	POFUT2	
BEND4	KIAA1549L	PDGFRA	UBE2D1	ATG9B	ITGAV	PPM1L	
BIVM	KIAA2022	PELI1	UBE2E1	ATP10A	JKAMP	PPP1R3C	
BLOC1S4	KIF16B	PEX13	UBE2E3	ATXN1	KANK4	PQLC2L	
BMPR1B	KIF21A	PGK1	UBE2K	B3GALT2	KANSL1L	PRDM5	
BRD1	KIZ	PHC1	UBE2N	B3GNT9	KATNAL1	PRDX6	
BRD7	KLHL13	PHF14	UBE2QL1	B4GALT2	KDELC2	PRKAG2	
BRWD1	KLHL23	PHF5A	UBE3A	B4GALT4	KDELR3	PRRG4	
BTBD9	KRAS	PHF6	UBN2	BAG2	KIF26B	PRSS12	
BTF3L4	KRIT1	PHIP	UBR5	BAG3	KIFC3	PSMD13	
BTRC	KRR1	PHKB	UBXN7	BAK1	KLC4	PTCHD3	
C10H15orf61	LAS1L	PHOSPHO2	UCN3	BASP1	KLF10	PTK2	
C1H12orf4	LDHB	PIBF1	ULK4	BBS9	KLHL36	PTPN9	
C1H7orf60	LEF1	PIH1D3	URB1	BDNF	KTN1	PTPRN	
C2H8orf88	LEO1	PKNOX1	URB2	BICC1	LACTB	PTPRN2	
C8H1orf27	LGI2	PLEKHA5	USP16	BIN1	LBH	PTTG1IP	
CACHD1	LIN28B	PLEKHH1	UST	BMPER	LCMT2	PTX3	
CACNB4	LIN52	PLRG1	UTP20	BORCS5	LEPROTL1	PUDP	
CAPRIN1	LNX1	PLXNA4	VBP1	BTBD10	LETM1	RAB10	
CASP10	LOC107306757	PNISR	VEZT	C1GALT1C1	LIMA1	RAB15	
CASP3	LOC107306862	PNPT1	VTA1	C1H2orf40	LIMD1	RAB23	
CCDC173	LOC107306866	POLR2B	WASF1	C1QTNF4	LIMK2	RAP1GDS1	
CCDC59	LOC107308079	POMT2	WASF3	C8H1orf21	LIN7A	RAPH1	
CCDC61	LOC107308218	POU2F1	WDR11	CACNB2	LMF2	RARG	
CCDC88C	LOC107308222	PPHLN1	WDR19	CALD1	LMO7	RASGRP1	
CCNG2	LOC107308341	PPP1R3G	WDR48	CAMK2D	LOC107305692	RASL11B	
CCNJ	LOC107308547	PPP2R3B	WDR61	CAPN11	LOC107306754	RBPMS	
CCNT2	LOC107308730	PPP4R3B	WDR75	CAPN2	LOC107306800	RCN1	
CCT2	LOC107308740	PREP	WRB	CAPRIN2	LOC107306825	REST	
CCT5	LOC107308746	PRMT7	WRN	CASC4	LOC107306834	RFTN1	
CD24	LOC107308812	PROSER1	XPNPEP1	CBLB	LOC107306841	RGS4	
CDC40	LOC107308919	PRPF4B	XPO5	CCDC102A	LOC107306871	RGS9BP	
CDC5L	LOC107308983	PRTG	XRCC3	CCDC113	LOC107307951	RIC3	
CDK8	LOC107309187	PSMA4	XRCC5	CCDC80	LOC107308888	RINT1	
CDKN2AIP	LOC107309251	PSMD14	XRN2	CCM2	LOC107308912	RIOK3	
CECR2	LOC107309267	PSPC1	YAE1D1	CCSER2	LOC107308948	RNF149	
CELF1	LOC107309281	PTBP2	YAP1	CD99	LOC107309052	ROBO1	
CEP162	LOC107309305	PTCH2	YEATS2	CDC42BPB	LOC107309196	ROR1	
CEP295	LOC107309667	PTER	YME1L1	CDKL1	LOC107309558	RPS6KA5	
CEP97	LOC107309750	PTGDR	ZBTB18	CDV3	LOC107309564	RPS6KL1	
CFAP36	LOC107309809	PTGR2	ZBTB42	CELSR1	LOC107309567	RRAS2	
CFAP61	LOC107310641	PTMA	ZC3H14	CEP126	LOC107309618	RRBP1	
CGGBP1	LOC107310649	PTN	ZFAND4	CEP170	LOC107309619	RTN4	
CHD7	LOC107310657	PTPN4	ZHX2	CEP170B	LOC107309896	RYK	

CHGB	LOC107310798	PWP2	ZIC3	CEP68	LOC107309997	SAMD4A	
CHIC2	LOC107310951	PYGO1	ZMAT3	CERS2	LOC107310048	SCAP	
CHRFAM7A	LOC107311283	QRFPR	ZMYM2	CFL2	LOC107310176	SCCPDH	
CKAP5	LOC107311432	RAB28	ZNF248	CHADL	LOC107310367	SDC1	
CLASP1	LOC107311492	RAB30	ZNF277	CHPF	LOC107310519	SDC2	
CLSTN3	LOC107311494	RBFA	ZNF280D	CHPF2	LOC107310722	SEC23A	
CMTR1	LOC107311554	RBM33	ZNF292	CHRDL2	LOC107310842	SEC23IP	
CNIH4	LOC107311589	RBM45	ZNF318	CHRM4	LOC107310852	SEC24D	
CNKSR3	LOC107312091	RBMX	ZNF326	CHRNA2	LOC107310998	SEC31B	
COQ3	LOC107312144	RBP5	ZNF648	CHST10	LOC107311125	SEC61A2	
CPSF2	LOC107312240	RCAN2	ZNF654	CHST14	LOC107311256	SEMA3C	
CPSF6	LOC107312272	RCOR3	ZRANB1	CHST2	LOC107311290	SEMA3E	
CREBZF	LOC107312446	RFWD2	ZRANB2	CHST3	LOC107311795	SEMA4B	
CRIPT	LOC107312514	RGMA		CKAP4	LOC107312120	SERPINH1	
CRNKL1	LOC107312724	RHOU		CLCC1	LOC107312420	SERPINI1	
CRYZL1	LOC107312753	RLIM		CNIH1	LOC107312697	SERTAD2	
CSDC2	LOC107312790	RNF2		CNN3	LOC107312843	SGCE	
CSNK1G1	LOC107312794	RNF4		COL11A1	LOC107313248	SH3BP4	
CTDSPL2	LOC107312844	RNF8		COL12A1	LOC107313342	SH3D19	
CTH	LOC107312851	RNGTT		COL1A2	LOC107313525	SH3PXD2B	
CTNND2	LOC107313038	RNMT		COL3A1	LOC107313631	SHC4	
CTR9	LOC107313438	RPAP3		COL5A2	LOC107313922	SHF	
CTTNBP2	LOC107313528	RPGR		COL6A1	LOC107314317	SIPA1L1	
CUL1	LOC107313656	RPP40		COL6A2	LOC107314438	SLC12A2	
CUL4B	LOC107313717	RPS6KA6		COL6A3	LOC107314540	SLC12A4	
CXADR	LOC107313962	RPUSD2		COPG1	LOC107314557	SLC16A2	
DACT2	LOC107314132	RRP15		CREB3L1	LOC107314568	SLC20A2	
DCLK2	LOC107314371	RRP1B		CSNK2A2	LOC107314574	SLC25A4	
DCX	LOC107314487	RSPO3		CSPG4	LOC107314696	SLC27A4	
DDX1	LOC107314585	RTCA		CSRP2	LOC107314709	SLC2A3	
DDX24	LOC107314723	RTF1		CTBP1	LOC107314914	SLC30A7	
DDX3X	LOC107314750	RTN4IP1		CTHRC1	LOC107314917	SLC35A5	
DENND5B	LOC107314806	RWDD1		DAGLA	LOC107314954	SLC37A3	
DESI2	LOC107314815	SACM1L		DAP	LOC107315268	SLC38A7	
DET1	LOC107314847	SALL1		DCAF5	LOC107315292	SLC41A2	
DHTKD1	LOC107315283	SALL3		DCN	LOC107316274	SLC4A2	
DHX15	LOC107315722	SBF2		DCTN1	LOC107316275	SLC4A3	
DIEXF	LOC107315759	SCAF11		DDAH1	LOC107316423	SLC5A6	
DIS3	LOC107315955	SCAF8		DES	LOC107316440	SLC9A6	
DIS3L2	LOC107316090	SEC24B		DFNA5	LOC107316593	SLIT2	
DLG1	LOC107316524	SEMA7A		DHRSX	LOC107316691	SLX4IP	
DLG3	LOC107316637	SENP6		DHX36	LOC107316692	SMIM18	
DLG5	LOC107316774	SEPSECS		DLC1	LOC107316866	SNAI2	
DLL1	LOC107316860	SETD2		DLGAP4	LOC107316896	SNX12	
DNAJC17	LOC107317002	SF3B1		DLK2	LOC107316904	SNX33	
DNMT3A	LOC107317006	SF3B3		DNAH7	LOC107316980	SNX7	

DOCK7	LOC107317009	SF3B6		DNAJB4	LOC107317105	SOGA3	
DPYSL4	LOC107317050	SH3BGR		DNAJC10	LOC107317449	SORCS2	
DROSHA	LOC107317055	SH3RF1		DNAJC25	LOC107317468	SOX4	
DTNA	LOC107317315	SHOC2		DNER	LOC107317657	SPATA4	
DYNC2H1	LOC107317372	SIPA1L2		DNPEP	LOC107317794	SPEG	
DYRK1A	LOC107317472	SIX4		DOK4	LOC107318056	SPRED1	
EAF1	LOC107317506	SKAP2		DOK7	LOC107318415	SPTLC3	
EDA	LOC107317605	SLAIN1		DPYSL5	LOC107318625	SREBF2	
EFCAB1	LOC107317619	SLC16A14		DRD5	LOC107318809	SRF	
EIF4A3	LOC107318221	SLC24A5		DSEL	LOC107319543	SRP72	
EIF4E	LOC107318346	SLC35F3		DSG2	LOC107320701	SSR1	
ELAVL4	LOC107318438	SLC38A8		DSTN	LOC107321229	SSR3	
ENY2	LOC107318449	SLC39A6		DUSP10	LOC107321238	ST3GAL1	
EPCAM	LOC107318520	SLC8A3		DUSP5	LOC107322296	ST5	
EPHA4	LOC107318577	SLITRK3		DYNC1H1	LOC107323964	STARD13	
EPHA5	LOC107318655	SLK		DYNC1LI1	LOC107324203	STIM1	
ERI3	LOC107318989	SMARCAD1		DYNC1LI2	LOXL3	STXBP6	
ESF1	LOC107320437	SMARCC1		EDA2R	LPP	STYX	
EVX1	LOC107321247	SMARCD3		EDEM3	LRCH2	SULF1	
EXOSC9	LOC107321951	SMC3		EDNRA	LRFN5	SYDE2	
EXTL3	LOC107323918	SMCHD1		EFHC1	LRP12	SYNJ2	
EYA4	LPCAT1	SMOC1		EHHADH	LRP4	SYNPO2	
FAM103A1	LRIG3	SNCG		EMC2	LRP8	TBC1D2B	
FAM110B	LRPPRC	SNRPB2		EMILIN1	LRRC20	TBCE	
FAM155B	LRRC1	SNTG1		EMILIN2	LRRC6	TES	
FAM171B	LRRC40	SNW1		EML6	LRRFIP2	TG	
FARP2	LYPD6B	SNX6		ENAH	LRRN4	TGDS	
FARSB	LYPLA1	SOCS2		EPM2A	LTBP1	TGFB2	
FASTK	MAP1A	SOX11		ERBB4	MAP2K1	TGFB3	
FASTKD2	MAP3K4	SP4		ERC1	MAP3K5	THBS1	
FAXC	MAP4K3	SP5		ERP44	MAP4	THBS2	
FBXO3	MAP4K5	SPAG6		ETV1	MAP7D1	TIMP3	
FBXW7	MB21D2	SPATA13		EVA1A	MAPK8IP1	TMCO1	
FERD3L	MBD5	SPDYA		EXOSC8	MAPRE3	TMEM138	
FGF13	MBIP	SPICE1		EXT1	MED28	TMEM150A	
FGF4	MDN1	SPOPL		EXT2	MESDC2	TMEM150C	
FGF8	MEIS1	SPRTN		FAM114A1	MFHAS1	TMEM209	
FKBP4	METTL16	SPSB4		FAM13C	MICU3	TMEM255A	
FMNL2	METTL5	SRSF11		FAM150B	MINPP1	TMEM263	
FMR1	METTL6	SS18		FAM155A	MMP15	TMEM43	
FRMD1	MGEA5	SSB		FAM177A1	MMS19	TMEM61	
FUBP1	MIPOL1	SSBP3		FAM49A	MOB4	TMEM63B	
FXR1	MORF4L1	ST6GALNAC3		FAM81A	MORN4	TMSB15B	
FYTTD1	MPP4	ST8SIA2		FAT1	MREG	TMSB4X	
FZD3	MPP6	STAG1		FBN1	MSMO1	TMTC2	
FZD7	MRAP	STK3		FBXL15	MUC4	TNK2	

G3BP2	MRE11A	SUGT1		FERMT2	MYH15	TNKS1BP1	
GAB2	MRPL1	SUMO1		FEZ2	MYH9	TNS1	
GABPA	MRPL53	SUPV3L1		FGF2	MYO1B	TOMM7	
GALNT13	MRPS14	SUV39H2		FHL2	NACAD	TPM1	
GAS2	MRPS18A	SYNE2		FHL5	NALCN	TPM3	
GFRA4	MSANTD3	T		FKBP14	NAV1	TPM4	
GINS3	MSGN1	TAF1		FKBP7	NAV3	TRAM1	
GLI3	MSH6	TAF2		FKBP9	NBEA	TRAPPC10	
GLRX2	MTA3	TBC1D19		FN1	NEK9	TRAPPC2	
GOLT1B	MTF2	TBX15		FNDC1	NEURL1	TRIB2	
GPATCH2	MTMR9	TBX18		FOSL2	NFKB2	TRIM35	
GPC3	MTURN	TCP1		FRMPD4	NID1	TRIM54	
GPLD1	MYEF2	TDG		FURIN	NOTCH2	TROVE2	
GPR85	NAALADL2	TDRD9		FUT11	NPM3	TRPC4	
GREB1L	NAF1	TENM4		FZD8	NPNT	TSC22D2	
GRIP1	NAP1L4	TERT		GADD45A	NRBP1	TSG101	
GTF2B	NASP	TEX10		GALC	NRSN1	TUB	



DISCUSSION

CHAPTER 3

I. Molecular details of EHT: what we know and what we are still missing!

Endothelial to hematopoietic transition is driven by an incredibly complex yet to be fully elucidated interplay between developmental timing and tissue microenvironment factors. Despite having identified some of the key transcriptional and pathways playing a role in this transition, still there remain several unanswered questions surrounding the progression from embryonic mesoderm to endothelium to HSPCs and adult HCs, especially the morphological changes and restructuring occurrence during this transition. Efficient generation and differentiation of hematopoietic cells in vitro to date has been accompanied with technical difficulties, and these difficulties are almost certainly worsened by a need for a more complete understanding of in vivo endothelial and hematopoietic development.

In this study, I aimed at defining molecularly the hemogenic endothelium undergoing EHT. For this, I took advantage of the precise culture system I contributed to set up (Yvernogeau et al., 2016) to isolate discrete cell populations from the non-committed mesoderm to hematopoietic cells including hemogenic endothelium as an intermediate. I used the level of LDL uptake to discriminate the endothelial cells from the hemogenic ECs that display a lower LDL uptake. I confirmed this point by validating the expression of several lineage-specific genes in particular RUNX1, CD144, and CD31. I obtained the transcriptomes of the different cell types and initiated a bioinformatic approach aiming at revealing the molecular identity of HECs. The first step of the analysis was to control the expression of the known marker in our population, we found that mesodermal population expresses specific mesoderm genes i.e TBOX, ZIC3, MYC, the endothelial population expresses VWF, KDR, FLI1, CDH5, GATA2, Tal1 … Whereas HEC population in addition to expressing endothelial-specific genes, it expresses RUNX1, GFI1, KIT, and CD44, finally the hematopoietic population expresses CD45, SPI1 … Together the expression of these key genes validates the specificity of the isolated and sequenced populations.

My approach is driven by the fact that, despite efforts made by several groups worldwide, the precise identity of the hemogenic endothelium remains largely inaccessible. Indeed, several essential regulators such as Runx1, Gfi1, Gfi1b, CD44

Gpr56, cdca7… have been identified but the complete molecular identity is still escaping our scrutiny.

To reach this aim, I asked several questions:

- How could I define molecularly a hemogenic endothelium?
- What are the specific differences between endothelium and hemogenic endothelium?
- Could I reveal specific features that have not been previously described?

To this end, I undertook isolating a specific gene signature for each of the specific cell types compared to all the other ones. This strategy allows isolating gene sets that are specifically upregulated by a given cell type but excludes genes which expression is shared by several cell types. Indeed, several essential factors present in both the endothelium and hemogenic endothelium are erased as well as factors shared between the hemogenic endothelium and hematopoietic cells. Runx1, Gfi1, Flk1… disappear because they are not specific. Then I used this specific gene sets to create a unique expression matrix and reconstruct gene interaction networks using WGCNA. Interestingly, WGCNA reconstructed two HEC networks of up-regulated genes and one HEC network of down-regulated genes. Cytoscape representation of these gene networks reveals gene interactions within the network and genes with multiple connections. These genes are called hubs and carry essential functions in the network (Charbord et al., 2014; Langfelder and Horvath, 2008).

In an attempt to decode the hidden identity of the hemogenic endothelium and the EHT, several bulk or single-cell transcriptomes have been produced over the past few years (For a review see (Ottersbach, 2019)). All the papers report isolation of EHT based on a combination of endothelial-specific markers genes such as VE-Cadherin or ESAM, the putative hemogenic endothelium-specific marker Ly6A, Runx1 expression driven by the +23 enhancer, the combination of the transcription factors Gfi1 and Gfi1b, the expression of the endo-hematopoietic marker CD31 and the presence or absence of several hematopoietic-specific markers such as CD41, c-Kit or CD45. In brief, despite common markers used between the different teams, it is not clear whether the same, or even close, cell types or cell fractions are isolated. Whatever the samples, the bioinformatics strategies are diverse for isolating specific differentially expressed gene sets.

Starting a few years ago, Swiers et al, 2013 were the first to establish the fact that HECs gradually lose their endothelial identity between E8.5-E10.5 as their hematopoietic potential increase and the endothelial and hematopoietic potential increase as well as that the endothelial and hematopoietic capacities are mutually exclusive. They took advantage of the RUNX1 +23 enhancer which is active in definitive HSPCs and HECs and performed the first functional and molecular analysis of embryonic HECs undergoing EHT. Therefore, they showed that the initiation of the ECs to become HECs is coordinated by RUNX1 (Swiers et al., 2013).

In 2005, Solaimani Kartalaei et al, took advantage of the *Ly6aGFP* (Sca1) mouse model, in which all HSCs throughout development are GFP^+ (de Bruijn et al., 2002; Joosten et al., 2002), to study the EHT. They present a set of RNA sequencing data obtained from highly enriched small numbers of EHT cells from *Ly6aGFP* embryos, aortic ECs, HECs, and emerging HSCs. Among the few (530) DEGs during EHT, they found *Gpr56* as the highest up-regulated gene encoding a cell surface receptor and then showed the functional involvement of *Gpr56* in HSC emergence during EHT. However, on the three experiments performed, one was clearly an outlier (Figure 2B) explained by lower sequencing depth. Nevertheless, the experiment was maintained for further analysis certainly inducing a bias in the final picture.

Using the same mouse model, this time li et al 2014, used another combination of marker (CD31, Cdh5, Esam and cKit) in order to compare GFP^+ and GFP^- endothelial cells and intra-aortic cluster cells. They focused on a specific trait of their microarray analysis which is the inflammatory pathway in HECs. They found that interferon signaling increased HSPC production (Li et al., 2014). The role of pro-inflammatory signals has also been shown to promote HSPC emergence from HECs in zebrafish (Espín-Palazón et al., 2014; He et al., 2015).

In 2012, and this time using the ES culture system, Lancrin et al aimed to look to downstream RUNX1 effectors that drive the EHT. To do so, they compared the transcriptomes of $RUNX1^{+/-}$ and $RUNX1^{-/-}$ hemogenic endothelium in ES culture (transcriptome generated by cDNA microarray). They found that the genes encoding Gfi1 and Gfi1b transcriptional repressors displayed significantly lower expression levels in the absence of RUNX1. Therefore, after chromatin immunoprecipitation, they demonstrated that RUNX1 directly regulates the expression of Gfi1 and Gfi1b at the

onset of hematopoietic development (Lancrin et al., 2012). These findings were further investigated by Thambryajah et al in 2015, who by combining mouse reporter lines and functional assays, they established that *Gfi1* expression is initiated in a rare population of HE whereas GFI1B expression is upregulated in emerging IAHC. And for the first time, they described that the repressive CoREST complex is implicated in the transition through the recruitment by GFI1 and GFI1b, of LSD1 a member of this complex (Thambyrajah et al., 2016).

In 2014, Guiu et al focused on the role of NOTCH effectors in EHT this time and used genome-wide analysis of RBPj-targeted genes during embryonic hematopoiesis, CHIP assay and ChIP on chip experiments, coupled with the analysis in the jag 1 deficient mice. They could demonstrate cdca7 gene as a new Notch target gene involved in early hematopoietic development. And confirmed the requirement of cdca7 for the HSC and hematopoietic progenitor emergence in zebrafish (Guiu et al., 2014).

Back to RUNX1, Gao et al in 2018 point to the differences between the transcriptomic profiles of HE depending on the hematopoietic site comparing YS and AGM at E9.5 and E10.5. They supposed that the drivers of the different lineage potential are likely to be the distinct molecular properties of the HE cells from which they are derived. Using the expression of GFP from the endogenous RUNX1 locus, they isolated and compared HE from different anatomic sites. They get a global view of the TFs expressed in the different compartment validating some already discovered genes implicated in EHT (Gao et al., 2018).

The same year Baron et al, used single-cell transcriptomics to reveal the dynamic of HSC production in the aorta. This elegant study in vivo focused on the analysis of the composition of the intra-aortic clusters and found these to contain type I and type II pre-HSCs as well as more committed HSPCs. They were the first to isolate single whole IAHCs in the ventral and dorsal side of the aorta of mice embryos and demonstrated that they have similar transcriptomes. Concerning the silencing of the endothelial program during EHT, they identified a set of genes and transcription factors activated in the initiation of the hematopoietic program during EHT. They showed that very few genes were differentially expressed between HE and non-HE cells at E10 and E11 (Baron et al., 2018).

In the most recent single-cell RNA-Seq study, CD44 was identified as a marker that becomes detectable specifically in the transitioning population in which endothelial and hematopoietic programs are co-expressed. In fact, together with ckit and Cdh5, CD44 allowed an accurate isolation of cell stages from non-HEC to mature HSPCs. Furthermore, CD44 appears to have a functional role, as a disruption of the interaction with its ligand hyaluronan interfered with the EHT (Oatley et al., 2018). In the same lab and using single cell transcriptomic techniques, Bergiers et al, studied the gene regulatory networks of the emergence of hematopoietic stem and progenitor cells from mouse embryonic endothelium at the single cell levels. They found that a heptad of transcription factors (Runx1, Gata2, Tal1, Fli1, Lyl1, Erg and Lmo2) is specifically co-expressed in the hemogenic population. They surprisingly suggest that even though Fli1 initially supports the endothelial cell fate, it acquires a pro-hematopoietic role when co-expressed with Runx1 (Bergiers et al 2018).

A close look at all these studies allows spotting flagrantly the heterogeneity in defining the HE population with different markers (review (Ottersbach, 2019)) as well as the time point chosen for the analyses. i.e Swiers et al looked to HE at E8.5 and E10.5 whereas Guiu et al looked at E11.5, Baron et al looked at E9 and E10. This points to a lack in the definition of a well-defined timeframe of the HE identification.

Baron et al and Gao et al discussed a close relationship between HE and EC in the major arteries with a shortlist of deferentially expressed genes in each population. Or in our analysis using WGCNA, we could spot 2 unique signatures of EC and HEC populations represented by different gene networks. The network of ECs composed of 600 genes and the two networks of HECs composed respectively of 260 and 200 genes showed distinct GO categories. EC module was enriched with Pleckstrin homology domain, membrane genes, lysosome, SH3 domain, cell adherens junctions and wnt pathway. Whereas HEC were enriched with cell migration, ECM, FA, phosphatase and Golgi apparatus. Hence in our culture system, we were able to define 2 distinct categories of HE and non-HE cells.

The difference in the number of DEGs in the previous analysis is due to the use of different bioinformatic methodology going from the isolation and comparison of populations, the exclusion of out layers or not, the definition of p-values and fold changes. In their analysis, many of the groups aimed to look for gene candidates,

surface markers and transcription factors. However, in our study, the use of WGCNA allowed us to have a global view of each step of the transition considering networks and not specific genes (5 specific networks going from M to EC to HEC and HC). It allowed us to define the implication of ECM, cytoskeleton and morphological changes in HE going through EHT. In this study, we focused on studying the HEC networks, but indeed in the future, it is interesting to look for the specification of the endothelium from the mesoderm and the switch endothelium to hemogenic endothelium.

Since the EHT is transient and since not all the HE cells go through the transition together, we decided to perform single-cell transcriptomics on our culture system. To do so we isolated 4 samples (1 M, 2 HEC, and 1 HC) and we performed single-cell transcriptomics using 10X genomics. The analysis I showed in the results are preliminary and will be more refined, but we believe that this analysis will help us to capture a specific signature of cells undergoing EHT.

II. The parallel between EMT and EHT: *what is the common side?*

At first glance, EHT seems to share similarities with another major cell fate change: the epithelial-to-mesenchymal transition (EMT). It has been demonstrated that TGFB signaling is shared between these two transitions. Activation of the TGFB pathway was shown to promote the EMT, while its role in the EHT has been seemingly controversial. While some reports suggested that activation of TGFB signaling promotes the EHT (Monteiro et al., 2016), others suggested that TGFB activity needs to be down-regulated for the EHT to proceed (Lempereur et al., 2018). Interestingly, in our HEC networks, we found some TGFB signaling members: TGFB1, TGFB3, TGFBR1, RBPMS… confirming that TGFB plays a role in the EHT.

More globally in our HEC networks we found some up-regulated genes known to play a role in the EMT as TESTIN, POFUT2, ITGAV, BMPER as well as down-regulated genes like CTBP2 and TEAD4. This finding suggests that effectively EHT and EMT share some molecular signature.

While looking to the connectivity in our HEC network, we found between the most connected genes POFUT2 a GDP-fucose protein O-fucosyltransferase 2. It is an enzyme responsible for adding fucose sugars in O linkage to serine or threonine residues in Thrombospondin Repeats TSRs. The majority of POFUT2 targets are

constitutive component of ECM or secreted matrix associated proteins able of influencing several processes as cell adhesion/migration, ECM synthesis, remodeling and finally modulating growth factor synthesis. In 2010, DU et al showed that the loss of POFUT2 in mice resulted in unrestricted EMT (Du et al., 2010).

Since POFUT2 is a hub in one of our HEC networks we thought that it could play an important role in the transition, so we decided to functionally validate it in our culture system. To do so we designed three siRNA against POFUT2 and performed transfection in our culture system. The knock-down of POFUT2 induces a decrease in the endothelial fraction and an increase in the hemogenic endothelial one suggesting that the loss of POFUT2 enhances the EHT. In addition, to further validate the gene network and its predictive value, we selected the POFUT2's first neighbors and examined if a decrease of POFUT2 mRNA expression was also associated with a modulation of several of its first neighbors. Based on gene connectivity, we selected 3 genes among the most connected to POFUT2. COL5A2 DNAJB4 and EMILIN1A. The three chosen genes showed a significant decrease upon POFUT2 siRNA knock-down. Taken together, this indicates that POFUT2 and, at least, 3 of their first neighbors are acting coordinately to regulate the HEC fate and the hematopoietic production.

The knock-down of TESTIN another hub of our HEC network didn't show any effect on the EHT this time. The loss of TESTIN altered one of the tested first neighbors but not the others. It was interesting to look for the half-life of this protein and indeed we found that proteins of the extracellular matrix and the cytoskeleton have a half-life of the order of a few days and even a few months whereas the majority of metabolic enzymes last only a few hours which can explain that even if we inhibit the mRNA of TESTIN we still have the protein which can compensate this loss. This experiment point to the importance of looking to the half-life of the protein before performing the knock-down experiments.

III. NOTCH and WNT pathway: Do our culture system recapitulate the known role of Notch and wnt during embryonic hematopoiesis?

The involvement of the Notch pathway in hematopoietic development has been described in different system and species. Notch expression in aortic ECs and HCs

displays a conserved signature between species. HC cluster production is accompanied by a decrease of Notch ligands. This pattern takes place as early as runx1 expression initiates in the hemogenic endothelium. This pattern is also prominent as HCs mature from CD41+ to CD45+ cells. Our transcriptomic data indicates that notch1 signaling (NOTCH1, JAG1, JAG2 …) is downregulated in our HC population. A requirement for Notch signaling and intraembryonic HSC production has been shown for mouse and zebrafish embryos. Loss-of-function experiments demonstrate that ablating Notch signaling suppresses definitive (embryo-derived), but not primitive (yolk sac-derived), hematopoiesis (Burns, 2005; Kumano et al., 2003; Robert-Moreno et al., 2008). Here, we show that the suppression of Notch signaling using DAPT enhances the hemogenic endothelial phenotype and the production of $CD45^{+}$ HCs in our culture system. However, this phenotype was already shown in our lab by the use of DAPT directly in the chicken embryo, and mouse aorta organotypic culture causing an increase in the HC production (Richard et al., 2013). Together these results suggest that the downregulation of Notch signaling in the hemogenic endothelium is required for aortic hematopoiesis to occur.

Concerning the Wnt pathway, when looking to the HEC networks, we find several actors of the Wnt pathway i.e CTNNB1, CTNNBIP1, MESD2, WLS, NPHP3, PORCN… Since it has been demonstrated that the generation of HSCs depends on β-catenin activity and becomes gradually β-catenin independent, we thought about testing an activator and an inhibitor of the β-catenin. We demonstrated that when using β-catenin inhibitor (PKF) we induced an increase of endothelial cells and a decrease of the hemogenic population. Whereas the use of β-catenin activator (SB) induces a decrease in the endothelial population and an increase in the hemogenic population. These results are in keeping with the results published by (Ruiz-Herguido et al., 2012) on the mouse AGM. Taken together we confirmed that the Wnt pathway plays a role in EHT.

IV. Conclusion and Perspectives

Finally, I could demonstrate that our culture system does recapitulate all the steps of the EHT occurring in the dorsal aorta of the embryo. It goes from a nonengaged mesoderm which will acquire endothelial phenotype and then hemogenic endothelial fate, those HEC will then go through EHT and will produce HCs.

Bioinformatics and statistical tests allowed me to identify specific DEGs in each population compared to the others, therefore, to determine a single and specific signature of each population. Those DEGs were used to perform WGCNA a powerful tool to generate gene networks, it generates 5 modules (1M, 1EC, 1HC, and 2 HEC) represented as networks.

We focused on studying the HEC population and looked at the GO of the HEC networks. It shows specifically an enrichment with ECM, cytoskeleton and EMT actors, categories that have not been fully studied in the EHT yet. We demonstrated that in keeping with the literature, NOTCH and wnt pathways play a role in the EHT. And identified a new actor of this transition POFUT2 known to play a role in the EMT and could demonstrate that the loss of POFUT2 enhances the EHT. Together these results confirmed the predictive value of our gene networks.

Nevertheless, we are aware of some potential drawbacks of our culture system recapitulated as:

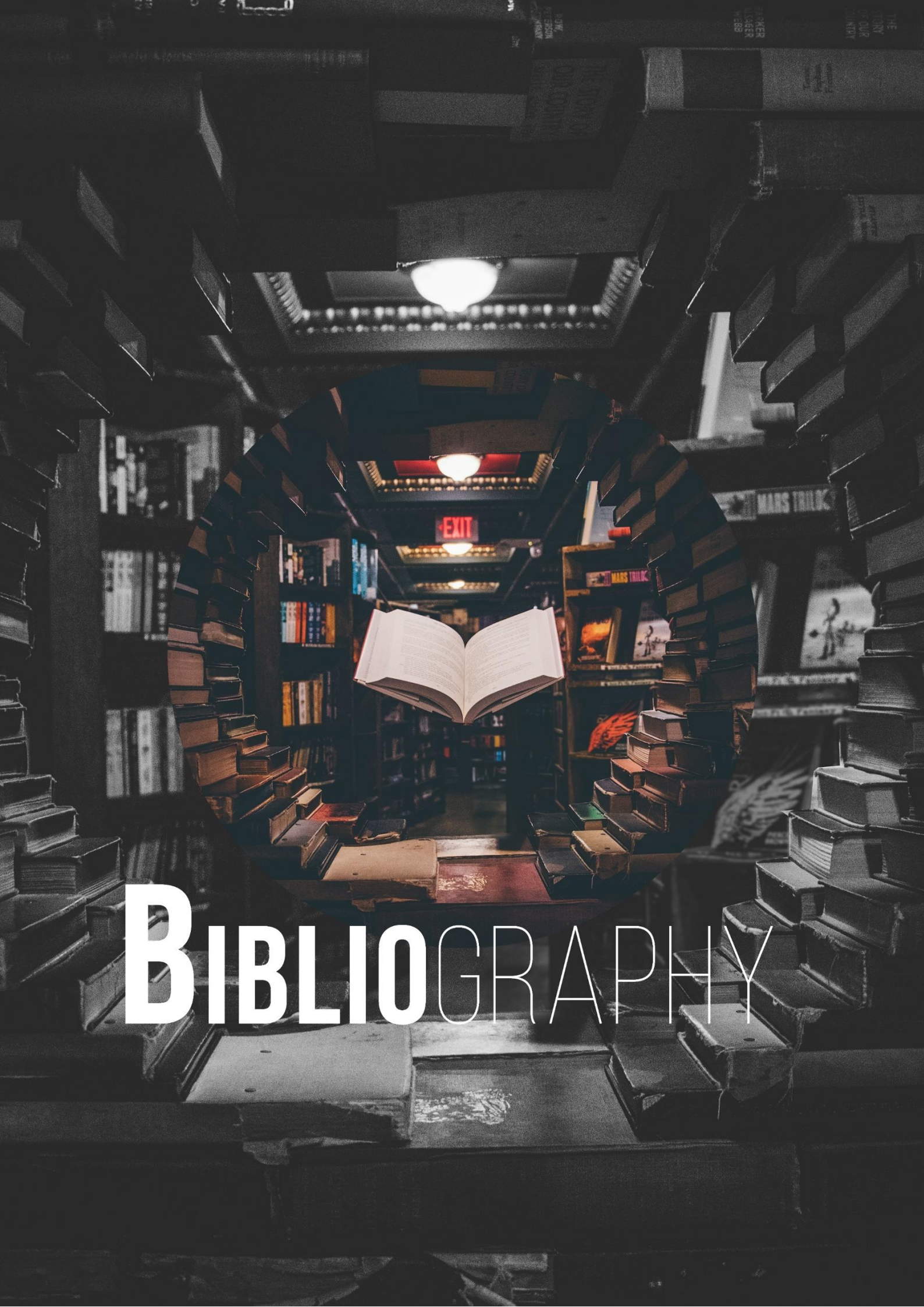
- Is the quail model relevant for the study of hematopoiesis?
- Is there any way to synchronize our culture system?
- Does the system recapitulate the EHT of the embryo?
- Does the EHT occurring in our culture system recapitulate YS or AGM hematopoiesis and what type of hematopoietic cells are we generating?
- Does our data set concur with the already published EHT *in vivo* and *in vitro* data sets?

As already described in the introduction, hematopoiesis is well conserved between species, the quail has been demonstrated as a key model for the study of this multistep process (quail-chicken chimeras…). We are convinced that hematopoiesis in quail is in keeping with the chicken and mammal ones. We faced different problems with this model and tried to find solutions. The first one is the absence of antibodies that can be used to identify the hematopoietic population. We managed to perform qPCR to validate the presence of the different hematopoietic cell categories using key lineage gene markers. The second one is that the culture system is not synchronized. Even if as described in Yvernogeau *et al*. 2016, almost 75% of the cells at day 2 are ECs, the EHT process is not synchronized and not all the HEC go through EHT and produce

HCs at the same time. In Yvernogeau *et al*., we identified a culture condition where in when the fetal serum is removed from the medium, 90% of the cells became endothelial. However, the culture is blocked at this stage and no EHT nor HC production occurs. This specific condition could be the key to synchronize the culture by removing the serum in the beginning of the culture and reintroducing it at day 4 when cells are ready to undergo EHT. The other one is to ensure that the hubs we have identified are actually expressed in the hematopoietic aorta. In order to assess that the culture system recapitulates cellular events occurring in the embryo, we are performing *in situ* hybridization *in toto* on chicken embryos to validate POFUT2 expression in the hematopoietic sites in particular the aorta.

To further test whether the culture system produces EMP or HSCs, I would perform grafts of hematopoietic cells produced by the culture into chicken, age-matched, recipients.

The 2 data sets we generated will be further studied in the lab. Many perspectives are established. We will start by a comparison of our transcriptome with transcriptomes of cells isolated from embryos and ES cell cultures by others. This can be done *in silico* by the use of different transcriptomes generated by colleagues in the field. (cite les articles) and can contribute to a global view of the EHT leading to a better understanding of this process. It is true that we focused in our study on the HEC population and identified new actors in the EHT. Since our culture system recapitulates all the EHT steps, we will look in details to the first step of the commitment between EC toward HEC, a key question in the EHT field that have implications in regenerative medicine. Finally, since our cells can be efficiently transfected, our culture system can be further used to the validation of the expression of novel genes playing role in the EHT.



BIBLIOGRAPHY

Abkowitz, J. L. (2002). Evidence that the number of hematopoietic stem cells per animal is conserved in mammals. *Blood* **100**, 2665–2667.
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., et al. (2009). Biomechanical forces promote embryonic haematopoiesis. *Nature* **459**, 1131–1135.
Adolfsson, J., Borge, O. J., Bryder, D., Theilgaard-Monch, K., Astrand-Grundstrom, I., Sitnicka, E., Sasaki, Y. and Jacobsen, S. E. (2001). Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. Immunity 15,659-69.
Akashi, K., Traver, D., Miyamoto, T. and Weissman, I. L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193–197.
Alvarez-Silva, M. (2003). Mouse placenta is a major hematopoietic organ. *Development* **130**, 5437–5444.
Alvarez-Silva, M., Belo-Diabangouaya, P., Salaun, J. and Dieterlen-Lievre, F. (2003) 'Mouse placenta is a major hematopoietic organ', Development 130(22): 5437-44.
Ambler, C. A., Nowicki, J. L., Burke, A. C. and Bautch, V. L. (2001). Assembly of Trunk and Limb Blood Vessels Involves Extensive Migration and Vasculogenesis of Somite-Derived Angioblasts. *Developmental Biology* **234**, 352–364.
Anjos-Afonso, F., Currie, E., Palmer, H. G., Foster, K. E., Taussig, D. C. and Bonnet, D. (2013). CD34– Cells at the Apex of the Human Hematopoietic Stem Cell Hierarchy Have Distinctive Cellular and Molecular Signatures. *Cell Stem Cell* **13**, 161–174.
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**, 257–267.
Barcena, A., Muench, M. O., Kapidzic, M. and Fisher, S. J. (2009). A New Role for the Human Placenta as a Hematopoietic Site Throughout Gestation. *Reprod Sci* **16**, 178–187.
Barker, J. E., Keenan, M. A. and Raphals, L. (1969). Development of the mouse hematopoietic system. II. Estimation of spleen and liver ?stem? cell number. *J. Cell. Physiol.* **74**, 51–55.
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. *Nat Commun* **9**, 2517.
Batta, K., Florkowska, M., Kouskoff, V. and Lacaud, G. (2014). Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells. *Cell Reports* **9**, 1871–1884.
Baum, C. M., Weissman, I. L., Tsukamoto, A. S., Buckle, A. M. and Peault, B. (1992). Isolation of a candidate human hematopoietic stem-cell population. *Proceedings of the National Academy of Sciences* **89**, 2804–2808.
Beaupain, D., Martin, C. and Dieterlen-Li, F. (1979). Related to the Origin of Stem Cells and Site of Erythropoiesis? *STEM CELLS* 15.
Bechtold, T. E., Smith, P. B. and Turpen, J. B. (1992). Differential stem cell contributions to thymocyte succession during development of Xenopus laevis. 9.
Begley, C. G. and Green, A. R. (1999). The SCL Gene: From Case Report to Critical Hematopoietic Regulator. 12.
Begley, C. G., Aplan, P. D., Denning, S. M., Haynes, B. F., Waldmann, T. A. and Kirsch, I. R. (1989). The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proceedings of the National Academy of Sciences* **86**, 10128–10132.
Bergiers, I., Andrews, T., Vargel Bölükbaşı, Ö., Buness, A., Janosz, E., Lopez Anguita, N., Ganter, K., Kosim, K., Celen, C., Itır Perçin, G., et al. (2018). Single-cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis. eLife 7, e29312.

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* **464**, 108–111.
Bertrand, J. Y., Giroux, S., Golub, R., Klaine, M., Jalil, A., Boucontet, L., Godin, I. and Cumano, A. (2005). Characterization of purified intraembryonic hematopoietic stem cells as a tool to define their site of origin. *Proceedings of the National Academy of Sciences* **102**, 134–139.
Böiers, C., Carrelha, J., Lutteropp, M., Luc, S., Green, J. C. A., Azzoni, E., Woll, P. S., Mead, A. J., Hultquist, A., Swiers, G., et al. (2013). Lymphomyeloid Contribution of an Immune-Restricted Progenitor Emerging Prior to Definitive Hematopoietic Stem Cells. *Cell Stem Cell* **13**, 535–548.
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* **464**, 116–120.
Bollerot, K., Romero, S., Dunon, D. and Jaffredo, T. (2005). Core binding factor in the early avian embryo: cloning of Cbfbeta and combinatorial expression patterns with Runx1. *Gene Expr. Patterns* **6**, 29–39.
Bos, F. L., Hawkins, J. S. and Zovein, A. C. (2015). Single-cell resolution of morphological changes in hemogenic endothelium. *Development* **142**, 2719–2724.
Boudreau, N., Andrews, C., Srebrow, A., Ravanpay, A. and Cheresh, D. A. (1997). Induction of the Angiogenic Phenotype by Hox D3. *The Journal of Cell Biology* **139**, 257–264.
Burns, C. E. (2005). Hematopoietic stem cell fate is established by the Notch-Runx pathway. *Genes & Development* **19**, 2331–2342.
Burns, C. E., Traver, D., Mayhall, E., Shepard, J. L. and Zon, L. I. (2005) 'Hematopoietic stem cell fate is established by the Notch-Runx pathway', Genes Dev 19(19): 2331-42.
Busch, K. and Rodewald, H.-R. (2016). Unperturbed vs. post-transplantation hematopoiesis: both in vivo but different. *Curr Opin Hematol* **23**, 295–303.
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**, 542–546.
Butko, E., Distel, M., Pouget, C., Weijts, B., Kobayashi, I., Ng, K., Mosimann, C., Poulain, F. E., McPherson, A., Ni, C.-W., et al. (2015). Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development* **142**, 1050–1061.
Buza-Vidas, N., Duarte, S., Luc, S., Bouriez-Jones, T., Woll, P. S. and Jacobsen, S. E. (2011) 'GATA3 is redundant for maintenance and self-renewal of hematopoietic stem cells', Blood 118(5): 1291-3.
Canon, J. and Banerjee, U. (2003). In vivo analysis of a developmental circuit for direct transcriptional activation and repression in the same cell by a Runx protein. *Genes Dev* **17**, 838–843.
Caprioli, A., Jaffredo, T., Gautier, R., Dubourg, C. and Dieterlen-Lièvre, F. (1998). Blood-borne seeding by hematopoietic and endothelial precursors from the allantois. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1641–1646.
Caprioli, A., Minko, K., Drevon, C., Eichmann, A., Dieterlen-Lièvre, F. and Jaffredo, T. (2001). Hemangioblast commitment in the avian allantois: cellular and molecular aspects. *Dev. Biol.* **238**, 64–78.
Capron, C., Lecluse, Y., Kaushik, A. L., Foudi, A., Lacout, C., Sekkai, D., Godin, I., Albagli, O., Poullion, I., Svinartchouk, F. et al. (2006). The SCL relative LYL-1 is required for fetal and adult hematopoietic stem cell function and B-cell differentiation. Blood 107, 4678-86.
Charbord, P. (1994). Hemopoietic stem cells: Analysis of some parameters critical for engraftment. *STEM CELLS* **12**, 545–562.
Charbord, P., Pouget, C., Binder, H., Dumont, F., Stik, G., Levy, P., Allain, F., Marchal, C., Richter, J., Uzan, B., et al. (2014). A systems biology approach for defining the molecular framework of the hematopoietic stem cell niche. *Cell Stem Cell* **15**, 376–391.
Charbord, P., Tavian, M., Humeau, L. and Peault, B. (1996). Early Ontogeny of the Human Marrow From Long Bones: An Immunohistochemical Study of Hematopoiesis and Its Microenvironment. 12.

Chasis, J. A. (2006). Erythroblastic islands: specialized microenvironmental niches for erythropoiesis: *Current Opinion in Hematology* **13**, 137–141.
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**, 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* **457**, 887–891.
Chen, Q., Cheng, J.-T., Tsai, L.-H., Schneider, N., Buchanan', G., Carroll', A., Crist', W., Ozanne, B., Siciliano, M. J. and Baer, R. (1990). The tat gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix - loop - helix protein. 10.
Cheng, P. and Gabrilovich, D. (2008). Notch signaling in differentiation and function of dendritic cells. *Immunol Res* **41**, 1–14.
Chisaka, O. and Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5. *Nature* **350**, 473.
Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C. and Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. 8.
Choi, K.-D., Vodyanik, M. A., Togarrati, P. P., Suknuntha, K., Kumar, A., Samarjeet, F., Probasco, M. D., Tian, S., Stewart, R., Thomson, J. A., et al. (2012). Identification of the Hemogenic Endothelial Progenitor and Its Direct Precursor in Human Pluripotent Stem Cell Differentiation Cultures. *Cell Reports* **2**, 553–567.
Christensen, J. L., Wright, D. E., Wagers, A. J. and Weissman, I. L. (2004). Circulation and Chemotaxis of Fetal Hematopoietic Stem Cells. *PLoS Biol* **2**, e75.
Chung, Y. S., Zhang, W. J., Arentson, E., Kingsley, P. D., Palis, J. and Choi, K. (2002). Lineage analysis of the hemangioblast as defined by FLK1 and SCL expression. *Development* **129**, 5511–5520.
Ciau-Uitz, A., Walmsley, M. and Patient, R. (2000). Distinct Origins of Adult and Embryonic Blood in Xenopus. *Cell* **102**, 787–796.
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. *Nat Cell Biol* **15**, 502–510.
Coffman, C., Harris, W. and Kintner, C. (1990). Xotch, the Xenopus homolog of Drosophila notch. *Science* **249**, 1438–1441.
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**, 9836–9841.
Corbel, C. (2002). Expression of αVβ3 integrin in the chick embryo aortic endothelium. 4.
Corbel, C., Salaün, J., Belo-Diabangouaya, P. and Dieterlen-Lièvre, F. (2007). Hematopoietic potential of the pre-fusion allantois. *Developmental Biology* **301**, 478–488.
Cormier, F. and Dieterlen-Lievre, F. (1988). The wall of the chick embryo aorta harbours M-CFC, G-CFC, GM-CFC and BFU-E. 10.
Costa, G., Kouskoff, V. and Lacaud, G. (2012). Origin of blood cells and HSC production in the embryo. *Trends in Immunology* **33**, 215–223.
Coulombel, L., Huyhn, A. and Izac, B. (1995). [Characterization of hematopoietic progenitor cells during the human embryonic development]. *C. R. Seances Soc. Biol. Fil.* **189**, 611–615.
Cuadros, M. A., Coltey, P., Nieto, M. C. and Martin, C. (1992). Demonstration of a phagocytic cell system belonging to the hemopoietic lineage and originating from the yolk sac in the early avian embryo. 12.
Cumano, A., Dieterlen-Lievre, F. and Godin, I. (1996). Lymphoid Potential, Probed before Circulation in Mouse, Is Restricted to Caudal Intraembryonic Splanchnopleura. *Cell* **86**, 907–916.

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**, 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**, 3862–3870.
Dancis, J., Jansen, V., Gorstein, F. and Douglas, G. W. (1968) 'Hematopoietic cells in mouse placenta', Am J Obstet Gynecol 100(8): 1110-21.
Dantschakoff, W. (1907) 'Uber das erste Aufretem des Blutelements im Huhnerembryo', Folia Hematol. 4: 159.
Davis, G. E. and Senger, D. R. (2005). Endothelial Extracellular Matrix: Biosynthesis, Remodeling, and Functions During Vascular Morphogenesis and Neovessel Stabilization. *Circulation Research* **97**, 1093–1107.
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**, 2465–2474.
de Bruijn, M. F. T. R., Ma, X., Robin, C., Ottersbach, K., Sanchez, M.-J. and Dzierzak, E. (2002). Hematopoietic Stem Cells Localize to the Endothelial Cell Layer in the Midgestation Mouse Aorta. *Immunity* **16**, 673–683.
de Pater, E., Kaimakis, P., Vink, C. S., Yokomizo, T., Yamada-Inagawa, T., van der Linden, R., Kartalaei, P. S., Camper, S. A., Speck, N. and Dzierzak, E. (2013). *Gata2* is required for HSC generation and survival. *J Exp Med* **210**, 2843–2850.
Delassus, S. and Cumano, A. (1996). Circulation of Hematopoietic Progenitors in the Mouse Embryo. *Immunity* **4**, 97–106.
Dieterlen-Lièvre, F. and Jaffredo, T. (2009). Decoding the hemogenic endothelium in mammals. *Cell Stem Cell* **4**, 189–190.
Dieterlen-Lièvre, F., Pouget, C., Bollérot, K. and Jaffredo, T. (2006). Are intra-aortic hemopoietic cells derived from endothelial cells during ontogeny? *Trends Cardiovasc. Med.* **16**, 128–139.
Dieterlen-Lievre, P. F. (1975). On the origin of haemopoietic stem cells in the avian embryo: an experimental approach. 13.
Dieterlen-Livre, F. and Martin, C. (1981). Diffuse lntraembryonic Hemopoiesis in Normal and Chimeric Avian Development. 12.
Dommergues, M., Aubeny, E., Dumez, Y., Durandy, A. and Coulombel, L. (1992) 'Hematopoiesis in the human yolk sac: quantitation of erythroid and granulopoietic progenitors between 3.5 and 8 weeks of development', Bone Marrow Transplant 9 Suppl 1: 23-7.
Dooley, K. A., Davidson, A. J. and Zon, L. I. (2005) 'Zebrafish scl functions independently in hematopoietic and endothelial development', Dev Biol 277(2): 522-36.
Down, J. D. and Ploemacher, R. E. (1993) 'Transient and permanent engraftment potential of murine hematopoietic stem cell subsets: differential effects of host conditioning with gamma radiation and cytotoxic drugs', Exp Hematol 21(7): 913-21.
Drake, C. J. and Fleming, P. A. (2000). Vasculogenesis in the day 6.5 to 9.5 mouse embryo. **95**, 9.
Drevon, C. and Jaffredo, T. (2014). Cell interactions and cell signaling during hematopoietic development. *Exp. Cell Res.* **329**, 200–206.
Du, J., Takeuchi, H., Leonhard-Melief, C., Shroyer, K. R., Dlugosz, M., Haltiwanger, R. S. and Holdener, B. C. (2010). O-Fucosylation of Thrombospondin Type 1 Repeats Restricts Epithelial to Mesenchymal Transition (EMT) and Maintains Epiblast Pluripotency During Mouse Gastrulation. *Dev Biol* **346**, 25–38.
Eich, C., Arlt, J., Vink, C. S., Kartalaei, P. S., Kaimakis, P., Mariani, S. A., Linden, R. van der, Cappellen, W. A. van and Dzierzak, E. (2018). In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. *Journal of Experimental Medicine* **215**, 233–248.
Eichmann, A., Corbel, C., Nataf, V., Vaigot, P., Breant, C. and Le Douarin, N. M. (1997). Ligand-dependent development of the endothelial and hemopoietic lineages from embryonic mesodermal

cells expressing vascular endothelial growth factor receptor 2. *Proceedings of the National Academy of Sciences* **94**, 5141–5146.
Eichmann, A., Marcelle, C., Bréant, C. and Le Douarin, N. M. (1993). Two molecules related to the VEGF receptor are expressed in early endothelial cells during avian embryonic development. *Mechanisms of Development* **42**, 33–48.
Eilken, H. M., Nishikawa, S.-I. and Schroeder, T. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457**, 896–900.
Elcheva, I., Brok-Volchanskaya, V., Kumar, A., Liu, P., Lee, J.-H., Tong, L., Vodyanik, M., Swanson, S., Stewart, R., Kyba, M., et al. (2014). Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. *Nat Commun* **5**, 4372.
Elefanty, A. G., Begley, C. G., Hartley, L., Papaevangeliou, B. and Robb, L. (1999). SCL Expression in the Mouse Embryo Detected With a Targeted lacZ Reporter Gene Demonstrates Its Localization to Hematopoietic, Vascular, and Neural Tissues. 11.
Eliades, A., Wareing, S., Marinopoulou, E., Fadlullah, M. Z. H., Patel, R., Grabarek, J. B., Plusa, B., Lacaud, G. and Kouskoff, V. (2016). The Hemogenic Competence of Endothelial Progenitors Is Restricted by Runx1 Silencing during Embryonic Development. *Cell Reports* **15**, 2185–2199.
Ema, H. and Nakauchi, H. (2000). Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. **95**, 6.
Ema, H., Douagi, I., Cumano, A. and Kourilsky, P. (1998). Development of T cell precursor activity in the murine fetal liver. *Eur. J. Immunol.* 7.
Ema, M. (2003). Combinatorial effects of Flk1 and Tal1 on vascular and hematopoietic development in the mouse. *Genes & Development* **17**, 380–393.
Ema, M., Takahashi, S. and Rossant, J. (2006). Deletion of the selection cassette, but not *cis* -acting elements, in targeted *Flk1-lacZ* allele reveals *Flk1* expression in multipotent mesodermal progenitors. *Blood* **107**, 111–117.
Endoh, M., Ogawa, M., Orkin, S. and Nishikawa, S. (2002). SCL/tal-1-dependent process determines a competence to select the definitive hematopoietic lineage prior to endothelial differentiation. *The EMBO Journal* **21**, 6700–6708.
Engel, P., Eck, M. J. and Terhorst, C. (2003). The SAP and SLAM families in immune responses and X-linked lymphoproliferative disease. *Nature Reviews Immunology* **3**, 813.
Espín-Palazón, R., Stachura, D. L., Campbell, C. A., García-Moreno, D., Cid, N. D., Kim, A. D., Candel, S., Meseguer, J., Mulero, V. and Traver, D. (2014). Proinflammatory signaling regulates hematopoietic stem cell emergence. *Cell* **159**, 1070–1085.
F. M., Leonard, E., Michael, S., Orschell-Traycoff, C. M., Yoder, M. C. and Srour, E. F. (2002) 'Roles of spleen and liver in development of the murine hematopoietic system', Exp Hematol 30(9): 1010-9.
Faloon, P., Arentson, E., Kazarov, A., Deng, C. X., Porcher, C., Orkin, S. and Choi, K. (2000). Basic fibroblast growth factor positively regulates hematopoietic development. *Development* **127**, 1931–1941.
Federici, H. (1926) Recherches experimentales sur les potentialités de l'îlot sanguin chez l'embryon Rana fusca. Arch Biol 36, 466-88.
Fehling, H. J. (2003). Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* **130**, 4217–4227.
Ferkowicz, M. J. (2003). CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development* **130**, 4393–4403.
Fitch, S. R., Kimber, G. M., Wilson, N. K., Parker, A., Mirshekar-Syahkal, B., Göttgens, B., Medvinsky, A., Dzierzak, E. and Ottersbach, K. (2012). Signaling from the Sympathetic Nervous System Regulates Hematopoietic Stem Cell Emergence during Embryogenesis. *Cell Stem Cell* **11**, 554–566.
Fontaine-Perus, J. and Jarno, V. (1995). Mouse chick chimera: a new model to study the in ovo developmental potentialities of mammalian somites. 14.
Fontaine-Perus, J. C., Calman, F. M., Kaplan, C. and Le, N. M. (1981). Seeding of the 10-day mouse embryo thymic rudiment by lymphocyte precursors in vitro. 8.

Frame, J. M., Fegan, K. H., Conway, S. J., McGrath, K. E. and Palis, J. (2016). Definitive Hematopoiesis in the Yolk Sac Emerges from Wnt-Responsive Hemogenic Endothelium Independently of Circulation and Arterial Identity: EMPs Arise from Arteriovenous Hemogenic Endothelia. *Stem Cells* **34**, 431–444.
Fraser, S. T., Ogawa, M., Yu, R. T., Nishikawa, S., Yoder, M. C. and Nishikawa, S.-I. (2002). Definitive hematopoietic commitment within the embryonic vascular endothelial-cadherinƨ population. *Experimental Hematology* 9.
Frelin, C., Herrington, R., Janmohamed, S., Barbara, M., Tran, G., Paige, C. J., Benveniste, P., Zuñiga-Pflücker, J.-C., Souabni, A., Busslinger, M., et al. (2013). GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. *Nature Immunology* **14**, 1037–1044.
Fujiwara, Y., Browne, C. P., Cunniff, K., Goff, S. C. and Orkin, S. H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. Proc Natl Acad Sci U S A 93, 12355-8.
Fujiwara, Y., Chang, A. N., Williams, A. M. and Orkin, S. H. (2004). Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. Blood 103, 583-5.
Fujiwara, Y., Chang, A. N., Williams, A. M. and Orkin, S. H. (2004). Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. Blood 103, 583-5.
Gama-Norton, L., Ferrando, E., Ruiz-Herguido, C., Liu, Z., Guiu, J., Islam, A. B. M. M. K., Lee, S.-U., Yan, M., Guidos, C. J., López-Bigas, N., et al. (2015). Notch signal strength controls cell fate in the haemogenic endothelium. *Nat Commun* **6**, 8510.
Gao, L., Tober, J., Gao, P., Chen, C., Tan, K. and Speck, N. A. (2018). RUNX1 and the endothelial origin of blood. *Experimental Hematology* **68**, 2–9.
Gao, X., Johnson, K. D., Chang, Y.-I., Boyer, M. E., Dewey, C. N., Zhang, J. and Bresnick, E. H. (2013). Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J. Exp. Med.* **210**, 2833–2842.
Gekas, C., Dieterlen-Lièvre, F., Orkin, S. H. and Mikkola, H. K. A. (2005). The Placenta Is a Niche for Hematopoietic Stem Cells. *Developmental Cell* **8**, 365–375.
Gering, M. and Patient, R. (2005). Hedgehog signaling is required for adult blood stem cell formation in zebrafish embryos. Dev Cell 8, 389-400.
Gilmour, J., Assi, S. A., Noailles, L., Lichtinger, M., Obier, N. and Bonifer, C. (2018). The Co-operation of RUNX1 with LDB1, CDK9 and BRD4 Drives Transcription Factor Complex Relocation During Haematopoietic Specification. *Sci Rep* **8**, 10410.
Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M. F., Conway, S. J., Ng, L. G., Stanley, E. R., et al. (2010). Fate Mapping Analysis Reveals That Adult Microglia Derive from Primitive Macrophages. *Science* **330**, 841–845.
Giroux, S., Kaushik, A. L., Capron, C., Jalil, A., Kelaidi, C., Sablitzky, F., Dumenil, D., Albagli, O. and Godin, I. (2007). lyl-1 and tal-1/scl, two genes encoding closely related bHLH transcription factors, display highly overlapping expression patterns during cardiovascular and hematopoietic ontogeny. Gene Expr Patterns 7, 215-26.
Godin, I. and Cumano, A. (2002). The hare and the tortoise: an embryonic haematopoietic race. *Nat Rev Immunol* **2**, 593–604.
Godin, I. E., Garcia-Porrero, J. A., Coutinho, A., Dieterlen-Lievre, F. and Marcos, M. A. (1993). Para aortic splanchnopleura from early mouse embryos contains B1a cell progenitors. Nature 364, 67-70.
Godin, I., Dieterlen-Lievre, F. and Cumano, A. (1995). Emergence of multipotent hemopoietic cells in the yolk sac and paraaortic splanchnopleura in mouse embryos, beginning at 8.5 days postcoitus. *Proceedings of the National Academy of Sciences* **92**, 773–777.
Godin, I., Garcia-Porrero, J. A., Dieterlen-Lièvre, F. and Cumano, A. (1999). Stem Cell Emergence and Hemopoietic Activity Are Incompatible in Mouse Intraembryonic Sites. *J Exp Med* **190**, 43–52.
Gordon, M. Y. (2002). Of mice and men ... and elephants. *Blood* **100**, 4679–4679.
Goyama, S., Yamaguchi, Y., Imai, Y., Kawazu, M., Nakagawa, M., Asai, T., Kumano, K., Mitani, K., Ogawa, S., Chiba, S., et al. (2004). The transcriptionally active form of AML1 is required for

hematopoietic rescue of the AML1-deficient embryonic para-aortic splanchnopleural (P-Sp) region. *Blood* **104**, 3558–3564.
Grimes, H. L., Chan, T. O., Zweidler-McKay, P. A., Tong, B. and Tsichlis, P. N. (1996). The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. *Molecular and Cellular Biology* **16**, 6263–6272.
Gritz, E. and Hirschi, K. K. (2016). Specification and function of hemogenic endothelium during embryogenesis. *Cell. Mol. Life Sci.* **73**, 1547–1567.
Growney, J. D., Shigematsu, H., Li, Z., Lee, B. H., Adelsperger, J., Rowan, R., Curley, D. P., Kutok, J. L., Akashi, K., Williams, I. R. et al. (2005) 'Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype', Blood 106(2): 494- 504.
Guiu, J., Bergen, D. J. M., De Pater, E., Islam, A. B. M. M. K., Ayllón, V., Gama-Norton, L., Ruiz-Herguido, C., González, J., López-Bigas, N., Menendez, P., et al. (2014). Identification of Cdca7 as a novel Notch transcriptional target involved in hematopoietic stem cell emergence. *J Exp Med* **211**, 2411–2423.
Gunji, Y., Sudo, T., Suda, J., Yamaguchi, Y., Nakauchi, H., Nishikawa, S., Yanai, N., Obinata, M., Yanagisawa, M., Miura, Y., et al. (1991). Support of Early B-Cell Differentiation in Mouse Fetal Liver by Stromal Cells and Interleukin-7. 7.
Hamidi, S. and Sheng, G. (2018). Epithelial-mesenchymal transition in haematopoietic stem cell development and homeostasis. *J Biochem* **164**, 265–275.
Hann, I. M., Bodger, M. P. and Hoffbrand, A. V. (1983) 'Development of pluripotent hematopoietic progenitor cells in the human fetus', Blood 62(1): 118-23.
Hao, Q.-L. (2001). Identification of a novel, human multilymphoid progenitor in cord blood. *Blood* **97**, 3683–3690.
Haynes, B. F., Martin, M. E., Kay, H. H. and Kurtzberg, J. (1988). Early events in human T cell ontogeny. Phenotypic characterization and immunohistologic localization of T cell precursors in early human fetal tissues. *Journal of Experimental Medicine* **168**, 1061–1080.
Hayward, C. P. M., Hassell, J. A., Denomme, G. A., Rachubinski, R. A., Brown, C. and Kelton, J. G. (1995). The cDNA Sequence of Human Endothelial Cell Multimerin A UNIQUE PROTEIN WITH RGDS, COILED-COIL, AND EPIDERMAL GROWTH FACTOR-LIKE DOMAINS AND A CARBOXYL TERMINUS SIMILAR TO THE GLOBULAR DOMAIN OF COMPLEMENT C1q AND COLLAGENS TYPE VIII AND X. *J. Biol. Chem.* **270**, 18246–18251.
He, Q., Zhang, C., Wang, L., Zhang, P., Ma, D., Lv, J. and Liu, F. (2015). Inflammatory signaling regulates hematopoietic stem and progenitor cell emergence in vertebrates. *Blood* **125**, 1098–1106.
Henninger, J., Santoso, B., Hans, S., Durand, E., Moore, J., Mosimann, C., Brand, M., Traver, D. and Zon, L. (2017). Clonal fate mapping quantifies the number of haematopoietic stem cells that arise during development. *Nature Cell Biology* **19**, 17–27.
Herbomel, P., Thisse, B. and Thisse, C. (1999). Ontogeny and behaviour of early macrophages in zebrafish. 11.
Hodgson, G. S. and Bradley, T. R. (1979). Properties of haematopoietic stem cells surviving 5-fluorouracil treatment: evidence for a pre-CFU-S cell? *Nature* **281**, 381–382.
Hollyfield, J. G. (1966). The origin of erythroblasts in Rana pipiens tadpoles. Dev Biol 14, 461-80.
Hoogenkamp, M., Lichtinger, M., Krysinska, H., Lancrin, C., Clarke, D., Williamson, A., Mazzarella, L., Ingram, R., Jorgensen, H., Fisher, A., et al. (2009). Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. *Blood* **114**, 299–309.
Houssaint, E. (1981). Differentiation of the mouse hepatic primordium. II. Extrinsic origin of the haemopoietic cell line. Cell Differ 10, 243-52.
Huber, T. L., Kouskoff, V., Joerg Fehling, H., Palis, J. and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* **432**, 625–630.
Huyhn, A., Dommergues, M., Izac, B., Croisille, L., Katz, A., Vainchenker, W. and Coulombel, L. (1995). Characterization of Hematopoietic Progenitors From Human Yolk Sacs and Embryos. 13.

Iacovino, M., Chong, D., Szatmari, I., Hartweck, L., Rux, D., Caprioli, A., Cleaver, O. and Kyba, M. (2011). HoxA3 is an apical regulator of haemogenic endothelium. *Nature Cell Biology* **13**, 72–78.
Imai, K., Kobayashi, M., Wang, J., Shinobu, N., Yoshida, H., Hamada, J., Shindo, M., Higashino, F., Tanaka, J., Asaka, M., et al. (1999). Selective secretion of chemoattractants for haemopoietic progenitor cells by bone marrow endothelial cells: a possible role in homing of haemopoietic progenitor cells to bone marrow. *Br J Haematol* **106**, 905–911.
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. *Journal of Experimental Medicine* **208**, 2417–2427.
Jacobsen, S. E. W. and Nerlov, C. (2019). Haematopoiesis in the era of advanced single-cell technologies. *Nature Cell Biology* **21**, 2–8.
Jacobson, L. O., Simmons, E. L., Marks, E. K. and Eldredge, J. H. (1951). Recovery from radiation injury. Science 113, 510-11.
Jaffredo, T., Bollerot, K., Sugiyama, D., Gautier, R. and Drevon, C. (2005). Tracing the hemangioblast during embryogenesis: developmental relationships between endothelial and hematopoietic cells. *Int. J. Dev. Biol.* **49**, 269–277.
Jaffredo, T., Gautier, R., Brajeul, V. and Dieterlen-Lièvre, F. (2000). Tracing the progeny of the aortic hemangioblast in the avian embryo. *Dev. Biol.* **224**, 204–214.
Jaffredo, T., Gautier, R., Eichmann, A. and Dieterlen-Lièvre, F. (1998). Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125**, 4575–4583.
Johnson, G. R. and Metcalf, D. (1977). Pure and mixed erythroid colony formation in vitro stimulated by spleen conditioned medium with no detectable erythropoietin. *Proceedings of the National Academy of Sciences* **74**, 3879–3882.
Johnson, G. R. and Moore, M. A. (1975). Role of stem cell migration in initiation of mouse foetal liver haemopoiesis. Nature 258, 726-8.
Jones, R. J., Wagner, J. E., Celano, P., Zicha, M. S. and Sharkis, S. J. (1990). Separation of pluripotent haematopoietic stem cells from spleen colony-forming cells. *Nature* **347**, 188–189.
Joosten, M., Valk, P. J. M., Jordà, M. A., Vankan-Berkhoudt, Y., Verbakel, S., Broek, M. V. D., Beijen, A., Löwenberg, B. and Delwel, R. (2002). Leukemic predisposition of pSca-1/Cb2 transgenic mice. *Experimental Hematology* **30**, 142–149.
Jordan, H. E. (1916). Evidence of hemogenic capacity of endothelium. *The Anatomical Record* **10**, 417–420.
Kalev-Zylinska, M. L. (2002). Runx1 in zebrafish hematopoiesis. 16.
Kang, H., Mesquitta, W.-T., Jung, H. S., Moskvin, O. V., Thomson, J. A. and Slukvin, I. I. (2018). GATA2 Is Dispensable for Specification of Hemogenic Endothelium but Promotes Endothelial-to-Hematopoietic Transition. *Stem Cell Reports* **11**, 197–211.
Kau, C. L. and Turpen, J. B. (1983). Dual contribution of embryonic ventral blood island and dorsal lateral plate mesoderm during ontogeny of hemopoietic cells in Xenopus laevis. J. Immunol. 131, 2262-2266
Kawabata, K., Ujikawa, M., Egawa, T., Kawamoto, H., Tachibana, K., Iizasa, H., Katsura, Y., Kishimoto, T. and Nagasawa, T. (1999). A cell-autonomous requirement for CXCR4 in long-term lymphoid and myeloid reconstitution. *Proceedings of the National Academy of Sciences* **96**, 5663–5667.
Kennedy, M., Awong, G., Sturgeon, C. M., Ditadi, A., LaMotte-Mohs, R., Zúñiga-Pflücker, J. C. and Keller, G. (2012). T Lymphocyte Potential Marks the Emergence of Definitive Hematopoietic Progenitors in Human Pluripotent Stem Cell Differentiation Cultures. *Cell Reports* **2**, 1722–1735.
Kennedy, M., D'Souza, S. L., Lynch-Kattman, M., Schwantz, S. and Keller, G. (2007). Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. *Blood* **109**, 2679–2687.

Kidd, S., Kelley, M. R. and Young, M. W. (1986). Sequence of the notch locus of Drosophila melanogaster: relationship of the encoded protein to mammalian clotting and growth factors. *Mol. Cell. Biol.* **6**, 3094–3108.
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**, 1109–1121.
Kinder, S. J., Tsang, T. E., Quinlan, G. A., Hadjantonakis, A. K., Nagy, A. and Tam, P. P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. Development 126, 4691-701.
Kissa, K. and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* **464**, 112–115.
Kondo, M., Weissman, I. L. and Akashi, K. (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell* **91**, 661–672.
Kumano, K., Chiba, S., Kunisato, A., Sata, M., Saito, T., Nakagami-Yamaguchi, E., Yamaguchi, T., Masuda, S., Shimizu, K., Takahashi, T., et al. (2003). Notch1 but Not Notch2 Is Essential for Generating Hematopoietic Stem Cells from Endothelial Cells. *Immunity* **18**, 699–711.
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): 4891-9.
Kurata, H., Mancini, G. C., Alespeiti, G., Migliaccio, A. R. and Migliaccio, G. (1998). Stem cell factor induces proliferation and differentiation of fetal progenitor cells in the mouse. *British Journal of Haematology* **101**, 676–687.
La Motte-Mohs, R. N. (2005). Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. *Blood* **105**, 1431–1439.
Labastie, M.-C. (1998). Molecular Identity of Hematopoietic Precursor Cells Emerging in the Human Embryo. 13.
Lacaud, G., Gore, L., Kennedy, M., Kouskoff, V., Kingsley, P., Hogan, C., Carlsson, L., Speck, N., Palis, J. and Keller, G. (2002) 'Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development in vitro', Blood 100(2): 458-66.
Lam, E. Y. N., Hall, C. J., Crosier, P. S., Crosier, K. E. and Flores, M. V. (2010). Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. *Blood* **116**, 909–914.
Lam, E. Y., Hall, C. J., Crosier, P. S., Crosier, K. E. and Flores, M. V. (2010) 'Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells', Blood 116(6): 909-14.
Lampisuo, M., Liippo, J., Vainio, O., McNagny, K. M., Kulmala, J. and Lassila, O. (1999). Characterization of prethymic progenitors within the chicken embryo. *International Immunology* **11**, 63–69.
Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N. K., Moroy, T., Bonifer, C., et al. (2012). GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood* **120**, 314–322.
Lancrin, C., Sroczynska, P., Serrano, A. G., Gandillet, A., Ferreras, C., Kouskoff, V. and Lacaud, G. (2010). Blood cell generation from the hemangioblast. *J Mol Med* **88**, 167–172.
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**, 892–895.
Lane, M. C. and Sheets, M. D. (2002). Primitive and Definitive Blood Share a Common Origin in Xenopus: A Comparison of Lineage Techniques Used to Construct Fate Maps. *Developmental Biology* **248**, 52–67.
Lane, M. C. and Smith, W. C. (1999). Primitive blood origins. 12.

Langfelder, P. and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559.
Lassila, O., Eskola, J., Toivanen, P., Martin, C. and Dieterlen-Lievre, F. (1978). The origin of lymphoid stem cells studied in chick yolk sac–embryo chimaeras. *Nature* **272**, 353–354.
Lempereur, A., Canto, P. Y., Richard, C., Martin, S., Thalgott, J., Raymond, K., Lebrin, F., Drevon, C. and Jaffredo, T. (2018). The TGFβ pathway is a key player for the endothelial-to-hematopoietic transition in the embryonic aorta. *Dev. Biol.* **434**, 292–303.
Lessard, J., Faubert, A. and Sauvageau, G. (2004). Genetic programs regulating HSC specification, maintenance and expansion. *Oncogene* **23**, 7199–7209.
Li D., Xue W., Li M., Dong M., Wang J., Wang X., et al. (2018). VCAM-1(+) macrophages guide the homing of HSPCs to a vascular niche. Nature 564 119–124. 10.1038/s41586-018-0709-7
Li, W., Ferkowicz, M. J., Johnson, S. A., Shelley, W. C. and Yoder, M. C. (2005). Endothelial cells in the early murine yolk sac give rise to CD41-expressing hematopoietic cells. *Stem Cells Dev.* **14**, 44–54.
Li, Y., Esain, V., Teng, L., Xu, J., Kwan, W., Frost, I. M., Yzaguirre, A. D., Cai, X., Cortes, M., Maijenburg, M. W., et al. (2014). Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev.* **28**, 2597–2612.
Liakhovitskaia, A., Gribi, R., Stamateris, E., Villain, G., Jaffredo, T., Wilkie, R., Gilchrist, D., Yang, J., Ure, J. and Medvinsky, A. (2009). Restoration of Runx1 expression in the Tie2 cell compartment rescues definitive hematopoietic stem cells and extends life of Runx1 knockout animals until birth. *Stem Cells* **27**, 1616–1624.
Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S. A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M. S., Wu, M., et al. (2012). RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis: RUNX1 shifts transcription factor binding patterns. *The EMBO Journal* **31**, 4318–4333.
Lie-A-Ling, M., Marinopoulou, E., Lilly, A. J., Challinor, M., Patel, R., Lancrin, C., Kouskoff, V. and Lacaud, G. (2018). Regulation of RUNX1 dosage is crucial for efficient blood formation from hemogenic endothelium. *Development* **145**, dev149419.
Lieschke, G. J., Oates, A. C., Paw, B. H., Thompson, M. A., Hall, N. E., Ward, A. C., Ho, R. K., Zon, L. I. and Layton, J. E. (2002). Zebrafish SPI-1 (PU.1) Marks a Site of Myeloid Development Independent of Primitive Erythropoiesis: Implications for Axial Patterning. *Developmental Biology* **246**, 274–295.
Lilly, A. J., Mazan, A., Scott, D. A., Lacaud, G. and Kouskoff, V. (2017). SOX7 expression is critically required in FLK1-expressing cells for vasculogenesis and angiogenesis during mouse embryonic development. *Mech Dev* **146**, 31–41.
Ling, K.-W., Ottersbach, K., van Hamburg, J. P., Oziemlak, A., Tsai, F.-Y., Orkin, S. H., Ploemacher, R., Hendriks, R. W. and Dzierzak, E. (2004). GATA-2 Plays Two Functionally Distinct Roles during the Ontogeny of Hematopoietic Stem Cells. *J Exp Med* **200**, 871–882.
Lis, R., Karrasch, C. C., Poulos, M. G., Kunar, B., Redmond, D., Duran, J. G. B., Badwe, C. R., Schachterle, W., Ginsberg, M., Xiang, J., et al. (2017). Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature* **545**, 439–445.
Lizama, C. O., Hawkins, J. S., Schmitt, C. E., Bos, F. L., Zape, J. P., Cautivo, K. M., Borges Pinto, H., Rhyner, A. M., Yu, H., Donohoe, M. E., et al. (2015). Repression of arterial genes in hemogenic endothelium is sufficient for haematopoietic fate acquisition. *Nature Communications* **6**, 7739.
Loges, S., Fehse, B., Brockmann, M. A., Lamszus, K., Butzal, M., Guckenbiehl, M., Schuch, G., Ergün, S., Fischer, U., Zander, A. R., et al. (2004). Identification of the Adult Human Hemangioblast. *Stem Cells and Development* **13**, 229–242.
Lorenz, E., Uphoff, D., Reid, T. R. and Shelton, E. (1951). Modification of irradiation injury in mice and guinea pigs by bone marrow injections. J Natl Cancer Inst 12, 197-201.
Lu SJ, Feng Q, Park JS, Lanza R. (2010) 'Directed differentiation of red blood cells from human embryonic stem cells'. Methods Mol Biol 636: 105-21.

Lu, S.-J., Feng, Q., Caballero, S., Chen, Y., Moore, M. A. S., Grant, M. B. and Lanza, R. (2007). Generation of functional hemangioblasts from human embryonic stem cells. *Nature Methods* **4**, 501–509.
Luckett, W. P. (1978) 'Origin and differentiation of the yolk sac and extraembryonic mesoderm in presomite human and rhesus monkey embryos', Am J Anat 152(1): 59-97.
Lugus, J. J., Park, C., Ma, Y. D. and Choi, K. (2009). Both primitive and definitive blood cells are derived from Flk-1+ mesoderm. *Blood* **113**, 563–566.
Ma, Q., Jones, D. and Springer, T. A. (1999). The Chemokine Receptor CXCR4 Is Required for the Retention of B Lineage and Granulocytic Precursors within the Bone Marrow Microenvironment. *Immunity* **10**, 463–471.
Maetzig, T., Ruschmann, J., Lai, C. K., Ngom, M., Imren, S., Rosten, P., Norddahl, G. L., von Krosigk, N., Sanchez Milde, L., May, C., et al. (2017). A Lentiviral Fluorescent Genetic Barcoding System for Flow Cytometry-Based Multiplex Tracking. *Molecular Therapy* **25**, 606–620.
Manaia, A. (2000). Lmo2 and GATA-3 hemopoietic expression. 11.
Mandal, L., Banerjee, U. and Hartenstein, V. (2004). Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. *Nature Genetics* **36**, 1019.
Manz, M. G. (2001). Dendritic cell potentials of early lymphoid and myeloid progenitors. *Blood* **97**, 3333–3341.
Manz, M. G., Traver, D., Akashi, K., Merad, M., Miyamoto, T., Engleman, E. G. and Weissman, I. L. (2001). Dendritic Cell Development from Common Myeloid Progenitors. 8.
Marcelo, K. L., Goldie, L. C. and Hirschi, K. K. (2013). Regulation of Endothelial Cell Differentiation and Specification. *Circ Res* **112**, 1272–1287.
Marshall, C. J., Moore, R. L., Thorogood, P., Brickell, P. M., Kinnon, C. and Thrasher, A. J. (1999). Detailed characterization of the human aorta-gonad-mesonephros region reveals morphological polarity resembling a hematopoietic stromal layer. *Developmental Dynamics* **215**, 139–147.
MARTIN, C. (1972). Technique d'explantation in ovo de blastodermes d'embryons d'oiseaux. C. R. Seances Soc. Biol. (Paris) 166: 283-285.
MARTIN, C., BEAUPAIN, D. and DIETERLEN-LIEVRE, F. (1978). Developmental relationships between vitelline and intra-embryonic haemopoiesis studied in avian 'yolk sac chimaeras'. Cell Differ 7: 115 30.
Mathieson, B. J. and Fowlkes, B. J. (1984). Cell Surface Antigen Expression on Thymocytes: Development and Phenotypic Differentiation of Intrathymic Subsets. *Immunol Rev* **82**, 141–173.
Maximow, A. (1909) 'Untersuchungen uber Blut und Bindegewebe. I. Die Fruhesten Entwicklongsstadien des Blut- und Bindewegebezellen beim Saugetierembryo, bis zum Anfang des Blutbildung in der Leber', Arch. Mikr. Anat 73(444).
McCulloch, E. A. and Till, J. E. (1960). The Radiation Sensitivity of Normal Mouse Bone Marrow Cells, Determined by Quantitative Marrow Transplantation into Irradiated Mice. *Radiation Research* **13**, 115.
McGrath, K. E., Frame, J. M. and Palis, J. (2015). Early hematopoiesis and macrophage development. *Seminars in Immunology* **27**, 379–387.
Medvinsky, A. and Dzierzak, E. (1996). Definitive Hematopoiesis Is Autonomously Initiated by the AGM Region. *Cell* **86**, 897–906.
Medvinsky, A. L., 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**, 64–67.
Melchers, F. (1979). Murine embryonic B lymphocyte development in the placenta. *Nature* **277**, 219–221.
Menzel, O., Vellai, T., Takacs-Vellai, K., Reymond, A., Mueller, F., Antonarakis, S. E. and Guipponi, M. (2004). The Caenorhabditis elegans ortholog of C21orf80, a potential new protein O-fucosyltransferase, is required for normal development. *Genomics* **84**, 320–330.

Micklem, H. S., Ford, C. E., Evans, E. P., Ogden, D. A. and Papworth, D. S. (1972). Competitivein vivo proliferation of foetal and adult haematopoietic cells in lethally irradiated mice. *J. Cell. Physiol.* **79**, 293–298.
Mikkola, H. K. A. (2003). Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* **101**, 508–516.
Millauer, B. (1993). High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of vasculogenesis and angiogenesis. *Cell* **72**, 835–846.
Miller, J., Horner, A., Stacy, T., Lowrey, C., Lian, J. B., Stein, G., Nuckolls, G. H. and Speck, N. A. (2002). The core-binding factor β subunit is required for bone formation and hematopoietic maturation. *Nat Genet* **32**, 645–649.
Minko, K., Bollerot, K., Drevon, C., Hallais, M.-F. and Jaffredo, T. (2003). From mesoderm to blood islands: patterns of key molecules during yolk sac erythropoiesis. *Gene Expr. Patterns* **3**, 261–272.
Monteiro, R., Pinheiro, P., Joseph, N., Peterkin, T., Koth, J., Repapi, E., Bonkhofer, F., Kirmizitas, A. and Patient, R. (2016). Transforming Growth Factor β Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells. *Developmental Cell* **38**, 358–370.
Moore, M. A. and Owen, J. J. (1967). Experimental studies on the development of the thymus. J Exp Med 126, 715-26.
Moore, M. A. S. and Metcalf, D. (1970). Ontogeny of the Haemopoietic System: Yolk Sac Origin of In Vivo and In Vitro Colony Forming Cells in the Developing Mouse Embryo. *Br J Haematol* **18**, 279–296.
Morrison, S. J., Hemmati, H. D., Wandycz, A. M. and Weissman, I. L. (1995). The purification and characterization of fetal liver hematopoietic stem cells. *Proceedings of the National Academy of Sciences* **92**, 10302–10306.
Mukouyama, Y., 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**, 27–36.
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**, 963–975.
Murray (1932) 'The development in vitro of the blood of the early chick embryo', P Roy Soc
Murray, P. D. F. (1932). The Development in vitro of the Blood of the Early Chick Embryo. *Proceedings of the Royal Society B: Biological Sciences* **111**, 497–521.
Myers, C., Charboneau, A. and Boudreau, N. (2000). Homeobox B3 Promotes Capillary Morphogenesis and Angiogenesis. *J Cell Biol* **148**, 343–352.
Nagasawa, T., Hirota, S., Tachibana, K., Takakura, N., Nishikawa, S., Kitamura, Y., Yoshida, N., Kikutani, H. and Kishimoto, T. (1996). Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. *Nature* **382**, 635–638.
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**, 3329–3334.
Nakajima-Takagi, Y., Osawa, M., Oshima, M., Takagi, H., Miyagi, S., Endoh, M., Endo, T. A., Takayama, N., Eto, K., Toyoda, T., et al. (2013). Role of SOX17 in hematopoietic development from human embryonic stem cells. *Blood* **121**, 447–458.
Nakano, H., Liu, X., Arshi, A., Nakashima, Y., van Handel, B., Sasidharan, R., Harmon, A. W., Shin, J.-H., Schwartz, R. J., Conway, S. J., et al. (2013). Haemogenic endocardium contributes to transient definitive haematopoiesis. *Nat Commun* **4**, 1564.
Nishikawa, S.-I., Nishikawa, S., Hirashima, M., Matsuyoshi, N. and Kodama, H. (1998a). Progressive lineage analysis by cell sorting and culture identifies FLK1+VE- cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. 11.

Nishikawa, S.-I., Nishikawa, S., Kawamoto, H., Yoshida, H., Kizumoto, M., Kataoka, H. and Katsura, Y. (1998b). In Vitro Generation of Lymphohematopoietic Cells from Endothelial Cells Purified from Murine Embryos. *Immunity* **8**, 761–769.

North, T. E., Goessling, W., Peeters, M., Li, P., Ceol, C., Lord, A. M., Weber, G. J., Harris, J., Cutting, C. C., Huang, P., et al. (2009). Hematopoietic Stem Cell Development Is Dependent on Blood Flow. *Cell* **137**, 736–748.

North, T., Gu, T.-L., Stacy, T., Wang, Q., Howard, L., Binder, M., Marín-Padilla, M. and Speck, N. A. (1999). Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. 13.

Oatley, M., Bölükbasi, Ö. V., Svensson, V., Shvartsman, M., Ganter, K., Zirngibl, K., Pavlovich, P. V., Milchevskaya, V., Foteva, V., Natarajan, K. N., et al. (2018). *Single-cell transcriptomics identifies CD44 as a new marker and regulator of haematopoietic stem cells development*. Developmental Biology.

Ogawa, M. (1993). Differentiation and proliferation of hematopoietic stem cells. *Blood* **81**, 2844–2853.

Ogawa, M., Kizumoto, M., Nishikawa, S., Fujimoto, T., Kodama, H. and Nishikawa, S. I. (1999).Expression of alpha4-integrin defines the earliest precursor of hematopoietic cell lineage diverged from endothelial cells. Blood 93, 1168-77

Okada, S., Nakauchi, H., Nagayoshi, K., Nishikawa, S., Miura, Y. and Suda, T. In Vivo and In Vitro Stem Cell Function of c-kit- and Sca-1-Positive Murine Hematopoietic Cells. 8.

Ottersbach, K. (2019). Endothelial-to-haematopoietic transition: an update on the process of making blood. *Biochemical Society Transactions* BST20180320.

Ottersbach, K. and Dzierzak, E. (2005). The Murine Placenta Contains Hematopoietic Stem Cells within the Vascular Labyrinth Region. *Developmental Cell* **8**, 377–387.

Owen, J. J. and Ritter, M. A. (1969) 'Tissue interaction in the development of thymus lymphocytes', J Exp Med 129(2): 431-42.

Padrón-Barthe, L., Temiño, S., Villa del Campo, C., Carramolino, L., Isern, J. and Torres, M. (2014). Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo. *Blood* **124**, 2523–2532.

Paige, C. J., Gisler, R. H., McKearn, J. P. and Iscove, N. N. (1984). Differentiation of murine B cell precursors in agar culture. Frequency, surface marker analysis and requirements for growth of clonable pre-B cells. *Eur. J. Immunol.* **14**, 979–987.

Palis, J. (1999a). Development of hematopoietic progenitors. 12.

Palis, J. (1999b). Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. 12.

Palis, J. (2004). No red cell is an island. *Nature* **432**, 964–965.

Palis, J. and Yoder, M. C. (2001). Yolk-sac hematopoiesis: The first blood cells of mouse and man. *Experimental Hematology* **29**, 927–936.

Palis, J., Chan, R. J., Koniski, A., Patel, R., Starr, M. and Yoder, M. C. (2001) 'Spatial and temporal emergence of high proliferative potential hematopoietic precursors during murine embryogenesis', Proc Natl Acad Sci U S A 98(8): 4528-33.

Pandolfi, P. P., Roth, M. E., Karis, A., Leonard, M. W., Dzierzak, E., Grosveld, F. G., Engel, J. D. and Lindenbaum, M. H. (1995) 'Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis', Nat Genet 11(1): 40-4.

Pardanaud, L. and Dieterlen-Lievre, F. (1993). Emergence of endothelial and hemopoietic cells in the avian embryo. Anat Embryol (Berl) 187, 107-14.

Pardanaud, L. and Dieterlen-Lievre, F. (1999). Manipulation of angiopoietc/hematopoietic potentials in the avian embryo. J Soc Biol 193, 171-9.

Pardanaud, L., Luton, D., Prigent, M., Bourcheix, L.-M., Catala, M. and Dieterlen-Lièvre, F. (1996). Two distinct endothelial lineages in ontogeny, one of them related to hemopoiesis. Development 122, 1363-1371.

Pardanaud, L., Yassine, F. and Dieterlen-Lievre, F. (1989). Relationship between vasculogenesis, angiogenesis and haemopoiesis during avian ontogeny. 13.

Patterson, L. J., Gering, M. and Patient, R. (2005) 'Scl is required for dorsal aorta as well as blood formation in zebrafish embryos', Blood 105(9): 3502-11.
Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G. and Kouskoff, V. (2015). In Vivo Repopulating Activity Emerges at the Onset of Hematopoietic Specification during Embryonic Stem Cell Differentiation. *Stem Cell Reports* **4**, 431–444.
Pearson, S., Lancrin, C., Lacaud, G. and Kouskoff, V. (2010). The Sequential Expression of CD40 and Icam2 Defines Progressive Steps in the Formation of Blood Precursors from the Mesoderm Germ Layer. *STEM CELLS* **28**, 1089–1098.
Peled, A., Grabovsky, V., Habler, L., Sandbank, J., Arenzana-Seisdedos, F., Petit, I., Ben-Hur, H., Lapidot, T. and Alon, R. (1999). The chemokine SDF-1 stimulates integrin-mediated arrest of CD34+ cells on vascular endothelium under shear flow. *J. Clin. Invest.* **104**, 1199–1211.
Peled, A., Kollet, O., Ponomaryov, T., Petit, I., Franitza, S., Grabovsky, V., Slav, M. M., Nagler, A., Lider, O., Alon, R., et al. (2000). The chemokine SDF-1 activates the integrins LFA-1, VLA-4, and VLA-5 on immature human CD34ɛ cells: role in transendothelial/stromal migration and engraftment of NOD/SCID mice. **95**, 8.
Pereira, C.-F., Chang, B., Gomes, A., Bernitz, J., Papatsenko, D., Niu, X., Swiers, G., Azzoni, E., de Bruijn, M. F. T. R., Schaniel, C., et al. (2016). Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells. *Developmental Cell* **36**, 525–539.
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., et al. (2013). Induction of a Hemogenic Program in Mouse Fibroblasts. *Cell Stem Cell* **13**, 205–218.
Petit-Cocault, L., Volle-Challier, C., Fleury, M., Peault, B. and Souyri, M. (2007). Dual role of Mpl receptor during the establishment of definitive hematopoiesis. *Development* **134**, 3031–3040.
Petrenko, O., Beavis, A., Klaine, M., Kittappa, R., Godin, I. and Lemischka, I. R. (1999). The Molecular Characterization of the Fetal Stem Cell Marker AA4. *Immunity* **10**, 691–700.
Plett, P. A., Frankovitz, S. M., Wolber, F. M., Abonour, R. and Orschell-Traycoff, C. M. (2002) 'Treatment of circulating CD34(+) cells with SDF-1alpha or anti-CXCR4 antibody enhances migration and NOD/SCID repopulating potential', Exp Hematol 30(9): 1061-9.
Ploemacher, R. E. and Brons, N. H. (1988) 'Isolation of hemopoietic stem cell subsets from murine bone marrow: II. Evidence for an early precursor of day-12 CFU-S and cells associated with radioprotective ability', Exp Hematol 16(1): 27-32.
Ploemacher, R. E., van Os, R., van Beurden, C. A. and Down, J. D. (1992) 'Murine haemopoietic stem cells with long-term engraftment and marrow repopulating ability are more resistant to gammaradiation than are spleen colony forming cells', Int J Radiat Biol 61(4): 489-99.
Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F. W. and Orkin, S. H. (1996). The T Cell Leukemia Oncoprotein SCL/tal-1 Is Essential for Development of All Hematopoietic Lineages. *Cell* **86**, 47–57.
Pouget, C., Gautier, R., Teillet, M.-A. and Jaffredo, T. (2006). Somite-derived cells replace ventral aortic hemangioblasts and provide aortic smooth muscle cells of the trunk. *Development* **133**, 1013–1022.
Pouget, C., Pottin, K. and Jaffredo, T. (2008). Sclerotomal origin of vascular smooth muscle cells and pericytes in the embryo. *Dev. Biol.* **315**, 437–447.
Psychoyos, D. and Stern, C. D. (1996). Fates and migratory routes of primitive streak cells in the chick embryo.Development 122, 1523-34.
Pui, J. C., Allman, D., Xu, L., DeRocco, S., Karnell, F. G., Bakkour, S., Lee, J. Y., Kadesch, T., Hardy, R. R., Aster, J. C., et al. (1999). Notch1 Expression in Early Lymphopoiesis Influences B versus T Lineage Determination. *Immunity* **11**, 299–308.
Radtke, F., Wilson, A., Stark, G., Bauer, M., van Meerwijk, J., MacDonald, H. R. and Aguet, M. (1999). Deficient T Cell Fate Specification in Mice with an Induced Inactivation of Notch1. *Immunity* **10**, 547–558.

Rafii, S., Kloss, C. C., Butler, J. M., Ginsberg, M., Gars, E., Lis, R., Zhan, Q., Josipovic, P., Ding, B.-S., Xiang, J., et al. (2013). Human ESC-derived hemogenic endothelial cells undergo distinct waves of endothelial to hematopoietic transition. *Blood* **121**, 770–780.

Rebel, V. I., Miller, C. L., Eaves, C. J. and Lansdorp, P. M. (1996a). The Repopulation Potential of Fetal Liver Hematopoietic Stem Cells in Mice Exceeds That of Their Adult Bone Marrow Counterparts. 9.

Rebel, V. I., Miller, C. L., Thornbury, G. R., Dragowska, W. H., Eaves, C. J. and Lansdorp, P. M. (1996b) 'A comparison of long-term repopulating hematopoietic stem cells in fetal liver and adult bone marrow from the mouse', Exp Hematol 24(5): 638-48.

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**, 252–263.

Richard, C., Drevon, C., Canto, P.-Y., Villain, G., Bollérot, K., Lempereur, A., Teillet, M.-A., Vincent, C., Rosselló Castillo, C., Torres, M., et al. (2013). Endothelio-mesenchymal interaction controls runx1 expression and modulates the notch pathway to initiate aortic hematopoiesis. *Dev. Cell* **24**, 600–611.

Robb, L., Elwood, N. J., Elefantyl, A. G., Kontgen', F., Li', R., Barnettl, L. D. and Begley', C. G. (1996). The sci gene product is required for the generation of all hematopoietic lineages in the adult mouse. 7.

Robb, L., Lyons, I., Metcalf, D. and Begley, C. G. (1995). Absence of yolk sac hematopoiesis from mice with a targeted disruption of the sci gene. *Proc. Natl. Acad. Sci. USA* 5.

Robert-Moreno, A., Espinosa, L., de la Pompa, J.L., and Bigas, A. (2005). RBPj{kappa}- dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. Development 132, 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., et al. (2008). Impaired embryonic haematopoiesis yet normal arterial development in the absence of the Notch ligand Jagged1. *EMBO J* **27**, 1886–1895.

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

Roch, A., Giger, S., Girotra, M., Campos, V., Vannini, N., Naveiras, O., Gobaa, S. and Lutolf, M. P. (2017). Single-cell analyses identify bioengineered niches for enhanced maintenance of hematopoietic stem cells. *Nat Commun* **8**, 221.

Royer-Pokora, B., Loos, U. and Ludwig, W. D. (1991). TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11;14)(p13;q11). Oncogene 6, 1887-93.

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

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

Sabin, F. R. (1920) 'Studies on the Origin of Bllod-Vessels and of Rde Blood-Corpuscules as seen in the Living Blastoderm of Chicks during the Second Day of incubation', Carnegie Institute, Washington

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**, 312–318.

Sanghez, V., Luzzi, A., Clarke, D., Kee, D., Beuder, S., Rux, D., Osawa, M., Madrenas, J., Chou, T.-F., Kyba, M., et al. (2017). Notch activation is required for downregulation of HoxA3-dependent endothelial cell phenotype during blood formation. *PLOS ONE* **12**, e0186818.

Schoenwolf, G. C., Garcia-Martinez, V. and Dias, M. S. (1992). Mesoderm movement and fate during avian gastrulation and neurulation. *Developmental Dynamics* **193**, 235–248.

Shahriyari, L. and Komarova, N. L. (2013). Symmetric vs. Asymmetric Stem Cell Divisions: An Adaptation against Cancer? *PLoS ONE* **8**, e76195.
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**, 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* **376**, 62.
She, Z.-Y. and Yang, W.-X. (2015). SOX family transcription factors involved in diverse cellular events during development. *European Journal of Cell Biology* **94**, 547–563.
Shook, D. and Keller, R. (2003). Mechanisms, mechanics and function of epithelial–mesenchymal transitions in early development. *Mechanisms of Development* **120**, 1351–1383.
Sidney, L. E., Branch, M. J., Dunphy, S. E., Dua, H. S. and Hopkinson, A. (2014). Concise Review: Evidence for CD34 as a Common Marker for Diverse Progenitors. *Stem Cells* **32**, 1380–1389.
Silver, L. and Palis, J. (1997). Initiation of murine embryonic erythropoiesis: a spatial analysis. Blood 89, 1154-64.
Sroczynska, P., Lancrin, C., Kouskoff, V. and Lacaud, G. (2009). The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. *Blood* **114**, 5279–5289.
Srour, E. F., Brandt, J. E., Grigsby, S., Leemhuis, T. and Hoffman, R. (1993). Long-Term Generation and Expansion of Human Primitive Hematopoietic Progenitor Cells In Vitro. 10.
Steiniger, B., Ulfig, N., Rieße, M. and Barth, P. J. (2007). Fetal and early post-natal development of the human spleen: from primordial arterial B cell lobules to a non-segmented organ. *Histochem Cell Biol* **128**, 205–215.
Sugimura, R., Jha, D. K., Han, A., Soria-Valles, C., da Rocha, E. L., Lu, Y.-F., Goettel, J. A., Serrao, E., Rowe, R. G., Malleshaiah, M., et al. (2017). Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature* **545**, 432–438.
Sugiyama, D. (2003). Erythropoiesis from acetyl LDL incorporating endothelial cells at the preliver stage. *Blood* **101**, 4733–4738.
Sugiyama, D., Ogawa, M., Nakao, K., Osumi, N., Nishikawa, S., Nishikawa, S., Arai, K., Nakahata, T. and Tsuji, K. (2007). B cell potential can be obtained from pre-circulatory yolk sac, but with low frequency. *Developmental Biology* **301**, 53–61.
Sutherland, H. J., Eaves, C. J., Eaves, A. C., Dragowska, W. and Lansdorp, P. M. (1989) 'Characterization and partial purification of human marrow cells capable of initiating long-term hematopoiesis in vitro', Blood 74(5): 1563-70.
Suzuki, T., Shimizu, Y., Furuhata, E., Maeda, S., Kishima, M., Nishimura, H., Enomoto, S., Hayashizaki, Y. and Suzuki, H. (2017). RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. *Blood Adv* **1**, 1699–1711.
Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K. D., et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nature Communications* **4**, 2924.
Takahashi, S., Shimizu, R., Suwabe, N., Kuroha, T., Yoh, K., Ohta, J., Nishimura, S., Lim, K. C., Engel, J. D. and Yamamoto, M. (2000). GATA factor transgenes under GATA-1 locus control rescue germline GATA-1 mutant deficiencies. Blood 96, 910-6.
Taniguchi, H., Toyoshima, T., Fukao, K. and Nakauchi, H. (1996) 'Presence of hematopoietic stem cells in the adult liver', Nat Med 2(2): 198-203.
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**, 9399–9403.
Tavian, 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. 7.

Tavian, M., Hallais, M. F. and Peault, B. (1999) 'Emergence of intraembryonic hematopoietic precursors in the pre-liver human embryo', Development 126(4): 793-803.
Tavian, M., Robin, C. and Coulombel, L. (2001). The Human Embryo, but Not Its Yolk Sac, Generates Lympho-Myeloid Stem Cells: Mapping Multipotent Hematopoietic Cell Fate in Intraembryonic Mesoderm. 9.
Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Möröy, T., et al. (2016). GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nature Cell Biology* **18**, 21–32.
Theodore, L. N., Hagedorn, E. J., Cortes, M., Natsuhara, K., Liu, S. Y., Perlin, J. R., Yang, S., Daily, M. L., Zon, L. I. and North, T. E. (2017). Distinct Roles for Matrix Metalloproteinases 2 and 9 in Embryonic Hematopoietic Stem Cell Emergence, Migration, and Niche Colonization. *Stem Cell Reports* **8**, 1226–1241.
Thompson, M. A., Ransom, D. G., Pratt, S. J., MacLennan, H., Kieran, M. W., Detrich, H. W., Vail, B., Huber, T. L., Paw, B., Brownlie, A. J., et al. (1998). TheclocheandspadetailGenes Differentially Affect Hematopoiesis and Vasculogenesis. *Developmental Biology* **197**, 248–269.
Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. and Jones, J. M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* **282**, 1145–1147.
Ting, C. N., Olson, M. C., Barton, K. P. and Leiden, J. M. (1996) 'Transcription factor GATA-3 is required for development of the T-cell lineage', Nature 384(6608): 474-8.
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**, 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**, 3765–3776.
Tompkins, R., Volpe, E. P. and Reinschmidt, D. (1980) Origin of hemopoietic stem cells in amphibian ontogeny. In : Developement and differentiation of Vertebrate Lymphocytes. J. D. Horton, ed. Amsterdam : Elsevier/north Holland, 25-34.
Tong, B., Grimes, H. L., Yang, T.-Y., Bear, S. E., Qin, Z., Du, K., El-Deiry, W. S. and Tsichlis, P. N. (1998). The Gfi-1B Proto-Oncoprotein Repressesp21WAF1 and Inhibits Myeloid Cell Differentiation. *Molecular and Cellular Biology* **18**, 2462–2473.
Traver, D., Akashi, K., Manz, M., Merad, M., Miyamoto, T., Engleman, E. G. and Weissman, I. L. (2000). Development of CD8 -Positive Dendritic Cells from a Common Myeloid Progenitor. *Science* **290**, 2152–2154.
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* **371**, 221–226.
Tsukada, M., Ota, Y., Wilkinson, A. C., Becker, H. J., Osato, M., Nakauchi, H. and Yamazaki, S. (2017). In Vivo Generation of Engraftable Murine Hematopoietic Stem Cells by Gfi1b, c-Fos , and Gata2 Overexpression within Teratoma. *Stem Cell Reports* **9**, 1024–1033.
Turpen, J. B. and Knudson, C. M. (1982). Ontogeny of hematopoietic cells in Rana pipiens: precursor cell migration during embryogenesis. Dev Biol 89, 138-51.
Turpen, J. B., Knudson, C. M. and Hoefen, P. S. (1981). The early ontogeny of hematopoietic cells studied by grafting cytogenetically labeled tissue anlagen: Localization of a prospective stem cell compartment. *Developmental Biology* **85**, 99–112.
Turpen, J. B., Marrion, R. M. and Williams, K. (1983). PERITONEAL EXUDATE IN LARVAL RANA PIPIENS CONTAINS CELLS THAT ARE EMBRYOLOGICALLY DERIVED FROM DORSAL LATERAL PLATE ~SODERM. **7**, 8.
Ueno, H. and Weissman, I. L. (2006). Clonal Analysis of Mouse Development Reveals a Polyclonal Origin for Yolk Sac Blood Islands. *Developmental Cell* **11**, 519–533.
Upadhaya, S., Reizis, B. and Sawai, C. M. (2018). New genetic tools for the in vivo study of hematopoietic stem cell function. *Experimental Hematology* **61**, 26–35.

Van Handel, B., Montel-Hagen, A., Sasidharan, R., Nakano, H., Ferrari, R., Boogerd, C.J., Schredelseker, J., Wang, Y., Hunter, S., Org, T. et al. (2012) 'Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium', Cell 150(3): 590-605.
Vellguth, S., von Gaudecker, B. and Müller-Hermelink, H. K. (1985). The development of the human spleen. Ultrastructural studies in fetuses from the 14th to 24th week of gestation. *Cell Tissue Res.* **242**, 579–592.
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 Dev.* **12**, 473–479.
Vodyanik, M. A., Bork, J. A., Thomson, J. A. and Slukvin, I. I. (2005). Human embryonic stem cell–derived CD34+ cells: efficient production in the coculture with OP9 stromal cells and analysis of lymphohematopoietic potential. *Blood* **105**, 617–626.
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**, 337.
Vosshenrich, C. A., Garcia-Ojeda, M. E., Samson-Villeger, S. I., Pasqualetto, V., Enault, L., Richard-Le Goff, O., Corcuff, E., Guy-Grand, D., Rocha, B., Cumano, A. et al. (2006) 'A thymic pathway of mouse natural killer cell development characterized by expression of GATA-3 and CD127', Nat Immunol 7(11): 1217-24.
Wadman, I. A., Osada, H., Grutz, G. G., Agulnick, A. D., Westphal, H., Forster, A. and Rabbitts, T. H. (1997). The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/ NLI proteins. Embo J 16, 3145-57.
Walmsley, M. E., Guille, M. J., Bertwistle, D., Smith, J. C., Pizzey, J. A. and Patient, R. K. (1994). Negative control of Xenopus GATA-2 by activin and noggin with eventual expression in precursors of the ventral blood islands. Development 120, 2519-29.
Wang, J. C. Y., Doedens, M. and Dick, J. E. (1997). Primitive Human Hematopoietic Cells Are Enriched in Cord Blood Compared With Adult Bone Marrow or Mobilized Peripheral Blood as Measured by the Quantitative In Vivo SCID-Repopulating Cell Assay. 7.
Wang, L., Li, L., Shojaei, F., Levac, K., Cerdan, C., Menendez, P., Martin, T., Rouleau, A. and Bhatia, M. (2004). Endothelial and Hematopoietic Cell Fate of Human Embryonic Stem Cells Originates from Primitive Endothelium with Hemangioblastic Properties. *Immunity* **21**, 31–41.
Wang, Q., Stacy, T., Binder, M., Marin-Padilla, M., Sharpe, A. H. and Speck, N. A. (1996). Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proceedings of the National Academy of Sciences* **93**, 3444–3449.
Wasteson, P., Johansson, B. R., Jukkola, T., Breuer, S., Akyurek, L. M., Partanen, J. and Lindahl, P. (2008). Developmental origin of smooth muscle cells in the descending aorta in mice. *Development* **135**, 1823–1832.
Weissman, I. L., Baird, S., Gardner, R. L., Papaioannou, V. E. and Raschke, W. (1977). Normal and Neoplastic Maturation of T-lineage Lymphocytes. *Cold Spring Harbor Symposia on Quantitative Biology* **41**, 9–21.
Weng, W., Sukowati, E. W. and Sheng, G. (2007). On Hemangioblasts in Chicken. *PLoS ONE* **2**, e1228.
Whitlock, C. A. and Witte, O. N. (1982). Long-term culture of B lymphocytes and their precursors from murine bone marrow. *Proceedings of the National Academy of Sciences* **79**, 3608–3612.
Wiegreffe, C., Christ, B., Huang, R. and Scaal, M. (2007). Sclerotomal origin of smooth muscle cells in the wall of the avian dorsal aorta. *Dev. Dyn.* **236**, 2578–2585.
Wiegreffe, C., Christ, B., Huang, R. and Scaal, M. (2009). Remodeling of aortic smooth muscle during avian embryonic development. *Dev. Dyn.* **238**, 624–631.
Wong, P. M., Chung, S. W., Chui, D. H. and Eaves, C. J. (1986). Properties of the earliest clonogenic hemopoietic precursors to appear in the developing murine yolk sac. *Proceedings of the National Academy of Sciences* **83**, 3851–3854.
Wright, D. E., Bowman, E. P., Wagers, A. J., Butcher, E. C. and Weissman, I. L. (2002). Hematopoietic Stem Cells Are Uniquely Selective in Their Migratory Response to Chemokines. *J Exp Med* **195**, 1145–1154.

Yamada, Y., Pannell, R., Forster, A. and Rabbitts, T. H. (2000). The oncogenic LIM-only transcription factor Lmo2 regulates angiogenesis but not vasculogenesis in mice. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 320–324.

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. 10.

Yang, L. (2005). Identification of Lin-Sca1+kit+CD34+Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* **105**, 2717–2723.

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

Yokomizo, T., Ogawa, M., Osato, M., Kanno, T., Yoshida, H., Fujimoto, T., Fraser, S., Nishikawa, S., Okada, H., Satake, M. et al. (2001) 'Requirement of Runx1/AML1/PEBP2alphaB for the generation of haematopoietic cells from endothelial cells', Genes Cells 6(1): 13-23.

Yokomizo, T., Ogawa, M., Osato, M., Kanno, T., Yoshida, H., Fujimoto, T., Fraser, S., Nishikawa, S., Okada, H., Satake, M., et al. (2001). Requirement of Runx1/AML1/PEBP2alphaB for the generation of haematopoietic cells from endothelial cells. *Genes Cells* **6**, 13–23.

Yoon, K.-J., Koo, B.-K., Im, S.-K., Jeong, H.-W., Ghim, J., Kwon, M., Moon, J.-S., Miyata, T. and Kong, Y.-Y. (2008). Mind Bomb 1-Expressing Intermediate Progenitors Generate Notch Signaling to Maintain Radial Glial Cells. *Neuron* **58**, 519–531.

Yoshimoto, M., Montecino-Rodriguez, E., Ferkowicz, M. J., Porayette, P., Shelley, W. C., Conway, S. J., Dorshkind, K. and Yoder, M. C. (2011). Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. *Proceedings of the National Academy of Sciences* **108**, 1468–1473.

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

Yvernogeau, L., Auda-Boucher, G. and Fontaine-Perus, J. (2012). Limb bud colonization by somite-derived angioblasts is a crucial step for myoblast emigration. *Development* **139**, 277–287.

Yvernogeau, L., Gautier, R., Khoury, H., Menegatti, S., Schmidt, M., Gilles, J.-F. and Jaffredo, T. (2016). An in vitro model of hemogenic endothelium commitment and hematopoietic production. *Development* **143**, 1302–1312.

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**, dev158162.

Zambidis, E. T., Peault, B., Park, T. S., Bunz, F. and Civin, C. I. (2005). Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. *Blood* **106**, 860–870.

Zambidis, E. T., Soon Park, T., Yu, W., Tam, A., Levine, M., Yuan, X., Pryzhkova, M. and Péault, B. (2008). Expression of angiotensin-converting enzyme (CD143) identifies and regulates primitive hemangioblasts derived from human pluripotent stem cells. *Blood* **112**, 3601–3614.

Zanetti, M., Braghetta, P., Sabatelli, P., Mura, I., Doliana, R., Colombatti, A., Volpin, D., Bonaldo, P. and Bressan, G. M. (2004). EMILIN-1 Deficiency Induces Elastogenesis and Vascular Cell Defects. *Mol Cell Biol* **24**, 638–650.

Zapata, A. (1979). Ultrastructural study of the teleost fish kidney. *Developmental & Comparative Immunology* **3**, 55–65.

Zeisberg, M. and Neilson, E. G. (2009). Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest* **119**, 1429–1437.

ZIEGLER B.M., SUGIYAMA D., CHEN M., GUO Y., DOWNS K.M., and SPECK N.A. (2006). The allantois and chorion, when isolated before circulation or chorioallantoic fusion, have hematopoietic potential. Development 133:4183-4192.

Zon, L. I., Mather, C., Burgess, S., Bolce, M. E., Harland, R. M. and Orkin, S. H. (1991). Expression of GATA-binding proteins during embryonic development in Xenopus laevis. Proc Natl Acad Sci U S A 88, 10642-6.
Zou, Y.-R., Kottmann, A. H., Kuroda, M., Taniuchi, I. and Littman, D. R. (1998). Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* **393**, 595–599.
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., et al. (2008). Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells. *Cell Stem Cell* **3**, 625–636.
Zweidler-Mckay, P. A., Grimes, H. L., Flubacher, M. M. and Tsichlis, P. N. (1996). Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Mol. Cell. Biol.* **16**, 4024–4034.

APPENDIX

I. Article 1

An in vitro model of hemogenic endothelium commitment and hematopoietic production

In this article, I performed and analysed the experiments in order to answer to the reviewers demand. I did the PSM dissection, the flow cytometry experiments as well as the analysis on different culture conditions.

© 2016. Published by The Company of Biologists Ltd | Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

An *in vitro* model of hemogenic endothelium commitment and hematopoietic production

Laurent Yvernogeau[1,2,*], Rodolphe Gautier[1,2], Hanane Khoury[1,2], Sara Menegatti[1,2,‡], Melanie Schmidt[1,2], Jean-Francois Gilles[3] and Thierry Jaffredo[1,2,§]

ABSTRACT

Adult-type hematopoietic stem and progenitor cells are formed during ontogeny from a specialized subset of endothelium, termed the hemogenic endothelium, via an endothelial-to-hematopoietic transition (EHT) that occurs in the embryonic aorta and the associated arteries. Despite efforts to generate models, little is known about the mechanisms that drive endothelial cells to the hemogenic fate and about the subsequent molecular control of the EHT. Here, we have designed a stromal line-free controlled culture system utilizing the embryonic pre-somitic mesoderm to obtain large numbers of endothelial cells that subsequently commit into hemogenic endothelium before undergoing EHT. Monitoring the culture for up to 12 days using key molecular markers reveals stepwise commitment into the blood-forming system that is reminiscent of the cellular and molecular changes occurring during hematopoietic development at the level of the aorta. Long-term single-cell imaging allows tracking of the EHT of newly formed blood cells from the layer of hemogenic endothelial cells. By modifying the culture conditions, it is also possible to modulate the endothelial cell commitment or the EHT or to produce smooth muscle cells at the expense of endothelial cells, demonstrating the versatility of the cell culture system. This method will improve our understanding of the precise cellular changes associated with hemogenic endothelium commitment and EHT and, by unfolding these earliest steps of the hematopoietic program, will pave the way for future *ex vivo* production of blood cells.

KEY WORDS: Hemogenic endothelium, Hematopoiesis, Endothelium, Aorta, Avian, Quail

INTRODUCTION

The first hematopoietic stem/progenitor cells (HSPCs) emerge during the initial phase of embryonic development in the form of cell aggregates, designated hematopoietic clusters, that are intimately associated with endothelial cells (ECs) in the floor of the dorsal aorta and in the umbilical and vitelline arteries (Dzierzak and Speck, 2008; Medvinsky et al., 2011). These hematopoietic clusters were shown to be pivotal in the formation of the adult blood system by providing the first adult-type hematopoietic stem cells. Aorta-associated hematopoietic cluster formation has been documented in a variety of vertebrate embryos and thoroughly characterized by immunohistological studies on sections (Jaffredo et al., 2005b) and, more recently, through high-resolution 3D visualization techniques (Yokomizo and Dzierzak, 2010). Tracing experiments *in vivo*, including live imaging approaches, revealed a developmental relationship between ECs and hematopoietic clusters (Jaffredo et al., 1998; de Bruijn et al., 2002; Zovein et al., 2008; Chen et al., 2009; Bertrand et al., 2010; Boisset et al., 2010; Kissa and Herbomel, 2010; Lam et al., 2010). Cluster emergence relies on the presence of specialized ECs, termed hemogenic endothelial cells, which, upon appropriate signaling (Richard et al., 2013), lose their endothelial phenotype and acquire hematopoietic traits. This complex, multi-step developmental process was designated the endothelial-to-hematopoietic transition (EHT) (Kissa and Herbomel, 2010). During EHT, hematopoiesis was shown to occur *de novo* at the expense of the hemogenic endothelium compartment that is progressively lost (Pouget et al., 2006; Kissa and Herbomel, 2010).

Since the hemogenic endothelium represents a small fraction of the ECs and is found only during the early stages of development in the aorta and the associated arteries, a number of efforts have been devoted to design *in vitro* models that faithfully recapitulate hemogenic endothelium commitment and hematopoietic production. The most robust and reliable system was obtained with the use of embryonic stem cells (ESCs). Using appropriate culture conditions, ESCs undergo differentiation and form cellular structures containing progenitors, termed blast colony-forming cells, that are able to give rise to colonies comprising both blood cells and ECs (Choi et al., 1998). Comparing *in vitro* cultures with the *in vivo* situation in the early mouse embryo, the blast colony-forming cell was shown to be the *in vitro* equivalent of the nascent mesoderm ingressing through the primitive streak (Huber et al., 2004) and was characterized by the expression of the receptor tyrosine kinase Flk1 (also known as Kdr and Vegfr2) and the T-box transcription factor brachyury (Robertson et al., 2000). Using well-established culture systems, it was possible to drive this early mesodermal progenitor to differentiate into hemogenic ECs able to give rise to blood cells through EHT and to document several aspects of blood formation from hemogenic ECs (Eilken et al., 2009; Lancrin et al., 2009), including some aspects of EHT (Ditadi et al., 2015). However, the number of hemogenic ECs generated in ESC cultures was reported to be low (1/3000) (Eilken et al., 2009), hampering the establishment of a versatile culture system that could be employed to address questions concerning hemogenic endothelium commitment from non-hemogenic ECs and the cellular and molecular changes that occur during EHT.

The pre-somitic mesoderm (PSM), or paraxial mesoderm, gives rise to the somites and contains numerous progenitor cells able to differentiate into the various lineages produced by the somites. In

[1]Sorbonne Universités, UPMC Univ Paris 06, IBPS, UMR 7622, Laboratoire de Biologie du Développement, Paris 75005, France. [2]CNRS, UMR 7622, Inserm U 1156, IBPS, Laboratoire de Biologie du Développement, Paris 75005, France. [3]Institute of Biology Paris-Seine, Sorbonne Universités, UPMC Univ Paris 06, Cellular Imaging Facility, Paris 75005, France.
*Present address: Hubrecht Institute, Uppsalalaan 8, Utrecht 3584 CT, The Netherlands. [‡]Present address: Cancer Research UK Manchester Institute, The University of Manchester, Wilmslow Road, Manchester M20 4BX, UK.

[§]Author for correspondence (thierry.jaffredo@upmc.fr)

Received 22 May 2015; Accepted 21 February 2016

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

addition to the axial skeleton, skeletal muscles and dermis, the PSM has also been shown to produce a small cohort of ECs that give rise to vascularization of the body wall and limbs and contribute to the aortic roof (Pardanaud et al., 1996; Pouget et al., 2006; Yvernogeau et al., 2012). Of note, somite-derived ECs, which are not initially endowed with hemogenic potential, can be turned into hemogenic cells following exposure to endoderm or to growth factors mimicking the endoderm (Pardanaud and Dieterlen-Lièvre, 1999). Hence, the PSM appears to be a source of naïve mesodermal progenitors that can be orientated towards various cell types upon exposure to the appropriate signals. Given the accessibility of PSM, we reasoned that it could be used as an *in vitro* model in which to trigger the differentiation of large numbers of ECs and hemogenic ECs able to recapitulate the cellular events occurring in the aorta at the time of hematopoietic production.

Here, we report the development of such an *in vitro*, feeder-free culture system that allows the commitment of naïve mesodermal cells *en masse* into hemogenic ECs that are able to undergo robust and long-lasting hematopoiesis over a period of at least 2 weeks. The approach uses pieces of PSM that are submitted to a cocktail of growth factors that drives them to differentiate into ECs. Hemogenic ECs, characterized by the expression of the transcription factor RUNX1, are specified from cells already expressing endothelial markers. Single-cell tracking of hemogenic ECs allows capture of the ephemeral flat-to-round cell transition and revealed unexpected traits of EHT. Moreover, by modulating the combination of growth factors it is possible to preclude EC commitment, accelerate EHT or to promote smooth muscle cell versus EC differentiation. The design of this robust, highly reproducible *in vitro* model will improve our understanding of the precise cellular and molecular changes associated with EHT and will pave the way for the future *ex vivo* production of blood cells and potentiate the discovery of blood cell modulators.

RESULTS

PSM as a naïve mesodermal tissue

Somites are segmental mesoderm derivatives known to be pivotal in the formation of vertebrate embryos by producing a variety of cell types, including skeletal muscles, dermis of the back, tendon, axial skeleton and endothelium of the body wall and limbs (Christ et al., 2007). The somites differentiate from the rostral end of the PSM, a band of loose mesenchyme that does not exhibit any segmental pattern but expresses an ensemble of oscillatory genes to form the somites (Pourquie, 2003). Taking advantage of previous experience with the *in vivo* manipulation of pieces of mesoderm from early embryos (Pardanaud and Dieterlen-Lièvre, 1999), we hypothesized that this tissue could serve as a source for naïve mesoderm that could be orientated towards the endothelial lineage and, more specifically, towards the hemogenic endothelial lineage following exposure to the appropriate signals.

Quail PSM was isolated as described (Pouget et al., 2006; Yvernogeau et al., 2012), free of surrounding contaminating tissues (Fig. 1A, Movie 1). We chose this species because our previous work was based on the use of quail PSM (Pouget et al., 2006) and because

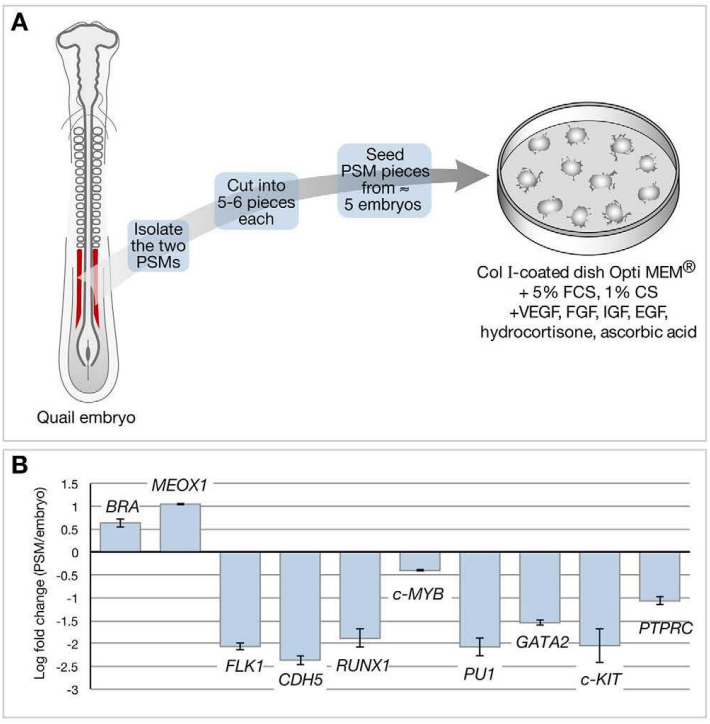


Fig. 1. Culture system and molecular identification of the PSM. (A) The two quail PSMs (red) encompassing a length of about ten somites were isolated, cut into five or six pieces each, and placed in 35-mm culture dishes coated with collagen I. Ten PSMs were plated per dish in Opti-MEM in the presence of 5% fetal calf serum, 1% chicken serum, VEGF, FGF, IGF, EGF, hydrocortisone and ascorbic acid. Cultures were followed over a period of 12 days and time-lapse imaged at different time points. (B) qRT-PCR analysis showing the identity of the PSM through the expression of mesodermal, endothelial and hematopoiesis-specific genes. A positive/negative fold change represents upregulation/downregulation, respectively, of expression in the PSM relative to the same stage embryo. RNA for qPCR was obtained by pooling ten PSMs from five embryos of the same stage. $n=2$ independent experiments. Error bars indicate s.d.

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

quail cells are known to display greater multiplication potential in culture than chicken cells. PSM was submitted to qRT-PCR analysis to probe the expression of mesoderm genes as well as endothelial and hematopoiesis-specific genes (Fig. 1B). Indeed, the PSM expressed brachyury (*BRA*), which encodes a T-box molecule that is expressed by epiblast cells and by the nascent mesoderm following gastrulation (Wilkinson et al., 1990; Kispert et al., 1995), and *MEOX1*, which is specific for uncommitted mesoderm and trunk paraxial mesoderm, but was free of endothelial and hematopoiesis-specific gene expression indicating that none of these lineage commitments has occurred at the time of PSM isolation.

Design of culture conditions and the observation of cultures

We tested different culture media and extracellular matrices that could favor adhesion and the differentiation of PSM-derived cells towards the endothelial lineage. The pieces of PSM exhibited a poor capacity to spread in the absence of extracellular matrix (not shown). Since collagen was reported to positively regulate EC commitment and migration (Whelan and Senger, 2003; Lamalice et al., 2007), we compared collagen I and IV in their efficacy to support endothelial differentiation. Based on cell morphology (Hirashima et al., 2003; Guo et al., 2007; Eilken et al., 2009), type I collagen-coated dishes offered a better differentiation of PSM cells than type IV collagen-coated dishes (not shown). We initially used a cocktail of growth factors from Lonza (SingleQuots Kit) reported to improve the growth of primary EC cultures. However, owing to lack of information regarding the concentrations of the individual growth factors, we decided to replace this cocktail by one of the same composition but using individually purchased growth factors of known concentration. Having chosen the medium (Opti-MEM) and the support (35-mm collagen I-coated dish), the following human growth factors and supplements known to work on avian cells (our unpublished work) were added to Opti-MEM/fetal calf serum (5%)/chicken serum (1%)/penicillin-streptomycin (100 units/ml): VEGF (2 ng/ml), FGF (4 ng/ml), IGF (3 ng/ml), EGF (10 ng/ml), hydrocortisone (200 ng/ml) and ascorbic acid (75 ng/ml).

The culture conditions being defined, we then followed morphological aspects of the culture over a period of 12 days (D). Retrospectively, the culture period could be divided into three phases: (1) spreading and EC differentiation from D0 to D3-4; (2) hematopoietic cell (HC) emergence and production from D4-5 to D8-9; (3) the continuation of HC production and multiplication from D8-9 to D12. During the first 24-48 h of culture (D0 to D2), the pieces of PSM spread onto the collagen I-coated dishes, forming flat layers of tightly adherent cells that displayed EC-like morphology (Fig. 2A-C, Movie 2). From D4, isolated cells among the layer of ECs began to produce round cells that detached from the culture (Fig. 2D). Their number increased over the next few days to form large areas covering the EC layer beneath (Fig. 2E,F). From this period, their number increased further to invade the whole culture dish within the next 3-4 days (Fig. 2G). Flow cytometry analysis using a monoclonal antibody recognizing the QH1 antigen (hereafter referred as to QH1), as a marker for endothelial and hematopoietic cells (Pardanaud et al., 1987) indicated that, at D4, ~70% of the cells were $QH1^+$ (Fig. 2H), as confirmed by QH1 immunocytological staining of the culture (Fig. 2I,J). The EC phenotype was also confirmed using the uptake of human AcLDL as readout (Fig. 2K,L).

Molecular characterization of the culture

We monitored the culture over a period of 12 days and analyzed the expression of gene sets representative of naïve mesoderm, endothelium, hemogenic endothelium and hematopoietic cells using semi-quantitative RT-PCR (Fig. 3A), with validation of some key genes using qPCR (Fig. 3B-D). From D0 to D6, cultures were analyzed daily, and every 2 days from D6 to D12.

BRA was detected immediately following culture onset and during D1, but was no longer expressed at D2. Initially absent at the onset of culture, *FLK1* expression appeared at D0.5-D1. Co-expression of *BRA* and *FLK1*, reported to be associated with the hemangioblast stage (Huber et al., 2004; Vogeli et al., 2006), was detected between D0.5 and D1. The transcription factor *SCL* (also known as *TAL1*), which is crucial for the emergence of the hematopoietic and endothelial lineages during embryoid body differentiation (Lancrin et al., 2009), was found to be expressed from D0.5, as was *GATA2*, which is known to be expressed by mesoderm and nascent ECs (Elefanty et al., 1997; De Val and Black, 2009). Expression of the genes encoding the endothelial and hematopoietic progenitor cell antigen CD34 (Tavian et al., 1996; Wood et al., 1997) and, to a weaker extent, the endothelial-specific calcium-dependent cell adhesion molecule CD144 (also known as cadherin 5) (see also Fig. 3B), was activated at the same time, indicating that endothelial commitment has occurred. This was accompanied by the onset of expression of *CD31* (also known as *PECAM1*), which is expressed on endothelial and hematopoietic cells during development (Newman et al., 1990), and of *vWF*, a key marker of EC function (Wagner et al., 1982), supporting a dynamic EC commitment. Of note, *RUNX1*, a key gene in the generation of blood from the hemogenic endothelium, was expressed at low levels from D0.5 and was significantly upregulated from D3, suggesting that hemogenic endothelium differentiation occurred shortly after the onset of culture. Taken together, ECs differentiated from D0.5, shortly followed by hemogenic ECs.

The culture persisted free of hematopoietic cells until D3. From D3 onwards, the myeloid- and B lymphoid-specific transcriptional activator *PU1* (also known as *PU.1* or *SPI1*) was expressed, shortly followed by *CD45* (also known as *PTPRC*), a pan-hematopoietic marker during development, indicating that hematopoietic commitment had occurred (see also Fig. 3C,D for a more quantitative analysis). *PU1* and *CD45* expression were significantly reinforced from D5, which is when the production of round cells first became prominent. From D6, most of the markers maintained their expression until the end of the culture period at D12.

In order to confirm the commitment into hemogenic endothelium, cultures were co-stained for AcLDL and with an antibody against RUNX1 from D0 to D3. RUNX1 expression was detectable by immunocytochemistry by D2. At this time, only a few $AcLDL^+$ cells expressed RUNX1. The number of $RUNX1^+$ cells dramatically increased at D3 (Fig. 3E-G, Fig. S1A-C), 1 day before the first conspicuous EHT events.

Characterization of the non-adherent cells

We characterized the non-adherent cell fraction from D4, the onset of their production, to the end of the culture period at D12, when non-adherent cells were present in large numbers. Non-adherent cells are hereafter designated as the floating (F) fraction followed by a number indicating their day of retrieval from the culture, i.e. F6 refers to non-adherent cells collected on D6. Non-adherent cells were collected daily from D4 to D12 by thoroughly rinsing the culture dishes, and were counted. A mean of $2\text{-}3\times10^4$ cells were produced daily from F4 to F10. By F10, the production increased to reach 1.04×10^5 cells by F12 (Fig. 4A). This increase is at least partly due to the multiplication of the non-adherent cells, as documented in Movie 3.

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

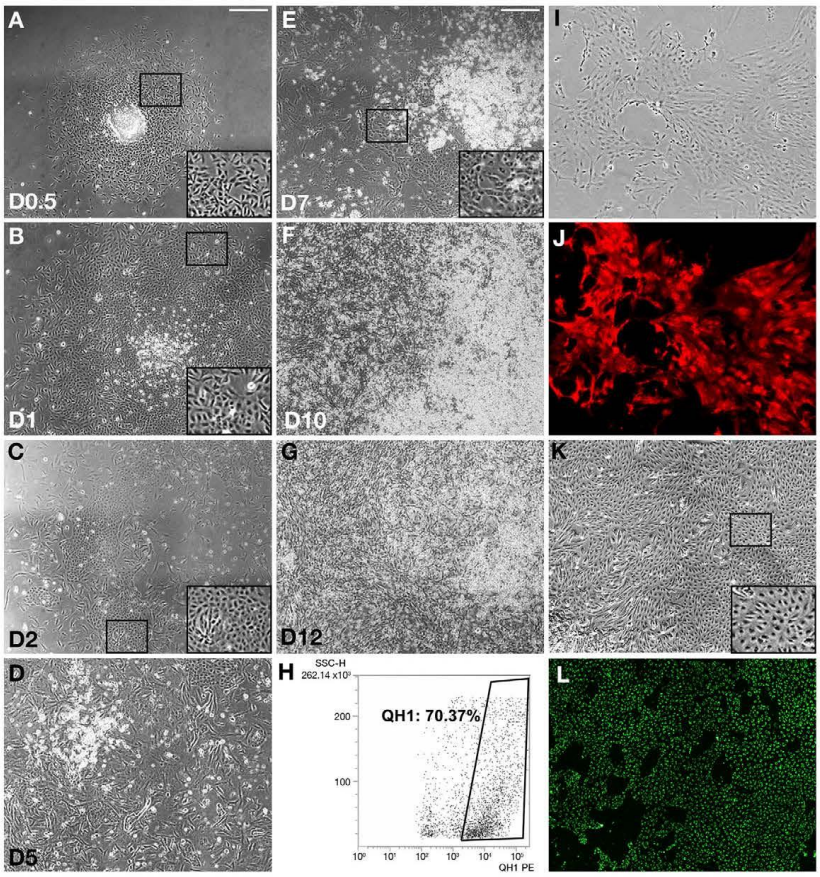


Fig. 2. Morphology and characterization of the culture over time. (A-G) Observation of the culture over a period of 12 days. (A) D0.5, showing initial spreading of a PSM piece. (B) D1, showing the spread of cells to form a flat layer of tightly attached cells, except for the center where spreading is still in progress. (C) D2, showing that cells unable to spread in the center have died. Large areas of adherent cells displaying EC sheet morphology are clearly visible. (D) D5, the beginning of the budding process. Some round, hematopoietic-like floating cells are visible either as single cells or as aggregates of refracting cells. (E) D7, when large aggregates of refracting cells are visible, with the flat layer of cells displaying EC sheet morphology beneath. (F) D10, the number of round, hematopoietic-like cells has significantly increased to cover large areas of the culture dish. (G) D12, the culture is now covered with round cells. The presence of flat EC-like cells has significantly decreased. Scale bar: 100 µm. (H) Flow cytometry analysis of the culture at D4 with the monoclonal antibody QH1. About 70% of the cells, which are mostly flat, attached cells at that stage, are QH1$^+$, testifying to the prominent presence of endo-hematopoietic cells. (I,J) Immunofluorescent characterization of the flat cells using the QH1 monoclonal antibody at D4. Flat cells clearly display cell surface QH1 (J, red), indicative of their endothelial phenotype. (I) Phase contrast. (K,L) AcLDL uptake. Living cells were submitted to AcLDL-A488 uptake for 3 h at 37°C. Endocytic vesicles appear green under UV transillumination. Most of the cells displayed AcLDL uptake (L, green), indicative of their endothelial phenotype. (K) Phase contrast. Insets show higher magnifications of boxed regions.

Since QH1 marks quail endothelial and hematopoietic cells (Pardanaud et al., 1987), we FACS analyzed the floating fraction for QH1 at D7. More than 98% of the cells were QH1$^+$, indicating that they probably exhibit a hematopoietic phenotype (Fig. 4B). FACs analysis results were confirmed by QH1 immunostaining of the non-adherent fraction (Fig. 4C,D).

With the aim of further characterizing these cells, we performed qRT-PCR on the non-adherent fraction from D4 to D12 using key markers of endothelial-to-hematopoietic cell commitment. As expected, *BRA* expression was never detected in floating cells (not shown). *CD144*, a key marker of ECs, was downregulated (Fig. 4E) in keeping with the endothelial-to-hematopoietic commitment analyzed *in vivo*. By contrast, *PU1* (Fig. 4F) and *CD45* (Fig. 4G) were upregulated with time, indicating that a progressive hematopoietic commitment was occurring. Interestingly, *PU1* expression preceded that of *CD45* in accordance with a progressive hematopoietic maturation in culture.

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

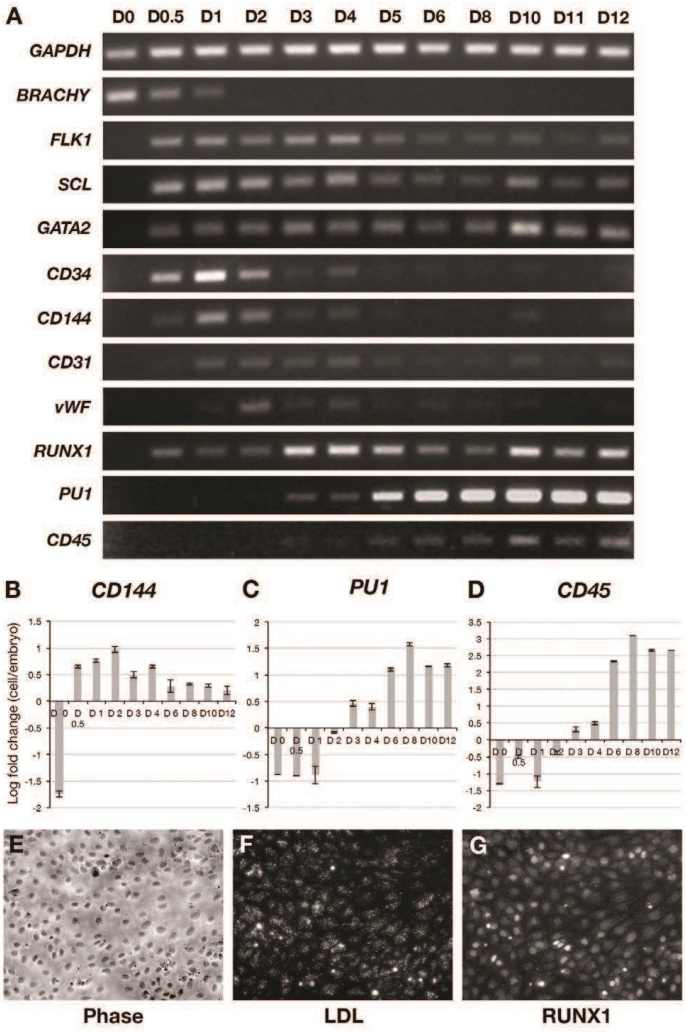


Fig. 3. Molecular characterization of the culture. (A) Semi-quantitative RT-PCR revealing the sequential expression of sets of genes required for mesoderm induction and endothelial and hematopoietic commitment. PCR reactions were run side by side on a 1% agarose gel. Cells from two or three culture dishes were harvested daily from D0 to D6 and every 2 days from D6 to D12. *GAPDH* serves as a control for mRNA amount. A progression from mesoderm to hematopoietic commitment is clearly visible through the expression patterns of several lineage-specific genes (see Results for details). (B-D) qPCR analysis of endothelial, hematopoietic commitment and hematopoietic differentiation-specific genes. The results are shown as the log-transformed fold change in expression (PSM compared with same stage embryo). (B) *CD144*, a recognized endothelial-specific gene, displays a transient increase in expression from D0.5 to D4, then drops and remains low but stable from D6 to D12. (C) *PU1*, one of the earliest genes expressed during hematopoietic commitment, becomes expressed by D3, indicating the onset of EHT. Its expression increased with time and remained high at D12. (D) *CD45*, a hematopoiesis-specific gene, becomes expressed at low level by D3, then is strongly increased by D6 indicative of substantial hematopoietic differentiation and remains high until D12. Error bars indicate s.d. (E-G) Uptake of AcLDL and expression of the transcription factor RUNX1 in D4 cultures. (E) Phase contrast. The flat endothelial-specific phenotype is clearly visible. (F) Cells display prominent uptake of AcLDL. (G) Several cells display high RUNX1 immunostaining in the nucleus, while weaker staining is also visible in most of the cells indicative of their hemogenic endothelial commitment.

To further identify the non-adherent fraction, we collected F6 cells and performed cytospin followed by May–Grünwald Giemsa staining. Thorough characterization of the cells revealed their hematopoietic phenotype and the presence of cells from the granulocyte, erythroblast and monocyte lineages (Fig. 4H).

Tracking EHT

Based on previous characterization, the production of round cells from flat cells faithfully corresponds to an EHT (Jaffredo et al., 1998; Kissa and Herbomel, 2010). Given the high number of hemogenic ECs generated in the culture and the large number of HCs produced, we decided to track EHT using live imaging. Cultures were imaged every 10-15 min over periods from 14-18 h during D3 to D4, when the first EHT events are initiated. To better track EHT, we developed a script running under ImageJ that allows: (1) the labeling of a cell undergoing EHT and to retrospectively identify the flat cell giving rise to the round cell; and (2) to follow the bright, newly formed round cell, over time (see Materials and Methods).

EHT was rapid and left a cell-free area indicating that the fate change occurred at the single-cell level. The passage from flat, adherent cell, to non-adherent cell took between 15 and 30 min. In general, no cell division was detected prior to the passage from flat to round, ruling out asymmetric cell division as a prerequisite for EHT. Round cell production thus caused a progressive exhaustion of the layer of flat cells. Upon detachment, the cell underwent dynamic movements, emitting cellular processes during a period of 30-45 min.

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

We then analyzed the role of serum in triggering EC commitment. Since the cultures were supplemented with 5% FCS and 1% CS, we withdrew one or other serum and examined the effect on the formation of flat cell layers and on AcLDL uptake by flow cytometry. Imaging and FACS analyses were performed at D4. Interestingly, withdrawal of FCS enhanced EC commitment compared with standard conditions, the percentage of $AcLDL^+$ cells reaching 91.5±0.5% compared with a mean of 40±13% for the PSM in standard conditions (Fig. 5A-D,I). By contrast, removal of CS resulted in a slight, non-significant decrease in the cell population taking up AcLDL at D4 (Fig. 5E,F). However, CS absence significantly impaired cell survival after D4, resulting in substantial cell death at D6-D7 (not shown). As a baseline, we also compared somite tissue in the standard conditions (Fig. 5G,H). Taken together, this analysis revealed that the removal of FCS has a dramatic effect on EC differentiation, strongly promoting the EC phenotype.

Since VEGF and FGF are reported to be potent inducers of angiogenesis and to synergize to promote vascular differentiation (Pepper et al., 1992; Asahara et al., 1995; Seghezzi et al., 1998), we focused on these two factors by analyzing their influence on EC commitment from the PSM. We removed one or other factor or both; we also removed all of the growth factors except VEGF. A morphological analysis was performed at D1 and D4 (Fig. S3). The absence of VEGF caused a phenotype very similar to that when FCS is removed, with a robust emergence of flat, tightly adherent cells at D1 and the presence of very large areas of flat cells even at D3 (Fig. S3C,D, compare with Fig. S3A,B). Removal of FGF, or both VEGF and FGF, resulted in poor EC differentiation at D5 with no visible flat cell area (Fig. S3E-H). When VEGF was added as the sole growth factor, EC differentiation occurred but was readily followed by the substantial production of round, hematopoietic cells as soon as D2 (Fig. S3I,J).

Finally, since the PSM generates several different cell types, we questioned whether it is possible to direct the PSM cells towards the smooth muscle cell lineage. TGFβ, which is known to promote smooth muscle cell differentiation in culture, was added at 25 ng/ml to PSM cultures at D0 and replaced the initial set of growth factors. The layer of cells readily displayed a fibroblast-like morphology, with elongated cells forming a compact layer (Fig. 5J,K). No budding cell was found. Expression of *FLK1*, *SCL*, *PU1* and *CD45* was absent from the culture, and *CD31* was barely detected at D2 and decreased thereafter (Fig. 5L), indicating a blockade in EC commitment. Surprisingly, *RUNX1* expression was still detected but was not associated with *PU1* expression, one of its target genes. By contrast, alpha smooth muscle actin (*αSMA*) mRNA was strongly detected (Fig. 5L), suggesting firm commitment towards the smooth muscle cell lineage. These results demonstrated the phenotypic plasticity of the PSM culture according to the culture conditions.

DISCUSSION

Here, using an easily amenable source of mesoderm, we report the establishment and tuning of a versatile culture system using PSM cells that, with appropriate culture conditions, are able to faithfully recapitulate the cellular events that take place during the formation of the aorta and HSPC production. A notable feature is the bias towards the hemogenic endothelium lineage and the subsequent production of hematopoietic cells through EHT; another is the fact that a large number of cells undergo EHT, making it possible to easily track cell fate changes at a single-cell level using time-lapse video microscopy.

The production of HSPCs from ECs was first shown in the chicken embryo in dye-marking studies (Jaffredo et al., 1998). Using interspecies grafting experiments, it was shown that HSPCs are produced at the expense of the hemogenic endothelial population, which progressively disappeared from the aortic floor (Pouget et al., 2006). In mammalian embryos, phenotypic and genetic approaches also demonstrated that the first HSPCs are derived from vascular ECs during a short period of time (de Bruijn et al., 2002; North et al., 2002; Zovein et al., 2008; Chen et al., 2009). Given the key role of the hemogenic endothelium and EHT in the production of HSPCs, an in-depth investigation of the cellular and molecular processes associated with these traits is needed.

A major hurdle in dissecting hemogenic endothelium commitment and EHT is the low number of cells per embryo coupled with the difficulties in isolating discrete steps associated with the cellular progression. Significant advances have been made, using transgenic mouse lines carrying the +23 *Runx1* hematopoietic enhancer (Swiers et al., 2013), in understanding the molecular control of the endothelial-hematopoietic balance and the timing of these changes. However, hemogenic EC isolation remains difficult, especially if one is to dissect discrete steps from non-hemogenic EC to EHT. Our culture system allows 50-70% of the cells to be directed towards the endothelial and the hemogenic endothelial fates. This is an important advance compared with the number of cells exhibiting these or similar phenotypes in ESC cultures (Eilken et al., 2009) or in embryos *in vivo* (Kissa and Herbomel, 2010; Yokomizo and Dzierzak, 2010). In addition, the stromal cell lines sometimes used in ESC differentiation protocols into EC (Guo et al., 2007) or hemogenic EC (Eilken et al., 2009) are not required in our system, which facilitates cell isolation if needed. Another interesting feature is the fact that EHT and hematopoietic commitment occur without any modification to the composition of the medium. This indicates that the culture conditions faithfully recapitulate the molecular events occurring during endothelial and hemogenic endothelial commitment in the embryonic aorta. Indeed, somitic ECs are not hemogenic, but it has been shown that when these cells are placed in appropriate conditions they are able to give rise to hemogenic ECs and to blood (Pardanaud and Dieterlen-Lièvre, 1999). In our case, we strongly bias the PSM cells towards the EC lineage and turn these non-hemogenic ECs into hemogenic ECs, as testified by the acquisition of endothelial traits followed by *RUNX1* expression during the first 4 days of culture.

Our daily molecular analysis indicates a progressive switch from the mesoderm to the hematopoietic state, with a passage through mesoderm, endothelium, hemogenic endothelium and hematopoietic fates. This is consistent with the proposed model of blood cell formation deduced from ESC cultures (Lancrin et al., 2010). We found that the initial commitment of mesodermal cells into EC from D0 to D2 is not associated with *RUNX1* expression. However, *RUNX1* is expressed in culture from D2, and its expression increases and extends to most, if not all, ECs thereafter. This is in keeping with the changes in *RUNX1* expression shown to occur *in vivo* during the formation of the aorta (Richard et al., 2013), thereby demonstrating that our culture conditions accurately reproduce the cellular and molecular events taking place *in vivo*. When they undergo EHT, each culture produces *de novo* at least 2×10^4 cells per 35-mm dish per day from D4 to D8, and this number increases further to reach 1×10^5 cells per dish at D12, corresponding to a considerable hematopoietic production considering the relatively low number of cells seeded at culture onset. At the molecular level, qPCR analysis of *CD144*, *PU1* and *CD45* reveals an early loss of endothelial traits in floating cells, in keeping with the changes that occur during EHT *in vivo*, and a progressive increase in CD45 with time consistent with

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

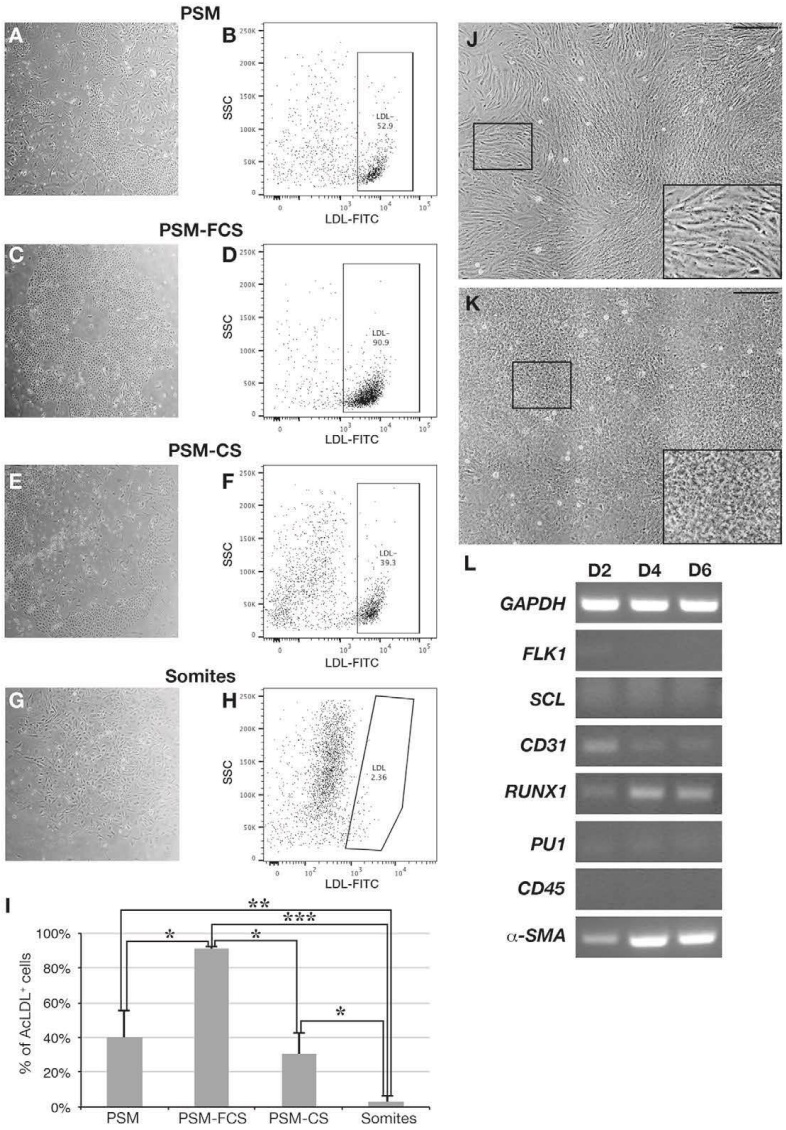


Fig. 5. EC differentiation depends on the tissue of origin and on medium composition. (A-I) Role of serum in EC differentiation. PSM in (A,B) standard conditions, (C,D) without FCS, (E,F) without CS, and (G,H) somites in standard conditions. Note the enhanced EC differentiation in the absence of FCS, with up to 90% of the cells positive for AcLDL uptake, and the poor differentiation that is shown by the somite cells. Dot plots are from individual representative experiments and are the result of 2000 events analyzed. (I) The percentage of AcLDL$^+$ cells in the different experimental conditions. Data are mean±s.e.m. n=3. *P<0.01, **P<0.001, ***P<0.0001, Student's t-test. Analyses are the result of three independent experiments with three independent wells per experiment. (J-L) Commitment to smooth muscle cells with TGFβ. (J,K) Phenotype of the culture at D4 (J) and D8 (K) following replacement of the initial set of growth factors by TGFβ. No endothelial-specific phenotype was visible during the culture period (12 days). No hematopoietic-like floating cell was visible. The rare floating cells are dead cells. Boxed regions are magnified in insets. Scale bars: 30 µm. (L) Semi-quantitative PCR analysis of the culture submitted to TGFβ. Cells were collected at D2, D4 and D6 and analyzed for the expression of several genes specific for endo-hematopoietic cells and for the alpha smooth muscle actin isoform (αSMA) that is specific for smooth muscle cells. Low to nil expression of endo-hematopoietic genes was found, whereas cells displayed increased expression of αSMA, indicative of smooth muscle cell commitment.

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

maturation of the recently produced HCs (Jaffredo et al., 2005a; Zape and Zovein, 2011). This sustained hematopoietic production is indeed due to the EHT that persists with time in culture, and also to hematopoietic cell multiplication. Future work will be needed to quantify the relative importance of these two events.

The hematopoietic cells that are produced differentiate into several hematopoietic lineages, as testified by the presence of at least three morphologically distinct types of hematopoietic cell – granulocytes, monocytes and erythroblasts – originating from three distinct types of progenitor. The lack of avian recombinant cytokines precludes a thorough clonogenic identification of the progenitor cells produced in culture. Further work will be necessary to examine whether lymphocytes, and eventually HSPCs, could also be produced. The differential expression of *PU1*, which is a direct target of RUNX1 (Huang et al., 2008), and *CD45* between D4 and D6 indicates that the newly formed HSPCs undergo progressive maturation in culture.

Owing to the limitations mentioned above and the short duration of the process, capturing hemogenic ECs undergoing EHT remains a challenge in culture. A large number of cells experiencing EHT are visible from D4. This unique situation would allow the isolation of hemogenic ECs undergoing EHT for analysis by various approaches. Investigation of the culture conditions has shown that it is possible to reinforce EC differentiation or to anticipate EHT by modulating the presence of FCS and VEGF. Our aim is to exploit this versatility to find a way to collect sufficient numbers of cells undergoing EHT for further cellular and molecular analyses.

In addition to ECs and hemogenic ECs, we demonstrate that it is possible to direct the culture towards the smooth muscle cell lineage by modulating the combination of growth factors. This opens the way to study, in greater depth, the molecular choices made by mesoderm cells during differentiation. This will also help to more accurately identify key factors involved in the EHT process from future high-throughput data. Given the versatility of the culture system, one could also envision its use to manipulate the PSM cells to produce striated muscle or tendon precursors, two cell types also derived following differentiation of the somite.

Taken together, our results provide a new and versatile system with which to study commitment to hemogenic ECs and the EHT. Future in-depth analysis of the molecular pathways involved in these processes will have important implications for the understanding of EHT and the search for key hematopoietic inducing signals and molecular pathways that are crucial in directing the production of HSPCs from the hemogenic endothelium.

MATERIALS AND METHODS

PSM isolation

We used quail (*Coturnix coturnix japonica*) embryo PSM, handled according to Fig. 1A. Eggs were incubated for 36-45 h at 37±1°C in a humidified atmosphere to reach 10-18 somite pairs. Microsurgery was performed as previously described (Pardanaud et al.,1996). The PSM was removed over a length corresponding to ten somites from both sides of the embryo. Five embryos (i.e. ten PSMs) were used per culture dish. Each PSM was cut into five or six equal pieces and rinsed in Opti-MEM plus GlutaMAX I containing 5% fetal calf serum (FCS), 100 units/ml penicillin/streptomycin and 1% chicken serum (CS) (all Gibco Life Technologies) before culture. Avian embryo care and procedures were in accordance with national and European laws.

Culture conditions

PSM was cultured in Opti-MEM with GlutaMAX I supplemented with 5% FCS, 1% chicken serum, 100 units/ml penicillin/streptomycin and the following growth factors (PromoCell/PromoKine unless stated otherwise): human VEGF (C64410; 2 ng/ml), human FGF (C60240; 4 ng/ml), human IGF (C60840; 3 ng/ml), human EGF (C60170; 10 ng/ml), hydrocortisone (Sigma, H6909; 200 ng/ml) and ascorbic acid (Sigma, A4544; 75 µg/ml). To promote smooth muscle cell differentiation, the growth factors were replaced by TGFβ (PromoCell/PromoKine, C63500; 25 ng/ml). PSM was cultured in Corning BioCoat 35 mm collagen I- or IV-coated dishes (Corning-Dutscher). Medium was changed every 2 days unless otherwise specified.

RNA extraction and qRT-PCR

RNA extractions were performed using the RNeasy Kit (Qiagen). Freshly isolated PSM (ten) were resuspended in the RNeasy buffer solution (RLT). For RNA extraction from cultured cells, the cells were first centrifuged (300 ***g*** for 10 min) to remove the medium and resuspended in RLT. Adherent cells were first trypsinized, washed in PBS containing 10% FCS, centrifuged and then resuspended in RLT. Quality and quantity of the extracted RNA was evaluated using Nanodrop. The primers used for semi-quantitative RT-PCR and for qPCR are listed in Table S1. PCR was performed on an Eppendorf Mastercycler Epgradient S. qPCR was performed using the LightCycler 480 (Roche) real-time system according to the manufacturer's instructions. Relative expression was calculated as $2^{[Ct(\text{gene of interest})-Ct(\text{gene of reference})]}$.

Immunostaining and DAPI staining of *in vitro* culture

We used acetylated low-density lipoprotein from human plasma coupled to Alexa Fluor 488 (AcLDL-A488; 1 mg/ml; Life Technologies, L23380, batches 1291485 and 1696210), the QH1 monoclonal antibody developed by Pardanaud et al. (1987) (obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242, USA) and antibody against RUNX1 (Abcam, ab92336, batches GR107772-3 and 107772-5) to determine the endothelial phenotype of PSM-derived cells. AcLDL-A488 (1/100 in PBS) was incubated together with QH1 antibody (1/20 in PBS) for 30 min at 37°C in the culture dish.

After three washes in PBS, goat anti-mouse IgG1 secondary antibody coupled to Alexa Fluor 555 (Molecular Probes; 1/100 in PBS), which recognizes QH1 antibody, was incubated for 30 min at 37°C. Cultures were then washed three times in PBS before imaging.

When RUNX1 was revealed, the cells were fixed following AcLDL uptake with 3.7% formaldehyde in PBS, rinsed three times with PBS containing 0.1% Triton X-100 for 10 min, incubated with the anti-RUNX1 antibody for 1 h at room temperature, followed by incubation with a goat anti-rabbit secondary antibody coupled to biotin (Southern Biotech, 4050-08) followed by three rinses in PBS (5 min each) and an incubation with Streptavidin-Cy3 (Invitrogen, 43-8315) diluted 1/500 in PBS.

To visualize nuclei, cultures were first fixed in 4% paraformaldehyde in PBS for 20 min, washed three times in PBS, and then incubated with DAPI in PBS containing 0.4% Triton X-100 for 20 min. After three washes in PBS, cultures were mounted with a coverslip before imaging on a Leica DM6000 B inverted microscope.

FACS analysis

Cells were stained with either AcLDL-A488 or with the QH1 monoclonal antibody. When adherent cells were used, cells were trypsinized and rinsed three times in PBS. When floating cells were used, cells were gently removed from the culture dish by pipetting. For AcLDL analysis, the cells were suspended in PBS containing 7-aminoactinomycin (7AAD) to exclude dead cells and analyzed with a MacsQuant analyzer 10 (Miltenyi Biotec). For QH1, cells were centrifuged, resuspended in PBS, incubated with the QH1 antibody (1/20) for 20 min at 4°C, washed in PBS and centrifuged. Cells were then incubated with goat anti-mouse IgG1-A488 (Molecular Probes; 1/100 in PBS) for 20 min at 4°C, washed in PBS and centrifuged. Finally, cells were resuspended in PBS containing 7AAD to exclude dead cells and analyzed on a FACSAria III (BD Biosciences) or on a MacsQuant analyzer 10. Analyses were performed with FlowJo 10. Statistics were performed with GraphPad Prism.

May–Grünwald Giemsa staining

Floating cells were gently removed from the culture dish by pipetting, washed, centrifuged and suspended in PBS before proceeding to cytospin (Cytocentrifuge, Shandon-Elliot). Glass slides with the spot of cells were

covered with May–Grünwald solution (Merck) for 3 min at room temperature. Five or six drops of PBS were added to the May–Grünwald solution directly on the slide, mixed gently, and incubated for another 3 min. Slides were then rinsed with PBS and covered with diluted Giemsa (1/10; Merck) for 20 min. Slides were finally washed with distilled water and air dried. Slides were mounted with a coverslip and a few drops of Entellan (Sigma). Images were taken on a Nikon Eclipse E800 microscope.

EHT tracking

Movies were recorded on a Leica DM6000 B inverted microscope at 37°C and 5% CO_2 with a 10× objective. Images were acquired every 10 min using a CoolSnap HQ2 camera (1392×1040 imaging pixels; Photometrics) over a mean period of 24 h using Leica MMAF software v1.6.0. Films were analyzed using ImageJ (NIH). The tracking was made in two steps because of the two cell morphologies. In the first step, and because before undergoing EHT there is little movement in the flat layer of ECs, the cell undergoing phenotypic changes is marked at the time or immediately after the time (user choice) of EHT. A circle is drawn to localize the cell based on its Cartesian coordinates. From this point, a second window is opened that will allow the user to trace the same cell back in time to before EHT (the number of frames back is also a user choice). The macro is available upon request. In the second step, the algorithm automatically finds the brightest points around the previous cell localization and selects the closest coordinates. Movies have been assembled and labeled using Final Cut Pro (Apple).

Acknowledgements

We thank Drs Charles Durand and Cecile Drevon for critical reading of the manuscript; Laurence Petit for efficient help in flow cytometry; and Sophie Gournet for excellent photographic and drawing assistance.

Competing interests

The authors declare no competing or financial interests.

Author contributions

L.Y., R.G. and T.J. designed experiments. L.Y., R.G., H.K., S.M. and M.S. performed experiments. J.-F.G. developed the cell-tracking program. L.Y., R.G., H.K. and T.J. analyzed data. L.Y. and T.J. wrote the paper.

Funding

This study was supported by grants from the Fondation pour la Recherche Médicale [DEQ20100318258] and Agence Nationale pour la Recherche/California Institute for Regenerative Medicine [ANR/CIRM 0001-02].

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.126714/-/DC1

References

Asahara, T., Bauters, C., Zheng, L. P., Takeshita, S., Bunting, S., Ferrara, N., Symes, J. F. and Isner, J. M. (1995). Synergistic effect of vascular endothelial growth factor and basic fibroblast growth factor on angiogenesis in vivo. *Circulation* **92**, 365-371.

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* **464**, 108-111.

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* **464**, 116-120.

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

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

Christ, B., Huang, R. and Scaal, M. (2007). Amniote somite derivatives. *Dev. Dyn.* **236**, 2382-2396.

de Bruijn, M. F. T. R., Ma, X., Robin, C., Ottersbach, K., Sanchez, M.-J. and Dzierzak, E. (2002). Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* **16**, 673-683.

De Val, S. and Black, B. L. (2009). Transcriptional control of endothelial cell development. *Dev. Cell* **16**, 180-195.

Ditadi, A., Sturgeon, C. M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A. D., Azzola, L., Ng, E. S., Stanley, E. G., French, D. L. et al. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat. Cell Biol.* **17**, 580-591.

Dzierzak, E. and Speck, N. A. (2008). Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat. Immunol.* **9**, 129-136.

Eilken, H. M., Nishikawa, S.-I. and Schroeder, T. (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457**, 896-900.

Elefanty, A. G., Robb, L., Birner, R. and Begley, C. G. (1997). Hematopoietic-specific genes are not induced during in vitro differentiation of scl-null embryonic stem cells. *Blood* **90**, 1435-1447.

Guo, R., Sakamoto, H., Sugiura, S. and Ogawa, M. (2007). Endothelial cell motility is compatible with junctional integrity. *J. Cell. Physiol.* **211**, 327-335.

Hirashima, M., Ogawa, M., Nishikawa, S., Matsumura, K., Kawasaki, K., Shibuya, M. and Nishikawa, S.-I. (2003). A chemically defined culture of VEGFR2+ cells derived from embryonic stem cells reveals the role of VEGFR1 in tuning the threshold for VEGF in developing endothelial cells. *Blood* **101**, 2261-2267.

Huang, G., Zhang, P., Hirai, H., Elf, S., Yan, X., Chen, Z., Koschmieder, S., Okuno, Y., Dayaram, T., Growney, J. D. et al. (2008). PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat. Genet.* **40**, 51-60.

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**, 625-630.

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

Jaffredo, T., Bollerot, K., Sugiyama, D., Gautier, R. and Drevon, C. (2005a). Tracing the hemangioblast during embryogenesis: developmental relationships between endothelial and hematopoietic cells. *Int. J. Dev. Biol.* **49**, 269-277.

Jaffredo, T., Nottingham, W., Liddiard, K., Bollerot, K., Pouget, C. and de Bruijn, M. (2005b). From hemangioblast to hematopoietic stem cell: an endothelial connection? *Exp. Hematol.* **33**, 1029-1040.

Kispert, A., Ortner, H., Cooke, J. and Herrmann, B. G. (1995). The chick Brachyury gene: developmental expression pattern and response to axial induction by localized activin. *Dev. Biol.* **168**, 406-415.

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

Lam, E. Y. N., Hall, C. J., Crosier, P. S., Crosier, K. E. and Flores, M. V. (2010). Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. *Blood* **116**, 909-914.

Lamalice, L., Le Boeuf, F. and Huot, J. (2007). Endothelial cell migration during angiogenesis. *Circ. Res.* **100**, 782-794.

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**, 892-895.

Lancrin, C., Sroczynska, P., Serrano, A. G., Gandillet, A., Ferreras, C., Kouskoff, V. and Lacaud, G. (2010). Blood cell generation from the hemangioblast. *J. Mol. Med.* **88**, 167-172.

Medvinsky, A., Rybtsov, S. and Taoudi, S. (2011). Embryonic origin of the adult hematopoietic system: advances and questions. *Development* **138**, 1017-1031.

Newman, P. J., Berndt, M. C., Gorski, J., White, G. C., II, Lyman, S., Paddock, C. and Muller, W. A. (1990). PECAM-1 (CD31) cloning and relation to adhesion molecules of the immunoglobulin gene superfamily. *Science* **247**, 1219-1222.

North, T. E., de Bruijn, M. F. T. R., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E. and Speck, N. A. (2002). Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16**, 661-672.

Pardanaud, L. and Dieterlen-Lièvre, F. (1999). Manipulation of the angiopoietic/hemangiopoietic commitment in the avian embryo. *Development* **126**, 617-627.

Pardanaud, L., Altmann, C., Kitos, P., Dieterlen-Lièvre, F. and Buck, C. A. (1987). Vasculogenesis in the early quail blastodisc as studied with a monoclonal antibody recognizing endothelial cells. *Development* **100**, 339-349.

Pardanaud, L., Luton, D., Prigent, M., Bourcheix, L.-M., Catala, M. and Dieterlen-Lièvre, F. (1996). Two distinct endothelial lineages in ontogeny, one of them related to hemopoiesis. *Development* **122**, 1363-1371.

Pepper, M. S., Ferrara, N., Orci, L. and Montesano, R. (1992). Potent synergism between vascular endothelial growth factor and basic fibroblast growth factor in the induction of angiogenesis in vitro. *Biochem. Biophys. Res. Commun.* **189**, 824-831.

Pouget, C., Gautier, R., Teillet, M.-A. and Jaffredo, T. (2006). Somite-derived cells replace ventral aortic hemangioblasts and provide aortic smooth muscle cells of the trunk. *Development* **133**, 1013-1022.

Pourquie, O. (2003). Vertebrate somitogenesis: a novel paradigm for animal segmentation? *Int. J. Dev. Biol.* **47**, 597-603.

Richard, C., Drevon, C., Canto, P.-Y., Villain, G., Bollérot, K., Lempereur, A., Teillet, M.-A., Vincent, C., Rosselló Castillo, C., Torres, M. et al. (2013). Endothelio-mesenchymal interaction controls runx1 expression and modulates the notch pathway to initiate aortic hematopoiesis. *Dev. Cell* **24**, 600-611.

Development (2016) 143, 1302-1312 doi:10.1242/dev.126714

Robertson, S. M., Kennedy, M., Shannon, J. M. and Keller, G. (2000). A transitional stage in the commitment of mesoderm to hematopoiesis requiring the transcription factor SCL/tal-1. *Development* **127**, 2447-2459.

Seghezzi, G., Patel, S., Ren, C. J., Gualandris, A., Pintucci, G., Robbins, E. S., Shapiro, R. L., Galloway, A. C., Rifkin, D. B. and Mignatti, P. (1998). Fibroblast growth factor-2 (FGF-2) induces vascular endothelial growth factor (VEGF) expression in the endothelial cells of forming capillaries: an autocrine mechanism contributing to angiogenesis. *J. Cell Biol.* **141**, 1659-1673.

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K. D. et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat. Commun.* **4**, 2924.

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

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**, 337-339.

Wagner, D. D., Olmsted, J. B. and Marder, V. J. (1982). Immunolocalization of von Willebrand protein in Weibel-Palade bodies of human endothelial cells. *J. Cell Biol.* **95**, 355-360.

Whelan, M. C. and Senger, D. R. (2003). Collagen I initiates endothelial cell morphogenesis by inducing actin polymerization through suppression of cyclic AMP and protein kinase A. *J. Biol. Chem.* **278**, 327-334.

Wilkinson, D. G., Bhatt, S. and Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657-659.

Wood, H. B., May, G., Healy, L., Enver, T. and Morriss-Kay, G. M. (1997). CD34 expression patterns during early mouse development are related to modes of blood vessel formation and reveal additional sites of hematopoiesis. *Blood* **90**, 2300-2311.

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

Yvernogeau, L., Auda-Boucher, G. and Fontaine-Perus, J. (2012). Limb bud colonization by somite-derived angioblasts is a crucial step for myoblast emigration. *Development* **139**, 277-287.

Zape, J. P. and Zovein, A. C. (2011). Hemogenic endothelium: origins, regulation, and implications for vascular biology. *Semin. Cell Dev. Biol.* **22**, 1036-1047.

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. et al. (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* **3**, 625-636.

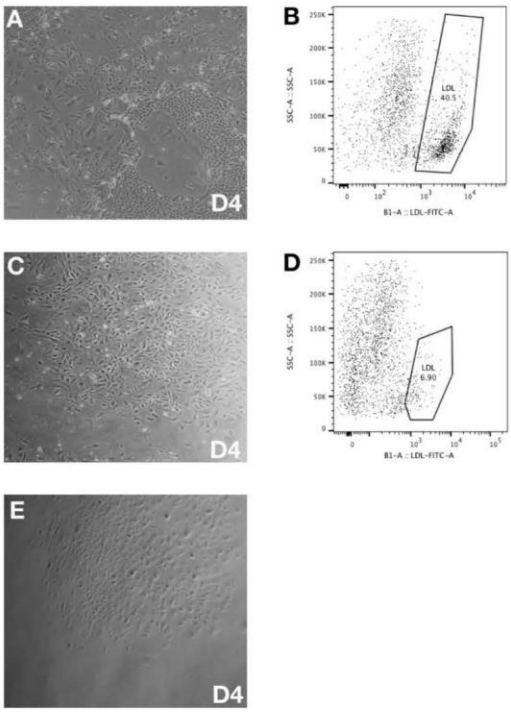


Fig. S2. Comparison between PSM, somites and lateral plate mesoderm isolated from embryos at the same somitic stage for their capacities to give rise to ECs. Dot plots are from one representative experiment. FACS analysis could not be performed for the lateral plate due to the low number of cells obtained. The dot plots are the result of 2000 events analyzed.

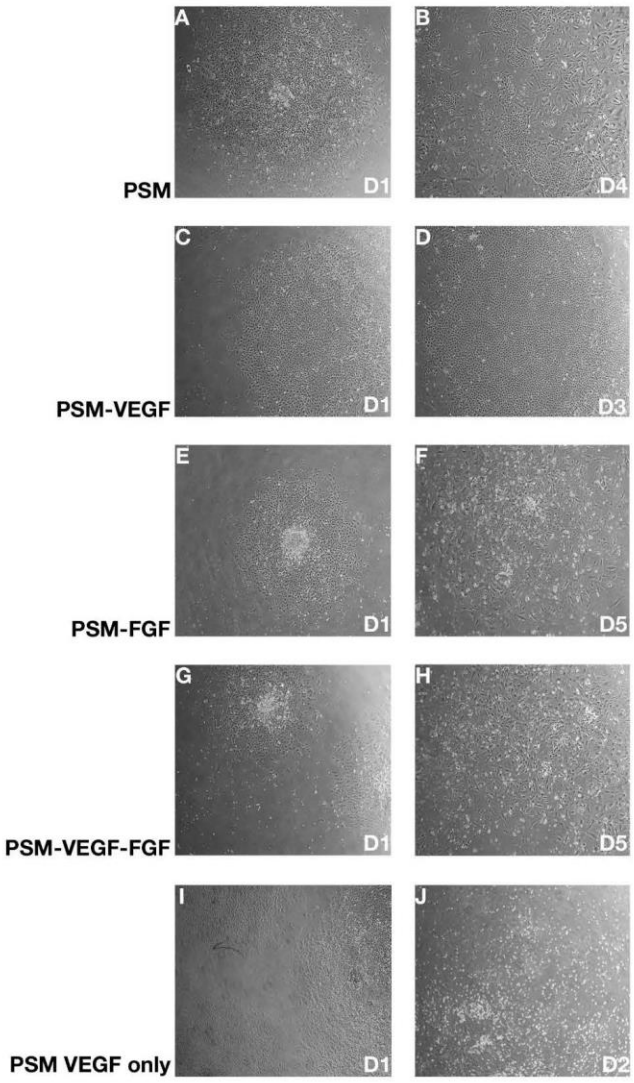


Fig. S3. PSM cultures in different growth factor conditions. In the absence of VEGF, PSM cells readily differentiate into flat, EC-like cells that are reminiscent of the cells obtained when PSM cells are placed in the absence of FCS (Figure 5C,D). When VEGF alone was added, an accelerated EC differentiation was observed quickly followed by a round, hematopoietic cell production. Cultures have been run from day (D) one to either D2 to 5 depending on the development of the culture.

II. Article 2

In vivo haematopoietic cell production by bone marrow-derived hemogenic endothelium

Accepted manuscript in Nature Cell Biology.

In this article I analyzed the YFP^+ vs YFP^- transcriptome and suggested the use of our culture conditions to validate the endothelial and hematopoietic populations in this study.

In vivo haematopoietic cell production by bone marrow-derived hemogenic endothelium

Yvernogeau L[1,¥], Gautier R[1#], Petit, L[1#], Khoury, H[1], Relaix F[2], Ribes V[3], Sang H[4], Charbord, P[1], Souyri M[1,7], Robin C[5,6] and Jaffredo T[1*]

[1] *Sorbonne Universités, UPMC Univ Paris 06, IBPS, CNRS UMR7622, Inserm U 1156, Laboratoire de Biologie du Développement; 75005 Paris*

[2] *Institut Mondor de recherches biomédicales, Inserm U955, Créteil France*

[3] *Institut Jacques Monod, Paris*

[4] *The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush ; Midlothian EH25 9RG ; U.K*

[5] *Hubrecht Institute, Uppsalalaan 8 ; 3584 CT Utrecht*

[6] *Regenerative medicine center, University Medical Center Utrecht, 3584 EA Utrecht*

[7] *Present address : UMRS 1131 1 Ave Claude Vellefaux, 75010 Paris*

¥ present address : Hubrecht Institute, Uppsalalaan 8 ; 3584 CT Utrecht

 # equal contribution.

**Correspondence to : thierry.jaffredo@upmc.fr*

One of the major challenges in regenerative medicine is to reproduce tailor-made tissue-specific stem cells or early committed progenitors for cell replacement therapies. Despite recent progress[1,2], controlled production of *bona fide* Haematopoietic Stem/Progenitor Cells (HSPCs) from pluripotent precursors still requires a better understanding of where, when and how these cells are physiologically produced during development. It is now well established that HSPCs derive from a specialized subset of Endothelial Cells (ECs) designated as haemogenic through an Endothelial-to-Haematopoietic Transition (EHT)[3-8]. EHT occurs early during embryogenesis in the main arteries of the embryo such as the aorta[9,10]. Whether EHT occurs at later stages and in other organs (e.g. bone marrow (BM)) is unknown. Here, we discovered the existence of a haemogenic endothelium capable of *de novo* generating HSPCs in the forming BM of chicken and mouse late foetuses and newborns. We identified the somite as the source of BM vascularization by tracing the BM-forming ECs through experimental embryology, genetic labelling and live imaging approaches and subsequently by analysing their haematopoietic progeny in adults. We demonstrated that the late foetal/newborn BM haemogenic ECs produce a small cohort of HSPCs capable of circulating and colonizing the secondary haematopoietic organs. Phenotypic and functional analyses disclosed that BM endothelium-derived HSPCs are mainly Multipotent Progenitors (MPPs) and a few Haematopoietic Stem Cells (HSCs). These cells harbour a very specific molecular signature close to the one of embryonic ECs with a prominent Notch pathway, endothelial-specific genes and transcription factors involved in EHT. Our results demonstrate that HSPCs are *de novo* generated past embryonic stages. We therefore disclose a new haematopoietic wave occurring in late foetal/newborn stage that fills the gap between the completion of embryonic haematopoiesis and the beginning of adult BM production. Identifying all steps of haematopoietic production and the molecular events controlling this process is of fundamental interest and might help to devise innovative stem cell therapies in the future.

To determine whether haemogenic ECs exist in the BM vascular network and their potential origin, we used a combination of experimental embryology, genetic approaches, transcriptomic and functional analyses on chicken and mouse models. Based on previous research[11-14], the somite appeared as a promising site of BM endothelial origin. The somite is a mesoderm-derived tissue that produces axial skeleton, muscles of the body wall and limbs, dermis and endothelium[15]. Somite-derived ECs form the vascular network of the body wall, kidney and limbs[11-13]. To trace the endothelial derivatives of the somite in the chicken we performed isotopic and isochronic grafts of GFP^+ chicken[16] presomitic mesoderm (PSM), the somite anlage, into wild-type (WT) chicken embryos at embryonic day (E)2 (15 somite pairs) (Fig. 1a). GFP^+ somite-derived cells were followed in individual grafted chickens (n=34) from the time of grafting (E2) to the adult stage by performing immunostaining and flow cytometry analyses on circulating blood, embryo sections or whole mount organs (Extended Data Table 1). A collection of endothelial/haematopoietic antibodies (MEP21)[17] or haematopoietic-specific antibodies against CD45 (or PTPRC, pan-haematopoietic marker), CD41 (or ITGA2B, HSPC marker)[18], c-Kit (or CD117, HSPC marker), KUL1 (macrophage/monocyte marker) and CD3 (T cell marker) were used. EHT was examined by live confocal imaging of BM thick slices (Extended Data Table 1). In agreement with our previous work[12], PSM removal did not affect the aortic integrity (Extended Data Fig.1a).

Importantly, PSM did not contain neither HCs nor ECs, as demonstrated here by immunostainings and flow cytometry analyses on dissected PSMs (Extended Data Fig.1b-d) or previously by qRT-PCR[19]. It thereby rules out graft contamination by these two cell types later on during development.

Two days after grafting (at E4, n=5), PSM-derived GFP$^+$ cells (including emigrating ECs and myoblasts) were visible in the limb bud (Extended Data Fig. 2a). GFP$^+$ cells were found around the neural tube, in limbs and the aorta. Of note, no GFP$^+$ cells were found within the intra-aortic clusters that harbour the first HSPCs (Extended Data Fig. 2b, c, d), as expected from previous reports[11,12]. By E8, the grafted tissue has contributed to the body wall and limb muscle masses and to vascular ECs (Extended Data Fig. 2e, f, g). No GFP$^+$ cells were found in the para-aortic foci (n=2) known to contain CD41$^+$ HSPCs[20] (Extended Data Fig. 2h, i). These data therefore proved that GFP$^+$ cells do not contribute to the formation of the haematopoietic system at that stage, which is in keeping with the absence of GFP$^+$ circulating cells and with the established relationship between intra-aortic clusters and para-aortic foci[21]. The limb BM remained avascular until E12 (n=2), a stage when the first blood vessels associated with circulating CD41$^+$ and CD45$^+$ cells penetrated the cartilage tissue (Extended Data Fig. 2j-m). No GFP$^+$ circulating cell was detected from E4 to E12 (not shown). The contribution of GFP$^+$ cells to the BM was extensively analysed at E16/E17, a developmental stage when sufficient BM cells could be observed or collected for flow cytometry analysis (n=16). Whole BM mounts showed the colonization of the grafted limb by GFP$^+$ cells (Fig. 1b) and the presence of a dense GFP$^+$ vascular network in the forming BM (Fig. 1c, d and Supplementary Movie 1). Cross section of the BM confirmed the presence of GFP$^+$ sinusoids hence demonstrating the somite origin of the BM vascular network (Fig. 1e, f). Numerous GFP$^-$CD45$^+$ blood cells outside the GFP$^+$ vascular network were observed (Extended Data Fig. 3a-c). However, a subpopulation of GFP$^+$CD41$^+$ and to a less extent of GFP$^+$CD45$^+$ cells was tightly associated with the GFP$^+$ sinusoids (Fig. 1g-j, Extended Data Fig. 3). The co-expression of blood cell markers and GFP (as shown on confocal sections) clearly demonstrated that a fraction of CD45$^+$ cells (Extended Data Fig. 3d-l and Supplementary Movie 2) and CD41$^+$ cells (Extended Data Fig. 4a-f and Supplementary Movie 3) arose from GFP$^+$ somite-derived ECs in the BM. To ascertain the haematopoietic production *in situ,* we adapted the previously described time- lapse live confocal imaging[8] to BM thick slices of E16/E17 grafted embryos that were stained with anti- CD41 antibodies. We were able to witness flat GFP$^+$CD41$^-$ ECs becoming round while expressing CD41$^+$ (as shown by CD41 staining after overnight imaging), therefore demonstrating the EHT (Supplementary Movie 4 and Extended Data Fig. 5). At E16/E17, a small definite fraction (0.1 to 1%) of GFP$^+$ blood cells was found circulating in line with the presence of round GFP$^+$CD41$^+$ or GFP$^+$CD45$^+$ cells in the contralateral left wing and hind limb BMs. In the grafted limbs, GFP$^+$ cells represented an average of 1.10±0.53% of the total BM mononucleated cells (Fig.1k) that were found unevenly distributed among the endothelial and haematopoietic lineages. Strikingly, the GFP$^+$ cells significantly contributed to the BM haematopoietic production, representing 13.7±7.3% of the CD41$^+$ cells, 3.0±1.3% of c-Kit$^+$ cells, 7.4±2.9% of CD3$^+$ cells and also to 14.2±5.1% of the MEP21$^+$ population and 0.97±0.52% of CD45$^+$ cells (Fig. 1l). On the other hand, 42.9±5.4% of the GFP$^+$ cells were MEP21$^+$ endothelial/haematopoietic cells whereas 32.4±10.4%, 28.4±8.7%, 17.2±4.7% and 5.0±2.8% of the GFP$^+$ cells were CD45$^+$, CD41$^+$, c-Kit$^+$ or CD3$^+$, respectively (Fig. 1m and Extended

Data Fig. 6a, b). Moreover, the MEP21^{+}CD45^{+} population, known to be enriched in multipotent progenitors[17] and representing 2.4% of the mononucleated cells, harboured 15.2% of GFP^{+} cells (Fig. 1n). Of note, no obvious difference was found in the haematopoietic composition of the BM between grafted (Extended Data Fig. 6a, b) and WT embryos (Extended Data Fig. 6c, d).

Secondary haematopoietic organs were colonized by GFP^{+} cells. For example, GFP^{+} cells were visible in the thymic lobules of E16/E17 grafted recipients. GFP^{+} cells were CD4^{+} (1.4%), indicative of T cell differentiation (Extended Data Fig. 7a-e). Contrary to BM, the vascular network of the thymus was not made by GFP^{+} cells (Supplementary Movie 5). A rather small, but clear GFP^{+} population was also present in the spleen (Extended Data Fig. 7f-h) and the bursa of Fabricius (Fig. 7i-k) therefore supporting the multilineage contribution by the BM endothelium-derived haematopoietic cell population.

GFP chimeras were also examined at different time points post-hatching (up to 41 weeks, Fig. 2a, b and Supplementary Table 1) and analysed for the presence of GFP^{+} cells associated with endothelial and haematopoietic lineages in the BM (n=7). The continued presence of GFP^{+} cells in the BM (Fig. 2c) and a rather similar distribution of GFP among the endothelial and haematopoietic fractions (Fig. 2d, e and Extended Data Fig. 8a-d) compared to E16/E17 embryos (Fig. 1l, m and Extended Data Fig. 6) was found in the growing chicken. The MEP21^{+}CD45^{+} population harboured 6.1% of GFP^{+} cells (Fig. 2f). FACS sorting followed by cytospin and May-Grünwald Giemsa staining of the GFP^{+}MEP21^{-} CD45^{-} cell fraction disclosed an erythroblast/erythrocyte population (Fig. 2g) whereas the GFP^{+}MEP21^{-}CD45^{+} fraction displayed eosinophils, heterophils, monocytes and thrombocytes (Fig. 2h).

To assess whether the process of haematopoietic production by somite-derived BM haemogenic ECs disclosed in Birds were conserved in Mammals, two mouse models were used. The contribution of the somite to the perinatal BM was traced by using Pax3$^{CRE/+}$ mice crossed to ROSA$^{mT/mG}$ mice. We first assessed that no Pax3^{+} cells were present in the embryonic aorta at E10.5, when the first HSPCs emerge (Extended Data Fig. 9a). Similar to the chicken, endothelial and haematopoietic cells associated with GFP expression were found in E19 to E21 foetal BM. The presence of mature cells from different haematopoietic lineages (CD45^{+}, CD3^{+}, B220^{+}, Gr1^{+}, CD11b^{+}, Ter119^{+}), HSPCs (Sca1^{+}c-Kit^{+}, CD144^{+}CD45^{+}) and endothelial cells (CD144^{+}CD45^{-}) were analysed in the perinatal mouse BM (Extended Data Fig. 9b). Moreover, the distribution of GFP^{+} cells in each population was measured (Extended Data Fig. 9c) (n=3). GFP^{+} cells contributed to all lineages. However, the highest contribution was found in the CD3^{+} T cell population, the CD144^{+}CD45^{+} population which is similar to the first HSPCs emerging from the aorta[22], and the Sca1^{+}c-Kit^{+} population (Extended Data Fig. 9c-f). These results were significant compared to control samples (Extended Data Fig. 9g). BM mononucleated cells contained an average of 0.30±0.02% GFP^{+} cells (Extended Data Fig. 9d) in line with the numbers found in birds (Fig. 2c and 1k), with 4.2±1.2% of CD144^{+}CD45^{+} cells expressing GFP (Extended Data Fig. 9c and e). The rare Sca1^{+}c-Kit^{+} population (known to harbour HSPCs) displayed a conspicuous fraction of GFP^{+} cells (10.0±0.3%) (Extended Data Fig. 9f), indicating the production of HSPCs by somite-derived BM endothelium.

To further specifically trace the progeny of ECs in the long term after birth, we used an indelible marking system to target the endothelium. For this purpose, floxed reporter YFP transgene mice[23] were crossed to mice with a tamoxifen-inducible Cre driven via the complete endothelial- specific VE-cadherin (CDH5) upstream regulatory sequences[24]. Day (D)1 pups received one injection of 50µg tamoxifen every day during 3 days (Fig. 3A). BMs were then analysed at different time points up to 15 weeks (105 days) for the presence of YFP^+ endothelial and haematopoietic cells (n=2- 6). Importantly, no $CD45^+$ cells displayed YFP expression immediately following the last tamoxifen injection ruling out contamination from haematopoietic progenitors expressing CD144 (Extended Data Fig. 10a). D27 BM analysis uncovered the presence of 0.096±0.011% YFP^+ cells in total mononuclear cells (Fig. 3b). Haematopoietic cells (4.1±0.4% of $CD144^+CD45^+$ and 0.0054±0.0013% of $CD144^-$ $CD45^+$) and ECs (42.4±6.8 % of $CD144^+CD45^-$) were YFP^+ (Fig. 3c). This staining is accurate regarding the absence of YFP staining observed in the non-induced animals (Extended Data Fig. 10b). Following-up the total number of YFP^+ cells, we observed that $CD144^+CD45^-$ ECs and the rare lineage-negative $Sca1^+c\text{-}Kit^+$ (LSK) HSPCs decreased over time to reach almost zero 105 days after birth, indicating a progressive disappearance of the haemogenic endothelium and haematopoietic production with age (Fig. 3d). Of note, inductions at later stages after birth (n=3-7) resulted in a reduced number of YFP^+ CD144+CD45- cells and LSK cells, confirming the progressive loss of haemogenic endothelium with age (Extended Data Fig. 11). Aiming to better identify the immature YFP^+ cell population, we analysed LSK cells using the SLAM family markers known to distinguish HSCs from MPPs[25] (Fig. 3e). We disclosed the presence of 7.8±1.6% of YFP^+ cells in the LSK population (Fig. 3e). YFP^+ cells were highly enriched in the $CD48^-$ $CD150^-$ MPP fraction[26] (23.1±4.7%) and also present in the $CD150^+CD48^-$ HSC fraction (5.0±1.6%) (Fig. 3e, f). Further analysis using CD229 and CD244 markers[25] to hierarchically discriminate the different haematopoietic progenitor populations revealed the presence of 65.3±8.9% of YFP^+ cells in the $CD150^-CD48^-$ $CD229^+CD244^+$ MMP3 fraction that represents transiently reconstituting haematopoietic progenitors and to a less extent in the MMP1 population that represents intermediate and transiently reconstituting progenitors (5.9±4.2%, Fig. 3g). Again, this staining was accurate when compared to the non-induced animals (Extended Data Fig. 10c).

To identify the HCs originating from a VE-cadherin ancestor at the molecular level, we performed transcriptomic analysis on YFP^- and YFP^+ LSK cells isolated from D27 BM. Principal component analysis (PCA) with the entire set of 30,922 genes separated YFP^+ from YFP^- LSK cells (Extended Data Fig. 12a and Table S3). Gene Ontology (GO) using database for annotation, visualization, and integrated discovery (DAVID, http://david.abcc.ncifcrf.gov) performed on differentially expressed genes (DEG) up-regulated in the two populations (i.e. 1,417 genes in the YFP^+ and 1,552 genes in the YFP^- LSK fraction, p-value= 0.03), disclosed differentially represented GO categories (Extended Data Fig. 12b). While GO characteristics of YFP^- LSK cells focused on transcription, translation and DNA repair, GO of the YFP^+ LSK cells were more diverse and included membrane, extra-cellular matrix (ECM), focal adhesion and actin-binding molecules. To more accurately determine the identity of the YFP^+ LSK cells, we included transcriptomes of several BM and AGM-derived endothelial and HC types (i.e. BM lin^+ HCs, BM $CD150^+CD48^-$ LSK HSCs, AGM $CD144^+CD45^-$ ECs, AGM $CD144^-CD45^+$ HCs and

AGM CD144^{+}CD45^{+} HSPCs to the YFP^{+} and YFP^{-} LSK cell transcriptomes, n=23 samples). PCA analysis on the whole gene set (18,920 genes) revealed three main groups. The first encompassed BM and AGM HCs, the second comprised adult and embryonic HSPCs including YFP^{-} LSK cells while the third group encompassed AGM ECs and YFP^{+} LSK cells (Fig. 4a, upper panel). Further analysis of DEGs identified by ANOVA (p-value= 0.001) on the 23 samples as observations and the gene set of 2,056 DEGs (986 up; 1,070 down) confirmed that the first principal component highlighted the difference between YFP^{+} and YFP^{-} LSK cells (Fig. 4a, lower panel). Hierarchical clustering (Extended Data Fig. 12c) now identified 4 groups that are circled in Fig. 4a (lower panel). To find out how DEGs are inter-related within the two populations, we resorted to weighted gene correlation network analysis (WGCNA)[27], which disclosed two major gene modules (sub networks). The magenta and black modules corresponded to genes up-regulated in YFP^{-} and YFP^{+} LSK cells, respectively (Fig. 4b). GO analysis of the two gene sets revealed strong differences in their biological functions. The magenta network contained genes implicated in cell proliferation (cell cycle, DNA replication, protein translation, mRNA splicing). The black module was enriched in molecules implicated in ECM, actin cytoskeleton and ECM-to-actin cytoskeleton interaction (Fig. 4c). Close-up view of the genes within the cyan network confirmed the presence of a core of transcription factors including key regulators of haematopoiesis, such as *Hoxa9, Meis1* and *Etv6* (Fig. 4d). The presence of several Notch genes, from canonical (Notch4, Dll4) and non-canonical (Mamld1), pathways and many endothelial genes such as *Pecam1*, *Kdr, Flt1, Flt4* and *Cdh5, Aplnr* (Fig. 4d) is a compelling argument for the endothelial origin of YFP^{+} LSK cells. We finally tested the HSC potential of the YFP^{+} and YFP^{-} LSK cells using transplantation experiments into irradiated recipients. YFP^{-} LSK cells contributed to a robust, multilineage haematopoiesis (Fig. 4e). Remarkably, YFP^{+} LSK cells also contributed, albeit to a less extend, to multilineage haematopoiesis with a preferential differentiation towards the lymphoid and erythroid lineages (Fig. 4e, f and Extended Data Fig. 13).

In conclusion, our cell tracing studies in both chicken and mouse species have identified a new wave of haematopoiesis (MMP3 and HSCs) occurring *de novo* in the BM of late embryonic/newborn and deriving from resident haemogenic ECs of somite origin. It is well known that the haematopoietic cells originating from the yolk sac will colonize the foetal liver with definitive erythroid, megakaryocyte, macrophage, neutrophil, and mast cell progenitors[28]. These progenitors will be sufficient for the mouse embryo to survive until birth, when the HSC-dependent wave will take over. One may speculate that the new transient haematopoietic production discovered in our study fills the gap between the end of the yolk sac haematopoiesis and the bone marrow HSC-dependent production. Interestingly, a transient HSC population harbouring robust adult haematopoietic reconstitution potential was recently uncovered during mouse embryonic life. This precise HSC population was shown to be lymphoid biased[29]. This, together with our finding reveals an unappreciated complexity of the ontogeny of the haematopoietic system. Indeed, the pool of HSCs that expended in the foetal liver starts to colonize the bone marrow at E17[30]. Since at birth bone marrow colonization is low and HSCs are present in very few numbers, some time might be required before HSCs find their final adult-type niches and start to differentiate and proliferate into more committed progenitors and mature blood cells. Alternatively, this transient haematopoietic wave might prepare the bone marrow niches to mature to an appropriate

state to accommodate the HSCs generated in the foetal liver. Such hypotheses will need to be verified in the future. A recent paper[31] indicates that yolk sac-derived erythromyeloid progenitors may contribute to intraembryonic vascular endothelium where they persist until adulthood, therefore unraveling an unexpected source of ECs. In our approaches, PSM does not contain erythromyeloid progenitors hence ruling out a contribution of this cell type to HSPC formation by the BM. A model of cell origins and lineage relationships in both chicken and mouse species is displayed in Fig. 5. Overall, our study proved that HSPCs could be produced through a haemogenic endothelium intermediate past the embryonic aortic and yolk sac stages, an important knowledge that should help designing novel therapies for haematopoietic disorders.

Methods

Antibodies for immunohistochemistry and flow cytometry.

See Extended Data Table 2.

Grafts of avian presomitic mesoderm (PSM).

Wild-type and transgenic GFP chicken (Gallus gallus) eggs were incubated at 38°C ± 1 °C in humidified atmosphere until embryos reached the appropriate stage established using Hamburger and Hamilton development tables[32]. GFP chicken were generated and kindly provided by the Roslin Institute[16]. An artificial dark field was obtained by injecting Indian ink, diluted 1:1 in phosphate-buffered saline (PBS), beneath the host embryos. Microsurgery was performed on the right side of the host. The PSM attached to the last formed somite was removed over a length corresponding to 5-10 somites. PSMs were rinsed in DMEM (PAA Laboratories)/10% foetal calf serum (FCS, Eurobio) and transplanted into a host submitted to the same ablation. Grafts were performed according to the original dorso-ventral and antero-posterior orientations. The chimeric embryos were incubated and analysed at several time points as indicated in Extended Data Table1.

After hatching, PSM grafted chickens were housed at the Gemeenschappelijk Dierenlaboratorium facility (Utrecht, Netherlands) until they reached adulthood. Animals were housed according to institutional guidelines, and procedures were performed in compliance with Standards for Care and Use of Laboratory Animals, with approval from the Dutch Animal Experiment Committee.

Chicken cell staining and analysis.

Chicken BM cells were flushed with PBS containing 5% FCS (Gibco) (PBS/FCS). The suspension was washed and passed through successive gauge needles (21, 23 and 26G, Terumo) to obtain a single cell suspension. Thymus, spleen and bursa of Fabricius were crushed; cell suspensions were filtered through a 40µm Nylon Cell Strainer (Falcon), and then washed twice in PBS/FCS.

Cell suspensions were then incubated with antibodies for 20 minutes on ice. Cells were washed and suspended in PBS supplemented with DAPI or 7-AAD (to exclude dead cells). Stained cells were analysed on a MACSquant (Miltenyi Biotec) or FACScalibur (Becton Dickinson).

Immunostainings on whole bone marrow, thymus, thick embryo slices and PSM.

Bone marrow and thymus were prepared as previously described in[31]. Tissues were stained with anti-GFP, anti-CD45 and anti-MEP21 antibodies (described above). A goat anti-rabbit IgG (H+L) coupled to Alexa Fluor 488, a goat anti-mouse IgG2a coupled to Alexa Fluor 555 and a goat anti-mouse IgG1 coupled to Alexa Fluor 647 (all from ThermoFisher Scientific) were used to reveal the anti-GFP, anti-CD45 and anti-MEP21 antibodies, respectively. Thick slices and dissected PSM were stained with anti-CD45 and anti-MEP21 antibodies. PSM were counterstained with DAPI. Transparent bone marrow, thymus, thick slice and PSM were imaged with a Zeiss LSM700 confocal microscope with a

10xPlanApo dry objective. Three-dimensional reconstructions were generated from z-stacks with Imaris software and converted to QuickTime files (.Mov).

Time-lapse confocal imaging of the chicken grafted bone marrow.

Transversal slices of E16 PSM grafted bone marrows were processed for time-lapse confocal imaging as described in[8,33]. Briefly, E16 grafted bone marrows were gently flushed out of the bones. Transversal slices (200µm thickness) of the marrows were obtained using a tissue chopper and stained with CD41-PE (clone 11C3, Bio-Rad) antibody for 20 minutes on ice. After 3 washes in PBS, slices were placed in a culture chamber, embedded in a 1% low melting point gel agarose. Five to 6 slices per experiments were over-night sequentially time-lapse imaged (every 20 minutes) on a Zeiss LSM700 confocal microscope. The following day, slices were stained again with CD41-PE (clone 11C3, Bio-Rad) antibody, washed and imaged to assess the emergence of new HSPCs during the overnight imaging.

Mice.

Experiments were carried out in accordance with the guidelines of the French Veterinary Department. Approval was obtained from the French Ministry of Agriculture institutional review board for these studies. ROSA26 mT/mG reporter mice contain a double-fluorescent Cre reporter that expresses membrane-targeted tandem dimer Tomato (mT) prior to Cre-mediated excision and membrane- targeted green fluorescent protein (mG) after excision[23]. ROSA26-stop-EYFP mutant mice have a *loxP*-flanked STOP sequence followed by the Enhanced Yellow Fluorescent Protein gene (EYFP) inserted into the *Gt(ROSA)26Sor* locus[34]. Pax3-Cre mice were generated by replacing the first Pax3- coding exon with a gene encoding Cre recombinase. Mice were grown as heterozygous[35]. Tg(Cdh5Cre/ERT2) transgenic mouse line was generated by insertion of a tamoxifen-inducible Cre (CreERT2) into a large PAC clone of the *Cadh5* (VE-Cadherin) gene resulting in endothelium-specific expression[24].

Pax3-mTmG mice generation

Homozygous mT/mG females were mated with heterozygous Pax3-Cre males. Foetuses were collected between E19 and E21, phenotyped based on the GFP expression in the central nervous system and analysed for their BM content.

Cdh5-YFP mice generation and induction

Homozygous ROSA-YFP females were mated with homozygous Tg(Cdh5-Cre/ERT2) males. After birth, inductions were made by intra-peritoneal injections of tamoxifen (T5648, Sigma) diluted in corn oil (C8267, Sigma).

Mouse bone marrow analysis and cell sorting.

Femurs and tibias were flushed with DMEM (Gibco) supplemented with 10% FCS (Eurobio). Cells were dissociated with Collagenase type 1 (C0130, Sigma) 1.25mg/ml in PBS+10% FCS for 30 min at

37°C, and washed twice with DMEM+10% FCS. 1 to 5.10^6 cells were stained with fluorescent antibodies (Extended Data Table 2) in PBS+0.5% BSA for at least 30 min at 4°C. Cells were washed, and suspended in PBS+0.5% BSA supplemented with 1.25 µg/ml DAPI (D9542, Sigma) for cell viability and analysed on a MACSQuant VYB (Miltenyi Biotec) or a LSR II (Becton-Dickinson) flow cytometer. Data were reanalysed with FlowJo X software.

For cell sorting, after collagenase treatment and washing, red blood cells were lysed using Gey's solution for 5 min at 4°C. After two washes with DMEM+10% FCS, Lin depletion was performed by incubation of cells with purified Gr-1, Ter119, B220, CD11b, CD4 and CD8 antibodies in PBS+0.5% BSA for 30 min at 4°C. After a wash, cells were incubated with Goat anti-rat microbeads (Miltenyi Biotec) for 30 min at 4°C. After washing, cells were applied on a LD column (Miltenyi Biotec) according to the manufacturer's guidelines and eluted Lin- cells were collected. Lin- cells were then stained with Sca1-PE and CD117(c-Kit)-PE-Cy7. LSK YFP^+ and YFP- cells were sorted using a FACS Aria III (Becton-Dickinson) cell sorter.

Analysis of transplantation capabilities.

Adult C57Bl/6-CD45.1 mice were exposed to a lethal single dose of 7.3 Gy (1.04 Gy/mn) from a X-ray generator (320 kV, Philips). 2,000 LSK YFP^+ or YFP^- (C57Bl/6-CD45.2) sorted cells were co-injected intravenously into the retro-orbital venous plexus with 2.10^5 normal C57Bl/6-CD45.1 bone marrow cells. Peripheral blood was collected at 2, 3, 4, 7 and 11 weeks after engraftment and nucleated cells were incubated with APC anti-mouse CD45.1, and PE anti-mouse CD45.2 (Biolegend, Ozyme), and analysed by flow cytometry for the presence of donor-derived ($CD45.2^+$) cells. Mice were sacrificed at 14 weeks and BMs were analysed for the grafting potential.

Transcriptomic analysis

Sorted LSK YFP^+ and YFP^- cells were lysed with 350µL RLT+ buffer (RNeasy Plus Micro kit, Qiagen) and conserved at -80°C. All RNA isolations were performed at the same time, according to the manufacturer's guidelines.

Affymetrix Mouse Gene 2.0 ST arrays were used. RNA concentration and integrity were evaluated with the Agilent bioanalysor 2100 (Genomi'C Facility platform, Cochin Institute, Paris). To correct from probe set definition inaccuracy, we used the version 17 of the custom Chip Definition File (CDF).

Design of the screening strategy and microarray analysis

Non-supervised analyses were performed on the global transcriptome of 6 samples (3 YFP^+ and 3 YFP^- replicates) and data were represented with Principal Component Analysis (PCA). Gene up-regulated in YFP^+ vs YFP^- cells were selected using ANOVA. Selected gene sets were then processed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (http://david.abcc.ncifcrf.gov)[36,37] to identify representative biological activities of YFP^+ vs YFP^- LSKs. For PCA, hierarchical clustering and other statistical analyses, we used Partek and JMP softwares. To find how our neonate bone marrow samples were positioned in relation to E11.5 AGM and 3 week-old bone marrow samples, we merged the present data with previously generated AGM and adult bone

marrow transcriptomes Dr. M. Souyri, unpublished; available upon request after publication). The AGM samples consisted in cells sorted according to the expression of CD144 (CDH5), CD34, CD45 (PTPRC) and c-Kit (CD117), enabling to discriminate immature HCs, broadly equivalent to HSC ($CD144^{med}CD45^{+}$ or $CD34^{+}$c-Kit^{+}) from mature HCs ($CD144^{med}CD45^{+}$ or $CD34^{med}$c-Kit^{med}) and endothelial cells ($CD144^{+}CD45^{-}$). The bone marrow samples consisted in cells sorted according to the expression of CD150 (SLAMF1), c-Kit, Sca-1 and markers of the different haematopoietic lineages (Lin), enabling to discriminate immature HCs, broadly equivalent to HSC ($CD150^{+}$c-Kit^{+}Sca-$1^{+}Lin^{-}$) from mature HCs (Lin^{+}) (Available upon request. A new matrix was therefore generated corresponding to 28 samples and including 18,920 genes. After removal of the batch effect (3 different experiments were merged), PCA was carried out. ANOVA was then performed contrasting YFP^{+} vs. YFP^{-} cells with $p < 0.001$, resulting in 2,479 differentially expressed genes (DEGs), 1,123 up-regulated and 1,356 down-regulated. Since on the PC1 vs. PC2 score plot of the PCA YFP^{+} cells were clearly discriminated from YFP^{-} cells, we resorted to Weighted Gene Correlation Network Analysis (WGCNA)[27] to unravel the different gene networks. The parameters were as follows: power: 18, adjacency: signed, correlation: Pearson's, linkage: average, minimum module size: 30. The chosen weight threshold for network representation using Netscape was 0.1.

Reporting Summary

Further information on experimental design is available in the Reporting Summary.

Data availability

YFP^{+} and YFP^{-} cells sequence data used in this study have been deposited in the NCBI Gene Expression Omnibus database (accession number GSE122857). Source data for Figs. 1k-m, 2c-e, 3d and 4a–l and Supplementary Figs. 6a,c, 8a, 9b, c and 11 are provided in Supplementary Table 4. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

References

1 Sugimura, R. *et al.* Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature* **545**, 432-438, doi:10.1038/nature22370 (2017).

2 Lis, R. *et al.* Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature* **545**, 439-445, doi:10.1038/nature22326 (2017).

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

4 Zovein, A. C. *et al.* Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* **3**, 625-636 (2008).

5 Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E. & Speck, N. A. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* **457**, 887-891, doi:10.1038/nature07619 (2009).

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

7 Swiers, G., Rode, C., Azzoni, E. & de Bruijn, M. F. A short history of hemogenic endothelium. *Blood Cells Mol Dis*, doi:S1079-9796(13)00209-X [pii] 10.1016/j.bcmd.2013.09.005 (2013).
8 Boisset, J. C. *et al.* In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* **464**, 116-120, doi:nature08764 [pii] 10.1038/nature08764 (2010).
9 Mikkola, H. K. & Orkin, S. H. The journey of developing hematopoietic stem cells. *Development* **133**, 3733-3744 (2006).
10 Dzierzak, E. & Speck, N. A. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat Immunol* **9**, 129-136 (2008).
11 Pardanaud, L. *et al.* Two distinct endothelial lineages in ontogeny, one of them related to hemopoiesis. *Development* **122**, 1363-1371 (1996).
12 Pouget, C., Gautier, R., Teillet, M. A. & Jaffredo, T. Somite-derived cells replace ventral aortic hemangioblasts and provide aortic smooth muscle cells of the trunk. *Development* **133**, 1013-1022 (2006).
13 Yvernogeau, L., Auda-Boucher, G. & Fontaine-Perus, J. Limb bud colonization by somite-derived angioblasts is a crucial step for myoblast emigration. *Development* **139**, 277-287, doi:10.1242/dev.067678 (2012).
14 Ambler, C. A., Nowicki, J. L., Burke, A. C. & Bautch, V. L. Assembly of trunk and limb blood vessels involves extensive migration and vasculogenesis of somite-derived angioblasts. *Dev Biol* **234**, 352-364 (2001).
15 Christ, B., Huang, R. & Scaal, M. Amniote somite derivatives. *Dev Dyn* **236**, 2382-2396, doi:10.1002/dvdy.21189 (2007).
16 McGrew, M. J. *et al.* Efficient production of germline transgenic chickens using lentiviral vectors. *EMBO Rep* **5**, 728-733, doi:10.1038/sj.embor.7400171 7400171 [pii] (2004).
17 McNagny, K. M. *et al.* Thrombomucin, a novel cell surface protein that defines thrombocytes and multipotent hematopoietic progenitors. *J Cell Biol* **138**, 1395-1407. (1997).
18 Thornton, M. A. & Poncz, M. Characterization of the murine platelet alphaIIb gene and encoded cDNA. *Blood* **94**, 3947-3950 (1999).
19 Yvernogeau, L. *et al.* An in vitro model of hemogenic endothelium commitment and hematopoietic production. *Development* **143**, 1302-1312, doi:10.1242/dev.126714 (2016).
20 Lassila, O., Eskola, J., Toivanen, P. & Dieterlen-Lievre, F. Lymphoid stem cells in the intraembryonic mesenchyme of the chicken. *Scand J Immunol* **11**, 445-448 (1980).
21 Jaffredo, T., Gautier, R., Brajeul, V. & Dieterlen-Lièvre, F. Tracing the progeny of the aortic hemangioblast in the avian embryo. *Developmental Biology* **224**, 204-214, doi:10.1006/dbio.2000.9799 (2000).
22 Taoudi, S. *et al.* Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. *Development* **132**, 4179-4191, doi:dev.01974 [pii] 10.1242/dev.01974 (2005).
23 Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593-605, doi:10.1002/dvg.20335 (2007).
24 Sorensen, I., Adams, R. H. & Gossler, A. DLL1-mediated Notch activation regulates endothelial identity in mouse fetal arteries. *Blood* **113**, 5680-5688, doi:10.1182/blood-2008-08-174508 (2009).
25 Oguro, H., Ding, L. & Morrison, S. J. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, 102-116, doi:10.1016/j.stem.2013.05.014 (2013).
26 Kiel, M. J., Yilmaz, O. H., Iwashita, T., Terhorst, C. & Morrison, S. J. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109-1121 (2005).
27 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).
28 Palis, J. Hematopoietic stem cell-independent hematopoiesis: emergence of erythroid, megakaryocyte, and myeloid potential in the mammalian embryo. *FEBS Lett* **590**, 3965-3974, doi:10.1002/1873-3468.12459 (2016).
29 Beaudin, A. E. *et al.* A Transient Developmental Hematopoietic Stem Cell Gives Rise to Innate-like B and T Cells. *Cell Stem Cell* **19**, 768-783, doi:10.1016/j.stem.2016.08.013 (2016).
30 Christensen, J. L., Wright, D. E., Wagers, A. J. & Weissman, I. L. Circulation and chemotaxis of fetal hematopoietic stem cells. *PLoS Biol* **2**, E75 (2004).
31 Plein, A., Fantin, A., Denti, L., Pollard, J. W. & Ruhrberg, C. Erythro-myeloid progenitors contribute endothelial cells to blood vessels. *Nature* **562**, 223-228, doi:10.1038/s41586-018-0552-x (2018).
32 Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49-92 (1951).

33 Boisset, J. C., Andrieu-Soler, C., van Cappellen, W. A., Clapes, T. & Robin, C. Ex vivo time-lapse confocal imaging of the mouse embryo aorta. *Nat Protoc* **6**, 1792-1805, doi:10.1038/nprot.2011.401 (2011).
34 Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* **1**, 4 (2001).
35 Engleka, K. A. *et al.* Insertion of Cre into the Pax3 locus creates a new allele of Splotch and identifies unexpected Pax3 derivatives. *Dev Biol* **280**, 396-406, doi:S0012-1606(05)00075-8 [pii] 10.1016/j.ydbio.2005.02.002 (2005).
36 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
37 Huang da, W. *et al.* Extracting biological meaning from large gene lists with DAVID. *Curr Protoc Bioinformatics* **Chapter 13**, Unit 13 11, doi:10.1002/0471250953.bi1311s27 (2009).
38 Chevallier, A., Kieny, M. & Mauger, A. Limb-somite relationship: origin of the limb musculature. *J Embryol Exp Morphol* **41**, 245-258 (1977).
39 Christ, B., Jacob, H. J. & Jacob, M. Experimental analysis of the origin of the wing musculature in avian embryos. *Anat Embryol (Berl)* **150**, 171-186 (1977).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature

Acknowledgements: The authors apologize to those investigators whose important work we were unable to cite or describe due to space constraints. We thank D. Traver for critical reading of the manuscript, and S. Gournet for help in image preparation. We are indebted to R. Adams for his sharing of the Tg(Cdh5Cre/ERT2) transgenic mouse line and to S. Germain for providing us with the founders. We thank the Cell Imaging and Flow Cytometry facility of the IBPS (Paris, France) for access and technical support in confocal image acquisition. This work was supported by CNRS, UPMC, Fondation Les Treilles (LY), EMBO short term fellowship (LY) and Grants from FRM (DEQ20100318258) and ANR/CIRM joint grant (ANR/CIRM 0001–02) in Jaffredo lab. The production of the GFP^+ chicken embryos was supported by grants from the BBSRC and the Wellcome Trust. Part of the work and LY were supported by an European Research Council grant (ERC, project number 220-H75001EU/HSCOrigin-309361) and a TOP-subsidy from NWO/ZonMw (912.15.017) in Robin lab.

Author Contributions L.Y and T.J. conceived the project. L.Y. and T.J. designed experiments on chicken embryos, performed chicken somite grafts and interpreted data. L.Y. performed chicken grafts, made immunohistological analysis, performed FACS analysis. R.G. performed chicken graft analysis. L.P. performed all mouse work including transplantation studies and analysed the data. H.S. provided GFP transgenic chickens. V.R. and F.R. provided the Pax3CRE X Rosa mTmG mice and critical associated materials. H.K. analysed YFP^+ vs YFP^- transcriptome. P.C. analysed transcriptomes, performed network analyses and interpreted data. M.S. designed Pax3CRE X ROSAmTmG approaches and the initial VECAD-Cre ERT X Rosa 26 approaches, analysed the data and provided the additional mouse embryonic and BM adult transcriptomes. C.R. and L.Y. designed and supervised the experiments on adult chickens and the live imaging on section. TJ wrote the manuscript. L.Y. and C.R. gave significant inputs on data analyses, findings, interpretations and manuscript writing.

Author information

The authors declare no financial competing interests.

Correspondence and request for materials should be addressed to thierry.jaffredo@upmc.fr.

Figures

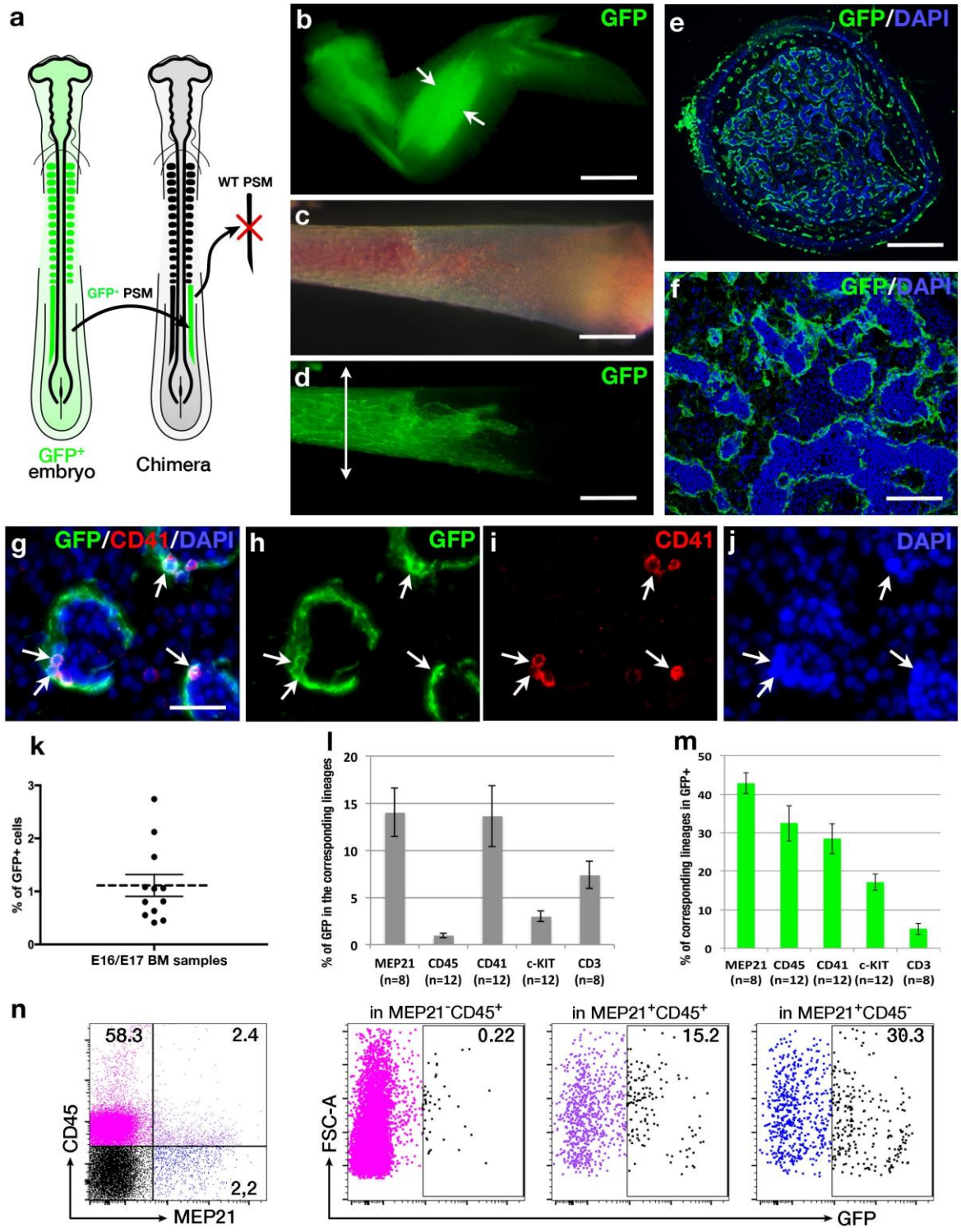

a
WT PSM
GFP+ PSM
GFP+ embryo
Chimera
b
GFP
c
d
GFP
e
GFP/DAPI
f
GFP/DAPI
g
GFP/CD41/DAPI
h
GFP
i
CD41
j
DAPI
k
% of GFP+ cells
E16/E17 BM samples
l
% of GFP in the corresponding lineages
MEP21 (n=8)
CD45 (n=12)
CD41 (n=12)
c-KIT (n=12)
CD3 (n=8)
m
% of corresponding lineages in GFP+
MEP21 (n=8)
CD45 (n=12)
CD41 (n=12)
c-KIT (n=12)
CD3 (n=8)
n
58.3
2.4
2,2
CD45
MEP21
in MEP21-CD45+
0.22
in MEP21+CD45+
15.2
in MEP21+CD45-
30.3
FSC-A
GFP

Figure 1. The chicken BM vascular network derives from the somite and produces a cohort of HSPCs during late embryonic life.

a, Scheme depicting the isotopic and isochronic graft of a PSM, the somite anlage, from a ubiquitous GFP transgenic chicken embryo into a wild-type chicken embryo host. PSM grafts were performed at E2, at the level of somites 14-21, which corresponds to the wing bud development[38,39]. **b**, The E16/E17 limb was colonized by GFP^+ somite-derived cells that formed visible muscle masses (arrows). **c, d,** Isolated radius showing conspicuous signs of BM haematopoiesis observed under visible (**c**) or UV illumination (**d**). GFP tubule-like structures were observed in the BM. The arrow indicates the level of section seen in (**e**). **e, f**, Representative cross section (arrow in (**d**)) through the bone stained for DAPI (cell nuclei) and GFP showing GFP^+ cells forming sinusoids. **g-j,** Representative fluorescent images showing the co-distribution of GFP^+ sinusoids (**h**) and $CD41^+$ HCs (**i**) (arrows), counterstained with DAPI (**j**). Merged picture is shown in (**g**). **k**, Percentage of GFP^+ cells in E16 grafted limbs. Each dot represents a grafted recipient. **l**, Distribution of GFP^+ cells in the endothelial and haematopoietic lineages. MEP21, endothelial/haematopoietic cell marker; CD45, pan-haematopoietic cell marker; CD41 and c-Kit, HSPC markers; CD3, lymphoid cell marker. **m**, Distribution of the endothelial and the different haematopoietic lineages within the GFP^+ population. Numbers in brackets under the x axis (**l** and **m)** are the numbers of grafted recipients. All data represent mean ± SEM from at least n=8 independent experimental animals. **n**, Representative flow cytometry analysis identifying the $MEP21^+CD45^+$ population known to be enriched in multipotent progenitors[17] and the contribution of the GFP^+ PSM-derived cells to the endothelial and/or haematopoietic (multipotent) progenitors. Scale bars in b, 1cm; in c, d, 300µm; in e, 200µm; in f, 80µm; in g, 15µm.

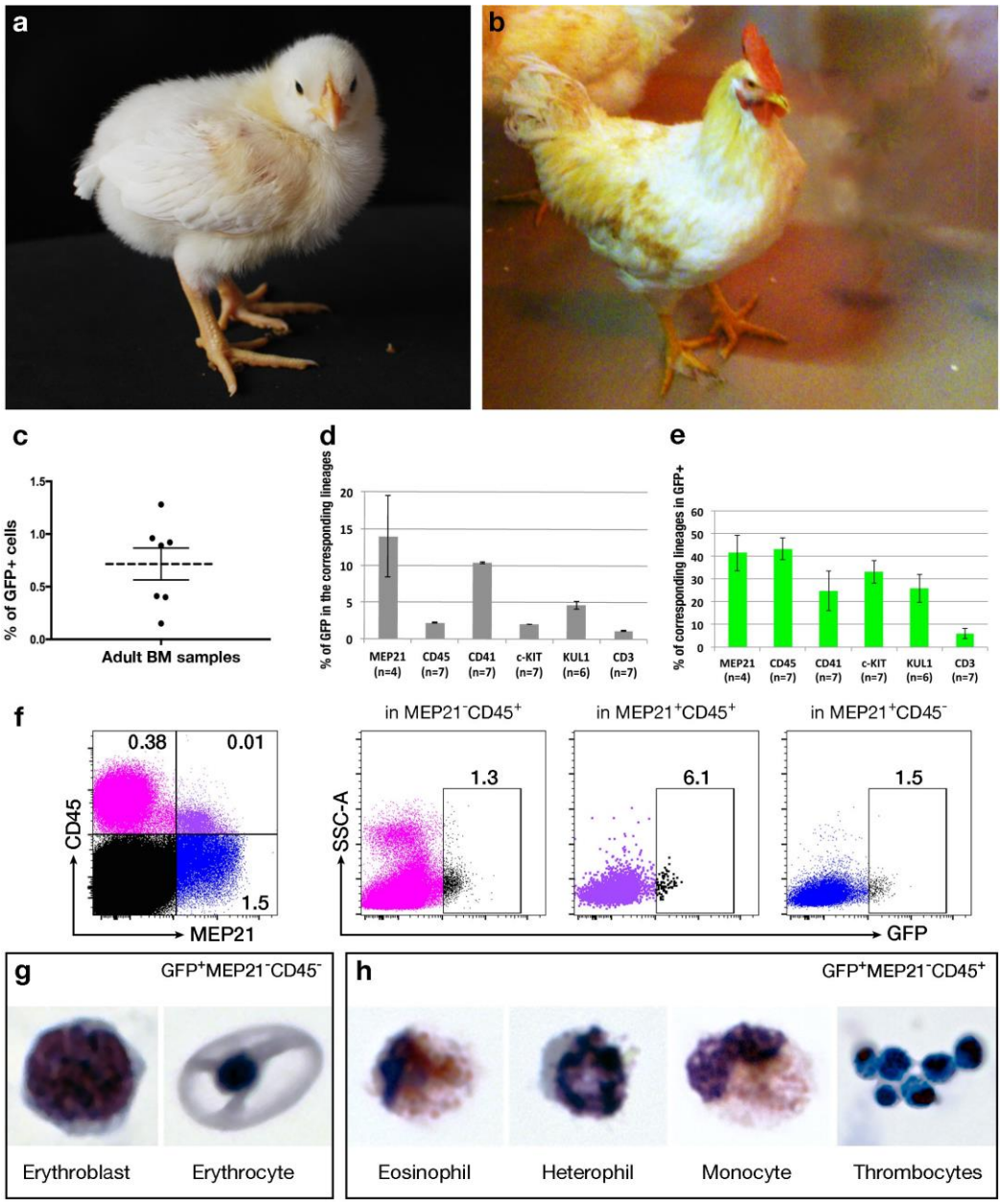


Figure 2. HSPCs derived from the late foetal BM endothelium persists after hatching.

a, b, PSM grafted chimeras at 8 days (**a**) and 5 months (b) post-grafting, respectively. **c**, Percentage of GFP$^+$ cells in grafted limbs at 5 months. Each dot represents a grafted recipient. **d**, Distribution of GFP$^+$ cells in the different BM endothelial and haematopoietic lineages of a 5.5 month-old recipient. MEP21, endothelial/haematopoietic marker; CD45, pan-haematopoietic marker; CD41, c-Kit, HSPC markers; KUL1, macrophage/monocyte marker; CD3, lymphoid marker. **e,** Distribution of the BM endothelial and haematopoietic lineages contained in the GFP$^+$ population. Numbers in brackets under the x axis (**d** and **e)** indicate the numbers of analysed transplanted recipients. All data represent mean ± SEM (n=4-7 independent experimental animals). **f**, Combined flow cytometry analysis of the adult grafted BM for MEP21 and CD45 populations showing the percentages of MEP21$^-$CD45$^+$, MEP21$^+$CD45$^+$ and MEP21$^+$CD45$^-$ populations and the contribution of GFP$^+$ cells within each

population. **g, h,** May-Grünwald Giemsa staining of sorted adult $GFP^{+}MEP21^{-}CD45^{-}$ (**g**) and $GFP^{+}MEP21^{-}CD45^{+}$ (**h**) cells.

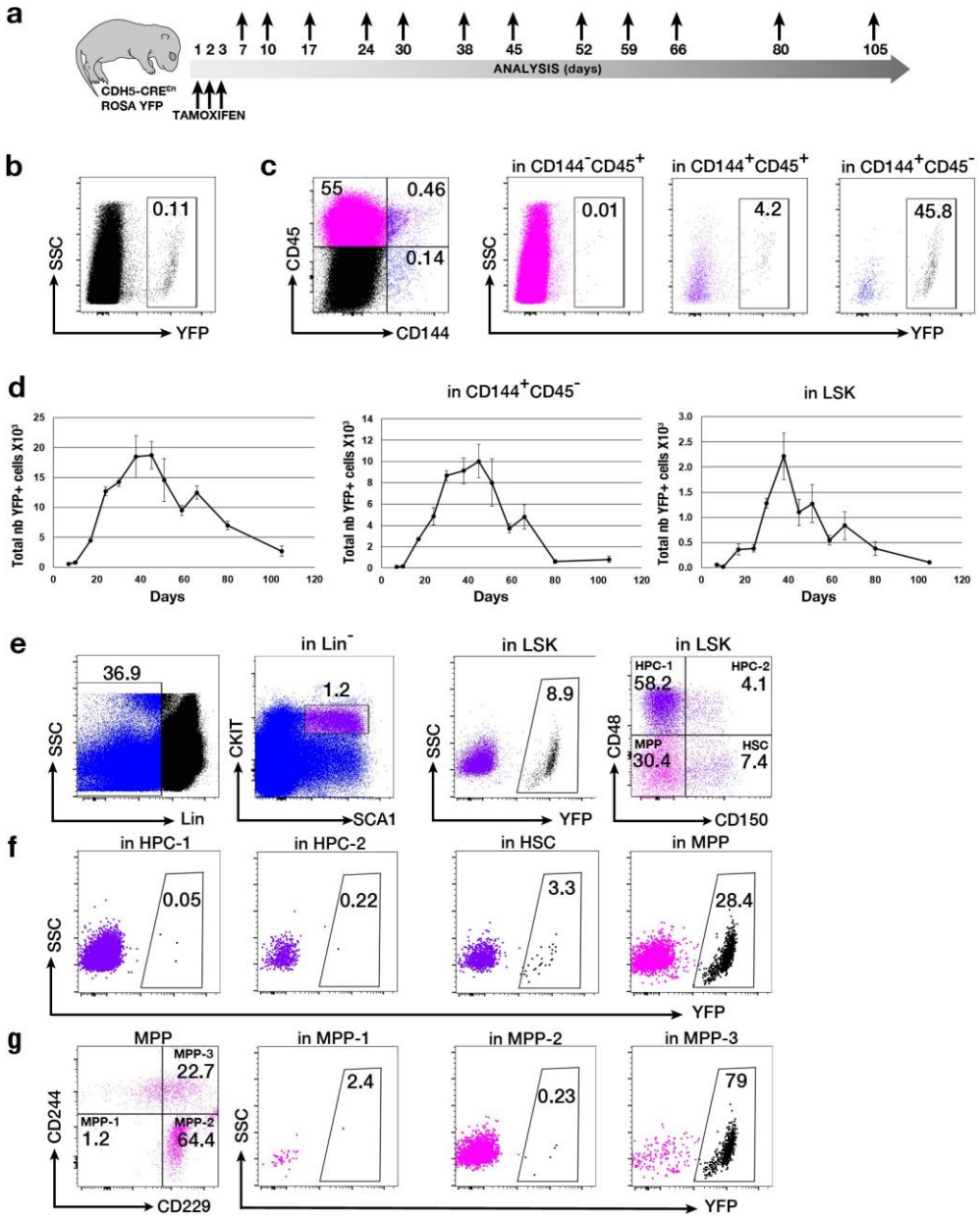


Figure 3. Conditional expression of YFP in VE-Cadherin cells in BM at birth discloses a contribution of vascular ECs to MPPs and Lin$^-$Sca1$^+$c-kit$^+$CD150$^+$CD48$^-$ HSCs.

a, Scheme showing the activation of VE-Cadherin (CD144, CDH5) by tamoxifen injection and the different time points of analysis. Bottom arrows, time of tamoxifen injections. Top arrows, time points of analyses. **b**, YFP expression in mononucleated BM cells at 27 days. **c**, Representative flow cytometry analysis of CD144$^-$CD45$^+$, CD144$^+$CD45$^+$ and CD144$^+$CD45$^-$ populations and YFP expression in each population. **d**, Time course analysis of total YFP$^+$ cells, and YFP$^+$ cells in CD144$^+$CD45$^-$ and LSK populations. n=2-5 individuals per time point. Error bars are mean ± SEM. **e**, **f, g,** Representative flow cytometry analyses of the SLAM/LSK cells (**e**), HSCs (**f**) and MPPs (**f, g**) revealed an enrichment of YFP$^+$ cells in the MPP3 fraction and HSCs.

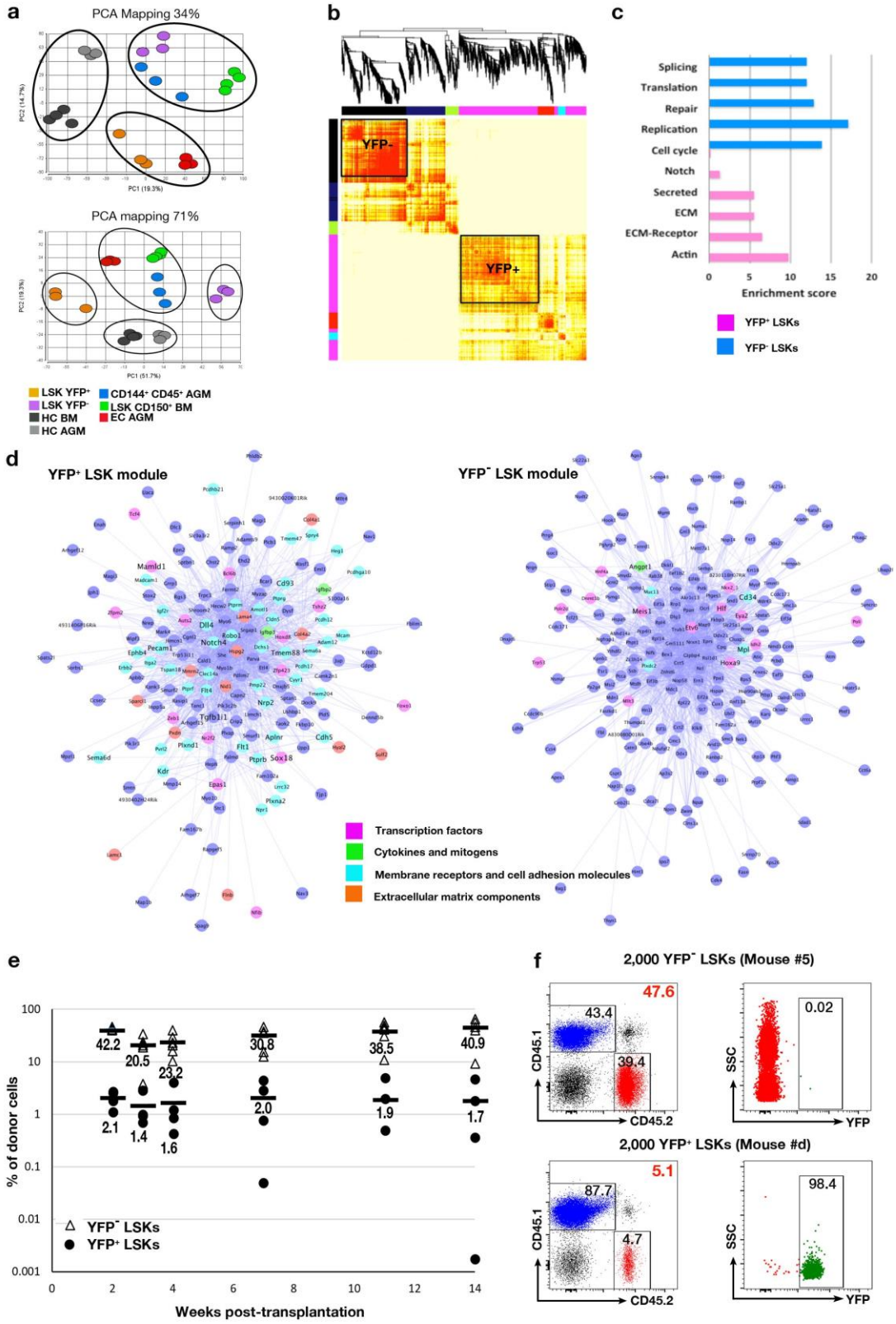

a
PCA Mapping 34%
PC1 (19.3%)
PC2 (14.7%)
PCA mapping 71%
PC1 (51.7%)
PC2 (19.3%)
LSK YFP+
LSK YFP-
HC BM
HC AGM
CD144+ CD45+ AGM
LSK CD150+ BM
EC AGM
b
YFP-
YFP+
c
Splicing
Translation
Repair
Replication
Cell cycle
Notch
Secreted
ECM
ECM-Receptor
Actin
0
5
10
15
20
Enrichment score
YFP+ LSKs
YFP- LSKs
d
YFP+ LSK module
YFP- LSK module
Transcription factors
Cytokines and mitogens
Membrane receptors and cell adhesion molecules
Extracellular matrix components
e
% of donor cells
100
10
1
0.1
0.01
0.001
42.2
20.5
23.2
30.8
38.5
40.9
2.1
1.4
1.6
2.0
1.9
1.7
YFP- LSKs
YFP+ LSKs
0
2
4
6
8
10
12
14
Weeks post-transplantation
f
2,000 YFP- LSKs (Mouse #5)
47.6
43.4
39.4
CD45.1
CD45.2
0.02
SSC
YFP
2,000 YFP+ LSKs (Mouse #d)
5.1
87.7
4.7
CD45.1
CD45.2
98.4
SSC
YFP

Fig. 4. BM vascular EC-derived HSPCs display an endo-haematopoietic transcriptomic profile and show long-term multilineage reconstitution.

a, PCA **(**upper panel): PC1 vs. PC2 score plot corresponding to the PCA of 23 samples as observations, and the whole gene set (18920 genes) as variables. PCA (lower panel): PC1 vs. PC2 score plot corresponding to the 23 samples as observations and the gene set of 2056 DEGs (986 up, 1070 down) as variables; the gene set was obtained by ANOVA (YFP^+ vs. YFP^-, $p < 0,001$, $f > 1$ or $f <$ -1); PC1 corresponds to the contrast between YFP+ and YFP- cells. **b**, Module detection by clustering tree cut. Same data as for PCA (lower panel). The parameters were as follows: power: 20; adjacency: signed; correlation: Pearson's; linkage: average; minimum module size: 30. Two major highlighted modules, black and magenta, are positively and negatively correlated to the YFP^+ to YFP^- cells contrast, respectively. **c**, Major GO categories corresponding to genes belonging to the black and magenta modules. Enrichment scores as given by DAVID database. Abbreviations: ECM-Receptor: ECM-receptor interaction; Actin: actin binding; ECM: extracellular matrix; Notch: notch binding/pathway; Replication: DNA replication; Repair: DNA repair; Splicing: mRNA splicing. **d**, Cytoscape connectivity plot of the black (up-regulated in YFP^+ cells) and magenta (up-regulated in YFP^- cells) modules. **e**, Comparison of levels of circulating blood cells analysed at different time points following engraftment between recipients transplanted with YFP^+ or YFP^- LSK cells. Empty triangles, YFP^- LSK cells grafted in six animals. Black circles, YFP^+ LSK cells grafted in 4 animals. Bars represent means. **f**, Flow cytometry analysis of representative C57Bl6-CD45.1 recipients injected with 2,000 YFP^- (upper plots) or YFP^+ (lower plots) LSK cells (C57Bl6-CD45.2), respectively (left plots). The right plots show YFP expression in the engrafted $CD45.2^+$ populations.

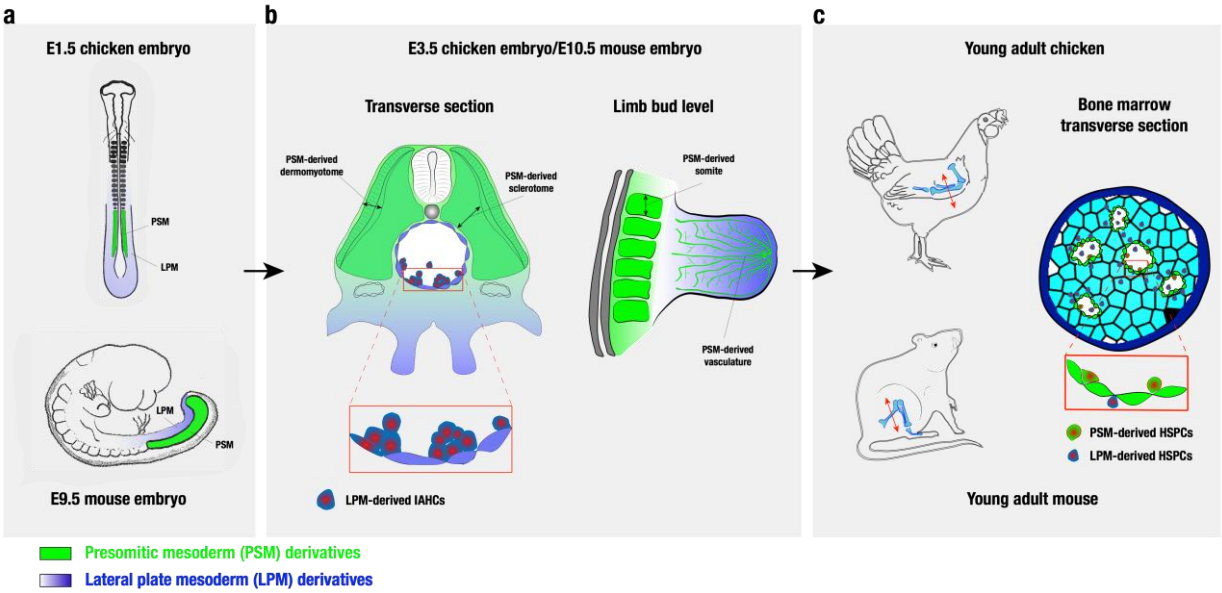
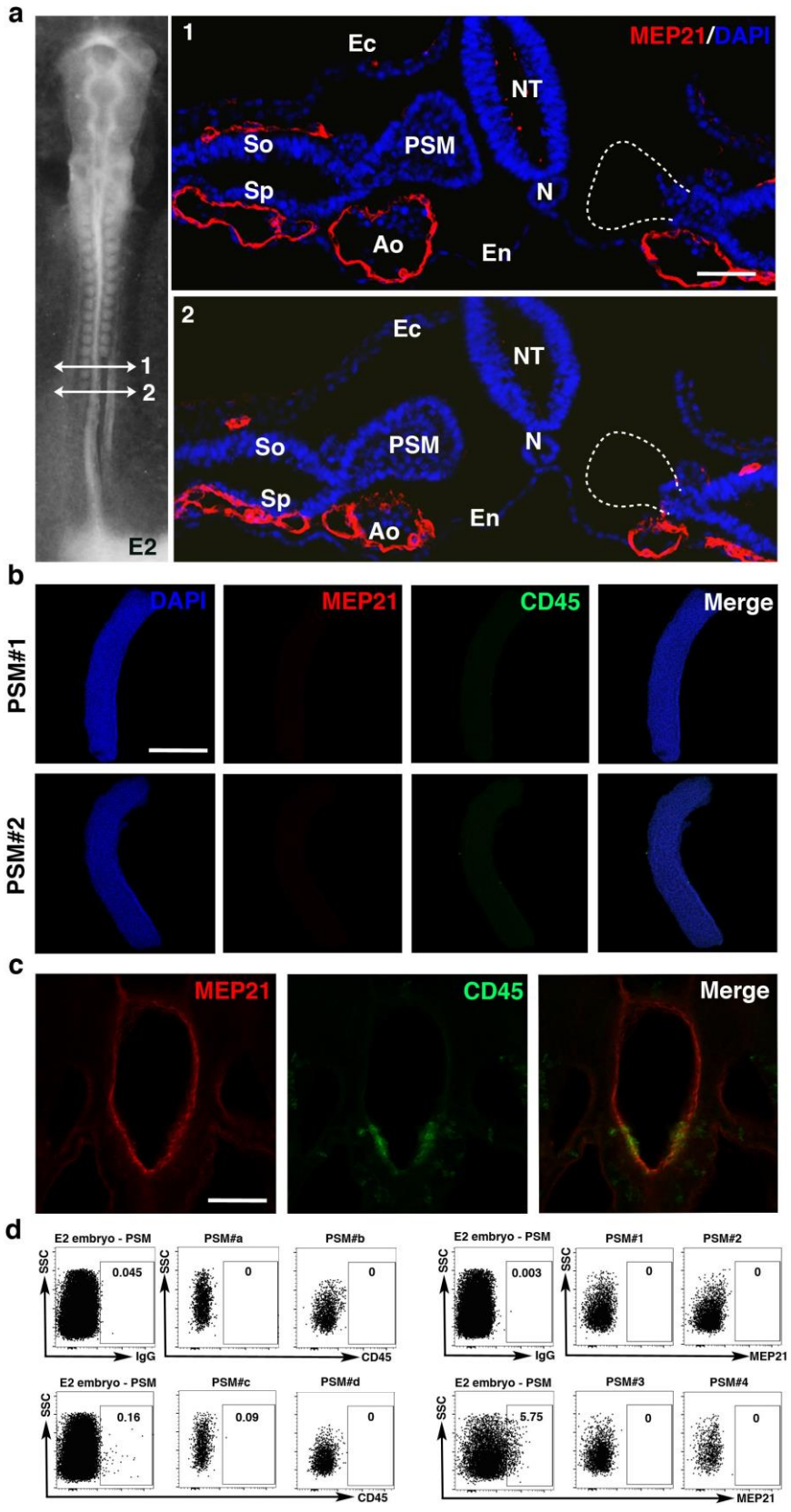


Fig. 5. Summary of presomitic and lateral plate mesoderm contribution to the endothelial and haematopoietic compartments of the bone marrow. a-c, Scheme summarizing the location and fate of presomitic mesoderm (PSM, green) and lateral plate mesoderm (LPM, blue) at three representative and equivalent time points of chicken and mouse development. **a**, Location of PSM and LPM in chicken (at E1.5, time point of graft isolation) and mouse (E9.5) embryos. PSM: unsegmented plate located at the caudal part of the developing embryo, on each side of the neural tube. LPM: mesoderm located laterally and ventrally to the PSM. **b**, The PSM progressively segments to form somites, which eventually differentiate into dermomyotome and sclerotome. They later give rise to muscles, vasculature and bones (i.e. from trunk, limbs). The haemogenic endothelial cells of the aorta and intra-aortic haematopoietic cluster (IAHC) cells derive exclusively from the LPM, which also gives rise to the limb mesenchyme. **c**, In young adults, bones and marrows are derived from the LPM while the vasculature is entirely derived from the PSM. Most haematopoietic stem and progenitor cells (HSPCs) come from LPM-derived HSCs (initially formed in the embryonic aorta). HSPCs are also *de novo* generated from the PSM-derived endothelium of the bone marrow (this study).

Extended Data Figures

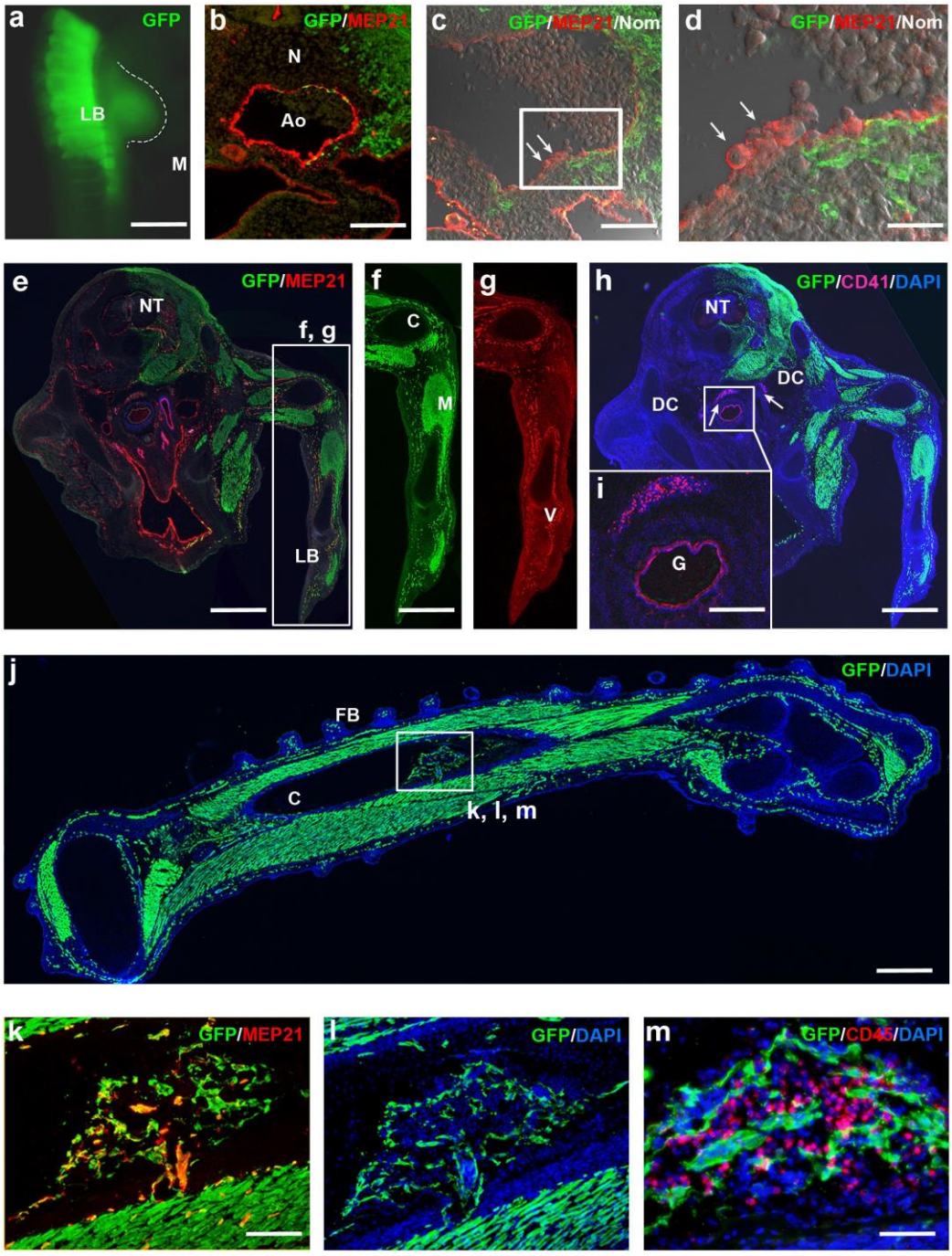

a
1
Ec
MEP21/DAPI
NT
So
PSM
Sp
N
Ao
En
2
Ec
NT
So
PSM
N
Sp
En
Ao
1
2
E2
b
DAPI
MEP21
CD45
Merge
PSM#1
PSM#2
c
MEP21
CD45
Merge
d
E2 embryo - PSM
PSM#a
PSM#b
SSC
0.045
0
0
IgG
CD45
E2 embryo - PSM
PSM#1
PSM#2
SSC
0.003
0
0
IgG
MEP21
E2 embryo - PSM
PSM#c
PSM#d
SSC
0.16
0.09
0
CD45
E2 embryo - PSM
PSM#3
PSM#4
SSC
5.75
0
0
MEP21

Extended Data Figure 1. PSM removal graft experiment.

a, E2 (15-somite stage, HH11/12) wild-type chicken embryo submitted to a PSM ablation. White arrows located the regions of transverse histological analysis at two representative anterior and posterior levels displayed in 1 and 2. MEP21 (red) and DAPI (blue) staining showed that only the PSM was removed without disturbing the surrounding tissues. **b**, Whole mounts of two PSMs isolated from E2 (15-somite stage, HH11/12) wild-type embryos and stained with DAPI (blue), MEP21 (red) and CD45 (green) antibodies, showing that PSMs do not contain any ECs or HCs after dissection. Two independent examples are shown (#1, #2). **c**, Mid trunk, thick transverse section of an E3 chicken embryo stained with MEP21 (red) and CD45 (green) antibodies shown as positive control of the immunostainings. Of note, PSMs (**b**) and thick slice (**c**) were stained at the same time, using the same antibody solutions. **d**, Representative FACS analysis on single PSMs isolated from 13-15 somite stage embryos showing no contamination by haematopoietic (CD45) or endothelial (MEP21) cells. Controls are two whole embryos (without PSMs) showing the presence of $CD45^{+}$ haematopoietic and $MEP21^{+}$ endothelial cells.

Ao, Aorta; Ec, Ectoderm; En, Endoderm; Nc, Notochord; NT, Neural Tube; PSM, Presomitic mesoderm; So, Somatopleura; Sp, Splanchnopleura.

Scale bars, 25µm in **a**; 50µm in **b** and **c**.

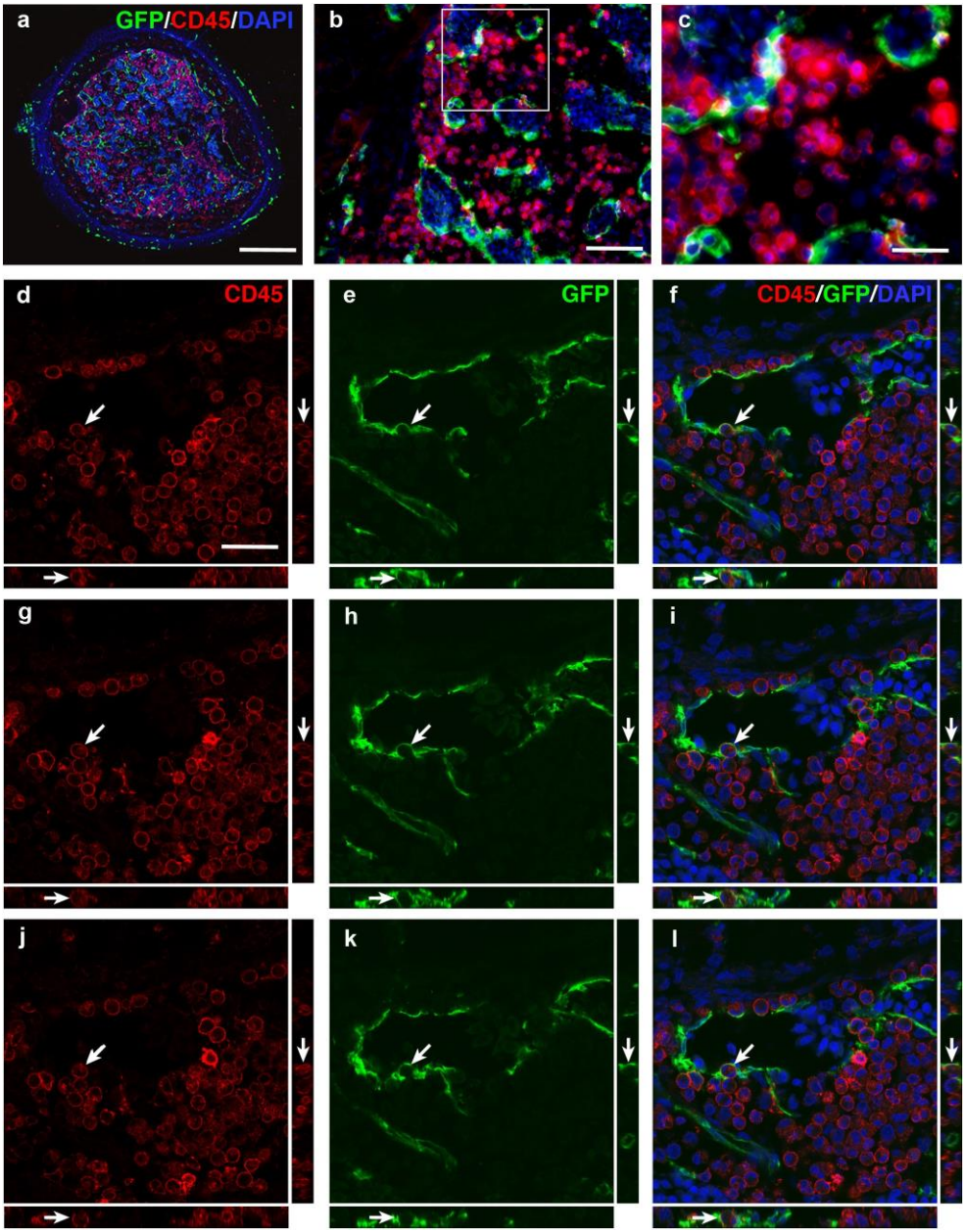


Extended Data Figure 2. Formation of the haematopoietic clusters of the aorta and vascularisation of limbs and bones by the somite-derived endothelial cells is not associated with the production of haematopoietic cells.

a, E4 grafted embryo showing GFP$^+$ somites and limb bud. **b-d**, Transverse section through the aorta of an E4 grafted embryo at the mid-trunk level. **b**, GFP and MEP21 merged signals. Overview of the aorta and the adjacent tissues. Some GFP$^+$ cells were integrated in the MEP21$^+$ endothelial roof of the aorta whereas some were scattered in the mesenchyme of the aortic floor in keeping with previous

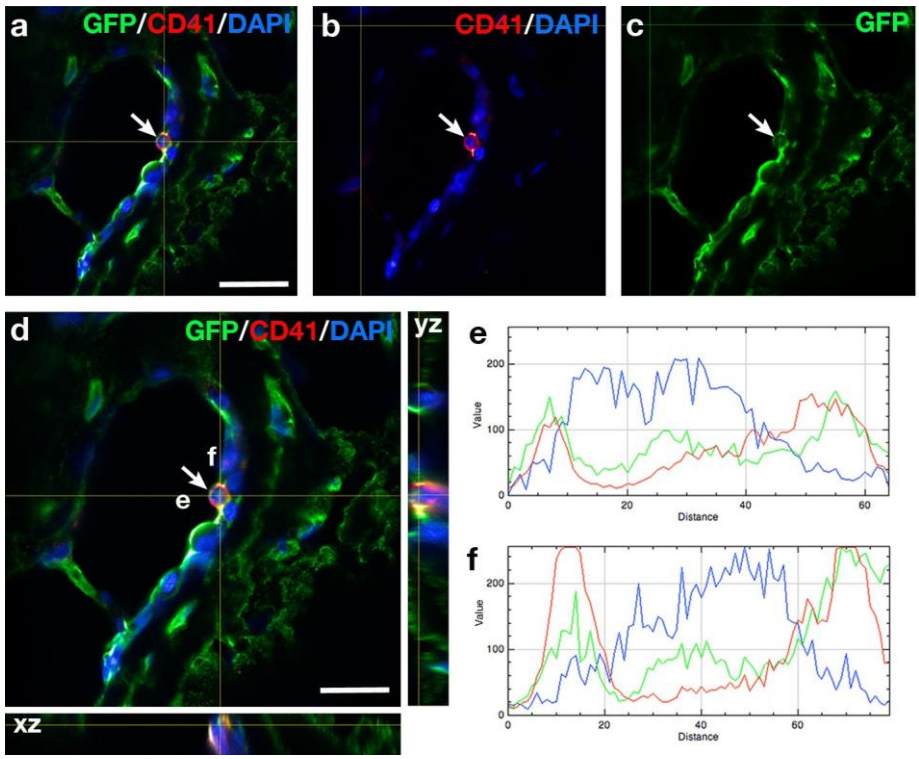
publications[11,12]. **c**, **d,** Immediate adjacent section of **b**. MEP21 and GFP immunohistochemistry and

Nomarski's interferential contrast merged signals. **c**, Close-up view of the ventral side of the aorta showing the haematopoietic clusters (white arrows) **d**, High magnification of the boxed area in **c** representing GFP^- haematopoietic clusters and the underlying GFP^+ cells originating from the grafted PSM. GFP^+ PSM-derived cells never contributed to the haematopoietic clusters in keeping with[12]. **e**, Transverse section through an E8 grafted embryo at the mid-trunk level. MEP21 (red) and GFP (green) double staining. **f, g,** Higher magnification of the frame shown in (**e**). GFP staining (**f**) and MEP21 vascular staining (**g**). Wing bud vessels displayed both MEP21 and GFP staining indicating that they originated from the graft. **h, i**, Transverse section through an E8 grafted embryo at the mid- trunk level. Triple staining for CD41 (red), GFP (green) and DAPI (blue). $CD41^+$ cells are found in the para-aortic foci (white arrows) close to the duct of Cuvier (DC) as described in[1]. **i**, higher magnification of the frame in (**h**). CD41 and GFP signals did not overlap indicating that, at that stage, intra- embryonic haematopoietic cells do not derive from the grafted PSM. **j-m**, Sagittal section through an E12 grafted limb embryo. Onset of bone marrow colonization. **j**, Double staining for GFP and DAPI. **k- m**, Higher magnifications of the framed area in (**j**). **k**, Double staining for MEP21 and GFP showing the partial overlap between the two markers. MEP21 was weakly expressed in angioblasts explaining the low fluorescent signal in some cells. **l**, Same area, from an immediate adjacent section. Double staining for GFP and DAPI. **m**, Triple staining for CD45, GFP and DAPI from an immediate adjacent section. $CD45^+GFP^-$ cells invaded the cartilage area. At this stage, only GFP^+ endothelial cells have colonized the limb. LB, Limb bud; Ao, Aorta; C, Cartilage; DC, Duct of Cuvier; FB, Feather bud; G, Gut; LB, Limb Bud; M, Mesonephros; NT, Neural tube; M, Muscle; V, Vessel. Scale bars in a, 200µm; in b, 70µm; in c, 30µm; in d, 20µm; in e, h, 250µm; in f, 350µm; in i, 150 µm; in j, 400µm; in k-l, 70µm; in m, 50µm.



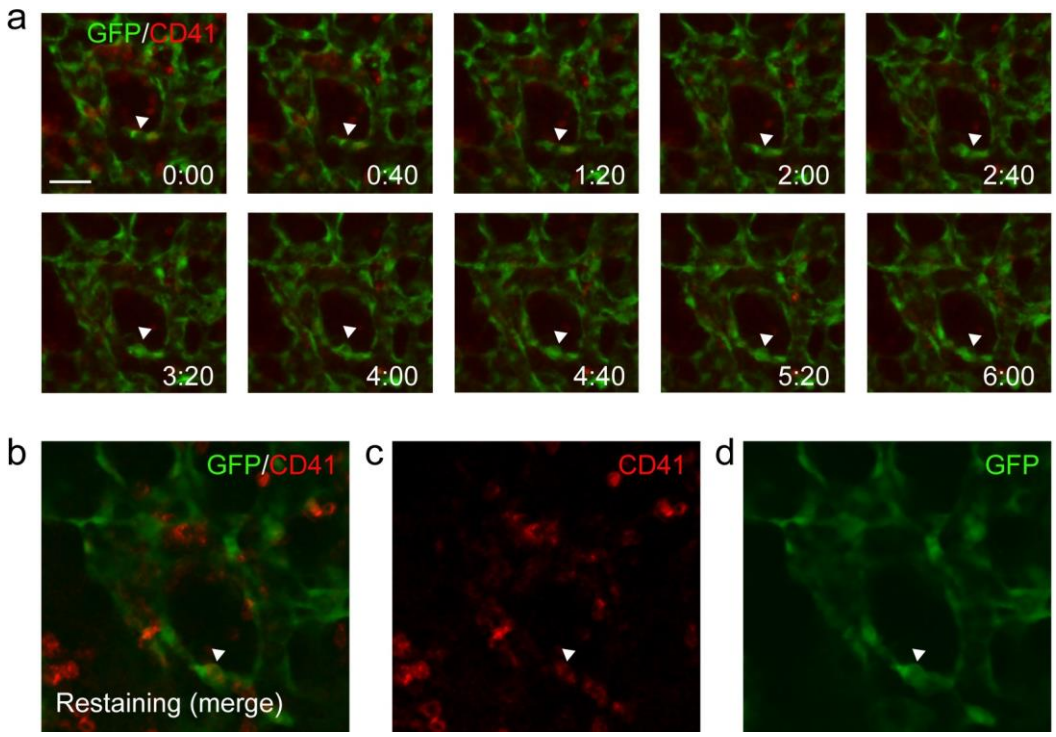
Extended Data Figure 3. GFP⁺ somite-derived cells contribute to E16/E17 bone marrow CD45⁺ haematopoietic production.

a-c, Representative fluorescent cross section showing the co-distribution of GFP⁺ sinusoid cells and CD45⁺ cells. **b**, Enlargement of the bone marrow showing GFP⁺CD45⁺ cells lining the sinusoids. **c**, Enlargement of the frame in **b**, showing the double positive cells. **d-l**, Series of confocal cross sections separated by 1µm, showing a GFP⁺CD45⁺ cell (white arrow). **d, g, j**, CD45 staining. **e, h, k**, GFP signal. **f, i, l**, Merged CD45, GFP and DAPI. Right and bottom banners correspond to YZ and XZ projections of the confocal image, respectively. Scale bars in a, 200µm; in b, 40µm; in c, 15µm; in d, 20µm.



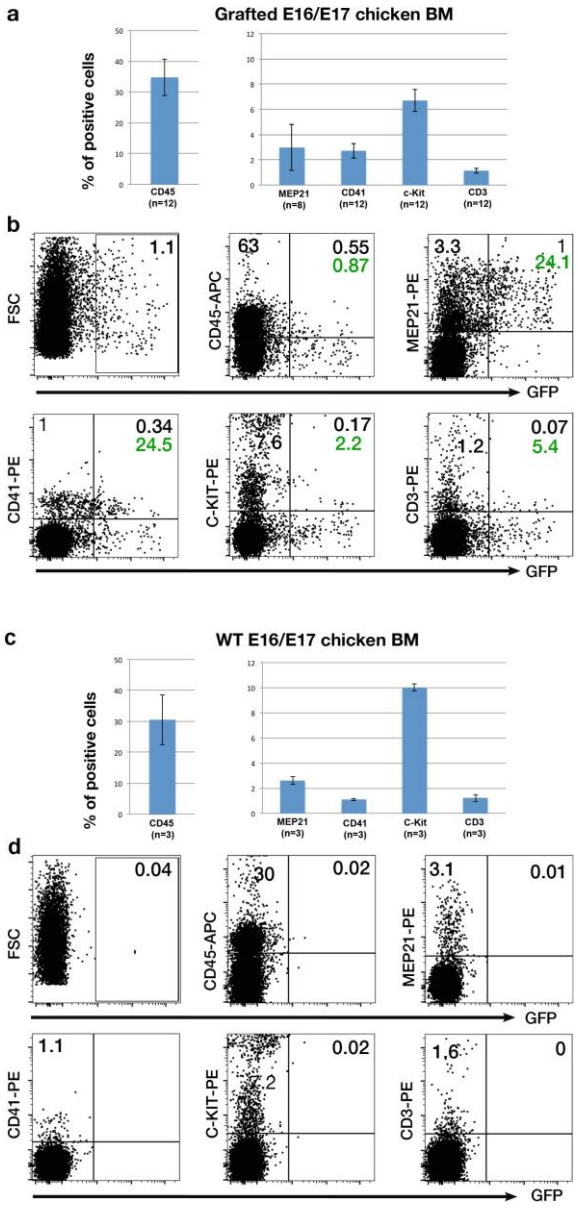
Extended Data Figure 4. GFP and CD41 co-expression in endothelial cells from an E16/E17 bone marrow analysed by confocal microscopy.

a, Confocal section of E16 BM after GFP, CD41 and DAPI (nucleus) staining. The arrow points to a $CD41^{+}GFP^{+}$ cell attached to the GFP^{+} vessel. **b**, CD41 and DAPI double staining. **c**, GFP fluorescence. **d**, The same image as in (**a**) shown enlarged. The two yellow lines indicate the x and y axes. The right stripe represents the yz view and the bottom stripe represents the xz view. **e, f**, RGB colour profiles of the yz and xz views, respectively. It shows that CD41 and GFP profiles are superposed indicating that both markers are within the same cell. The distance is in pixels. Scale bars in a, 20µm; in d, 25µm.



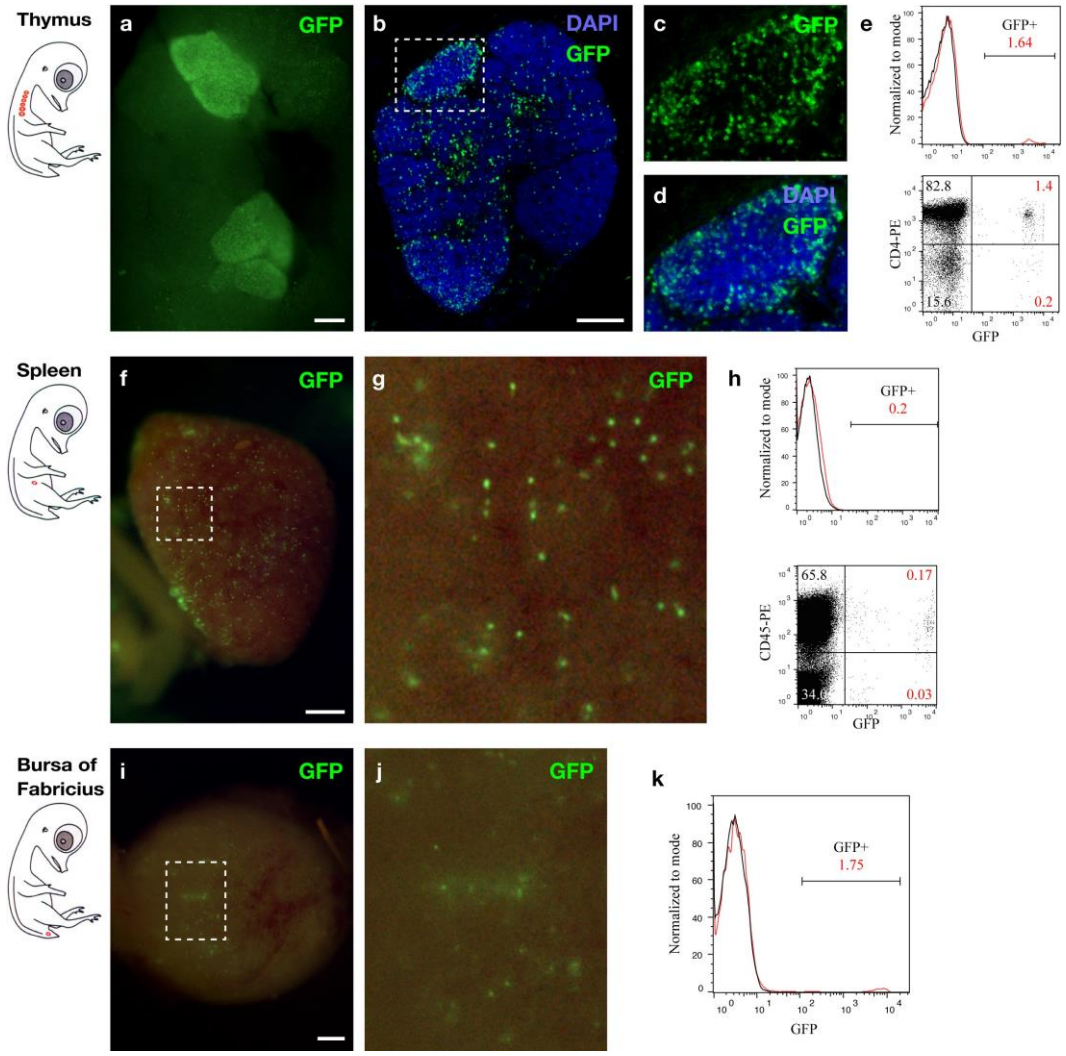
Extended Data Figure 5. Emergence of CD41$^+$ haematopoietic precursors from GFP$^+$ endothelial cells of the bone marrow.

a, Time-lapse imaging of a transversal section of an E16 grafted bone marrow showing the progressive emergence of a GFP$^+$CD41$^-$ cell from the endothelium of a blood vessel. **b**, After time- lapse imaging (**a**) the section was stained again with anti-CD41-PE antibodies. The newly emerged GFP$^+$ cell expressed CD41 (while it was not at the beginning of the imaging). Time in hours and minutes. Scale bar, 15µm.



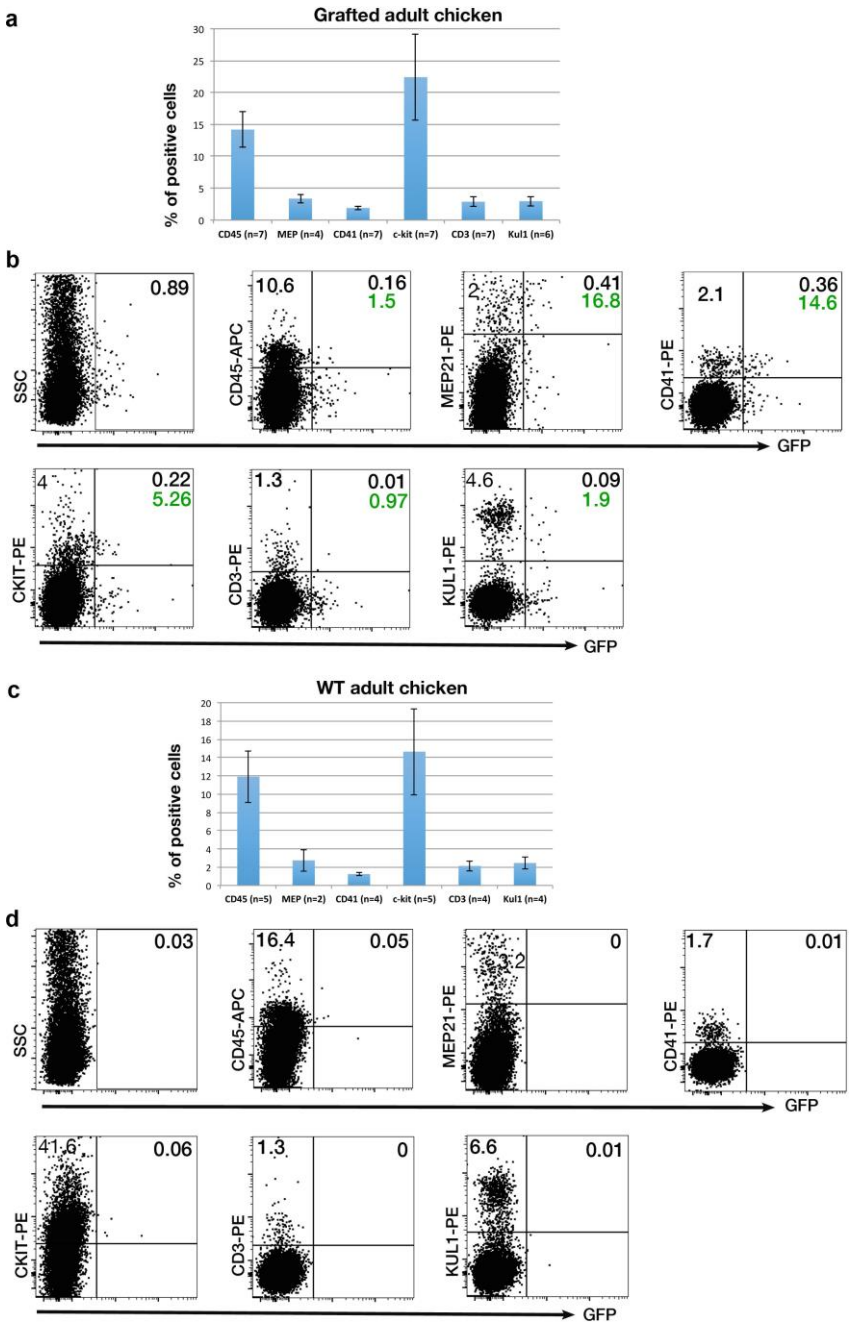
Extended Data Figure 6. Flow cytometry analysis reveals the presence of endothelial and haematopoietic cells expressing GFP in E16 BM.

Distribution of CD45, MEP21, CD41, c-KIT, and CD3 populations in grafted (**a, b**) or wild-type (**c**, **d**) BM mononucleated cells. Numbers within parenthesis indicate the number of samples analysed for each marker. **b**, Representative flow cytometry analysis of an E16/E17 grafted BM showing that GFP$^+$ cells derived from the PSM graft contributed to all lineages. Numbers in green indicate the percentage of GFP$^+$ cells in each population (i.e. in CD45$^+$, MEP21$^+$, CD41$^+$, c-KIT$^+$ and CD3$^+$ populations). **d**, Representative flow cytometry analysis of an E16/E17 wild-type BM. Error bars are mean ± SEM.



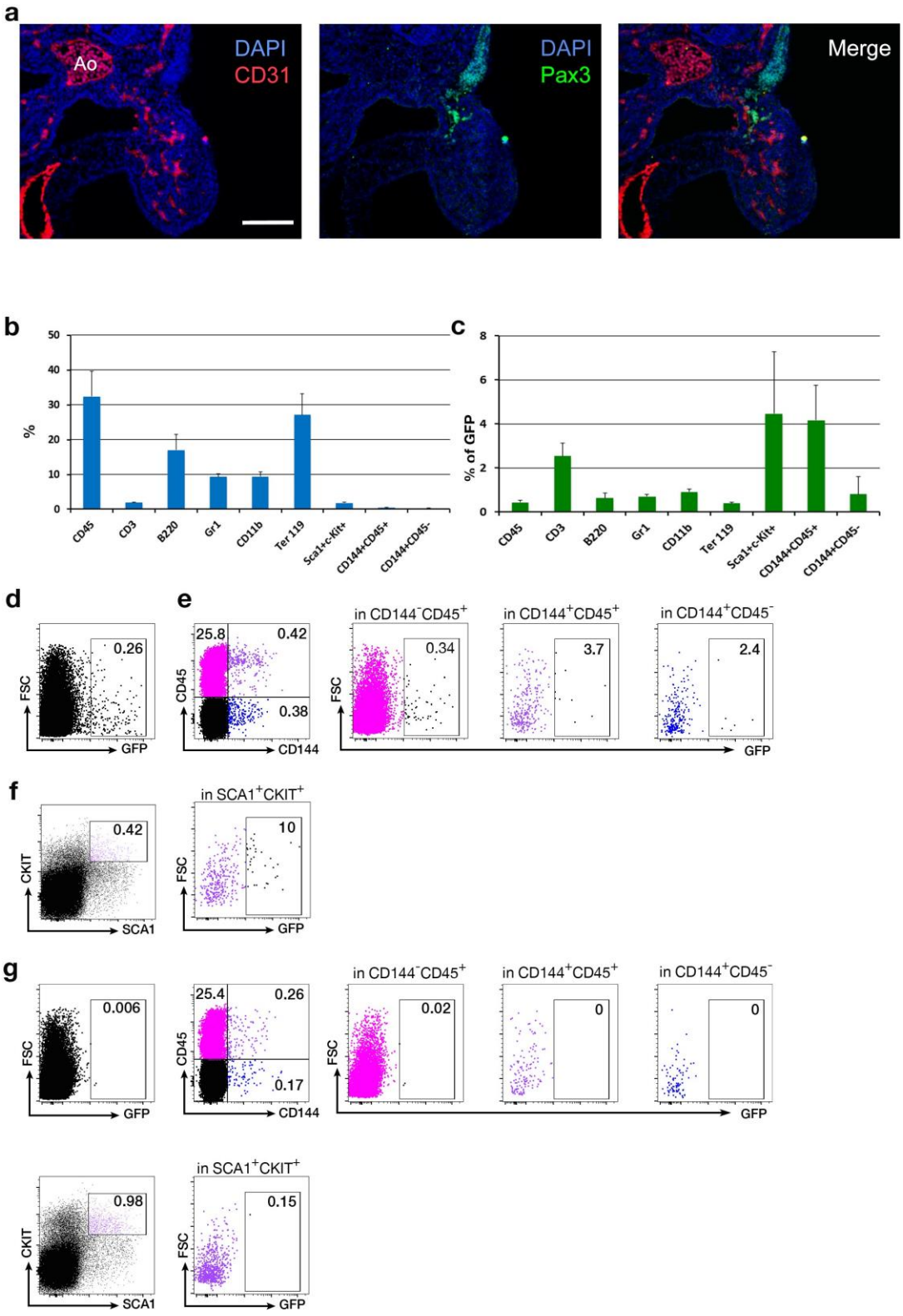
Extended Data Figure 7. PSM-derived cells provide haematopoietic precursors able to colonize the secondary haematopoietic organs of an E16/E17 chicken.

a, **f**, **i**, Global views of the thymus (**a**), spleen (**f**) and bursa of Fabricius (**i**) of an E16 chicken colonized by GFP$^+$ PSM-derived grafted cells. **a,** Most of the thymus lobes were colonized by GFP$^+$ cells. **b**, Immunostaining (with DAPI) of a thymus transverse section. **c-d**, Higher magnification of the frame in (**b**). GFP (**c**) and merge of DAPI and GFP (**d**). **e**, Flow cytometry analysis of an E16/E17 thymus revealed the contribution of GFP$^+$ PSM-derived cells to the whole thymus and more precisely to the T lineage (GFP$^+$CD4$^+$). **g**, Higher magnification of the frame in (**f**), showing dispersed GFP$^+$ cells in the spleen. **h**, Flow cytometry analysis of the spleen, showing the contribution of GFP$^+$ PSM-derived cells to the whole spleen and more precisely to the CD45$^+$ haematopoietic lineage. **j**, Higher magnification of the frame in (**i**) showing dispersed GFP$^+$ cells in the bursa. **k**, Flow cytometry analysis showing that the bursa of Fabricius, located far away from the PSM graft, was also colonized by PSM-derived haematopoietic progenitors that derived from the graft. Drawings are shown on the left side of the organ pictures to locate the thymus, spleen and bursa of Fabricius in an E16/E17 chicken. Scale bars in a, b, f, i: 100µm.



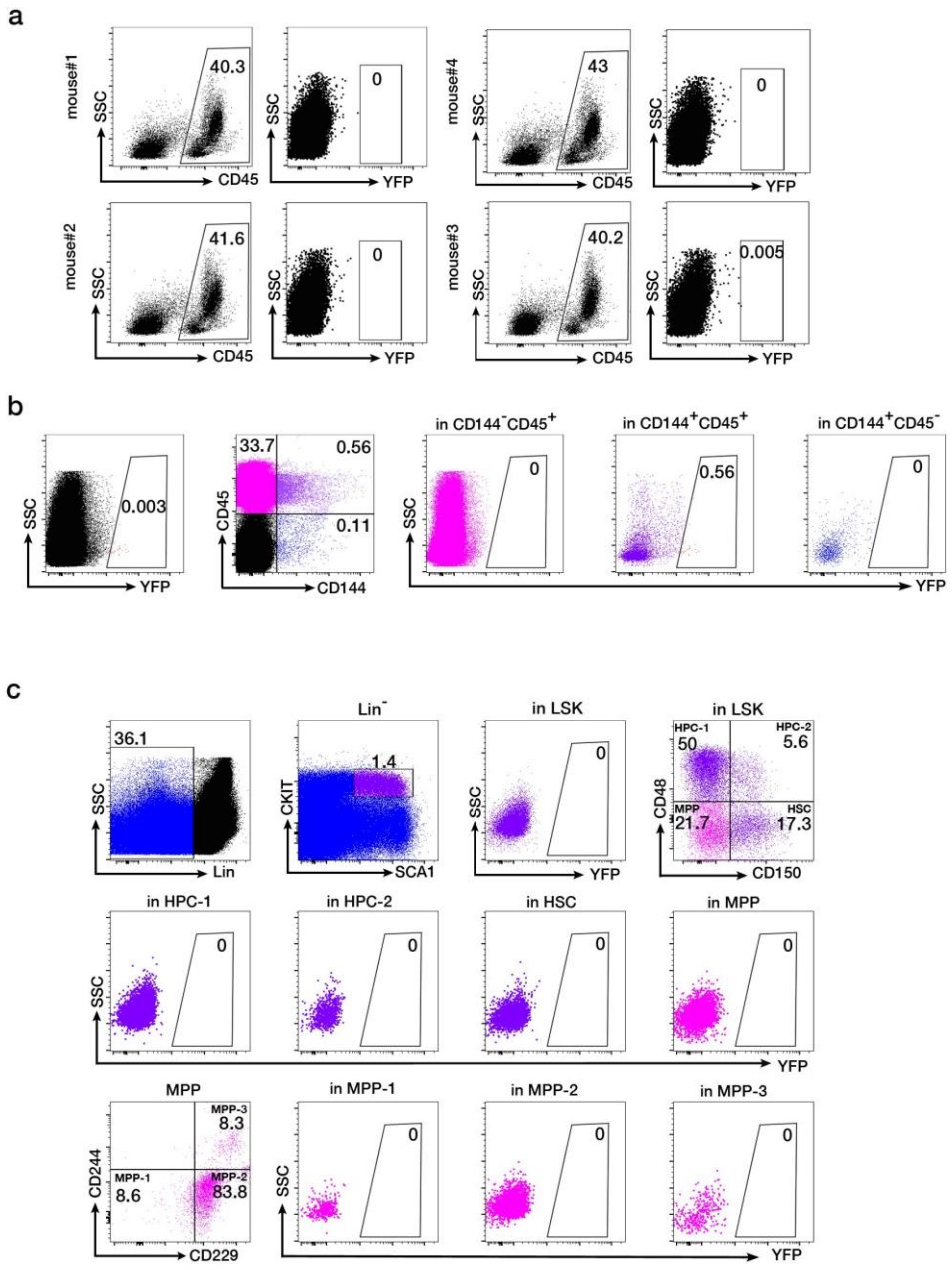
Extended Data Figure 8. Flow cytometry analysis reveals the presence of endothelial and haematopoietic cells expressing GFP in adult BM.

Distribution of CD45, MEP21, CD41, c-KIT, CD3 and KUL1 populations in grafted (**a, b**) and wild-type (**c, d**) adult BM mononucleated cells. Numbers within the parenthesis indicate the number of samples analysed for each marker. **b**, Representative multilineage flow cytometry analysis of a grafted adult BM showing that GFP$^+$ cells derived from the PSM graft contributed to all lineages. Numbers in green indicate the percentages of GFP$^+$ cells in each population (i.e. in CD45$^+$, MEP21$^+$, CD41$^+$, c-KIT$^+$, CD3$^+$ and KUL1$^+$ populations). **d**, Representative multilineage flow cytometry analysis of a wild-type adult BM showing the absence of GFP staining. Error bars represent mean ± SEM.



Extended Data Figure 9. Pax3⁺ somite-derived cells do not contribute to the aorta haematopoieisis but contribute to the endothelial and haematopoietic lineages in the late fœtal BM.

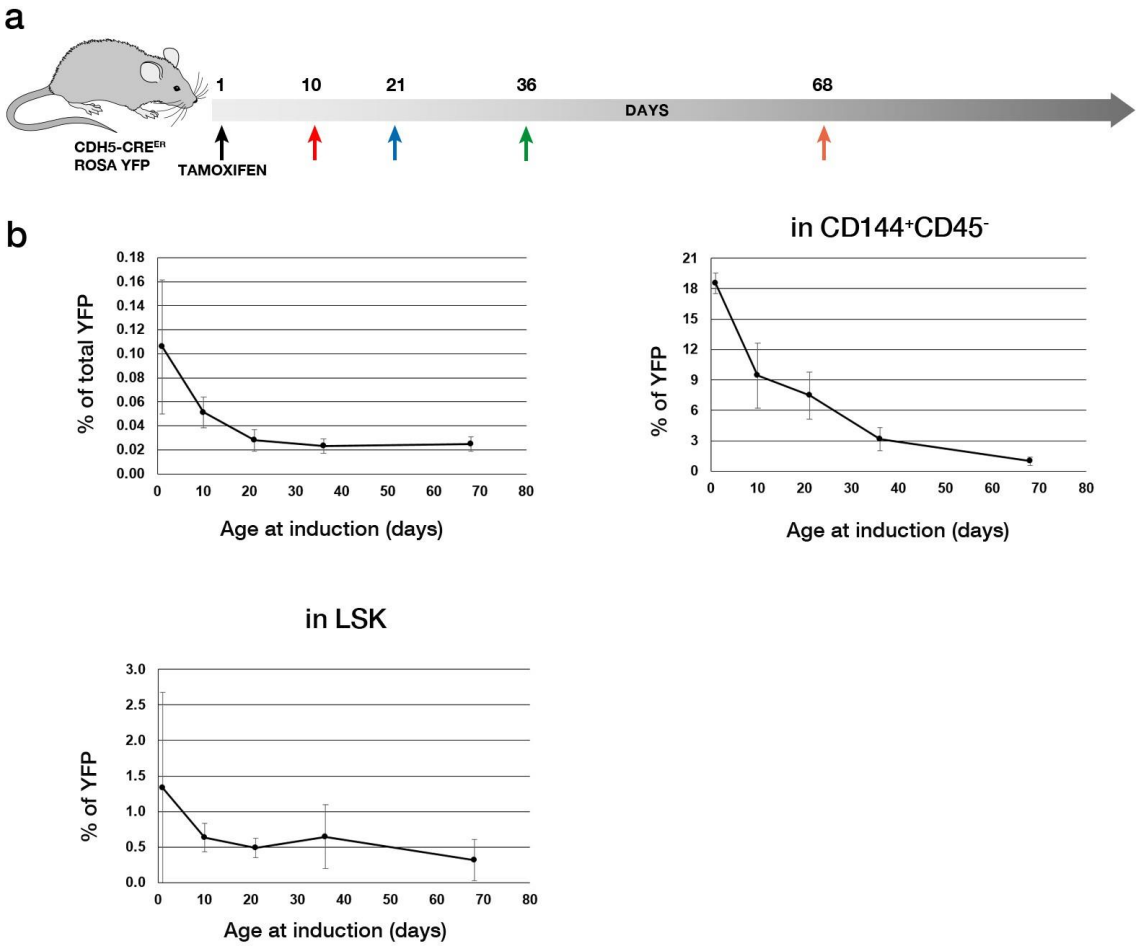
a, Transverse section through the aorta of an E10.5 Pax3 GFP mouse embryo at the mid-trunk level. CD31 (red, left), Pax3-GFP (green, middle) double staining counterstained with DAPI (blue). Right panel represents the merge. **b**, Flow cytometry analysis of the endothelial and haematopoietic lineages in the BM of Pax3-Cre/mTmG foetuses. **c**, Flow cytometry analysis showing GFP^+ cell (i.e. somite-derived cell) contribution to the different endothelial and haematopoietic populations. Mouse somite-derived cells mostly contributed to new-born BM $CD3^+$, $Sca1^+$c-Kit^+ and $CD144^+CD45^+$ cell populations. Data in **b**, **c** represent mean ± SEM from 3-5 pooled foetuses in 3 independent experiments. **d**, Representative GFP expression in mononucleated BM cells. **e**, Representative flow cytometry analysis of the mononucleated BM cells revealed GFP^+ cells mainly in the $CD144^+CD45^+$ population. **f**, Flow cytometry analysis showing the presence of GFP^+ cells in the $Sca1^+$c-KIT^+ population (i.e. HSPC population). **g**, Negative control, i.e. non-recombined foetuses.



Extended Data Figure 10. Flow cytometry analyses showing no YFP expression in haematopoietic cells shortly after tamoxifen induction and in non-induced BM mouse.

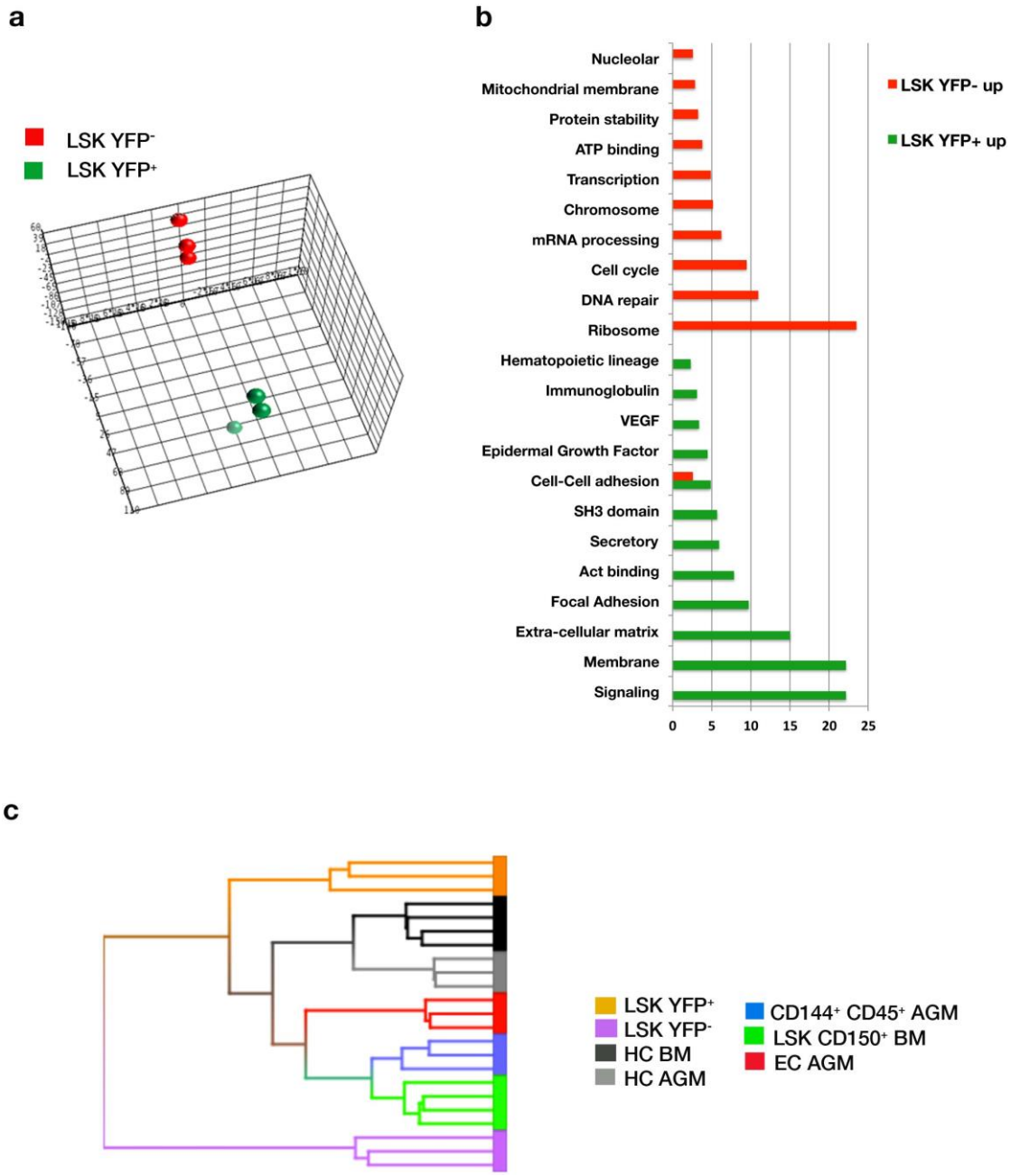
a, Flow cytometry analysis for the presence of YFP$^+$ cells in the CD45$^+$ fraction of the BM immediately after the last tamoxifen injection. Representative new-born mice injected with tamoxifen at post-natal days 1, 2, and 3 and analysed for the expression of CD45 and YFP at day 4. No CD45$^+$ YFP$^+$ cells were found indicating the absence of contamination by haematopoietic progenitors expressing VE- Cadherin. **b, c,** BM from a representative non-induced mouse at 27 post-natal days. **b**, Representative flow cytometry analysis showing the absence of YFP expression in mononucleated BM cells and in the CD144$^-$CD45$^+$, CD144$^+$CD45$^+$ and CD144$^+$CD45$^-$ cell fractions. **c**, Representative flow cytometry

analysis of the mononucleated BM cells showing the absence of YFP expression in the Sca1^{+}c-KIT^{+} population (i.e. HSPC population) and in the different HSC, HPC and MPP populations.



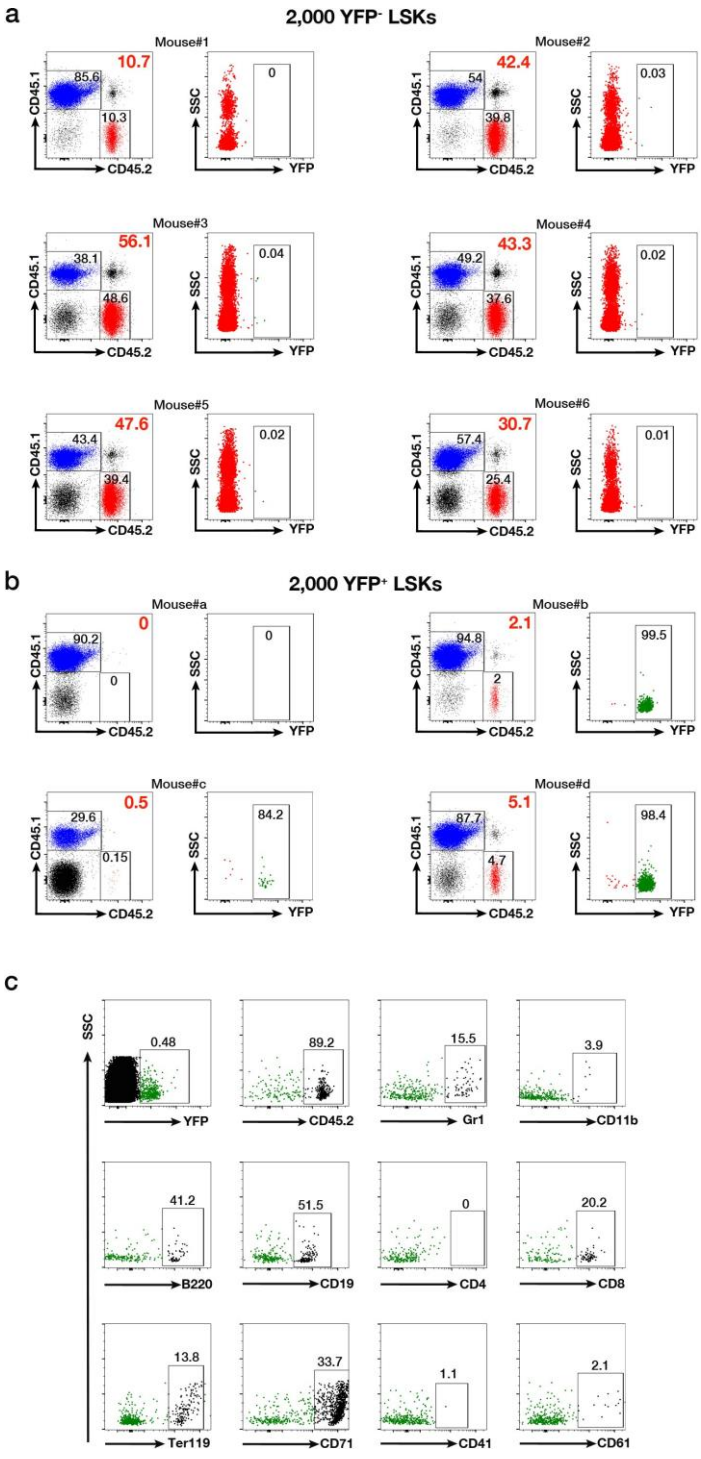
Extended Data Figure 11. YFP induction in VE-cadherin$^+$ cells at different time points after birth reveals a decrease in the number of YFP$^+$ endothelial and haematopoietic cells.

a, Scheme showing the activation of YFP in VE-Cadherin$^+$ (Cdh5) cells by tamoxifen injection at different time points after birth (coloured arrows). **b**, Analysis of the percentages of total YFP$^+$ cells, and YFP$^+$ cells in CD144$^+$CD45$^-$ and LSK populations at 21 days post tamoxifen injection. Of note, the x axis values represent the age of the mice from the time of tamoxifen injection. Error bars are mean ± SEM.



Extended Data Figure 12. PCA, Gene Ontology and hierarchical clustering analysis show major differences between YFP$^+$ and YFP$^-$ LSK cells.

a, PCA with the entire set of mRNAs (30,922 genes) as variables and the basic set of YFP$^+$ (green) and YFP$^-$ (red) LSK cells as observations. The two types of transcriptomes were strongly separated. **b**, Major GO categories given by DAVID for the gene sets up-regulated in YFP$^+$ (green) and YFP$^-$ LSK (red) cells. **c**, Hierarchical clustering obtained with the 23 samples as observations and the gene set of 2056 DEGs (986 up, 1070 down) as variables was generated from the PCA displayed in Fig. 4a, bottom panel. Branch organisation reflects the association between the different samples displayed on the PCA.


a
2,000 YFP- LSKs
Mouse#1
10.7
85.6
10.3
0
Mouse#2
42.4
54
39.8
0.03
Mouse#3
56.1
38.1
48.6
0.04
Mouse#4
43.3
49.2
37.6
0.02
Mouse#5
47.6
43.4
39.4
0.02
Mouse#6
30.7
57.4
25.4
0.01
b
2,000 YFP+ LSKs
Mouse#a
0
90.2
0
0
Mouse#b
2.1
94.8
2
99.5
Mouse#c
0.5
29.6
0.15
84.2
Mouse#d
5.1
87.7
4.7
98.4
CD45.1
CD45.2
SSC
YFP
c
SSC
0.48
YFP
89.2
CD45.2
15.5
Gr1
3.9
CD11b
41.2
B220
51.5
CD19
0
CD4
20.2
CD8
13.8
Ter119
33.7
CD71
1.1
CD41
2.1
CD61

Extended Data Figure 13. Long-term repopulation analysis in the peripheral blood of recipients transplanted with YFP^- or YFP^+ LSK cells.

a, b, Flow cytometry analysis of circulating blood from 6 (**a**) and 4 (**b**) C57Bl/6-CD45.1 recipient mice at 11 weeks injected with 2,000 YFP^- (**a**) or YFP^+ (**b**) LSK cells (C57Bl/6-CD45.2), respectively. Mice recipients transplanted with YFP^- LSK cells displayed a more robust reconstitution than the ones transplanted with YFP^+ LSK cells. Of note, two mice in the LSK YFP^+ series died before two weeks post-injection. **c,** Back-gating within the GFP^+ population of the mononucleated bone marrow cells from mouse #d at 14 weeks of reconstitution. Dot plot representation of flow cytometry multilineage analysis showing Gr1 (granulocytes), CD11b (macrophages/monocytes), B220 and CD19 (B cells), CD4 (T cells), CD8 (T cells), Ter119 and CD71 (red cells), CD41 and CD61 (megakaryocytes) staining. The YFP population was more prominent in B220, CD19, CD71, Gr1 and CD8 fractions.

Extended Data Tables

Time point of analysis	E4		E8	E12	E16		E17	E20	P2	P4	P17	P19	P35	P130	P144	P166
Total number of cases	5		2	2	13		3	1	1	1	1	1	1	1	1	1
Whole organism																
	1		1						1							
BM				1	10	2	2	1	1	1	1	1	1	1	1	1
					3	3	2	1								
Thymus					3	2	1	1	1	1				1	1	1
					3	1									1	1
Bursa of Fabricius															1	
					2											
Spleen														1	1	1
					1											
Blood	5	1	2	1	2	1	11		3					1	1	1
	5		2		2											

E : Embryonic day

P : Post-hatching

Number of cases for each type of analysis.

Legend of the Table box

FACS	Immunostaining
Picture organ *in toto*	Video

Not done

Not applicable

Extended Data Table 1. Number of grafted chickens, time points of analyses and type of analyses performed.

This table summarizes the number of chicken samples according to the age at analysis; the organs or tissues analysed, i.e., thymus, bone marrow, spleen; the kind of analysis applied i.e., FACS, immunostaining, confocal imaging or videos; and the number of cases analysed for each time point.

Antibody	Fluorochrome	Clone	Specificity	Supplier	Reference
CD4	purified	GK1.5	T cells	Biolegend	BLE100402
	PE			Biolegend	BLE100408
	APC			Biolegend	BLE100412
CD8a	purified	53-6.7	T cells	Biolegend	BLE100702
	APC			Biolegend	BLE100712
CD11b	purified	M1/70	Granulocytes, Macrophages	Biolegend	BLE101202
	APC			Biolegend	BLE101212
CD19	PE	6D5	B cells	Biolegend	BLE115508
CD41	APC	MWReg30	Megacaryocytes	Biolegend	BLE133914
CD45	APC	30-F11	Haematopoietic cells	Biolegend	BLE103112
CD45.1	APC	A20	Haematopoietic cells, Ly5.1 allele specific	Biolegend	BLE110714
CD45.2	PE	104	Haematopoietic cells Ly5.2 allele specific	Biolegend	BLE109808
CD45R/B220	purified	RA3-6B2	B cells	Biolegend	BLE103202
	APC			Biolegend	BLE103212
CD48	PerCP Cy5.5	HM48-1	SLAM family	Biolegend	BLE103421
CD61	PE	2C9.G2 (HMb3-1)	Megacaryocytes	Biolegend	BLE104308
CD71	PE	RI7217	Erythroid cells	Biolegend	BLE113807
CD117	APC Cy7	2B8	HSPC	Biolegend	BLE105826
	PE-Cy7			Biolegend	BLE105814
CD144	PE	11D4.1	Endothelial cells	BD Biosciences	BLE562243
CD150	BV510	TC15-12F12.2	SLAM family	Biolegend	BLE115929
CD229	Biotin	Ly9ab3	SLAM family	Biolegend	BLE122903
CD244	PE-Cy7	m2B4(B6)458.1	SLAM family	Biolegend	BLE133511
Gr-1	purified	RB6-8C5	Granulocytes, Monocytes	Biolegend	BLE108402
	APC			Biolegend	BLE108412
	PE			Biolegend	BLE108408
Sca-1	PE	D7	HSPC	Biolegend	BLE108108
Ter119	purified	TER-119	Erythroid cells	Biolegend	BLE116202
	APC			Biolegend	BLE116212
Streptavidin	PE			Biolegend	BLE405203
Streptavidin	PE-Dazzle594			Biolegend	BLE405247

Antibody	Fluorochrome	Clone	Specificity	Supplier	Reference
CD3	PE	CT-3	lymphoid cells	SouthernBiotech	8200-09
CD4	PE	CT-4	lymphoid cells	SouthernBiotech	8210-09
CD8	PE	CT-8	lymphoid cells	SouthernBiotech	8220-09
Bu-1	PE	AV20	B cells	SouthernBiotech	8395-09
CD45	PE	LT40	Haematopoietic cells	SouthernBiotech	8270-09
	APC			SouthernBiotech	8270-11
KUL01*	PE	/	Monocytes/macrophages	Abcam	ab25441
CD41/61	PE	/	Haematopoietic progenitor cells	AbD serotec	MCA2240PE
c-kit	PE	/	HSPC	SouthernBiotech	8380-09
MEP21	purified	Gift from McNagny's Lab			
GFP	purified	/	recognize GFP+ cells	Abcam	Ab290

Secondary antibodies					
Goat anti-Mouse IgG1	Alexa-555		recognize MEP21	Life technologies	A21127
Goat anti-Rabbit (polyclonal)	Alexa-488		recognize anti-GFP	Life technologies	A11008

* ***not available anymore***

Extended Data Table 2. List of anti-chicken and anti-mouse antibodies.

List of anti-mouse (first) and anti-chicken (second) antibodies used in the study, including the conjugated fluorochrome when applicable, clone of origin, cell specificity, supplier and catalogue number.

Extended Data Table 3. Microarray expression data. RNA expression analyses of the LSK YFP$^+$ vs YFP$^-$ cells

Extended Data Table 4. Statistics source data

Supplementary movies

Movie S1. Whole-mount staining of an E16 chicken bone marrow after GFP$^+$ PSM graft.

The bone marrow was stained with anti-MEP21 (endothelial marker, red), anti-CD45 (haematopoietic marker, blue) and anti-GFP (green, which revealed grafted PSM-derived cells) antibodies. The marrow was heavily colonized by GFP$^+$ PSM-derived cells that contributed to the whole bone marrow vascularization. Extended data movie 1 is visible at https://figshare.com/s/1f221351c0a9b80350b7.

Movie S2. A GFP$^+$ endothelial cell co-expressing CD45 as revealed by confocal analysis of an E16 BM.

Series of confocal sections separated by 1µm, showing a GFP^+CD45^+ cell (white arrow). Panels: up left, DAPI; up right, CD45; down left, GFP; down right, merge. Right and bottom banners correspond to YZ and XZ projections of the confocal image, respectively. Extended data movie 2 is visible at https://figshare.com/s/1889888fbdaf12b9bf0a.

Movie S3. GFP^+ endothelial cells co-expressing CD41 are integrated in the vascular endothelium of E16 BM.

Series of confocal sections of 1µm, showing a GFP^+CD41^+ cell integrated in the GFP^+ vascular endothelium. Extended data movie 3 is visible at https://figshare.com/s/b13431a688936f65013a.

Movie S4. Emergence of $CD41^+$ haematopoietic precursors from GFP^+ endothelial cells in the bone marrow.

a, Time-lapse live imaging of a transversal E16 bone marrow slice showing the emergence of a GFP^+ cell from the endothelium of a blood vessel. **b**, After time-lapse imaging (**a**) the section was stained again with anti-CD41-PE antibodies. The newly-emerged GFP^+ cell expressed the haematopoietic marker CD41. Of note, this cell was $CD41^-$ at the beginning of imaging. Time is in hours and minutes. Scale bar, 15µm. Extended data movie 4 is visible at https://figshare.com/s/0aa2fd1e81e09918fd80.

Movie S5. Whole-mount staining of an E16 chicken thymus after GFP^+ PSM graft.

The thymus was stained with anti-MEP21 (endothelial marker, red), anti-CD45 (haematopoietic marker, blue) and anti-GFP (green, which revealed grafted PSM-derived cells) antibodies. The thymus was colonized by GFP^+ PSM-derived cells, some of the lobes being more colonized than others. GFP^+ cells could colonize secondary haematopoietic organs that are located far from the grafting site. Extended data movie 5 is visible at https://figshare.com/s/d8e15dcd3ab9e509ce13.

1 Dieterlen-Lievre, F. & Martin, C. Diffuse intraembryonic hemopoiesis in normal and chimeric avian development. *Dev Biol* **88**, 180-191 (1981).

III. Article 3

The quail as an avian model system: its genome provides insights into social behaviour, seasonal biology and infectious disease response

Article in revision in BMCB and available on biorxiv

https://www.biorxiv.org/content/10.1101/575332v1

In this article, I contributed to the manual annotation of around 8000 unannotated genes of the quail genome.

You can find the LOC assignment list I annotated in the appendix 4.

The quail as an avian model system: its genome provides insights into social behaviour, seasonal biology and infectious disease response

Katrina M Morris[1]*, Matthew M Hindle[1], Simon Boitard[2], David W Burt[3], Angela F Danner[4], Lel Eory[1], Heather L Forrest[4], David Gourichon[5], Jerome Gros[6,7], LaDeana Hillier[8], Thierry Jaffredo[9], Hanane Khoury[9], Christine Leterrier[10], Andrew Loudon[11], Andrew S Mason[1], Simone L Meddle[1], Francis Minvielle[12], Patrick Minx[8], Frédérique Pitel[2], J Patrick Seiler[4], Tsuyoshi Shimmura[13], Chad Tomlinson[8], Alain Vignal[2], Robert G Webster[4], Takashi Yoshimura[14], Wesley C. Warren[8], Jacqueline Smith[1*]

1. The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, United Kingdom
2. GenPhySE, Université de Toulouse, INRA, ENVT, 31326 Castanet Tolosan, France
3. The John Hay Building, Queensland Biosciences Precinct, 306 Carmody Road, The University of Queensland, St Lucia QLD 4072, Australia
4. St. Jude Children's Research Hospital, Virology Division, Department of Infectious Diseases, 262 Danny Thomas Place, Memphis, TN 38105, USA.
5. PEAT Pôle d'Expérimentation Avicole de Tours. Centre de recherche Val de Loire, INRA, UE 1295, Nouzilly, France
6. Department of Developmental and Stem Cell Biology, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris, Cedex 15, France
7. CNRS URA2578, 25 rue du Dr Roux, 75015 Paris, France
8. McDonnell Genome Institute, Washington University School of Medicine, 4444 Forest Park Blvd., St Louis, MO 63108, USA
9. Sorbonne Université, IBPS, CNRS UMR7622, Inserm U 1156, Laboratoire de Biologie du Développement; 75005 Paris
10. UMR85 Physiologie de la Reproduction et des Comportements, INRA, CNRS, Université François Rabelais, IFCE, INRA Centre Val de Loire, 37380 Nouzilly, France
11. Centre for Biological Timing, Faculty of Biology, Medicine and Health, School of Medical Sciences, 3.001, A.V. Hill Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK
12. GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France
13. Department of Biological Production, Tokyo University of Agriculture and Technology, 3-8-1 Harumi-cho, Fuchu, Tokyo 183-8538, Japan
14. Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan

*- corresponding authors

ABSTRACT

The Japanese quail (*Coturnix japonica*) is a popular domestic poultry species and an increasingly significant model species in avian developmental, behavioural and disease research. We have produced a high-quality quail genome sequence, spanning 0.93 Gb assigned to 33 chromosomes. In terms of contiguity, assembly statistics, gene content and chromosomal organization, the quail genome shows high similarity to the chicken genome. We demonstrate the utility of this genome through three diverse applications. First, we identify selection signatures and candidate genes associated with social behaviour in the quail genome, an important agricultural and domestication trait. Second, we investigate the effects and interaction of photoperiod and temperature on the transcriptome of the quail medial basal hypothalamus, revealing mechanisms of photoperiodism. Finally, we investigate the response of quail to H5N1 influenza infection. In quail lung, many key immune genes and pathways were downregulated, and this may be key to the susceptibility of quail to H5N1. This genome will facilitate further research into diverse research questions using the quail as a model avian species.

INTRODUCTION

Japanese quail (*Coturnix japonica*) is a popular domestic poultry species raised for meat and eggs in Asia and Europe. Quail have been used in genetics research since 1940[1], and are an increasingly important model in developmental biology, behaviour and biomedical studies[2]. Quail belong to the same family as chickens (Phasianidae) but have several advantages over chickens as a research model. They are small and easy to raise, have a rapid growth rate and a short life cycle, becoming sexually mature only seven to eight weeks after hatching[3].

Quail have become a key model in several research fields. The avian embryo has long been a popular model for studying developmental biology due to the accessibility of the embryo. The quail embryo survives manipulation and culture better than chicken embryos making them ideal for this type of research[3]. Quail have been used as model for stem cell differentiation, for example a culture system that mimics the development of hematopoietic stem cells has been recently developed, as quail show greater cell multiplication in these cultures than chickens[4]. Quail are also used to study the genetics underlying social behaviours[5], sexual behaviour[6,7], pre- and post-natal stress programming[8], and emotional reactivity[9-12]. Japanese quail have a fast and reliable reproductive response to increased photoperiod, making them an important model species for investigation into seasonal behaviour and reproduction in birds[13-15]. The molecular mechanisms behind seasonality including metabolism and growth, immunity, reproduction, behaviour and feather moult is poorly understood despite its importance in the management of avian species.

Quail are also important in disease research. Strains of quail have been developed as models of human disease such as albinism[16] or necrotizing enterocolitis in neonates[17], and quail lines have been selected on their immunological response[18]. There are key differences in the immunogenetics of quail and chicken - particularly in the major histocompatibility complex (MHC)[19,20]. Investigating the immunology quail is important for understanding infectious disease spread and control in poultry. For example they are a key species for influenza transmission, with previous research showing that quail may play a key role as an intermediate host in evolution of avian influenza[21-23]. Zoonotic H5N1 influenza strains have crossed from quail to human causing mortality in the past[24,25], making them a potential pandemic source.

We have produced a high-quality annotated genome of the Japanese quail (*Coturnix japonica*). Here we describe the assembly and annotation of the quail genome and demonstrate key uses of the quail genome in immunogenetics, disease, seasonality and behavioural research demonstrating its utility as an avian model species.

RESULTS

Genome assembly and annotation

We sequenced a male *Coturnix japonica* individual from an inbred quail line using an Illumina HiSeq 2500 instrument. Total sequence genome input coverage of Illumina reads was ~73x, using a genome size estimate of 1.1 Gb. Additionally, 20x coverage of long PacBio reads were sequenced and used to close gaps. The assembled male genome *Coturnix japonica* 2.0 is made up of a total of 2,531 scaffolds (including single contigs with no scaffold association) with an N50 scaffold length of 2.9 Mb (N50 contig length is 511 kb). The assembly sequence size is 0.927 Gb with only 1.7 % (16 Mb) not assigned to 33 total chromosomes. *Coturnix japonica* 2.0

assembly metrics were comparable to previous assemblies of Galliformes, and superior to other quail genomes[26,27] in ungapped (contigs) sequence length metrics (**Table 1**). Specifically, in comparison to recently published genomic data from the Japanese quail[27], our genome is substantially less fragmented (contig N50 of 0.511 Mb vs 0.027 Mb), has been assigned to more chromosomes, and has more complete annotation with ncRNA, mRNA and pseudogenes predicted. Our estimate of total interspersed repetitive elements was 19% genome-wide based on masking with Windowmasker[28]. In the genomes of other quail species the estimated repeat content was much lower, ~10% less in both species[26].

To improve the quantity and quality of data used for the annotation of the genome, we sequenced RNA extracted from seven tissues sampled from the same animal used for the genome assembly. Using the same inbred animal increases the alignment rate and accuracy. The amount of data produced for annotation from the 7 tissues is respectively in giga bases: 18.9 in brain, 35.6 in heart, 19.3 in intestine, 27.8 in kidney, 39.0 in liver, 18.8 in lung and 34.0 in muscle. High sequencing depth was aimed at, to help detect low expression genes, including tissue-specific ones, in these various tissues. In total we predicted 16,057 protein-coding genes and 39,075 transcripts in the *Coturnix japonica* genome (**Table 2**). In comparison to other assembled and annotated Galliformes, transcript and protein alignments of known chicken RefSeq proteins to *Coturnix japonica* suggest the gene representation is sufficient for all analyses described herein (**Table 3**). However, we find ~1000 fewer protein coding genes in the Japanese quail than the northern bobwhite (*Colinus virginianus*) and scaled quail (*Callipepla squamata*) genomes[26]. We attribute this to the use of different gene prediction algorithms, and the slightly lower assembled size of Japanese quail, 927 Mb compared to 1 Gb in other quail genomes[26] (**Table 1**).

For further annotation, a set of genes unnamed by the automated pipeline were manually annotated. As part of an ongoing project to investigate hemogenic endothelium commitment and HSC production[4], transcriptomes were produced for two cultured cell fractions. Study of these cells is critical for developmental biology and regenerative medicine, and quail are an excellent model for studying these as they produce much more hematopoietic cells than similar chicken cultures. Approximately 8000 genes were expressed in these cells lines which lacked genes names or annotation from the automated annotation pipeline. Using BLAST[29] searches to identify homology to other genes, approximately 3000 of these were manually annotated (**Supplementary Data 1**).

Genome completeness was also quantitatively assessed by analyzing 4,915 single copy, orthologous genes derived from OrthoDB v7 and v9[30]. Presence and contiguity of these conserved, avian specific genes were tested with BUSCO v3.0.2[31]. A comparison with the chicken assembly (Gallus gallus 5.0[32]) indicates that 95% of these genes are present and full length in all three assemblies. The percentage of duplicated, fragmented and missing genes are also very similar between the assemblies (**Supplementary Fig. 1**). The quail genome has 10 more missing and 23 more fragmented genes than the Gallus gallus 5.0 assembly. However, relative to the total number of genes in the benchmarking set, these increases amount to just 0.2% and 0.5%, respectively. This indicates that the quail genome, like the chicken genome, is highly contiguous and in terms of its expected gene content, it is close to complete.

Galliforme genome synteny

Comparative mapping of the quail and chicken genomes revealed a high conservation of the chromosomal arrangement (**Fig. 1; Supplementary Data 2**), with no major rearrangements since the divergence of the two species approximately 23 MYA[33]. All identified quail chromosomes showed synteny conservation to their chicken chromosomal counterparts. By comparison, the turkey (*Meleagris gallopavo*) genome is more highly rearranged with two chromosomes having synteny conservation to each of chicken chromosomes 2 and 4[34]. No large intrachromosomal translocations were seen between chicken and quail chromosomes, compared to the two seen in the turkey[34]. Inversions and inter-chromosomal translocations were common, with 33 large (>1Mb) inversions or translocations occurring between chicken and quail chromosomes **(Fig. 2; Supplementary Data 2**). The quail chromosomes are more compact than their chicken counterparts (14% smaller on average).

Orthologous genes between quail and closely related species were identified through reciprocal BLAST searches. One-to-one orthologs in chicken were identified for 78.2% of all quail genes and 91.8% of protein-coding quail genes (**Supplementary Table 1**), indicating a high degree of genic conservation in the quail genome. Fewer orthologs were seen between turkey and quail genes (69.3%), although the number of orthologs of protein-coding genes was similar (91.7%), so the discrepancy is likely due to missing non-coding gene predictions in the turkey genome. As expected, conservation of one-to-one orthologs was lower with the mallard duck (*Anas platyrhynchos*), with duck orthologs identified for 64.5% of quail genes (78.9% protein-coding genes).

Endogenous retroviruses (ERVs)

ERVs represent retroviral integrations into the germline over millions of years and are the only Long Terminal Repeat (LTR) retrotransposons which remain in avian genomes[35,36]. Whilst the majority of ERVs have been degraded or epigenetically silenced, more recent integrations retain the ability to produce retroviral proteins, impacting the host immune response to novel exogneous infections[37,38]. A total of 19.4 Mb of the *Coturnix japonica* 2.0 assembly was identified as ERV sequence using the LocaTR pipeline[35] **(Supplementary Data 3 and 4**). ERVs therefore account for 2.1% of the quail genome sequence, levels similar to those in the chicken and turkey[32] (**Supplementary Table 2**), and similarly analysed passerine birds[35].

The majority of ERV sequences in all three genomes were short and fragmented, but 393 intact ERVs were identified in the quail, most of which were identified as alpha-, beta- or gamma-retroviral sequences by reverse transcriptase homology. It is possible that the smaller genome size of the quail compared to other birds reflects a more limited expansion of ERVs and other repeats (such as the LINE CR1 element; **Supplementary Table 2**) within the genome, following the basal avian lineage genome contraction[36,39]. However, ERV content is highly species-specific[35].

Despite variation in total and intact ERV content, the overall genomic ERV distribution in these three gallinaceous birds was highly similar. ERV sequence density was strongly correlated with chromosome length on the macrochromosomes and Z ($r > 0.97$; $P < 0.001$), but there was no significant correlation across the other smaller chromosomes. Furthermore, ERV density on each Z chromosome was at least 50% greater than would be expected on an autosome of equal length. These results support the depletion of repetitive elements in gene dense areas of the genome, and the persistence of insertions in poorly recombining regions, as was seen in the chicken[35]. This is further supported by the presence of clusters of intact ERVs (where density was five times the genome-wide level) on the macrochromosomes and sex chromosomes (**Supplementary Table 2**).

Immune gene repertoire

We investigated the immune genes in the quail genome in detail due to the importance of quail in disease research. The MHC-B complex of the quail has been previously sequenced and found to be generally conserved compared to chicken in terms of gene content and order[19,20]. However the quail MHC contains a higher copy number of several gene families within the MHC-B[19] and shows increased structural flexibility[20], as well as an inversion in the *TAP* region[19]. The MHC-B sequence in the quail genome extends from the previously sequenced scaffold, and this additional region also contains similar gene content and order to chicken, but with gene copy number variations. As in the chicken, the *CD1A* and *B* genes are found downstream of the MHC I region, while many *TRIM* family genes and *IL4I1* are encoded upstream. The BG region, which encodes a family of butrophylin genes known as *BG* genes in the chicken, was also present in the quail. Within this region, six *BG* genes were identified in the quail, compared to thirteen in the chicken[40]. At least five of these *BG* genes are transcribed in the quail lung and ileum. The chicken and turkey have an additional MHC locus known as the Rfp-Y or MHC-Y locus, which contains several copies of non-classical MHCI-Y and MHCIIB-Y genes. However, no MHC-Y genes have been previously identified in quail. BLAST searches of both the quail genome and quail transcriptomes, as well as the bobwhite and scaled quail genomes, failed to identify any MHC-Y genes, indicating this locus probably does not exist in the quail.

Cathelicidins and defensins are two families of antimicrobial peptides that have activities against a broad range of pathogens and exhibit immune-modulatory effects. Orthologs of all four chicken cathelicidins and of thirteen chicken defensins[41] were identified in the quail genome (**Supplementary Data 5**). Due to their high divergence, of the thirteen defensins only four were annotated through the annotation pipeline, with the remainder identified through BLAST and HMMER searches with chicken defensins. The only poultry defensin missing from the quail genome was *AvBD7*. The defensins were encoded in a 42 kb cluster on quail chromosome 3, as in chickens. A 4 kb gap in the scaffold in this region may explain the missing *AvBD7* sequence.

Several genes are thought to be key for influenza resistance in both humans and birds, including *RIG-I, TLR* and *IFITM* genes. *RIG-I* has not previously been identified in chicken, despite being present in ducks and many other bird orders, and is considered highly likely to be deleted from the chicken genome[42]. In addition, an important RIG-I binding protein RNF135 has also not been identified in chicken[43]. Likewise, an ortholog of *RIG-I* or *RNF135* could not be identified in the quail genome or transcriptomes through BLAST searches and therefore is likely missing in the quail also. Orthologs of all five chicken *IFITM* genes (*IFTIM1, 2, 3, 5* and *10*) were identified in the quail genome and transcriptomes. In addition, orthologs of each chicken TLR, including key TLRs for viral recognition, *TLR4* and *TLR7*, were identified in the quail genome, except for *TLR1A*.

Selection for social motivation

Quail has been used as a model to study the genetic determinism of behaviour traits such as social behaviours and emotional reactivity[11,12,44], these being major factors in animal adaptation. Moreover quail selected with a low social motivation behave in a way that can be related to autistic-like traits, so the genes and causal variants are of wider interest to the biomedical community. Here we use the new quail genome assembly to improve previous results on the

detection of selection signatures in lines selected for sociability. Due to the non-availability of a useable quail reference genome at the start of these studies, genomic sequence data produced from two DNA pools of 10 individuals each from two quail lines diverging for social motivation had been aligned to the chicken reference genome, GallusWU2.58[45]. As a result, only 55% of the reads had mapped in proper pairs, whereas by using our quail genome as a reference, this number increased to 92%. This corresponds to an improvement of the averaged coverage from 9x to 20x and of the number of analysed SNPs from 12,364,867 to 13,506,139.

The FLK[46] and local[45] score analysis led to the detection of 32 significant selection signature regions ($p < 0.05$), (**Supplementary Data 6**). **Supplementary Fig. 2** shows an example of such a region on Chr20. This represents a substantial improvement in the number of detected regions, compared with the 10 regions obtained when using the chicken genome as a reference[45]. Of the 32 detected regions, six may be merged in pairs due to their physical proximity, four regions map to new linkage groups absent in the previous analysis, and eight correspond with results obtained in the previous study (**Supplementary Data 6**). Altogether, 17 new regions were detected. Of these, eight could be seen in the previous analysis, but had not been considered as they did not reach the significance threshold, and nine are solely due to the availability of our quail assembly. Two very short selection signatures previously detected using the chicken assembly as reference are not recovered here and were most probably false positives.

These results confirm the selection signature regions harbouring genes involved in human autistic disorders or being related to social behaviour[45] (*PTPRE*, *ARL13B*, *IMPK*, *CTNNA2*). Among the genes localised in the newly detected genomic regions, several have also been shown to be implicated in autism spectrum disorders or synaptogenic activity (**Supplementary Data 6**): mutations in the *EEF1A2* gene (Eukaryotic elongation factor 1, alpha-2) have been discovered in patients with autistic behaviours[47]; *EHMT1* (Euchromatin Histone Methyltransferase 1) is involved in autistic syndrome and social behaviour disorders in human and mouse[47-50]; *LRRTM4* (Leucine Rich Repeat Transmembrane Neuronal 4) is a synapse organizing protein, member of the *LRRTM* family, involved in mechanisms underlying experience-dependent synaptic plasticity[51].

Autistic spectrum disorders are observed in several disorders that have very different aetiology, including fragile X Syndrome, Rett Syndrome or Foetal Anticonvulsant Syndrome. While these disorders have very different underlying etiologies, they share common qualitative behavioural abnormalities in domains particularly relevant for social behaviours such as language, communication and social interaction[52,53]. In line with this, several experiments conducted on high social (HSR) and low social (LSR) reinstatement behaviour quail indicate that the selection program carried out with these lines is not limited to selection on a single response, social reinstatement, but affect more generally the ability of the quail to process social information[8]. Differences in social motivation, but also individual recognition have been described between LSR and HSR quail[49,50]. Inter-individual distances are longer in LSR quail[49] and LSR young quail have decreased interest in unfamiliar birds[51] and lower isolation distress than HSR ones[12].

Further experiments will be required to examine the possible functional link between the selected genes and the divergent phenotype observed in these lines. Also, by analyses of genes known to be differentially expressed in the zebra finch during song learning we hope to comparatively understand molecular systems linked to behaviour in the avian brain.

A model for avian seasonal biology

Quail is an important model for studying seasonal biology. Seminal work in quail established that pineal melatonin[57,58] is regulated by the circadian clock[59]. In mammals, photosensing is dependent on a single retinal photoreceptor melanopsin (*OPN4*) that regulates pineal melatonin release. Nocturnal melatonin is critical for mammalian neuroendocrine response to photoperiod and is likely to target melatonin receptors in the *pars tubularis* (PT)[60]. Birds have a distinct non-retinal mechanism for photoreception through deep-brain photoreceptors[61] and melatonin does not appear to be critical for most avian seasonal cycles[62]. The medial basal hypothalamus (MBH) seems to be a critical region for avian perception of photoperiod[63]. There are currently three main candidates for avian deep-brain photoreceptors that communicate the photoperiod signal to seasonal cycles: *OPN4*[64], neuropsin[65] (*OPN5*) and vertebrate ancient[66] (*VA*).

While melatonin may not be a critical component to avian photoperiod signal transduction it may play a role. Photoperiodic regulation of Gonadotropin-inhibitory hormone (GnIH), first identified in quail, has been shown to be regulated by melatonin[67]. Melatonin receptors are also located in the quail PT[68] and like the mammalian PT[69] the expression of core clock genes in the quail PT[70] are phase-shifted with photoperiod. Previously, two studies[63,71] have examined temperature dependent effects of photoperiod on core clock genes, TSHβ in the PT and DIO2 and DIO3 in the MBH. Here, we leverage the new quail genome for genome-wide analysis to determine how photoperiod and temperature interact to determine the MBH transcriptome (**Fig. 2A**).

We examined the effect of short- (SD) and long-day (LD) photoperiod (SD, 6L18D & LD, 20L4D) and temperature (9°C and 23°C) (**Fig. 2A; Supplementary Fig. 3**) on genome-wide transcription and identified 269 significantly differentially expressed genes (DEGs; FDR<0.05, $\log^2$FC>1; **Supplementary Data 7**). 127 DEGs were regulated irrespective of temperature, 60 and 82 DEGs were specific to the contrast with SD 9°C and 23°C, respectively.

We identified 16 temperature dependent DEGs with a large modulating effect of temperature ($\log^2$FC>1) (**Fig. 2E**). With the exception of aldehyde dehydrogenase (*ALDH1A1*), the temperature-dependent photoperiod effected DEGs were down-regulated in LD. There was an equal division of genes between temperature dependent amplification and suppression of LD down-regulated genes.

The MBH shows strong TSHβ induction in LD (**Fig. 2C-D**, $\log^2$FC=7.96 at 9°C, 8.36 at 23°C), indicating the stamp contains the adjacent PT as well as the MBH. Ikegami et al.[71] in-situ data support the localisation of TSHβ in the quail PT. Consistent with the MBH findings of Ikegami et al.[71], we observed significant up-regulation of *DIO2* and down-regulation of *DIO3*, in LD. We also observed a significant effect of cold (9°C) in short days as an amplifier of *DIO3* LP down-regulation (**Fig. 2E,** $\log^2$FC=-3.86 at 9°C, -2.51 at 23°C). We were unable to confirm any significant effect of cold on *DIO2*. We note significant photoperiod-dependent down-regulation of the thyroid hormone specific transporter *SLC16A2* in LP that was amplified at 9°C ($\log^2$FC=-1.19 at 9°C, -1.63 at 23°C).

Differential regulation of G-protein coupled receptor (GPCR) signalling was the most enriched pathway regulated by photoperiod (**Fig. 2F; Supplementary Data 8**). It also emerged as the largest connecting component within the String interaction network of DEG genes (**Fig. 2G**).

TSHβ itself binds to the GPCR THR[72]. G-protein signalling is also critical for opsin signalling[73]. We also observed transcriptional regulation in other GPCR hormone receptors, including Relaxin, Vasopressin, LH, Prolactin, and GH. GnRH is associated with VA opsins in AVT neurones and has been suggested as a photoperiod sensor[66]. We also noted down-regulation of the neuronally important GPCR GPR20 (**Fig. 2G**). In mice, deficiency of GPR20 is associated with hyperactivity and may play a role in cAMP-dependent mitogenesis[74]. There was a strong enrichment of collagen biosynthetic processes and extracellular matrix organisation processes (**Fig. 2F)** and a large body of genes associated with cell differentiation and development (**Fig. 2H**).

We observed photoperiod-dependent regulation of a single clock gene, *CRY4*. *CRY4* is up-regulated in LP (log^2FC=0.85 at 23°C, 1.37 at 9°C). This is consistent with the finding of Yasuo et al.[63], that the expression of *PER2-3*, *CLOCK*, *BMAL1*, *CRY1-2* and *E4BP4* remain stable across photoperiods. Intriguingly, CRY4 has recently been proposed as a component of light-dependent magnetoreception in the avian retina[75].

We detected photoperiod effects on *OPN4* transcripts, which were up-regulated in LD. Photoperiod-dependent expression in *OPN4* may well play a role in the photoperiod-refractory response. Encephalopsin (*OPN3*) was found to be highly expressed in the MBH (2.31..2.42 log^2CPM) but without significant changes in expression. OPN3 has recently been identified in the hypothalamus of chick hatchlings[76] but not as yet to the MBH of adult birds. *OPN5* (-0.46..-0.89 log^2CPM) and *VA* (-0.11..0.31 log^2CPM) were also unchanging and expressed at a low level in the MBH sample.

In conclusion, we confirm the importance of temperature and photoperiod-dependent regulation of thyroid hormone metabolism in the avian MBH (**Fig. 3**). Temperature-dependent amplification and suppression of the photoperiod response may indicate qualitative differences in the MBH pathways or simply reflect different stages of progression through seasonally phased processes. This could be further investigates by contrasting across time series at different temperatures. We also observed concurrent regulation of multiple hormonal signalling pathways, this may reflects a diversity of pathways and cell types in the MBH or reflect a corrective mechanism to account for cross-talk with other GPCR pathways. We observed LH, PRL and GH receptor transcripts changes which may indicate modulation of a GnRH-anterior pituitary feedback mechanism. Intriguingly, we also note LD induction of CRY4 which raises the question, is there a role for seasonally active magneto-sensing in the MBH? In addition to observing high *OPN3* expression in the MBH, we also noted LD overexpression of *OPN4*, which could provide a potential component for an avian photoperiod-refractory mechanism. This study has demonstrated the utility of genome-wide transcriptome analysis in quail to provide valuable insights and novel hypothesis for avian seasonal biology.

Quail response to H5N1 influenza

Highly pathogenic influenza A viruses (HPAI), such as H5N1, are responsible for enormous economic losses in the poultry industry and pose a serious threat to public health. While quail can survive infection with low pathogenic influenza viruses (LPAI), they experience high mortality when infected with strains of HPAI[77]. Quail are more susceptible than chickens to infection by some strains of H5N1 including one that caused human mortality[25]. Previous research has shown that quail may play a key role as an intermediate host in the evolution of avian influenza, allowing

viral strains to spread from wild birds to chickens and mammals[21,22,25,78]. Unlike quail and chicken, aquatic reservoir species such as duck are tolerant of most HPAI strains[79]. The generation of a high-quality quail genome has enabled us to perform an RNA-seq differential analysis of gene expression in quail infected with LPAI and HPAI, to better understand the response of quail to influenza infection. Lung and ileum samples were collected at 1 day post infection (1dpi) and 3 days post infection (3dpi). We also reanalysed previous data collected from duck and chickens[80] and compare this to the quail response.

To provide an overview of the response to LPAI and HPAI in quail we examined pathway and GO term enrichment of DEGs (see **Supplementary Data 9, 10 and Supplementary Figs. 4-7**). In response to LPAI infection, pathways enriched in the ileum included metabolism, JAK/STAT signalling, IL6 signalling and regulation of T-cells (**Supplementary Fig. 4)**. In the lung, pathways upregulated included complement, IL8 signalling, and leukocyte activation (**Supplementary Fig. 5)**. In the lung at 3dpi highly enriched GO terms included 'response to interferon-gamma', 'regulation of NF-kappaB', 'granulocyte chemotaxis' and 'response to virus' (**Supplementary Data 9)**, which are key influenza responses. This indicates an active immune response occurs to LPAI infection in quail, involving both ileum and lung, but with the strongest immune response occurring in the lung.

Genes upregulated in response to HPAI in the ileum were related to metabolism and transport, while inflammatory response was downregulated at 1dpi (**Supplementary Fig. 6**). Downregulated pathways at 1dpi included IL-6, IL-9 and neuroinflammation signalling pathways (**Supplementary Fig. 6**). In the quail lung many genes were downregulated after HPAI infection (**Supplementary Data 9**). At 3dpi, most downregulated pathways and terms were linked to immune system processes. GO terms with the highest fold enrichment in downregulated genes at this time included T and B cell proliferation, TNF signalling pathway, TLR pathway and IFN-G production **(Supplementary Data 10**). Pathways downregulated included both Th1 and Th2 pathways, T cell, B cell and macrophage signalling pathways **(Supplementary Fig. 7**). This indicates that key immune responses in quail, are downregulated in ileum, and particularly in the lung at day 3, following HPAI infection.

To compare the response of quail, duck and chicken, clustering of gene counts was examined using BioLayout 3D[81]. This revealed a cluster of 189 genes that were strongly upregulated at 1dpi in the duck, which showed no response in chicken and quail (**Supplementary Table 3**). This cluster was dominated by RIG-I pathway and IFN response genes including *IFNG*, *DDX60*, *DHX58*, *IRF1*, *IRF2*, and *MX1*. Pathways associated with this cluster includes MHCI processing and death receptor signalling (**Fig. 4**). Thus, the lack of this early anti-viral response may be key to the susceptibility of Galliformes to HPAI.

To further compare the responses between the three species, enrichment of pathways in each species was examined (**Fig. 5**). This revealed very few commonly regulated pathways between the three species. However, at 1dpi in the ileum and 3dpi in the lung there were many pathways that were downregulated in the quail, not altered in chicken, and upregulated in the duck. In the ileum at 1dpi, this included pattern recognition and death receptor signalling. In the lung at 3dpi this involved host of immune related pathways including production of NOS by macrophages, pattern recognition, B and T cell signalling and NK-KB, IL8 and IL2 signalling.

The proportion of genes commonly regulated between quail, chicken and duck to HPAI infection was also examined (**Fig. 6**). Consistent with the heatmap comparison (**Fig. 5**), the response of chicken, quail and duck were largely unique, with few genes commonly differentially expressed. There was a large set of genes that were upregulated in duck, while being downregulated in quail at 3dpi, in both ileum and lung. In lung these genes were related primarily to innate immune system pathways, including pattern recognition pathways, cytokine production, leukocyte adhesion, TNF production, interferon production, B cell signalling and response to virus (**Supplementary Data 10**). Genes with the greatest differential expression included *RSAD2* which inhibits viruses including influenza, *IFIT5* which senses viral RNA and *OASL* which has antiviral activity. These differences further highlight that the anti-viral immune response is dysregulated in quail. Additionally in both ileum and lung the apoptosis pathway was enriched in duck, but not quail (**Supplementary Data 10**). Apoptosis is known to be a key difference in the response of chickens and ducks to HPAI infection[82].

Lastly, we examined the response of key families involved in influenza and immune response, focussing on the lung (**Supplementary Table 5, Supplementary Data 11**). *IFTIM* genes have previously been found to have a key role in HPAI resistance[80] and may block AIV from entering cells[83]. Consistent with previous findings in the chicken[80], quail showed no significant upregulation of *IFTIM* genes, while these genes in duck were strongly upregulated, (**Supplementary Table 5**). TLRs and MHC receptors are involved in recognition of foreign molecules and triggering either an innate (TLR) or adaptive (MHC) immune response. TLR3, 4 and 7, which bind viral RNAs, were upregulated in response to LPAI in quail. A reversal was seen in response to HPAI, with *TLR4*, and *7* substantially downregulated. Likewise, genes of both MHC class I and II were upregulated in response to LPAI and downregulated in response to HPAI. By comparison there was no perturbation of TLR and MHC genes in chicken and upregulation of class I genes in duck. The quail seems to have a highly dysfunctional response to HPAI infection with key innate and adaptive immune markers downregulated at 3dpi, which contrasts with the strong immune response mounted by the duck and minimal immune response in the chicken.

In conclusion, we found that quail have a robust immune response to infection with LPAI, allowing them to survive the infection. However, they show dysregulation of the immune response after infection with HPAI, and this may explain their susceptibility to HPAI strains. IFITM response was not seen to HPAI while genes associated with apoptosis were downregulated, and this may allow the virus to easily enter cells and spread early in infection. Antiviral and innate immune genes, including those involved in antigen recognition, immune system activation, and anti-viral responses were downregulated at 3dpi, which would prevent an effective immune response and viral clearance once infection is established. This study provides crucial data that can be used to understand the differing response of bird species to AIV, which will be critical for managing and mitigating these diseases in the future.

DISCUSSION

Here we describe the assembly, annotation and use of a high-quality quail genome, an important avian model in biological and biomedical research. We have demonstrated the utility of this genome in both infectious disease and behavioural research providing further confirmation of the importance of quail as a research model, and for its role in agricultural and animal health studies. Specifically, the availability of this genome has allowed us to make significant discoveries

in the unique response of quail to highly pathogenic avian influenza infection, helping elucidate the basis for extreme susceptibility seen in this species. It has also allowed us to identify and confirm genes and genomic regions associated with social behaviour, showing many similarities to genes associated with autism in humans and thus represents a possible biomedical model for autism. Furthermore, we have shown that genome-wide transcriptomics using this genomes facilitated further insights and hypothesis into the mechanism of photoperiodism in avian seasonal biology. Moving forward, the availability of a high-quality quail genome will facilitate the study of diverse topics in both avian and human biology, including disease, behaviour, comparative genomics, seasonality and developmental biology.

METHODS

Whole Genome Sequencing and Assembly

To facilitate genome assembly by avoiding polymorphism, we produced an individual as inbred as possible. We started with a quail line previously selected for early egg production and having a high inbreeding coefficient[84] and four generations of brother-sister matings produced a dedicated line "ConsDD" (PEAT, INRA Tours, France). A 15 week-old male *Coturnix japonica* (id. 7356) was then selected from this line for the sequencing project. Genomic DNA was extracted from a blood sample using a high-salt extraction method[85]. Our sequencing plan followed the recommendations provided in the ALLPATHS2 assembler[86]. This model requires 45x sequence coverage of each fragment (overlapping paired reads ~180 bp length) from 3 kb paired end (PE) reads as well as 5x coverage of 8 kb PE reads. These sequences were generated on the HiSeq2500 Illumina instrument. Long-reads used for gap filling were generated at 20x coverage on the same DNA source using a RSII instrument (Pacific Biosciences). The Illumina sequence reads were assembled using ALLPATHS2 software[86] using default parameter settings and where possible, and scaffold gaps were closed by mapping and local assembly of long-reads using PBJelly[87]. The Illumina long insert paired-end reads (3 kb and 8kb PE) were used to further extend assembled scaffolds using SSPACE[83]. The draft assembly scaffolds were then aligned to the genetic linkage map[44] and the Galgal4.0 chicken reference (Genbank accession: GCA_000002315.2) to construct chromosome files following previously established methods[32]. Finally, all contaminating contigs identified by NCBI filters (alignments to non-avian species at the highest BLAST score obtained), and all contigs < 200 bp were removed prior to final assembly submission.

Gene Annotation

Specific RNA-seq data for the genome annotation was produced from the same animal used for the genome assembly. RNA was extracted from heart, kidney, lung, brain, liver, intestine, and muscle using Trizol and the Nucleospin® RNA II kit (MACHEREY-NAGEL), following the manufacturer's protocol.

The *Coturnix japonica* assembly was annotated using the NCBI pipeline, including masking of repeats prior to *ab initio* gene predictions, for evidence-supported gene model building. We utilized an extensive variety of RNA-Seq data to further improve gene model accuracy by alignment to nascent gene models that are necessary to delineate boundaries of untranslated

regions as well as to identify genes not found through interspecific similarity evidence from other species. A full description of the NCBI gene annotation pipeline was previously described[89]. Around 8,000 lacked gene symbols from this pipeline, and these were further annotated manually by using BLAST searches using the corresponding sequences and extracting protein names from Uniprot.

Comparative analyses

A set of single copy, orthologous, avian-specific genes were selected from OrthoDB v. 9[30] and their status (present, duplicated, fragment or missing) were tested with BUSCO v.3.0.2[31] in the *Gallus gallus* 5.0 and *Coturnix japonica* 2.0 genomes. *Ab initio* gene predictions were done within the BUSCO framework using tBLASTn matches[90] followed by avian specific gene predictions with Augustus v. 3.3[91]. Gene status was assessed by running HMMER[92] with the BUSCO HMM profiles of the orthologous sequences. Comparative maps and breakpoint data were generated using AutoGRAPH[93] using chicken and quail gff annotation files, using default settings.

Endogenous retrovirus identification

Endogenous retroviruses (ERVs) were identified in the *Coturnix japonica* 2.0 and Turkey 5.0 genome assemblies using the LocaTR identification pipeline[35] and compared to a previous analysis of ERVs in the Gallus gallus 5.0 genome assembly[32]. LocaTR is an iterative pipeline which incorporates LTR_STRUC[94], LTRharvest[95], MGEScan_LTR[96] and RepeatMasker[97] search algorithms.

Sociability selection study

The data and methods used have been described previously[45]. Briefly, two quail lines were used, divergently selected on their sociability[9]: high social (HSR) and low social (LSR) reinstatement behaviour. A total of 10 individuals from generation 50 of each quail line were sequenced after equimolar DNA pooling. Sequencing was performed (paired-ends, 100 bp) on a HiSeq 2000 sequencer (Illumina), using one lane per line (TruSeq sbs kit version 3). The reads (190,159,084 and 230,805,732 reads, respectively, for the HSR and LSR lines) were mapped to the CoJa2.2 genome assembly using BWA[98], with the mem algorithm. Data are publicly available under SRA accession number SRP047364. Within each line, the frequency of the reference allele was estimated for all SNPs covered by at least 5 reads, using Pool-HMM[99]. This analysis provided 13,506,139 SNPs with allele frequency estimates in the two lines. FLK values[46] were computed for all these SNPs, and the local score method[45] was applied to the p-value on single-marker tests.

Photoperiod study

MBH tissue was collected as previously[71]. Male 4-week old quail were obtained from a local dealer in Japan and kept under SD conditions (6L18D) for 4 weeks. At 8 weeks of age, quail were

transferred to LD conditions (20L4D) and kept under LD conditions for 4 weeks to develop their testes. And then, 12 week-old LD quail were transferred to short day and low temperature (SL: 6L18D 9C) conditions for another 4 weeks to fully regress their testes. All samples were collected at ZT18. (Light onset is same for LD and SD and light offset was extended in LD group). RNA-Seq was performed using a TruSeq stranded mRNA prep (Revision E 15031047) with 125bp paired-end reads on a HiSeq Illumina 2500 with four replicates in each of the three conditions.

Reads were quality (Phred>25) and adapter trimmed with trim galore (version 0.4.5). Tophat (version 2.1.0)[100] with bowtie2 (version 2.2.6) was used to map reads to the quail genome (GCA_001577835.1 *Coturnix japonica* 2.0), using the NCBI annotation. We determined feature counts for gene loci using the featureCounts program[101] in the subread (version 1.5.0) package[102]. Statistical analysis was performed using the limma package[103] (version 3.36.1) in the R programming environment (version 3.5.0). The trimmed mean of M-values normalization method (TMM) was used for normalisation with Voom for error estimation (Sup. Table 2). We retained gene loci with more than 10x coverage in three replicates in at least two conditions. A categorical least squared regression model was fitted using LD 23°C, SD 23°C, and SD 9°C conditions. Statistics for pairwise comparisons were then recalculated by refitting contrasts to the model for LD 23°C vs SD 23°C, LD 23°C vs SD 9°C and SD 23°C vs SD. The Benjamini Hochberg[104] approach was used to estimate the false discovery rate. For reporting numbers of photoperiod significant genes, we applied thresholds of FDR <0.05, log2 CPM > 0, and absolute log2 fold change > 1. Temperature-dependent genes are reported as those with a photoperiod significant effect at either 23°C or 9°C and a significant effect when contrasting SD 9°C and SD 23°C at the same thresholds defined across photoperiods.

Influenza response study

All experiments involving animals were approved by the Animal Care and Use Committee of St. Jude Children's Research Hospital and performed in compliance with relevant policies of the National Institutes of Health and the Animal Welfare Act. All animal challenge experiments were performed in animal biosafety level 2 containment facilities for the LPAI challenges and in biosafety level 3 enhanced containment laboratories for the HPAI challenges. Viral challenges of quail, tissue collection, RNA extractions and sequencing were carried out as previously described for chicken[80]. Briefly, fifteen mixed-sex quail were challenged with 106 EID_{50} intranasally, intratracheally, and intraocularly of LPAI A/Mallard/British Columbia/500/2005 (H5N2) in phosphate buffered saline (PBS). Fifteen quail were challenged with 101.5 EID50 intranasally, intratracheally, and intraocularly of HPAI A/Vietnam/1203/2004 (H5N1) in PBS. Mock infection control groups (n=12) were also inoculated, receiving an equivalent volume and route of administration with PBS. Animals were monitored daily for clinical signs. Lung and ileum samples were collected from all birds on 1dpi and 3 dpi. RNA extractions were performed using Trizol and QIAGEN's RNeasy kit. For sequencing thirty-six cycle single-ended sequencing was carried out on the Genome Analyser IIx using Illumina v3 Sequencing by Synthesis kits.

All quail as well as duck and chicken RNA-seq reads from the previous study[80] were analysed as follows. Ileum and lung RNAs were analysed from PBS infected control (3 samples from each of 1dpi and 3dpi), H5N1-infected (3 samples from each of 1dpi and 3dpi, except quail ileum 1dpi which had 2 samples) and H5N2-infected (3 samples from each of 1dpi and 3dpi). 251

million reads of 36 nucleotides in length were generated in total for quail. Reads were quality checked using FastQC and trimmed for quality using Trim-galore. Mapping was performed to the quail genome (GCA_001577835.1 Coturnix japonica 2.0), chicken genome (GCA_000002315.3 Gallus_gallus-5.0) and duck (GCA_000355885.1 BGI_duck_1.0) using Tophat2[100] using default options. For quantification and differential analysis, the following pipeline was used. First transcripts were assembled and quantified using cufflinks[105], guided with the NCBI annotation for the relevant genome, and the multi-read correct option was used. The transcriptomes were merged using stringtie merge[106] and cuffdiff[105] was used for differential analysis using default settings. To determine orthology between quail, duck and chicken genes, reciprocal BLAST searches were performed. For analysis of GO term enrichment the PANTHER overrepresentation test[107] was used and for pathway analysis Ingenuity Pathway Analysis software (QIAGEN) was used. For clustering analysis BioLayout 3D[76] was used using default settings except 1.4 inflation for Markov clustering.

REFERENCES

1. Shimakura, K. Notes on the genetics of the Japanese quail: I. The simple, Mendelian, autosomal, recessive character, "brown-splashed white," of its plumage (in Japanese with English summary). *Jpn. J. Genet.* **16**, 106–112 (1940).
2. Minvielle, F. What are quail good for in a chicken-focused world? *World's Poult. Sci. J.* **65**, 601–608 (2009).
3. Huss, D., Poynter, G. & Lansford, R. Japanese quail (*Coturnix japonica*) as a laboratory animal model. *Lab. Anim. (NY)* **37**, 513–519 (2008).
4. Yvernogeau, L. et al. An in vitro model of hemogenic endothelium commitment and hematopoietic production. *Development* **143**, 1302–1312 (2016).
5. Mills, A. D., Crawford, L. L., Domjan, M. & Faure J. M. The behavior of the Japanese or domestic quail *Coturnix japonica*. *Neurosci. Biobehav. Rev.* **21,** 261–281 (1997).
6. Adkins-Regan, E. Hormones and sexual differentiation of avian social behavior. *Dev Neurosci.* **31**, 342–350 (2009).
7. Meddle, S. L. et al. Copulation activates Fos-like immunoreactivity in the male quail forebrain. *Behav. Brain Res*. **85**, 143–159 (1997).
8. Marasco, V., Herzyk, P., Robinson, J. & Spencer, K. A. Pre- and Post-Natal Stress Programming: Developmental Exposure to Glucocorticoids Causes Long-Term Brain-Region Specific Changes to Transcriptome in the Precocial Japanese Quail. *J. Neuroendocrinol.* **28**, (2016).
9. Mills, A. D. & Faure, J. M. Divergent selection for duration of tonic immobility and social reinstatement behavior in Japanese quail (*Coturnix coturnix japonica*) chicks. *J. Comp. Psychol.* **105**, 25–38 (1991).
10. Jones, R. B. & Mills, A. D. Divergent selection for social reinstatement behaviour in Japanese quail: Effects on sociality and social discrimination. *Avian Poultry Biol. Rev.* **10**, 213–223 (1999).
11. Beaumont, C. et al. A genome scan with AFLP markers to detect fearfulness-related QTLs in Japanese quail. *Anim. Genet*. **36**, 401–407 (2005).
12. Recoquillay, J. et al. Evidence of phenotypic and genetic relationships between sociality, emotional reactivity and production traits in Japanese quail. *PLoS One* **8**, e82157 (2013).
13. Robinson, J. E. & Follett, B. K. Photoperiodism in Japanese quail: the termination of seasonal breeding by photorefractoriness. *Proc. R. Soc. Lond. B Biol. Sci.* **215**, 95–116 (1982).
14. Nakane, Y. & Yoshimura, T. Deep brain photoreceptors and a seasonal signal transduction cascade in birds. *Cell Tissue Res*. **342**, 341–344 (2010).
15. Nakane, Y. & Yoshimura, T. Universality and diversity in the signal transduction pathway that regulates seasonal reproduction in vertebrates. *Front. Neurosci*. **8**, 115 (2014).
16. Homma, K., Jinno, M., Sato, K. & Ando, A. Studies on perfect and imperfect albinism in the Japanese quail (*Coturnix coturnix japonica*). *Jpn. J. Zootechnical. Sci.* **39**, 348–352 (1968).
17. Waligora-Dupriet, A. J. et al. (2009) Short-chain fatty acids and polyamines in the pathogenesis of necrotizing enterocolitis: Kinetics aspects in gnotobiotic quails. *Anaerobe* **15**, 138–144 (2009).
18. Watanabe, S. & Nagayama, F. Studies on the serum IgG levelin Japanese quail. *Jpn. Poult. Sci.* **16**, 59–64 (1979).

19. Shiina, T. et al. Comparative Genomic Analysis of Two Avian (Quail and Chicken) MHC Regions. *J. Immunol*. **172**, 6751–6763 (2004).
20. Hosomichi, K. et al. The major histocompatibility complex (Mhc) class IIB region has greater genomic structural flexibility and diversity in the quail than the chicken. *BMC Genomics* **7**, 322 (2006).
21. Makarova, N. V., Ozaki, H., Kida, H., Webster, R. G. & Perez, D. R. Replication and transmission of influenza viruses in Japanese quail. *Virology* **310**, 8–15 (2003).
22. Perez, D. R. et al. Role of quail in the interspecies transmission of H9 influenza A viruses: molecular changes on HA that correspond to adaptation from ducks to chickens. *J. Virol.* **77**, 3148–3156 (2003).
23. Wan, H. & Perez, D. R. Quail carry sialic acid receptors compatible with binding of avian and human influenza viruses. *Virology* **346**, 278–286 (2006).
24. Guan, Y. et al. Emergence of multiple genotypes of H5N1 avian influenza viruses in Hong Kong SAR. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 8950–8955 (2002).
25. Webster, R. G. et al. Characterization of H5N1 influenza viruses that continue to circulate in geese in southeastern China. *J. Virol*. **76**, 118–126 (2002).
26. Oldeschulte, D. L. et al. Annotated Draft Genome Assemblies for the Northern Bobwhite (*Colinus virginianus*) and the Scaled Quail (*Callipepla squamata*) Reveal Disparate Estimates of Modern Genome Diversity and Historic Effective Population Size. *G3 (Bethesda)* **7**, 3047–3058 (2017).
27. Wu, Y. et al. Population genomic data reveal genes related to important traits of quail. *Gigascience* **7**, giy049, (2018).
28. Morgulis, A., Gertz, E. M., Schäffer, A. A. & Agarwala, R. WindowMasker: window-based masker for sequenced genomes. *Bioinformatics* **22**, 134–141 (2006).
29. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol*. 215, 403–410 (1990).
30. Zdobnov, E. M. et al. OrthoDB v9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res.* **45**, D744–D749 (2017).
31. Waterhouse, R. M. et al. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Mol. Biol. Evol.* **35**, 543–548 (2017).
32. Warren, W. C. et al. A New Chicken Genome Assembly Provides Insight into Avian Genome Structure. *G3*, **7**, 109–117 (2017).
33. van Tuinen, M. & Dyke, G. J. Calibration of galliform molecular clocks using multiple fossils and genetic partitions. *Mol. Phylogenet. Evol.* **30**, 74–86 (2004).
34. Dalloul, R. A. et al. Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. *PLoS Biol.* **8**, e1000475 (2010).
35. Mason, A. S., Fulton, J. E., Hocking, P. M. & Burt D. W. A new look at the LTR retrotransposon content of the chicken genome. *BMC Genomics*, **17**, 688 (2016).
36. Kapusta, A. & Suh, A. Evolution of bird genomes-a transposons-eye view, *Ann. N. Y. Acad. Sci.* **1389**, 164–185 (2017).
37. Varela, M., Spencer, T. E., Palmarini, M. & Arnaud, F. Friendly viruses. *Ann. N. Y. Acad. Sci.* **1178**, 157–172 (2009).
38. Aswad, A. & Katzourakis, A. Paleovirology and virally derived immunity, *Trends Ecol. Evol.* **27**, 627–636 (2012).

39. Kapusta, A., Suh, A. & Feschotte, C. Dynamics of genome size evolution in birds and mammals. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1460–1469 (2017).
40. Salomonsen, J. et al. Sequence of a complete chicken BG haplotype shows dynamic expansion and contraction of two gene lineages with particular expression patterns. *PLoS Genet.* **10**, e1004417 (2014).
41. Cheng, Y., Prickett, M. D., Gutowska, W., Kuo, R., Belov, K. & Burt D. W. Evolution of the avian β-defensin and cathelicidin genes. *BMC Evol. Biol.* **15**, 188 (2015).
42. Barber, M. R., Aldridge, J. R. Jr., Webster, R. G. & Magor, K. E. Association of RIG-I with innate immunity of ducks to influenza. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 5913–5918 (2010)
43. Magor, K. E. et al. Defense genes missing from the flight division. *Dev. Comp. Immunol.* **41**, 377–388 (2013).
44. Recoquillay, J. et al. A medium density genetic map and QTL for behavioral and production traits in Japanese quail. *BMC Genomics* **16**, 10 (2015).
45. Fariello, M. I. et al. Accounting for linkage disequilibrium in genome scans for selection without individual genotypes: The local score approach. *Mol. Ecol.* **26**, 3700–3714 (2017).
46. Bonhomme, M. et al. Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics*, **186**, 241–262 (2010).
47. Nakajima, J. et al. De novo EEF1A2 mutations in patients with characteristic facial features, intellectual disability, autistic behaviors and epilepsy. *Clin. Genet.* **87**, 356–361 (2015).
48. Kleefstra, T. et al. Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am. J. Hum. Genet.* **79**, 370–377 (2006).
49. Balemans, M. C. et al. Reduced exploration, increased anxiety, and altered social behavior: Autistic-like features of euchromatin histone methyltransferase 1 heterozygous knockout mice. *Behav. Brain. Res.* **208**, 47–55 (2010).
50. Mitra, A. K., Dodge, J., Van Ness, J., Sokeye, I. & Van Ness, B. A de novo splice site mutation in EHMT1 resulting in Kleefstra syndrome with pharmacogenomics screening and behavior therapy for regressive behaviors. *Mol. Genet. Genomic Med.* **5**, 130–140 (2017).
51. Roppongi, R. T., Karimi, B. & Siddiqui, T. J. Role of LRRTMs in synapse development and plasticity. *Neurosci. Res.* **116**, 18–28 (2017).
52. Rutter, M. Diagnosis and definition of childhood autism. *J. Autism Child Schizophr.* **8**, 139–161 (1978).
53. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 4th edition (2000).
54. Francois, N., Mills, A. D. & Faure, J. M. Inter-individual distances during open-field tests in Japanese quail (*Coturnix japonica*) selected for high or low levels of social reinstatement behaviour. *Behav. Processes* **47**, 73–80 (1999).
55. Schweitzer, C., Houdelier, C., Lumineau, S., Levy, F. & Arnould, C. Social motivation does not go hand in hand with social bonding between two familiar Japanese quail chicks, *Coturnix japonica*. *Animal Behaviour* **79**, 571–578 (2009).
56. Francois, N., Decros, S., Picard, M., Faure, J. M. & Mills, A. D. Effect of group disruption on social behaviour in lines of Japanese quail (*Coturnix japonica*) selected for high or low levels of social reinstatement behaviour. *Behav. Processes* **48**, 171–181 (2000).

57. Ralph, C., Hedlund, L. & Murphy, W. A. Diurnal cycles of melatonin in bird pineal bodies. *Comp. Biochem. Physiol.* **22**, 591–599 (1967).
58. Lynch, H. J. Diurnal oscillations in pineal melatonin content. *Life Sci. Pt 1 Physiol. Pharmacol.* **10**, 791–795 (1971).
59. Cockrem, J. F. & Follett, B. K. Circadian rhythm of melatonin in the pineal gland of the Japanese quail (*Coturnix coturnix japonica*). *J. Endocrinol.* **107**, 317–324 (1985).
60. Wood, S. & Loudon, A. Clocks for all seasons: unwinding the roles and mechanisms of circadian and interval timers in the hypothalamus and pituitary. *J. Endocrinol*. **222**, R39–59 (2014).
61. Menaker, M. Extraretinal light perception in the sparrow. I. Entrainment of the biological clock. *Proc. Natl. Acad. Sci. U. S. A.* **59**, 414–421 (1968).
62. Yoshimura, T. Thyroid hormone and seasonal regulation of reproduction. *Front. Neuroendocrinol.* **34**, 157–166 (2013).
63. Yasuo, S., Watanabe, M., Okabayashi, N., Ebihara, S. & Yoshimura, T. Circadian Clock Genes and Photoperiodism: Comprehensive Analysis of Clock Gene Expression in the Mediobasal Hypothalamus, the Suprachiasmatic Nucleus, and the Pineal Gland of Japanese Quail under Various Light Schedules. *Endocrinology* **144**, 3742–3748 (2003).
64. Haas, R., Alenciks, E., Meddle, S. & Fraley, G. S. Expression of deep brain photoreceptors in the Pekin drake: a possible role in the maintenance of testicular function. *Poult. Sci*. **96**, 2908–2919 (2017).
65. Nakane, Y. et al. A mammalian neural tissue opsin (Opsin 5) is a deep brain photoreceptor in birds. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15264–15268 (2010).
66. García-Fernández, J. M. et al. The hypothalamic photoreceptors regulating seasonal reproduction in birds: a prime role for VA opsin. *Front. Neuroendocrinol.* **37**, 13–28 (2015).
67. Chowdhury, V. S. et al. Melatonin Stimulates the Release of Gonadotropin-Inhibitory Hormone by the Avian Hypothalamus. *Endocrinology* **151**, 271–280 (2010).
68. Cozzi, B. et al. Distribution and characterization of melatonin receptors in the brain of the Japanese quail, *Coturnix japonica*. *Neurosci. Lett.* **150**, 149–152 (1993).
69. Lincoln, G., Messager, S., Andersson, H. & Hazlerigg, D. Temporal expression of seven clock genes in the suprachiasmatic nucleus and the pars tuberalis of the sheep: evidence for an internal coincidence timer. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13890–13895 (2002).
70. Yasuo, S. et al. Photoinducible Phase-Specific Light Induction of Cry1 Gene in the Pars Tuberalis of Japanese Quail. *Endocrinology* **145**, 1612–1616 (2004).
71. Ikegami, K. et al. Low temperature-induced circulating triiodothyronine accelerates seasonal testicular regression. *Endocrinology* **156**, 647–659 (2015).
72. Millar, R. P., Newton, C. L. & Roseweir, A. K. Chapter 2 - Neuroendocrine GPCR Signaling. In *Handbook of Neuroendocrinology* (eds Fink, G., Pfaff, D. W. & Levine, J. E.) pp. 21–53 (Academic Press; 2012).
73. Shichida, Y. & Matsuyama, T. Evolution of opsins and phototransduction. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 2881–2895 (2009).
74. Hase, M., Yokomizo, T., Shimizu, T. & Nakamura, M. Characterization of an Orphan G Protein-coupled Receptor, GPR20, That Constitutively Activates Gi Proteins. *J. Biol. Chem.* **283**, 12747–12755 (2008).

75. Pinzon-Rodriguez, A., Bensch, S. & Muheim, R. Expression patterns of cryptochrome genes in avian retina suggest involvement of Cry4 in light-dependent magnetoreception. *J. R. Soc. Interface* **15**, 20180058 (2018).
76. Kato, M. et al. Two Opsin 3-Related Proteins in the Chicken Retina and Brain: A TMT-Type Opsin 3 Is a Blue-Light Sensor in Retinal Horizontal Cells, Hypothalamus, and Cerebellum. *PLoS One*. **11**, e0163925 (2016).
77. Bertran, K. et al. Pathobiology and transmission of highly and low pathogenic avian influenza viruses in European quail (*Coturnix c. coturnix*). *Vet. Res.* **44**, 23 (2013).
78. Nguyen, T. H. et al. Intersubtype Reassortments of H5N1 Highly Pathogenic Avian Influenza Viruses Isolated from Quail. *PLoS One* **11**, e0149608 (2016).
79. Cornelissen, J. B., Vervelde, L., Post, J. & Rebel, J. M. Differences in highly pathogenic avian influenza viral pathogenesis and associated early inflammatory response in chickens and ducks. *Avian Pathol.* **42,** 347–364 (2013).
80. Smith, J. et al. A comparative analysis of host responses to avian influenza infection in ducks and chickens highlights a role for the interferon-induced transmembrane proteins in viral resistance. *BMC Genomics* **16**, 574 (2015).
81. Theocharidis, A., van Dongen, S., Enright, A. J. & Freeman, T. C. Network visualization and analysis of gene expression data using BioLayout Express (3D). *Nat. Protoc.* **4**, 1535–1550 (2009).
82. Kuchipudi, S. V. et al. Rapid death of duck cells infected with influenza: a potential mechanism for host resistance to H5N1. *Immunol. Cell Biol.* **90**, 116–123 (2010).
83. Amini-Bavil-Olyaee, S. et al. The antiviral effector IFITM3 disrupts intracellular cholesterol homeostasis to block viral entry. *Cell Host Microbe* **13**, 452–464 (2013).
84. Minvielle, F., Monvoisin, J. L., Costa, J., Frénot, A. & Maeda, Y. Changes in heterosis under within-line selection or reciprocal recurrent selection: an experiment on early egg production in Japanese quail. *J. Anim. Breed. Genet*. **116**, 363–377 (1999).
85. Roussot, O. et al. AFLP linkage map of the Japanese quail *Coturnix japonica. Genet. Sel. Evol.* **35**, 559–572 (2003).
86. MacCallum, I. et al. ALLPATHS 2: small genomes assembled accurately and with high continuity from short paired reads. *Genome Biol.* **10**, R103 (2009).
87. English, A. C. et al. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One*. **7**, e47768 (2012).
88. Boetzer, M. & Pirovano, W. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* **15**, 211 (2014).
89. Thibaud-Nissen, F., Souvorov, A., Murphy, T., DiCuccio, M. & Kitts, P. Eukaryotic Genome Annotation Pipeline. In *The NCBI Handbook [Internet]*. 2nd edn. 133–156 (National Center for Biotechnology Information (US), 2013)
90. NCBI Resource Coordinators. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **45**, D12–D17 (2017).
91. Stanke, M. et al. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–W439 (2006).
92. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14**, 755–763 (1998).
93. Derrien, T., André, C., Galibert, F. & Hitte, C. AutoGRAPH: an interactive web server for automating and visualizing comparative genome maps. *Bioinformatics* **23**, 498–499 (2007).

94. McCarthy, E. M. & McDonald, J. F. LTR_STRUC: a novel search and identification program for LTR retrotransposons, *Bioinformatics* **19**, 362–367 (2003).
95. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).
96. Rho, M., Choi, J. H., Kim, S., Lynch, M. & Tang, H. De novo identification of LTR retrotransposons in eukaryotic genomes. *BMC Genomics* **8**, 90 (2007).
97. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0.3 (2013) (Available: http://repeatmasker.org).
98. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
99. Boitard, S., Kofler, R., Francoise, P., Robelin, D., Schlotterer, C. & Futschik, A. Pool-hmm: a Python program for estimating the allele frequency spectrum and detecting selective sweeps from next generation sequencing of pooled samples. *Mol. Ecol. Resour.* **13**, 337–340 (2013).
100. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. & Salzberg, S. L. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
101. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930 (2014).
102. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108, (2013)
103. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
104. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).
105. Trapnell, C. et al. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
106. Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T. & Salzberg, S. L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* (2015).
107. Thomas, P. D. et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* **13**, 2129–2141 (2003).

Abbreviations

AIV – avian influenza virus

bp – base pairs

CR1 – chicken repeat 1

DEG – differentially expressed gene

dpi – days post infection

EID_{50} - 50% embryo infectious dose

ERV – endogenous retrovirus

Gb – gigabases

GO - gene ontology

HPAI – highly pathogenic avian influenza

HSR - high levels of social reinstatement

IFITM – interferon induced transmembrane protein

IFN – interferon

Kb - kilobases

LINE - long interspersed nuclear element

LPAI – low pathogenic avian influenza

LSR - low levels of social reinstatement

Mb - megabases

MHC – major histocompatability complex

NA- Not available

PBS – phosphate buffered saline

SNP – single nucleotide polymorphism

TAP – transporter associated with antigen processing

TLR – toll-like receptor

TNF – tumour necrosis factor

Ethics statement

ConsDD, HSR and LSR animals were bred at INRA, UE1295 Pôle d'Expérimentation Avicole de Tours, F-37380 Nouzilly, in accordance with European Union Guidelines for animal care, following the Council Directives 98/58/EC and 86/609/EEC. Animals were maintained under standard breeding conditions and subjected to minimal disturbance. Furthermore, the ethics

committee approved the rearing protocol (authorization number 00915.02). The use of quail in photoperiod experiments were approved by the Animal Experiment Committee of Nagoya University. All experiments involving animals in the infection studies were approved by the Animal Care and Use Committee of St. Jude Children's Research Hospital and performed in compliance with relevant policies of the National Institutes of Health and the Animal Welfare Act.

Availability of supporting data

Data has been submitted to the public databses under the following accession numbers: GCA_001577835.1 [genome sequence data]; SRA SRR2968870 - SRR2968911 [transcription annotation data]; Array Express accession numbers E-MTAB-3311(quail), E-MTAB-2909 (duck), E-MTAB-2908 (chicken) [RNAseq data for infection studies]; SRA accession number SRP047364 [HSR/LSR sequencing]; SRA accession number PRJNA490454 [RNA-seq data for photoperiod study].

Acknowledgements

The authors would like to thank the Edinburgh Genomics sequencing facility (Edinburgh, UK) for carrying out the transcriptomic sequencing, RNA sequencing of the reference individual was performed on the GeT-Plage platform (http://get.genotoul.fr/en/) and funded by the INRA Genetics Division (QuailAnnot program). HSR and LSR lines sequencing was supported by the French Agence Nationale de la Recherche (SNP-BB project, ANR-009-GENM-008). We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) for providing help, computing and storage for these resources and to the ITAVI SeqVol program for financing gallo-anseriformes genome sequencing. KMM was supported by a National Health and Medical Research Council Overseas Postdoctoral Fellowship. This work was funded in part by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, under contract numbers HHSN266200700005C and HHSN272201400006C, by ALSAC and by BB/N015347/1 and an HFSP 2015 award (RGP0030/2015).

Authors' contributions

Inbred quail line for sequencing: FM, DG; Selection for social motivation: CL, DG; Genome sequencing and assembly: WW, JG, CT, PM, LH, DB; Transcriptome for annotation: AV, FP; Assembly quality assessment; LE; ERV analysis: AM; selection signature analyses: FP, SB, AV; photoperiod study: TY, DB, TS, SM, AL, MMH; avian flu studies: RW, JPS, KM, AD, HF, JS, DB; preparation of manuscript: KM, JS, TY, SM, AL, MMH, DB, AV, WW, LE, AM.

Competing interests

The authors declare there are no competing interests

TABLES

Table 1. Representative assembly metrics for sequenced Galliform genomes[1].

Common name	Assembled version	N50 contig (Mb)	N50 scaffold (Mb)	Total assembly size (Gb)	Assembled chromosomes
Japanese quail	Coturnix japonica 2.0	0.511	3.0	0.93	33
Japanese quail	Wu et al. PMID: 29762663	0.027	1.8	1.01	34
Chicken	Gallus gallus 5.0	2.895	6.3	1.20	30
Scaled quail	ASM221830v1	0.154	1.0	1.01	NA
Northern bobwhite	ASM59946v2	0.056	2.0	1.13	NA
Turkey	Turkey 5.0	0.036	3.8	1.12	33

[1]All species-specific assembly metrics derived from the NCBI assembly archive.

Table 2. Representative gene annotation measures for assembled Galliform genomes[1].

Common name	Assembled version	Protein coding genes	Total ncRNA	mRNAs
Japanese quail	Coturnix japonica 2.0	16,057	4,108	39,075
Japanese quail	Wu et al. PMID: 29762663	16, 210	NA	NA
Chicken	Gallus gallus 5.0	19,137	6,550	46,334

Turkey	Turkey 5.0	18,511	8,552	33,308

[1]All species-specific gene annotation metrics derived from the NCBI RefSeq database.

Table 3. Estimates of gene and protein representation for sequenced Galliform genomes.

		Transcript[1]		Protein[2]	
Common name	Assembled version	Average % identity	Average % coverage	Average % identity	Average % coverage
Japanese quail	Coturnix japonica 2.0	93.4	96.2	80	85
Chicken	Gallus gallus 5.0	90.4	84.3	78	84.6
Turkey	Turkey 5.0	99.1	93.8	80.7	80.1

[1] Predicted transcripts per species aligned to Aves known RefSeq transcripts (n=8,776); turkey aligned to same species Genebank (n=380).
[2] Predicted proteins per species aligned to Aves known RefSeq (n=7,733).

FIGURES

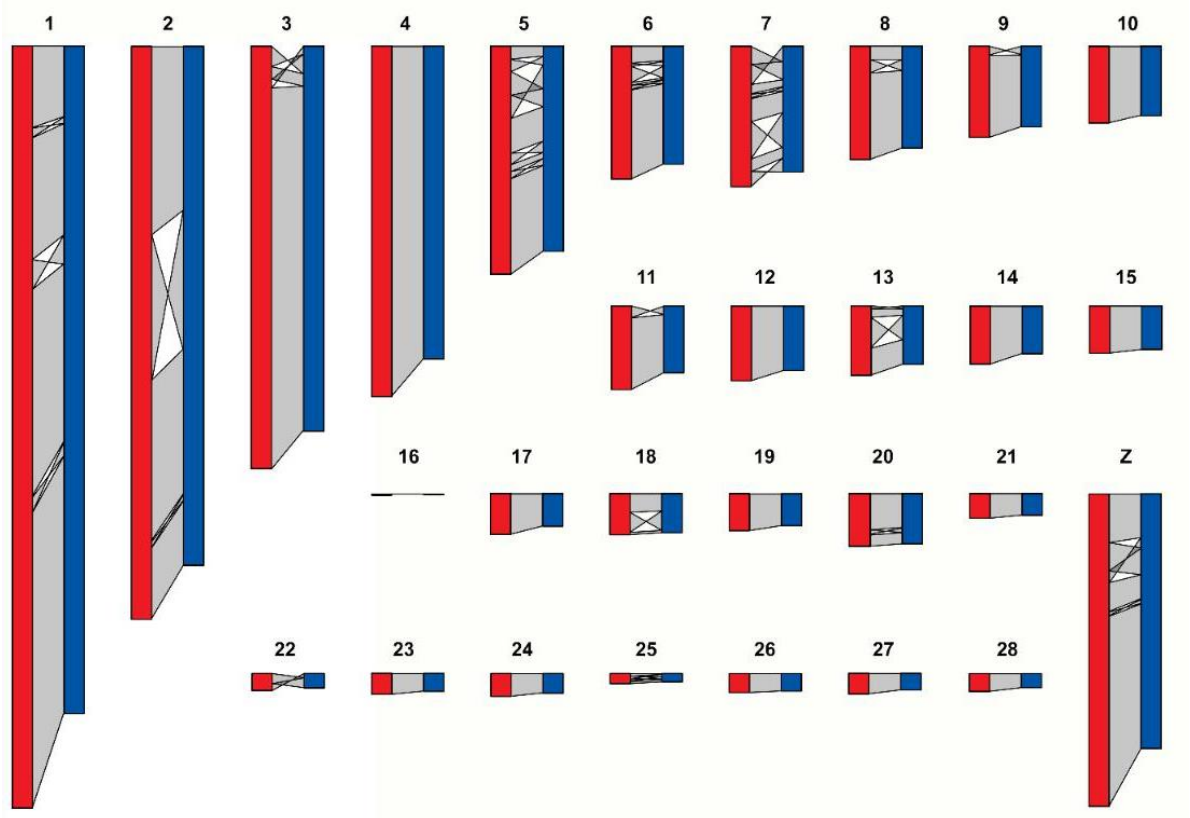


Figure 1: Synteny map of chicken (red) and quail (blue) chromosomes

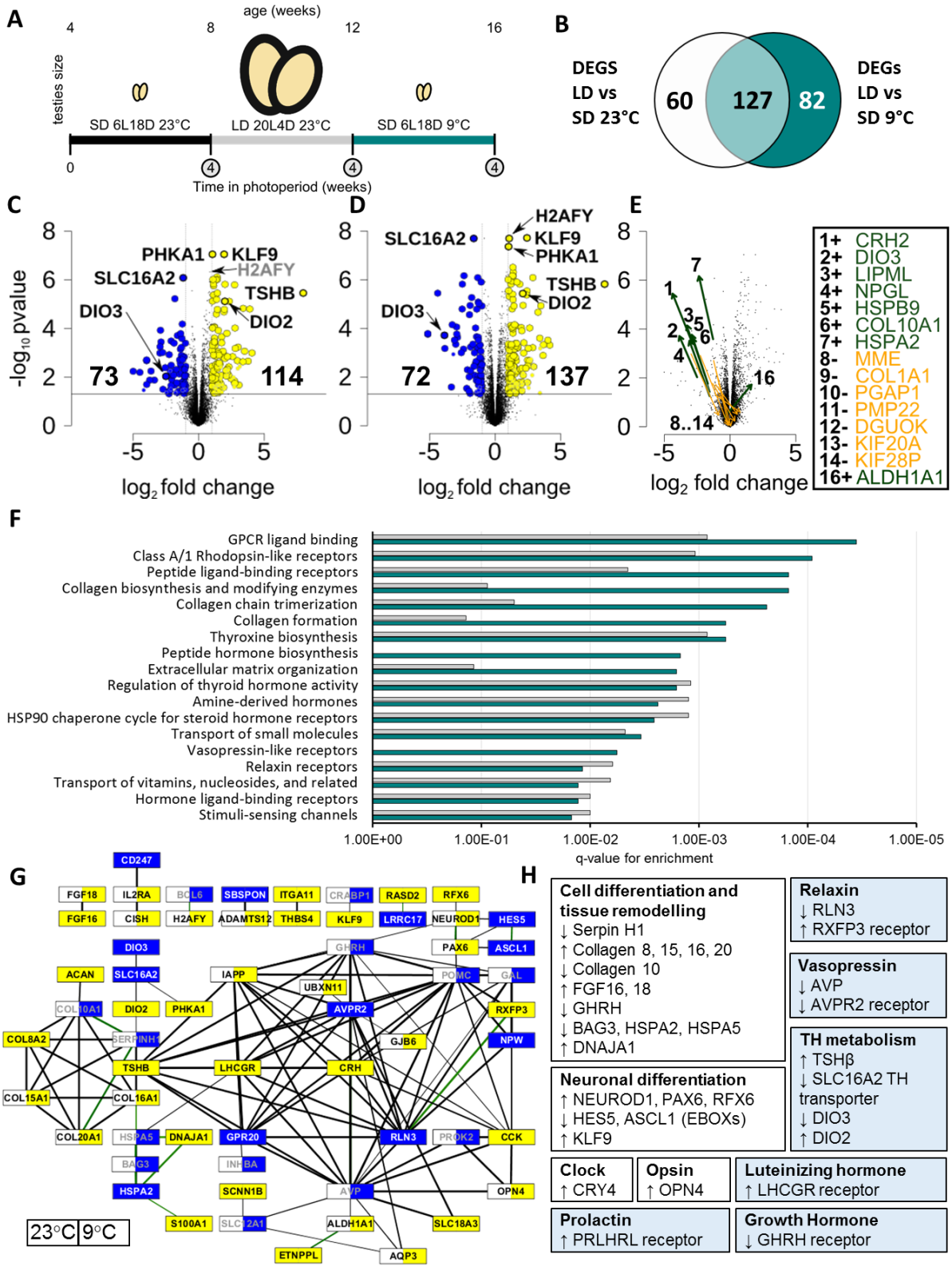


Figure 2: Genome-wide analysis of temperature-dependent transcriptome responses to photoperiod in quail.

Experimental design showing the 3 time-points each sampled after 4 weeks of the target photoperiod (circled) with RNASeq at n=4 **A.**Intersection of DEGs between LD 23°C vs SD 23°C and LD 23°C vs SD 9°C **B**. Volcano plots comparing LD 23°C vs SD 23°C showing 71 up (yellow) and 42 down (blue) DEGs **C** and LD 23°C vs SD 23°C **D**. Grey labels do not pass fold change threshold at 23°C. Temperature-dependent effects on fold change in DEGs when comparing SD at 23°C and SD 9°C. Arrows point from 23°C to 9°C and indicate a significant amplifying (green) or dampening (orange) effect of 9°C on

photoperiod response **E** significantly enriched pathways in DEG genes at LD vs SD 23°C (grey) and LD vs SD 9°C (teal) q-value thresholds **F**. Network of up (yellow), down (blue), and no significant change (white) regulated inter-connected genes (LD vs SD) using the String database. The left side of a node indicates the expression change at 23°C and right at 9°C. Edges are weighted by the combined score, and green edges represent experimental support **G**.

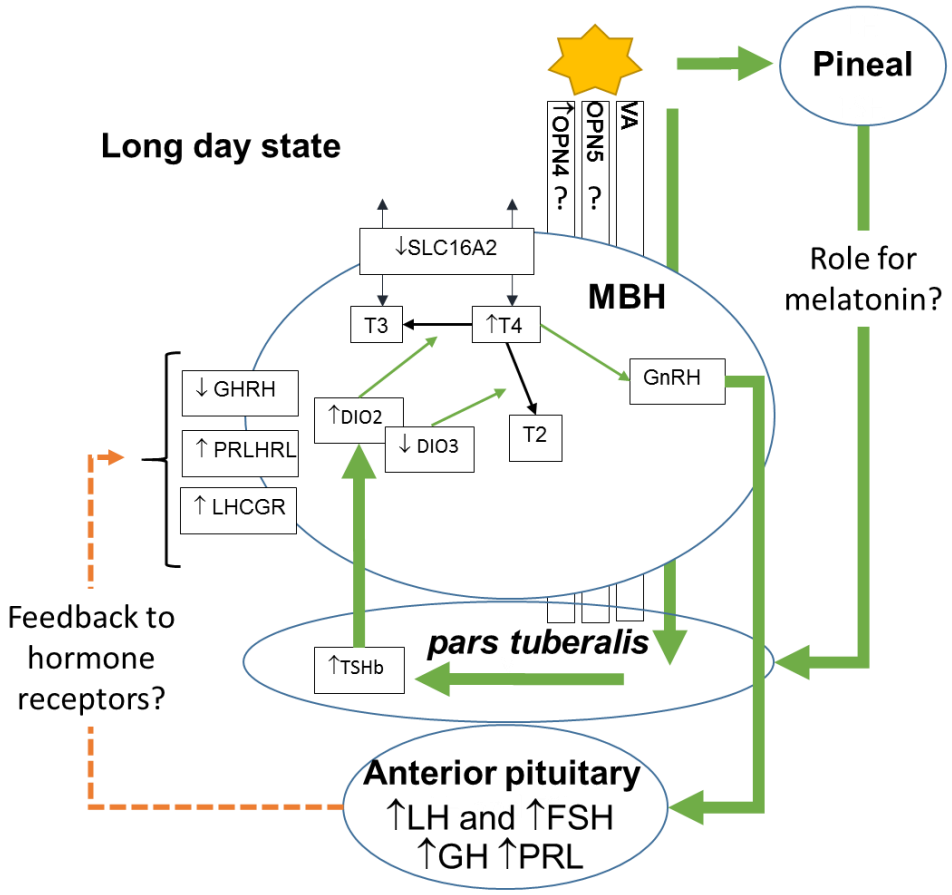


Figure 3: Photoperiod signalling in the MBH incorporating observations from RNASeq

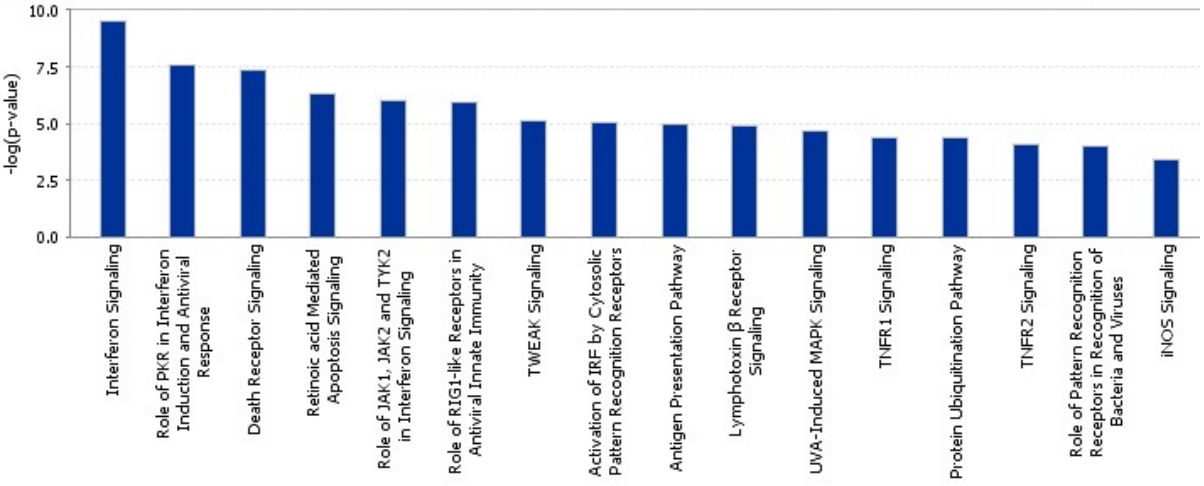


Figure 4: Enriched pathways in a cluster of genes highly expressed in duck lung after HPAI infection

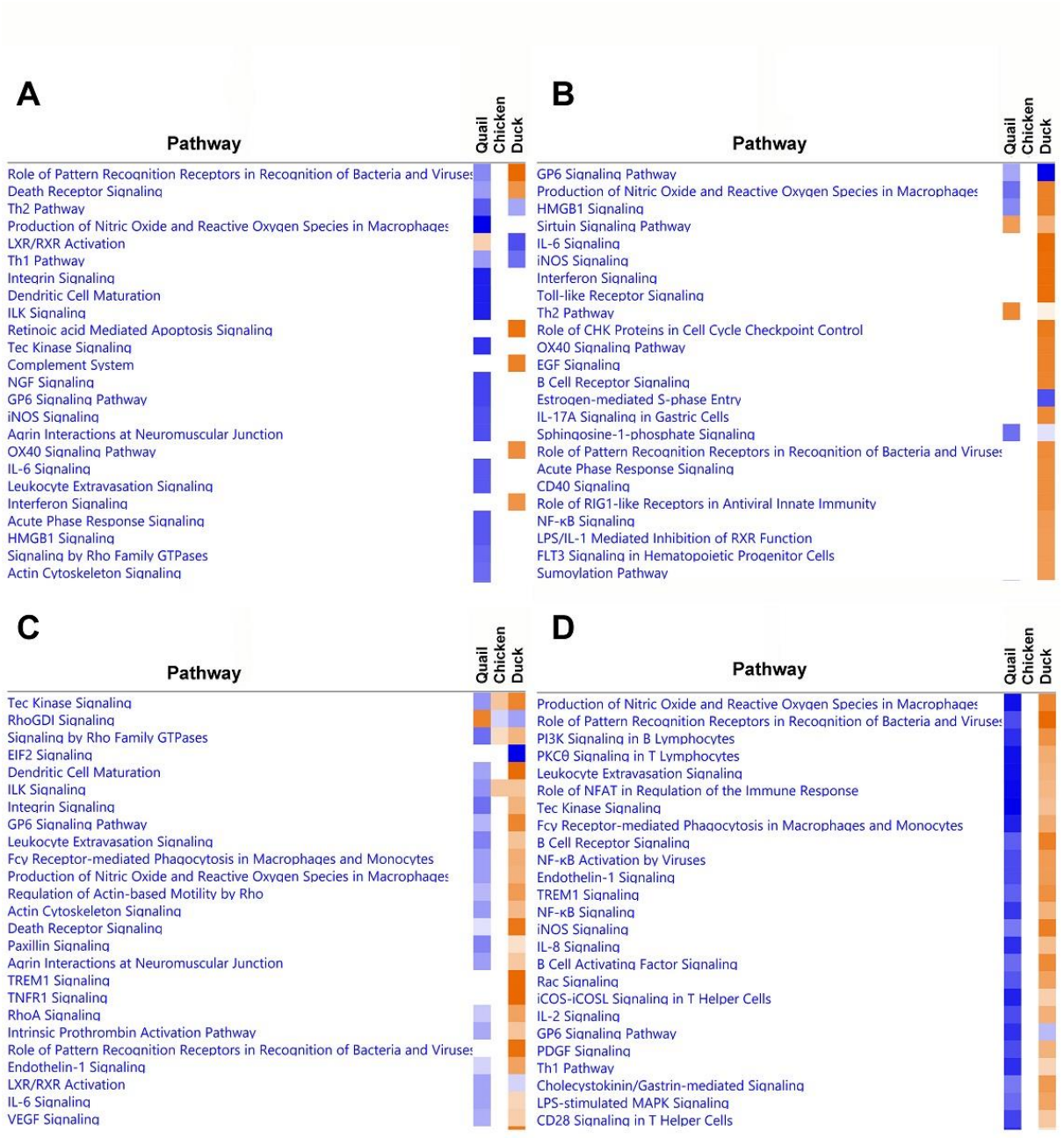


Figure 5: Heatmap comparison between pathways upregulated (orange) and downregulated (orange) in quail, chicken and duck, in ileum day 1 (A), ileum day 3 (B), lung day 1 (C) and lung day 3 (D).

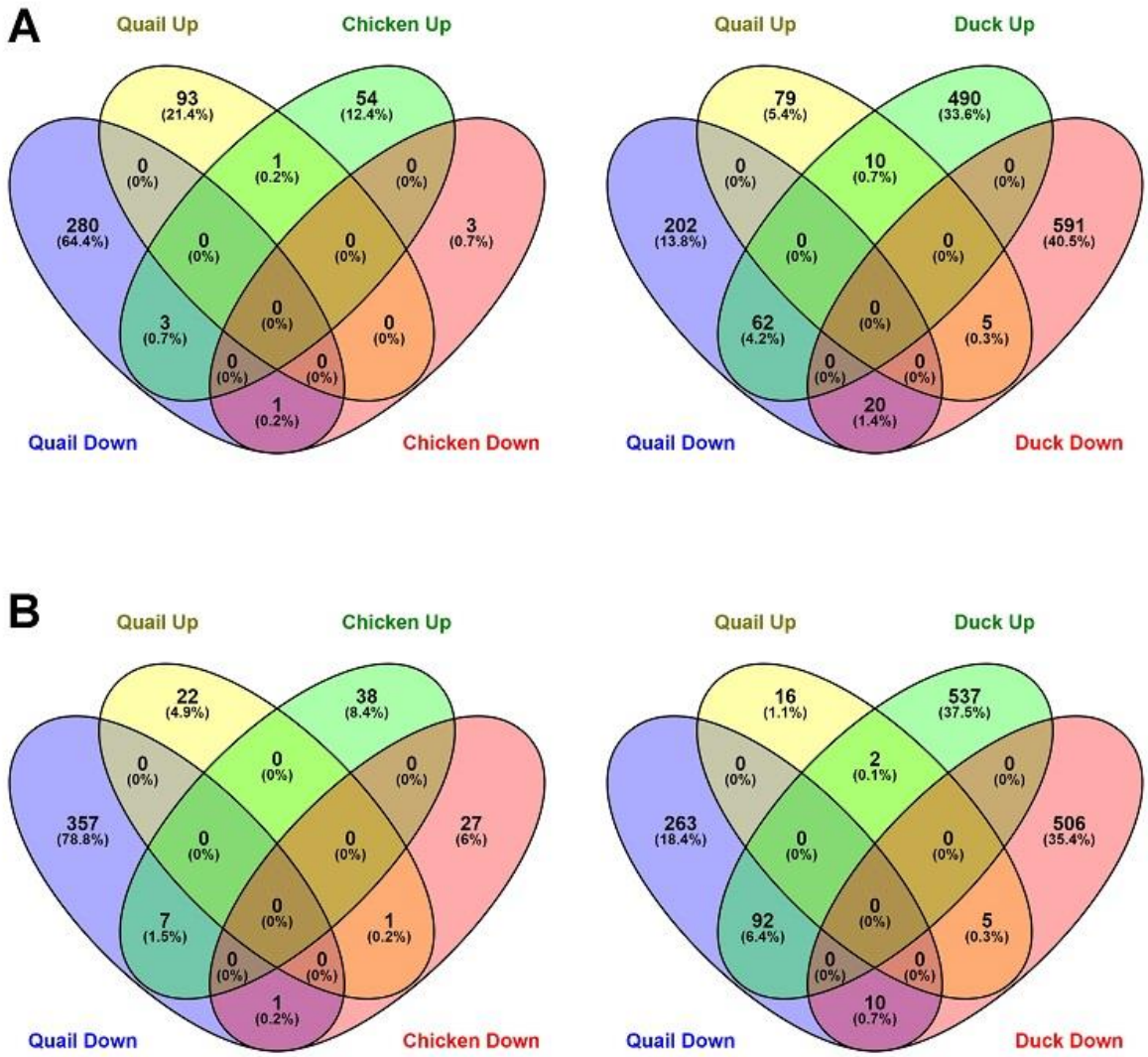


Figure 6: Proportion of genes commonly regulated between quail and chicken or duck to H5N1 infection on day 3 in ileum (A) and lung (B)

Supplementary Data Legends:

Supplementary Data 1
List of unannotated quail genes and their manual annotation
Supplementary Data 2
Location of breakpoints between chicken and quail chromosomes
Supplementary Data 3
BED file containing location of ERVs in quail genome
Supplementary Data 4
BED file containing location of ERVs in turkey genome
Supplementary Data 5
Sequences of cathelicidins and defensin genes identified in quail genome
Supplementary Data 6
List of selection signatures detected from the HSR / LSR lines
Supplementary Data 7
Statistics from photoperiod differential expression study
Supplementary Data 8
Pathway analysis from photoperiod study
Supplementary Data 9
List of differentially expressed genes, FDR < 0.05 and fold change > 1.6 in infection study
Supplementary Data 10
Overrepresented GO terms in differentially expressed genes in infection study
Supplementary Data 11
Significant upregulation and downregulation of IFITM, MHC and TLR genes

IV. Article 4

Molecular control of the endothelial-to-hematopoietic transition

This is the version C of the manuscript. We are still working on it.

Molecular control of the endothelial-to-hematopoietic transition

Khoury, H[1], Gautier R[1], Charbord, P[1], and Jaffredo T[1*]

[1] *Sorbonne Université, UPMC Univ Paris 06, IBPS, CNRS UMR7622, Inserm U 1156, Laboratoire de Biologie du Développement; 75005 Paris*

**Correspondence to : thierry.jaffredo@upmc.fr*

Introduction

Hematopoiesis is an essential biological process leading to the daily formation of blood cells able to maintain the homeostasis of the body. This production relies on rare cells designated as hematopoietic stem cells (HSCs) able to produce both all the hematopoietic lineages and to maintain their number roughly constant through self-renewal mechanisms. In all the vertebrate species described, HSCs are formed early during development at the level of the embryonic dorsal aorta and associated arteries through an endothelial-to-hematopoietic commitment (EHT) proposed more than a century ago (Jordan, 1916). In addition to the aorta, EHT also occurs earlier on in development at the level of the yolk sac where it was shown to produce a wave of Erythro-Myeloid Progenitors (EMPs) as well as lympho-myeloid progenitors. This unique type of cell emergence is characterized by the progressive loss of the endothelial transcriptional program, the switching on of the hematopoietic program accompanied by a morphological change that conduct to the rounding of the flat endothelial cells, the breaking of the tight junctions and the release into the vessel lumen. Thus, EHT appears as a unique method to produce most of the hematopoietic cells including HSCs. The use of mouse or human pluripotent stem cells has allowed to first demonstrate EHT (Lancrin et al., 2010; Schröder et al., 2010) and to further permit significant advances in the cellular and molecular control of EHT. Recent advances in live imaging using the zebrafish or murine embryos have allowed EHT to be observed in real time and hematopoietic cells to be produced unambiguously demonstrating EHT within the aorta. Due to the use of pluripotent stem cells and on transcriptomic approaches, progress have been made in the past few years on the molecular control of EHT. In addition to the RUNX1 transcription factor (TF), a master regulator of EHT (Okuda et al., 1996; Wang et al., 1996; Chen et al., 2009), the GATA2 TF was also shown to be a master regulator of EHT (Ottersbach 2019) in addition with other TFs called heptad (Wilson et al., Cell Stem Cell 2010). In addition, two downstream targets of Runx1, Gfi1, and Gfi1B were shown to be essential for hematopoietic production from the hemogenic endothelium (Tambyrajah et al., 2016; Lancrin et al., 2012). A recent study also indicates that Bmi1, a member of the polycomb complex contributes to Runx1 silencing to prevent premature EHT (Eliades et al., 2016). In addition to these TFs and their modulators, several signaling pathways were shown to control EHT either by acting directly on endothelial cells such as Notch or Wnt (Bigas and Espinoza, 2012) or by acting on the immediate microenvironment such as BMP or Hedgehog (Dzierzak and Bigas, 2018).

However, few is known about how these TFs and signaling pathways orchestrate the EHT and few essential regulators have been isolated. A recent study pointed out on the role of the blood flow and shear stress to modulate EHT and the existence of pulsatile cycles involving members of the cytoskeleton to trigger EHT (Lancino et al., 2018).

Recently we developed an *ex vivo* culture system that allows stepwise commitment of the mesoderm into hemogenic endothelial cells that undergo EHT to give rise to multilineage hematopoiesis (Yvernogeau et al., 2016). Here we have utilized this system to purify cell types belonging to key steps on the road to hematopoietic production and to isolate specific mRNA signatures using transcriptome analysis. Through comprehensive meta-analysis and isolation of specific gene expression signatures, 798 genes arranged in three networks representing the hemogenic endothelium state. These networks do not contain the standard hemogenic endothelium regulators since present in upstream or downstream cell types but reveal unique gene sets with previously unrecognized regulators of hemogenic endothelium. Two gene sets are positively regulated in hemogenic endothelium and one is negatively regulated. Within gene sets, genes are interconnected and contain members that are highly interconnected defining hubs as well as several genes belonging to specific signaling pathways. Experimental knock-down of one signaling pathway and two positively regulated hubs strongly modifies the culture and modulates EHT giving networks a strong predictive value. In addition to hubs, genes connected to the hubs also appeared deregulated following hub knock-down strengthening the existence and the robustness of the networks. These gene sets uncovered new, unique regulators of the hemogenic endothelium state and encompasses genes with strong predictive value.

Results

LDL^{int} cells represent an EC population upregulating RUNX1 and down-regulating CD31 and CD144.

One advantage of the culture system (Yvernogeau et al., 2016) is the stepwise progression of the mesoderm towards the hematopoietic fate that allows isolating discrete cellular steps over the culture period of 12 days. In addition, by removing serum or VEGF from the culture, it is possible to prevent hemogenic EC commitment and to orient the cells towards a more advanced vascular endothelium state. We previously showed that ECs emerging in culture uptake AcLDL in keeping with their endothelial phenotype including those expressing RUNX1. We reasoned to use the levels of AcLDL uptake to separate ECs from hemogenic ECs. This is based on the observation that RUNX1 initiated expression is found at D3 in only a few ECs, but in almost 100% of the cells at D5. RUNX1 expression is shortly followed by that of the earliest hematopoietic-specific transcription factors, in particular, PU1, a direct target of RUNX1. In this respect, hemogenic ECs i.e., RUNX1-expressing cells, should undergo a decrease of AcLDL uptake in keeping with the progressive loss of the endothelial program. Based on this assumption, D2 and D4 adherent cells were incubated with AcLDL over a period of two hours, trypsinized and Fluorescence-Activated Cell Sorted (FACS) based on the AcLDL level of expression. At D2, FACS analysis showed cells harboring a wide, albeit continuous, range of fluorescence from nil to high (Figure A1A). At D4, we separated two populations: one with a high level of expression of AcLDL (LDL^{hi}) and another with an intermediate level of expression (LDL^{int}) (Figure A1B). In the absence of serum, the majority of the population is within the LDL^{hi} (Figure A1C).

The different LDL populations displayed on Figures 1A-C were sorted and analyzed for the expression of two endothelial markers i.e., CD31 (PECAM) and CD144 (VE-Cadherin) and the canonical hemogenic endothelial marker RUNX1. Gene expression on the D0 pre-somitic mesoderm served as a baseline. Both CD31 mRNA was highly expressed in LDL^{hi} vs LDL^{int} cells with a mean fold change of two (Figure A1D) CD144 mRNA expression was high at D2 and D4 with no clear change on the level of expression between D2 and D4 and between D4 and D4 without SVF but was at least 8-fold decreased in D4 LDL^{int} cells (Figure A1E). However, RUNX1 mRNA expression was strongly increased in LDL^{int} compared with the other LDL-positive populations (Figure A1F).

Identification of the gene network characteristic of hemogenic ECs

Sample isolation strategy

Based on this result, we isolated the different cell fractions to perform NGS (Figure S1A). We used the level of LDL uptake to isolate ECs ($LDL^{high\ or\ +}$) from hemogenic ECs (LDL^{int}). Presomitic mesoderm was directly isolated from the embryo whereas hematopoietic cells were collected as the floating fraction, in keeping with Yvernogeau et al. (2016).

Bioinformatic analysis

We used high-throughput sequencing to compare the molecular signatures of the different cell fractions. We isolated 20 samples corresponding to the different cell fractions (Figure A2A; Figure S1A). mRNAs were isolated and the libraries were sequenced using next-generation sequencing Illumina technology. According to the quality control, we detected 20 332 genes. PCA on the entire set of mRNAs and samples (20 samples; 20 332 genes) indicated differences between the different cell fractions. PC1 (30.2% of the variance) and PC3 (9.85% of the variance) allowed to discriminate ECs from pre-somitic mesoderm (M) and hemogenic ECs (HEC) from hematopoietic cells (HC), respectively (Figure A2B). A first hierarchical clustering on the entire set of mRNAs and samples showed some discrepancies in the ordering with the M fraction found between the EHC and HC fractions (Figure S1C). To identify the genes that were specifically up- or down-regulated in the different cell populations we followed a two-step procedure. First, we compared pairwise using ANOVA (with p-values <0.06 and fold changes ≥1.45 or ≤−1.45) the M, EC, HC and HEC cell populations. Second, we considered using Venn diagrams the gene sets that were up-regulated (or down-regulated) in a given population according to the 4 comparisons (Figure A2C; Figure S1D). This procedure led to the identification of 2462 up-regulated and 2087 down-regulated genes (upregulated DEGs in the 4 populations Table S1; Table S2). We then performed a novel PCA that provided a clearcut discrimination between the 4 populations, according to PC1 (43% of the variance) corresponding to the contrast between M and EC and to PC3 (17% of the variance) corresponding to that contrast between HC and HEC (Figure A2D). And this time hierarchical clustering showed a perfect ordering going from M through EC and HEC fractions until HC population (Figure S1E). Analysis of the GO categories clearly confirmed the phenotypic differences between the 4 distinct populations. The M cell population was enriched in genes implicated in spliceosome, Hippo and hedgehog signaling pathways and shared the category adherent junction with the EC population. The EC population showed

enrichment in lysosome, FoxO, Wnt and Rap1 signaling pathways and glycosaminoglycan degradation. The HEC population was enriched in genes implicated in focal adhesion, ECM, regulation of actin cytoskeleton, MAPK and PI3K-Akt pathways, glycosaminoglycan biosynthesis and axon guidance. Finally, the HC population was enriched in genes implicated in metabolic pathways, oxidative phosphorylation, Jak-STAT and B and T cell receptor signaling pathways (Figure A2E).

WGCNA then performed using the matrix of up-regulated genes revealed the existence of 5 modules, one for EC, one for M, one for HC and two for HEC (Figure A2F). From this point, we decided to focus on the two HEC modules of up-regulated genes. The designated green (Figure A2G) and red (Figure S2A) modules corresponding to 192 and 253 genes respectively significantly up-regulated in HEC had a very particular extended topology consisting in relatively dense cores connected by a few genes; hence there were very prominent hubs: the ECM components EMILIN1, EMILIN2, COL5A1, FBN1 and COL1A2, COL6A2, COL6A3, CSPG4 in the red and green modules, respectively, the ECM enzyme PLOD2, the migration cue SLIT2, the transcription factor CREB3L1 in the red module, and the transcription factor NFKB2 and the cytokines TGFB3 and INHBA in the green module. The gene connectivity of the most connected genes in each module is displayed in (Figure A2H and Figure S2B). Analysis of the GO categories revealed shared and distinctive patterns between the two HSCs networks. Both networks showed an enrichment in molecules implicated in the ECM, with Pleckstrin homology domain, SH3 domain, components of the Golgi apparatus, cell adherent junction, membrane and cell signaling. On the contrary, genes implicated in Focal adhesion were enriched in the green module only, while the red module showed enrichment in genes implicated in microtubule, migration and phosphatase activity (Figure A2I).

Modulation of the Wnt pathway contributes to the endo-hematopoietic balance

To investigate whether ß-catenin activity influences the generation of HEC and the EHT, we tested the effect of inhibitors or activators of the pathway on culture. As an inhibitor, we used the small molecule PKF118-310 (hereafter referred as PKF118), an inhibitor of the TCF4 ß-CATENIN protein interaction (Hallet et al., 2012). We tested 3 different concentrations of PKF118 (0.66µM, 1µM and 1.5µM). The last concentration provoked toxic effects whereas the first one did not produce significant results (not shown). The concentration of 1µM caused drastic changes to the culture. Treated cells

were flatter and had fewer round cells than the non-treated samples (Figure A3A; Figure A3B; Figure A3D; Figure A3E). Flow cytometry analysis confirmed the increase of LDL^{hi} cells at the expense of LDL^{int} cells (Figure A3C; Figure A3F; Figure A3G; Figure A3H) and the strong decrease of the floating cells (Figure A3I) (n=5). QRT-PCR on adherent cells revealed the reinforcement of CD144 mRNA expression (Figure A3J) in keeping with the increase of the endothelial phenotype and the decrease of RUNX1 expression (Figure A3K) consistent with the lower number of LDL^{int}cells.
We also tested the Wnt activator SB216763 hereafter referred as SB, that inhibits GSK3 alpha and GSK3 beta. We used two different concentrations i.e., 10µM and 20µM with 20µM displaying the strongest effect. The result is an increase in the number of round hematopoietic cells produced compared with the mock- (DMSO) treated culture (Compare Figure S3A to Figure S3D and Figure S3B to Figure S3E). Flow cytometry analysis reveals a 1.6-fold change in the number of LDL^{int} cells at the expense of LDL^{hi} cells (Figure S3C; Figure S3F and Figure S3G; Figure S3H). Taken together the Wnt pathway plays a role in EHT. These results are completely in keeping with results published by Ruiz-Herguido et al., 2012 on the mouse AGM.

The Gamma secretase inhibitor DAPT stimulates the EHT

We used gamma secretase as an inhibitor of Notch to inquire into the role of this signaling pathway in hemogenic endothelium and EHT. DAPT 25µM was applied to the culture at D2. Analysis was performed at D4. DAPT readily produced a dramatic change in culture. Within the two days of DAPT contact, many cells rounded up, detached from the adherent layer and were found in the supernatant (Compare Figure S4A; Figure S4B). Cells in the supernatant were round and very bright suggestive of a hematopoietic production (Figure S4B) and Yvernogeau et al., 2016). FACS analysis showed a sharp increase in the number of LDL^{int} cells at the expense of LDL^{hi} cells (Figure S4C; Figure S4D and Figure S4G; Figure S4H) accompanied by a parallel increase of hematopoietic cells (Figure S4E; Figure S4F and Figure S4I). QRT-PCR analysis showed a highly significant increase of CD45 expression characteristic of hematopoietic differentiation (Figure S4J).

POFUT2 loss of function enhances the EHT

Based on the list of genes positively correlated to the HEC state obtained following WGCNA analysis and the measure of gene connectivity, we selected POFUT2 for

potentially playing a strong role in HEC. POFUT2 is upregulated in HEC compared to EC or HC. POFUT2 (Protein O-Fucosyltransferase 2) catalyzes the reaction that attaches fucose through an O-glycosidic linkage to a conserved serine or threonine residue in the consensus sequence C1-X(2,3)-S/T-C2-X(2)-G of thrombospondin type 1 repeats where C1 and C2 are the first and second cysteines, respectively. It fucosylates members of the ADAMTS family, the Trombosporin and Spondin families. This fucosylation is required to regulate EMT during formation of the mesoderm and definitive endoderm (ref). We used a knock-down strategy utilizing siRNA. We designed three different siRNAs against POFUT2 (Table A1) that were transfected at day 1 of culture incipience. The analysis was performed at day 4 (Figure A4A) and was compared to Lipofectamine alone or a scrambled POFUT2 siRNA. We tested 3 concentrations of siRNA against POFUT2 i.e. 30, 60 and 120nM (n=5) and compared their knock-down efficiency by measuring POFUT2 mRNA expression by q-PCR on transfected cells. The lowest concentration decreased POFUT2 mRNA expression by 80% whereas the highest displayed signs of toxicity and was not retained (not shown). The 60nM concentration decreased POFUT2 by 90% on transfected cells and was selected for further use (Figure A4B). POFUT2 mRNA knock-down at D1 resulted in the appearance of more round hematopoietic cells in the culture compared with Lipofectamine or scramble si controls (Figure A4C; Figure A4E; Figure A4G). Flow cytometry analysis revealed an increase in the $AcLDL^{int}$ population at the expense of the $AcLDL^{hi}$ population (Figure 4D, F, H). Taken together POFUT2 knock-down resulted in a decrease of the number of ECs ($p<0.001$) and an increase in the number of HEC ($p<0.08$) (Figure A4I; Figure A4J).

To further validate the gene network and its predictive value, we selected the POFUT2's first neighbors (Figure S2C) and examined if a decrease of POFUT2 mRNA expression was also associated with a modulation of several of its first neighbors. Based on gene connectivity (Figure S2C), we selected 3 genes among the most connected to POFUT2. COL5A2 encodes an alpha chain for one of the low abundance fibrillar collagens; DNAJB4, is a molecular chaperone, tumor suppressor, and member of the heat shock protein-40 family that binds and regulates E-Cadherin and EMILIN1A (Elastin Microfibril Interfacer 1) is thought to regulate smooth muscle cell adhesion and vessel assembly. COL5A2, the most connected gene to POFUT2 showed a significant decrease upon POFUT2 siRNA mediated knock-down when compared to the Lipofectamine control but not to the scramble ($p<0.001$and 0.3 respectively) (Figure A4K). DNAJB4 and EMILIN1A both showed a significant decrease upon POFUT2 siRNA knock-down whatever the control used (Figure A4L; Figure A4M). Taken together, this indicates that

POFUT2 and, at least, 3 of their first neighbors are acting coordinately to regulate the EHC fate and the hematopoietic production.

Material and Methods

Presomitic mesoderm isolation

We used quail (*Coturnix coturnix japonica*) embryo pre-somitic mesoderm according to Yvernogeau et al, (2016). Eggs were incubated for 36-45 hours at 37±1°C in a humidified atmosphere to reach 10-18 somite pairs. Microsurgery was performed as previously described Pardanaud, 1996 #2110]. The PSM was removed over a length corresponding to 10 somites from both sides of the embryo. Five embryos (i.e. 10 PSM) were used per culture dish. Each PSM was cut into 4 to 5 equal pieces, rinsed in Opti-MEM plus GlutaMAX™ I (Gibco Life Technologies) +5% Fetal Calf Serum (FCS, Gibco Life Technologies) +100 Units per ml penicillin/streptomycin (Gibco Life Technologies) + 1% Chicken Serum (Gibco Life Technologies) before culture.

Medium, culture dishes

PSM were cultured under the following conditions Opti-MEM® with GlutaMAX™ I supplemented with 5% FCS, 1% Chicken Serum, 100 Units per ml penicillin/streptomycin and the following growth factors (PromoCell/PromoKine): hVEGF (≠C64410); 2ng/ml), hFGF (≠C60240) ; 4ng/ml), hIGF (≠C60840) ; 3ng/ml), hEGF (≠C60170) 10ng/ml), hydrocortisone (200ng/ml SIGMA ≠H6909), ascorbic acid (75ug/ml SIGMA ≠A4544). PSM were cultured in 35mm collagen I- or IV -coated dishes from Corning BioCoat (Corning - Dutscher France). Medium was changed every two days unless specified.

RNA extraction, qRT-PCR, primers

All RNA extractions were performed using the RNeasy™ kit from Qiagen SAS France. Freshly isolated PSM (10) were re-suspended into the RNeasy™ buffer solution (RLT). For RNA extraction on cultured cells, the supernatant cells were first centrifuged to remove the medium and re-suspended into the RNeasy™ buffer solution. Adherent cells were first trypsinized, washed in PBS/FCS (10%), centrifuged and then re-suspended into the RNeasy™ buffer solution. Quality/quantity of the extracted RNA was evaluated using a Nanodrop.

Staining of in vitro culture

We used Low Density Lipoprotein from Human Plasma, Acetylated, coupled to Alexa-488 (AcLDL-A488 1mg/ml Life Technologies, Ref. L23380, batches N° 1291485

and 1696210) to determine the endothelial phenotype of PSM-derived cells. AcLDL-A488 (1/100 in PBS) was incubated for 2 hours at 37°C in the culture dish. Cultures were then washed 3 times in PBS before flow cytometry and cell sorting.

FACS analysis

Cells were stained with AcLDL coupled to Alexa-488. When adherent cells were used, cells were tripsinized and rinsed thrice in PBS. When floating cells were used, supernatant cells were gently removed from the culture dish by pipetting. For LDL analysis, cells were cells were suspended in PBS containing 7AAD to exclude dead cells and analyzed with a MacsQuant analyzer 10. Analyses were made with FlowJo™ 10. Statistics were performed with GraphPad Prism.

Lipofection-mediated siRNA transfer

Lipofectamine RNAiMax was used to transfect cells at day 1 or 2 with specific siRNA or Scramble si control following the manufacturer's protocols. The transfection efficiency was evaluated by analysing the mRNA expression of the targeted gene by qPCR.

RNA sequencing and bioinformatic analyses

RNA libraries were prepared by Fasteris Life Science (Switzerland). In brief, total RNA was submitted to a poly-A mRNA purification using oligodT magnetic beads. Supernatants were kept for small RNA library preparation, and poly-A–RNAs after elution from the beads were prepared according to the manufacturer's protocol (TruSeq RNA Sample Prep- Kit V2; Illumina). Transcripts were broken at 95°C in presence of zinc, and first-strand cDNA syntheses were performed using random primers. A second-strand cDNA synthesis was performed in the presence of deoxyuridine triphosphate, and after a 3′ A addition step, adapters were ligated, and an amplification by PCR was per- formed to generate the DNA colony template libraries. Small RNA libraries were performed according to the manufacturer's protocol (TruSeq Small RNA Library Prep kit; Illumina). After acrylamide gel purification of small RNA between 18–30 nt, single-stranded ligation of 3′ adapter and the 5′ adapter were performed before reverse transcription and PCR amplification to generate the DNA colonies template. All the samples were sequenced using 1 × 125–bp single reads high-throughput sequencing (RNA-Seq) in single lane on a HiSeq 2000 sequencing system (Illumina). RNA-Seq analysis was performed on Partek Flow™. For mRNA libraries, sequence reads in fastq format were aligned to the quail reference genome assembly RefSeq GCF_001577835.1 (Figure S1B workflow) using

the STAR aligner. The number of reads for all the features were then counted using the quantification to a model E/M and normalized for each library, and then ANOVA were performed to compare read values between different libraries.

Statistical analysis

Statistical analyses (*t* tests) were performed with GraphPad and Excel software. Data were considered statistically significant for $P < 0.05$.

Data analysis

The lists of differentially expressed genes were generated by ANOVA and Venn Diagram . GO was performed using DAVID (http://david.abcc.ncifcrf.gov). To find out the network structure of the gene set HEC, we used WGCNA (Langfelder and Horvath, 2008). Adjacencies were given according to signed Pearson correlation. The soft threshold power b that resulted in approximate scale-free topology was 36. Modules were constructed with average linkage hierarchical clustering. Minimum module size was 30.

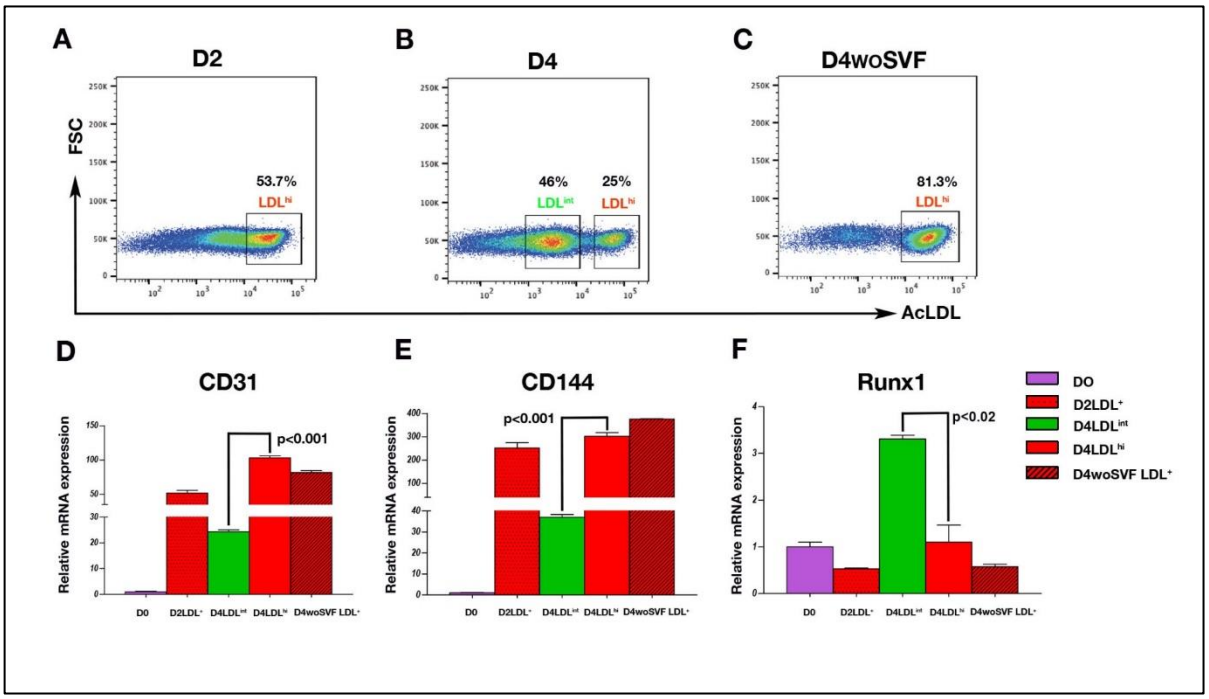


Figure A1: The level of AcLDL uptake separates endothelium from hemogenic endothelium

A-C: Flow cytometry analysis showing the expression of AcLDL uptake in the culture. A: analysis at D2. The cells display a continuum of AcLDL uptake from nil to high. Frame shows the sorted population. B: analysis at D4 in normal culture conditions. Two distinct populations with contrasted AcLDL uptake are seen. The left frame isolates the LDL^{int} population and the right frame, the LDL^{hi} population. C: analysis at D4 in the absence of fetal calf serum known to orientate the cells towards the vascular endothelium lineage (Yvernogeau et al., 2016). A large percentage of the population is contained within the LDL^{hi} population in keeping with their strong endothelial commitment. D-F: quantitative PCR on the different cell fractions for CD31 (D), CD144 (E) and RUNX1 (F). CD31 and CD144 are decreased in the LDL^{int} population compared with the other endothelial populations. RUNX1 is strongly increased in the LDL^{int} population compared with the other endothelial fractions (n=2). Significance is given compared with the D4 LDL^{hi} fraction. Data are shown as means ± SEM.

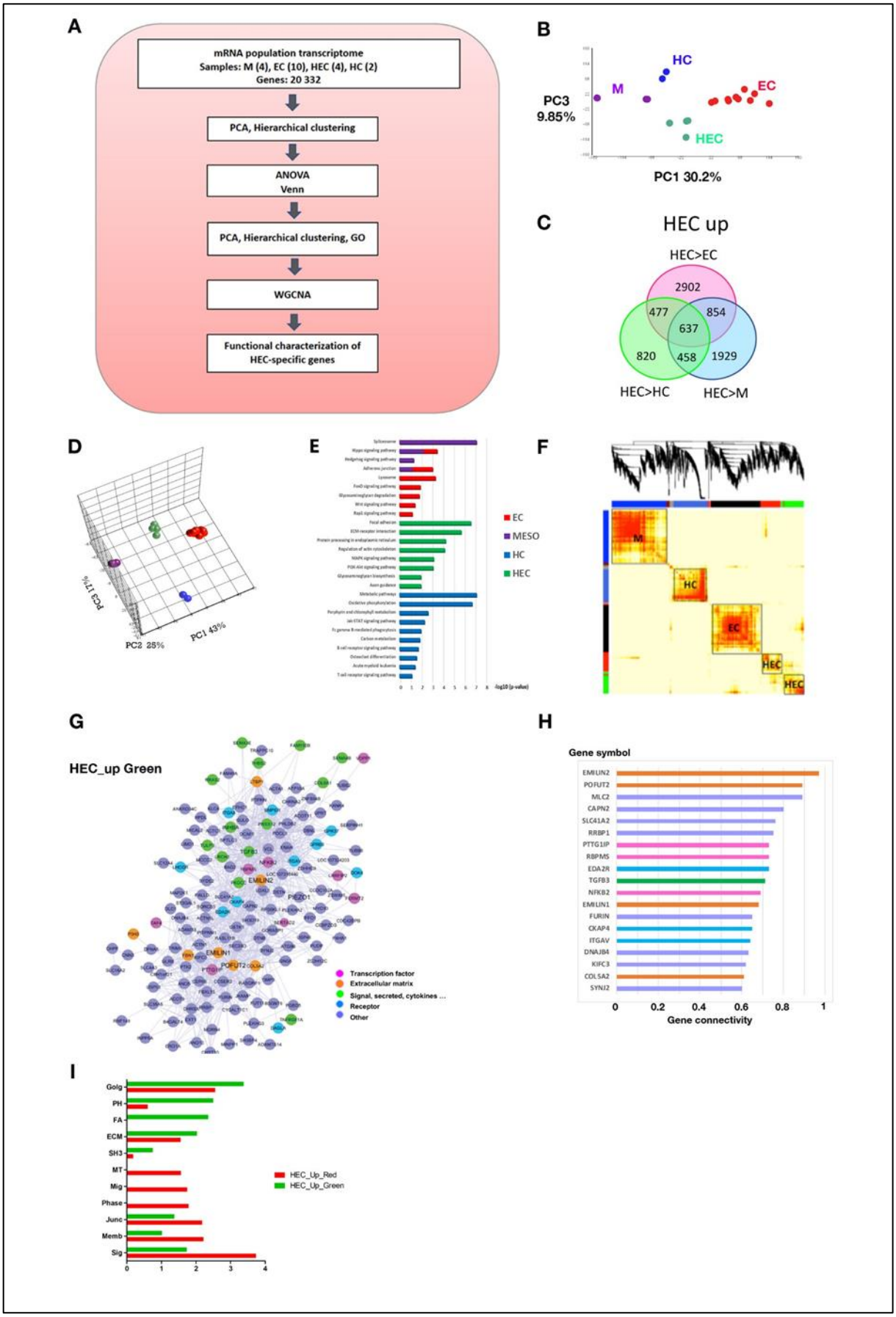

A
mRNA population transcriptome
Samples: M (4), EC (10), HEC (4), HC (2)
Genes: 20 332
PCA, Hierarchical clustering
ANOVA
Venn
PCA, Hierarchical clustering, GO
WGCNA
Functional characterization of
HEC-specific genes
B
HC
M
EC
HEC
PC3
9.85%
PC1 30.2%
C
HEC up
HEC>EC
2902
477
854
637
820
458
1929
HEC>HC
HEC>M
D
PC3 17%
PC2 25%
PC1 43%
E
EC
MESO
HC
HEC
F
M
HC
EC
HEC
HEC
G
HEC_up Green
EMILIN2
PIEZO1
EMILIN1
POFUT2
Transcription factor
Extracellular matrix
Signal, secreted, cytokines ...
Receptor
Other
H
Gene symbol
EMILIN2
POFUT2
MLC2
CAPN2
SLC41A2
RRBP1
PTTG1IP
RBPMS
EDA2R
TGFB3
NFKB2
EMILIN1
FURIN
OKAP4
ITGAV
DNAJB4
KIFC3
COL5A2
SYNJ2
0
0.2
0.4
0.6
0.8
1
Gene connectivity
I
Golg
PH
FA
ECM
SH3
MT
Mig
Phase
Junc
Memb
Sig
0
1
2
3
4
HEC_Up_Red
HEC_Up_Green

Figure A2: General outline of the study and transcriptome analysis

A: flow chart outlining the analyses on the different cell fraction transcriptomes. B: PCA on the 20322 genes expressed by the 4 cell fractions. Mesoderm (M) are in purple. Endothelial cells (ECs) are in red. Hemogenic ECs (HEC) are in green and Hematopoietic Cells (HC) are in blue. PC1 (30.2% of the variance) separates M from ECs whereas PC3 (9.85% of the variance) separates HEC from HC. C: Venn diagram of genes differentially upregulated in HEC vs. the other cell types. D: Major GO categories given by DAVID using the genes specifically up-regulated in the 4 cell populations. E: PCA using the genes specifically up-regulated in the 4 cell populations. PC1 (43% of the variance discriminates M from ECs whereas PC3 (17% of the variance) discriminates HEC from EC. F: Dendrogram and correlation map of the genes up-regulated revealing the cell-specific modules. G: Organization of the HEC-specific Green module of genes upregulated. H: Connectivity plot of the 19 most connected genes in the Green module. I: Major GO categories given by DAVID for genes belonging to the two HEC-specific modules i.e., Green and Red. Golg, golgi; PH, pleckstrin homology domain; FA, focal adhesion; ECM, extracellular matrix; SH3, SH3 domain; MT, microtubule; Mig, migration; Phase, phosphatase; Junc, cell adherens junctions; Memb, membrane; Sig, cell signaling.

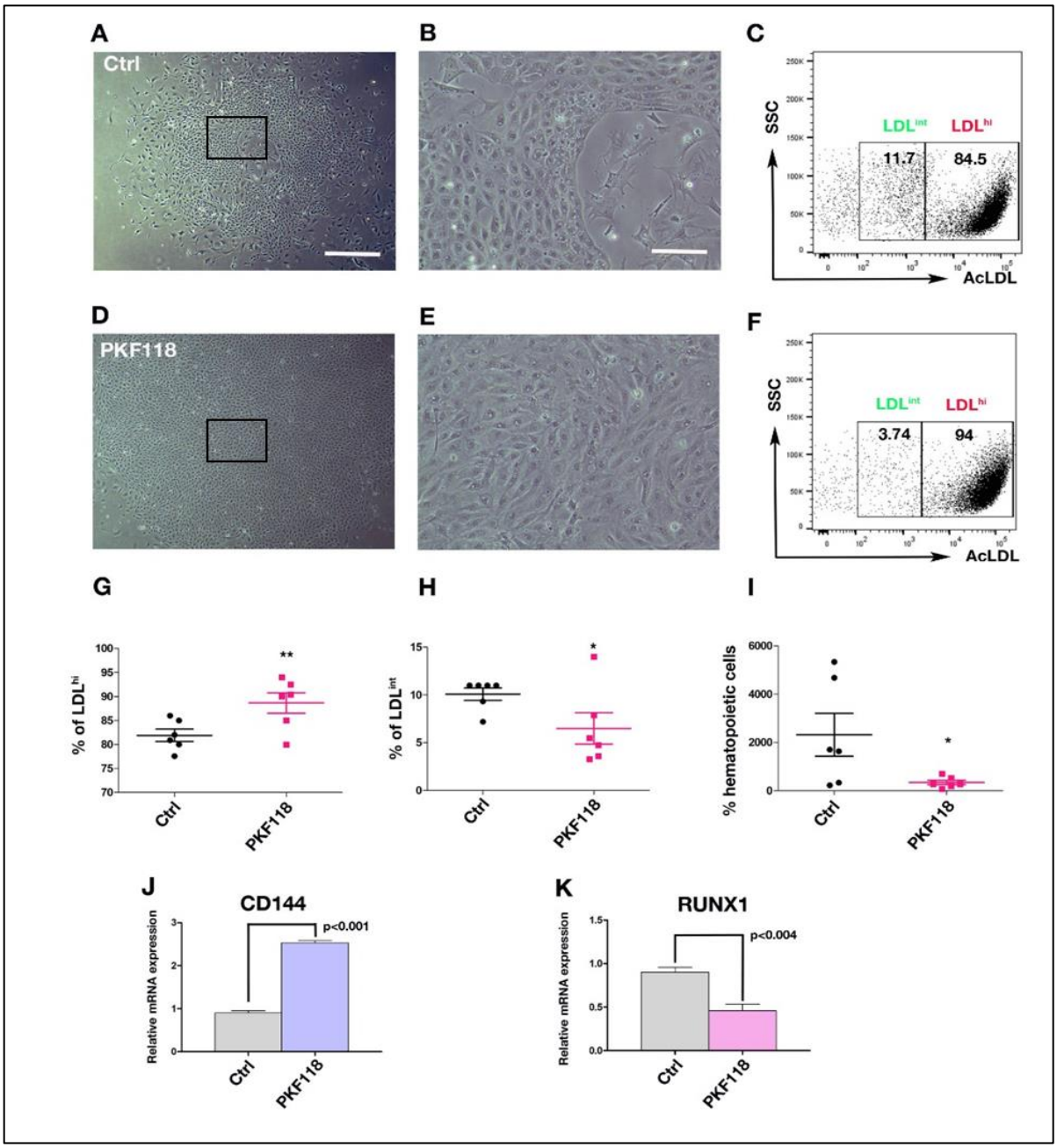


Figure A3: The Wnt antagonist, PKF118, inhibits EHT by reinforcing the endothelial phenotype

A: PSM culture mock treated at D1 and observed at D4. B: enlargement of the frame in A showing the flat endothelial layer and the few round cells resulting from the initiation of EHT at this stage. C: Representative FACS analysis showing the percentages of LDL^{hi} et LDL^{int} in the mock- (water) treated culture. D: PSM culture treated with PKF118 1µM at D1 and observed at D4. Note the flattening of the cell layer treated with PKF118 compared with the mock-treated culture. E: Enlargement of the frame in D showing the flattening of the cells and the presence of fewer round cells. F: representative FACS analysis showing the percentages of LDL^{hi} et LDL^{int} in the PKF118-treated culture (G). G: Percentage of the LDL^{hi} cells in mock- vs PKF118-treated cells (n=5, p<0.02). H: Percentage of the LDL^{int} cells in mock- vs PKF118-treated cells (n=5, p<0.1). I: Percentage of floating cells in mock- vs PKF118-treated cells (n=5, p<0.1). J: qRT-PCR for CD144 on mock- and PKF118-treated cells. A significant increase of CD144 mRNA expression is detected in the PKF118-treated cultures. K: qRT-PCR for RUNX1 on mock- and PKF118-treated cells. A significant decrease of RUNX1 mRNA expression is detected in the PKF118-treated cultures. Means are ± SEM. A, D, bar= 400µm; B, E, bar= 20µm.

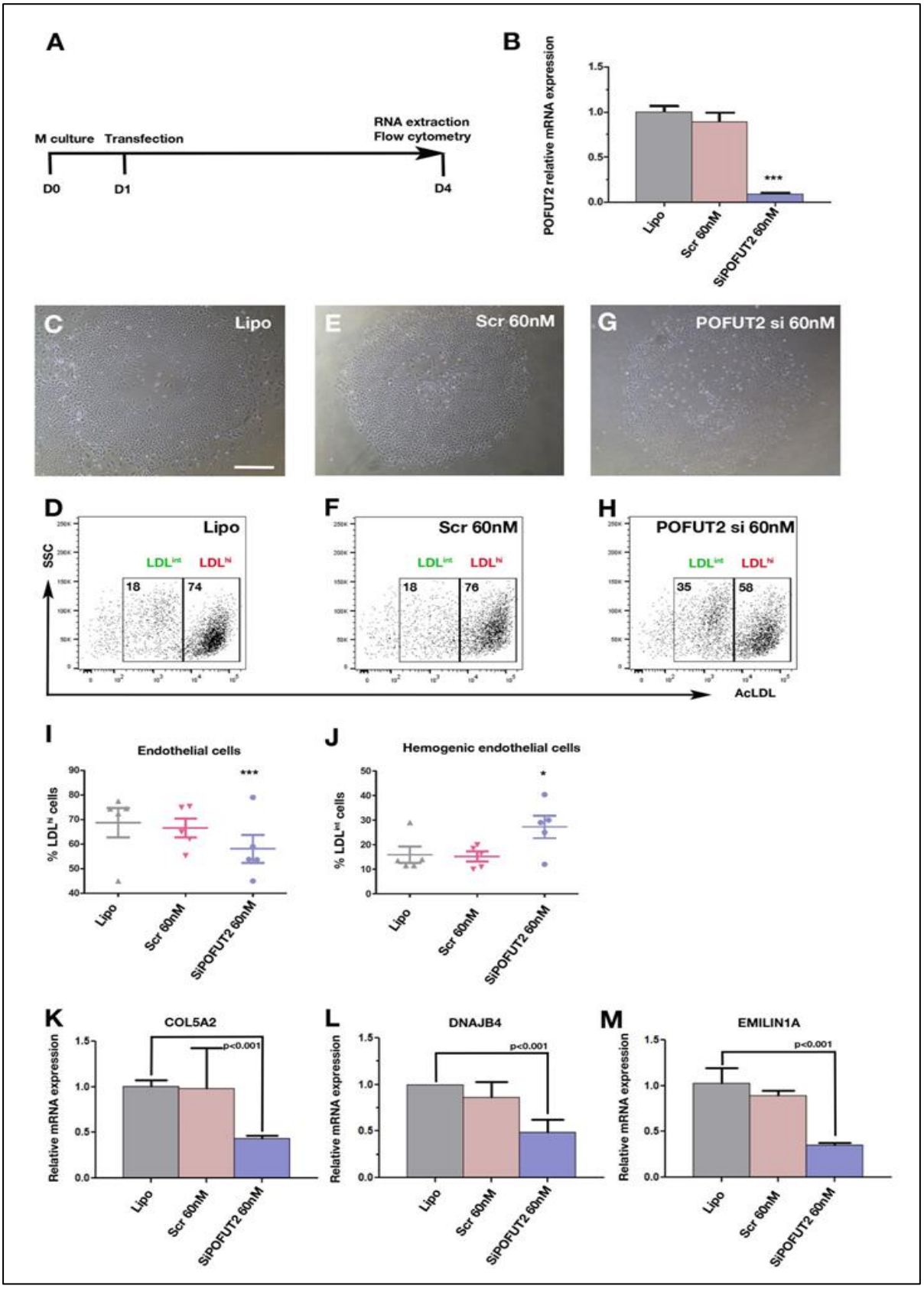

A
M culture
Transfection
RNA extraction
Flow cytometry
D0
D1
D4
B
POFUT2 relative mRNA expression
1.5
1.0
0.5
0.0
***
Lipo
Scr 60nM
SiPOFUT2 60nM
C
Lipo
E
Scr 60nM
G
POFUT2 si 60nM
D
Lipo
SSC
LDLint
LDLhi
18
74
F
Scr 60nM
LDLint
LDLhi
18
76
H
POFUT2 si 60nM
LDLint
LDLhi
35
58
AcLDL
I
Endothelial cells
% LDLhi cells
90
80
70
60
50
40
***
Lipo
Scr 60nM
SiPOFUT2 60nM
J
Hemogenic endothelial cells
% LDLint cells
50
40
30
20
10
0
*
Lipo
Scr 60nM
SiPOFUT2 60nM
K
COL5A2
Relative mRNA expression
1.5
1.0
0.5
0.0
p<0.001
Lipo
Scr 60nM
SiPOFUT2 60nM
L
DNAJB4
Relative mRNA expression
1.5
1.0
0.5
0.0
p<0.001
Lipo
Scr 60nM
SiPOFUT2 60nM
M
EMILIN1A
Relative mRNA expression
1.5
1.0
0.5
0.0
p<0.001
Lipo
Scr 60nM
SiPOFUT2 60nM

Figure A4: POFUT2 knock down enhances EHT.

A: experimental design. PSM cells were put in culture at D0, transfected at D1 and analyzed at D4 for AcLDL uptake by flow cytometry and q-PCR for the candidate gene and some of its immediate neighbors in the gene network. B: Effect of Lipofectamine (Lipo), scramble (Scr) and POFUT2 siRNA transduction on the expression of endogenous POFUT2. At 60nM, POFUT2 mRNA expression was decreased by more than 90% compared with the liposome and scramble siRNA controls.

C, E, G: culture at D4 that received Lipofectamine (C), scramble si (E) and POFUT2 si (F) at D1. Note that the number of round cells is more important in the siPOFUT2 condition than with the two other ones. D, F, H: representative flow cytometry experiment showing the balance between AcLDLint and AcLDLhi cells when cells received Lipofectamine (D), scramble si (F) and POFUT2 si (H). Note that the number of AcLDLint cells has increased at the expense of AcLDLhi cells in the POFUT2 si condition. I: percentage of AcLDLhi cells in the three conditions. A significant ($p<0,001$) decrease of LDLhi cells is visible in the POFUT2 si conditions (n=5).

J: percentage of AcLDLint cells in the three conditions. A low but significant ($p<0.08$) increase of AcLDLint cells is observed in the POFUT2 si condition compared with the two others. K: mRNA expression of COL5A2 is significantly decreased when POFUT2 is knock down by siRNA. L: DNAJB4 mRNA expression is decreased in POFUT2 knock down condition. M: EMILIN1A mRNA expression is significantly decreased in POFUT2 knock down condition. C, E, G, bar= 400µm.

Table A1: Sequences of the siRNA designed

siRNA	Sequence
Scr sense POFUT2	GGUUUCGUCUAUAGAUUGUTT
Scr antisense POFUT2	ACAAUCUAUAGACGAAACCTT
POFUT2 si1 sense	GCUUCACUUAUGAAGACUUTT
POFUT2 si1 antisense	AAGUCUUCAUAAGUGAAGCTT
POFUT2 si2 sense	GCAGAUUUAUGUCCUGCAATT
POFUT2 si2 antisense	UUGCAGGACAUAAAUCUGCTT
POFUT2 si3 sense	AAAUACCAUACUACGACGGTT
POFUT2 si3 antisense	CCGUCGUAGUAUGGUAUUUTT

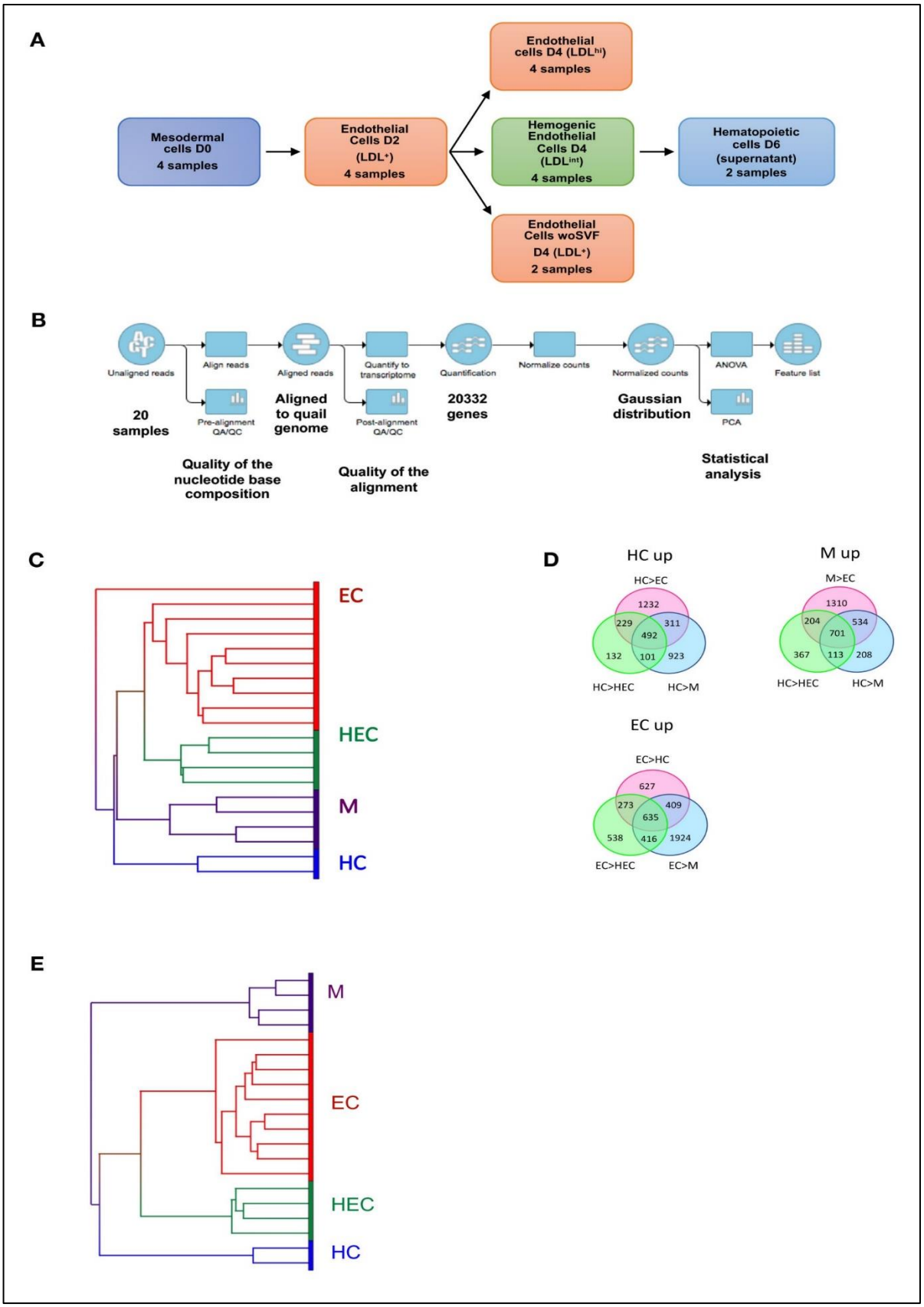

A
Mesodermal cells D0
4 samples
Endothelial Cells D2
(LDL+)
4 samples
Endothelial cells D4 (LDLhi)
4 samples
Hemogenic Endothelial Cells D4
(LDLint)
4 samples
Endothelial Cells woSVF
D4 (LDL+)
2 samples
Hematopoietic cells D6
(supernatant)
2 samples
B
Unaligned reads
Align reads
Aligned reads
Quantify to transcriptome
Quantification
Normalize counts
Normalized counts
ANOVA
Feature list
Pre-alignment QA/QC
Post-alignment QA/QC
PCA
20 samples
Aligned to quail genome
20332 genes
Gaussian distribution
Quality of the nucleotide base composition
Quality of the alignment
Statistical analysis
C
EC
HEC
M
HC
D
HC up
HC>EC
1232
229
311
492
132
101
923
HC>HEC
HC>M
M up
M>EC
1310
204
534
701
367
113
208
HC>HEC
HC>M
EC up
EC>HC
627
273
409
635
538
416
1924
EC>HEC
EC>M
E
M
EC
HEC
HC

Figure S1: Outlines of transcriptome and statistical analysis

A: Chart of the different cell fractions isolated, their LDL characteristics and their numbers. B: Partek flow pipeline of transcriptome analysis. C: Hierarchical clustering of the 20 samples based on the expression of the 20332 genes. Note that M is found in between HEC and HC D: Venn diagrams of the hematopoietic- (HC), mesoderm- (M) and endothelial- (EC) specific up-regulated genes. HC= 492, M= 701, EC= 635. E: Hierarchical clustering on the sets of differentially expressed genes. The series now starts from M and ends at HC.

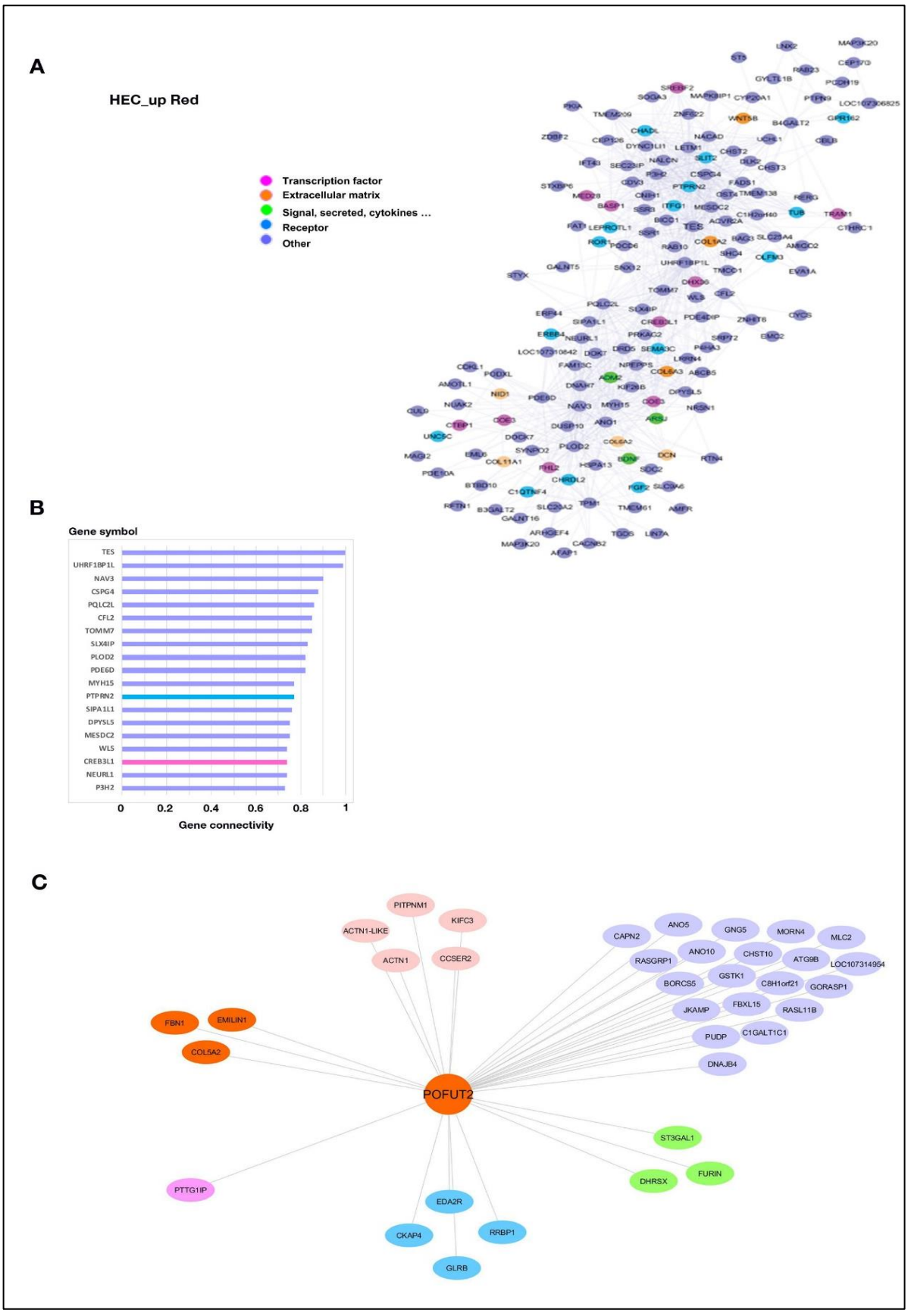

A
HEC_up Red
Transcription factor
Extracellular matrix
Signal, secreted, cytokines ...
Receptor
Other
B
Gene symbol
TES
UHRF1BP1L
NAV3
CSPG4
PQLC2L
CFL2
TOMM7
SLX4IP
PLOD2
PDE6D
MYH15
PTPRN2
SIPA1L1
DPYSL5
MESDC2
WLS
CREB3L1
NEURL1
P3H2
0
0.2
0.4
0.6
0.8
1
Gene connectivity
C
PITPNM1
KIFC3
ACTN1-LIKE
ACTN1
CCSER2
FBN1
EMILIN1
COL5A2
POFUT2
PTTG1IP
EDA2R
CKAP4
RRBP1
GLRB
ST3GAL1
FURIN
DHRSX
CAPN2
ANO5
GNG5
MORN4
MLC2
RASGRP1
ANO10
CHST10
ATG9B
LOC107314954
BORCS5
GSTK1
C8H1orf21
GORASP1
JKAMP
FBXL15
RASL11B
PUDP
C1GALT1C1
DNAJB4

Figure S2: HEC-specific red gene network and POFUT2 first neighbors

Cytoscape representation of the HEC-specific red network (up-regulated in HEC). The color code of the different molecule categories is indicated. B: Gene connectivity chart of the HEC-specific red network. The first 19 genes are represented. C: POFUT2 first neighbors. The color code is the same than in A.

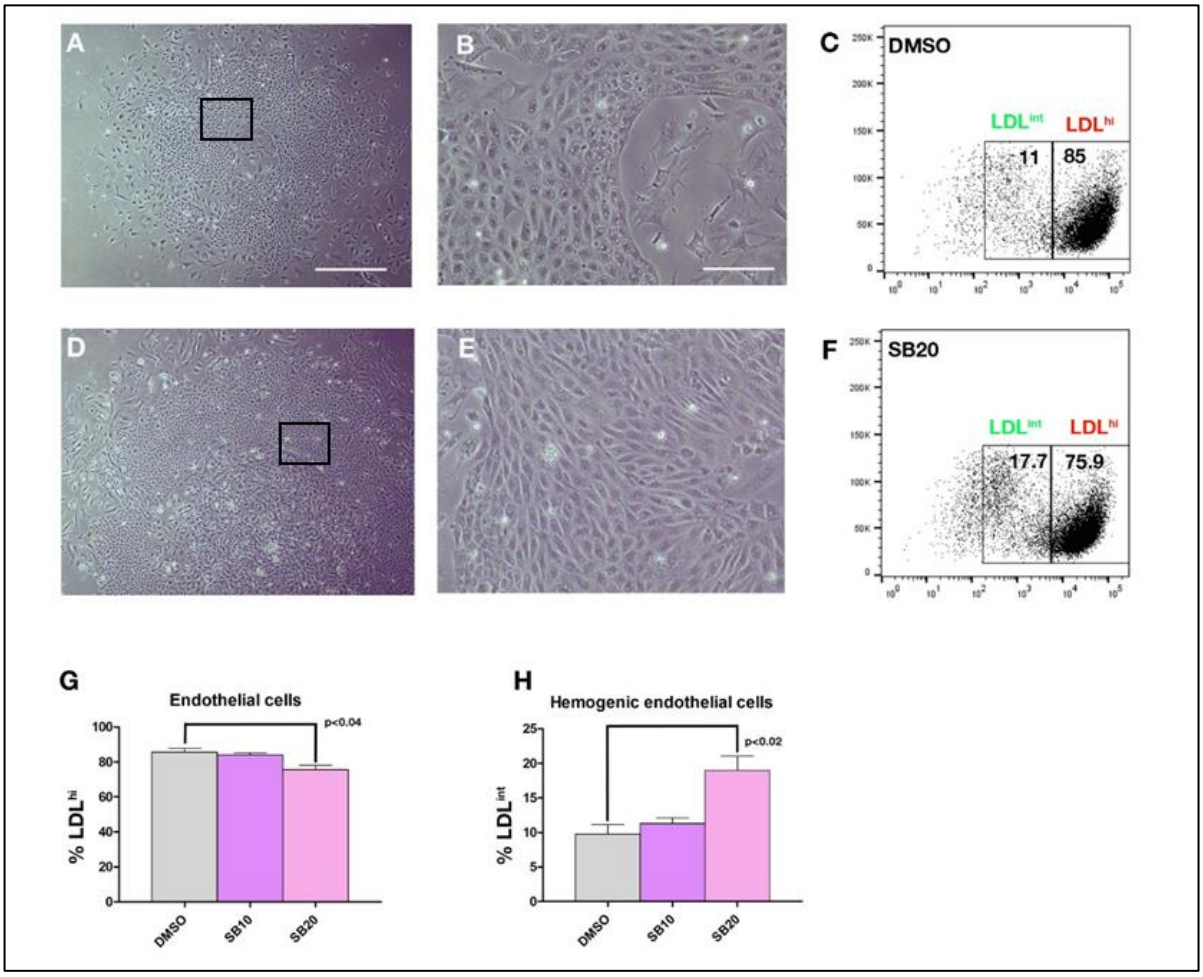


Figure S3: The Wnt agonist SB enhances EHT

A: PSM culture mock treated at D1 and observed at D4. B: enlargement of the frame in A showing the flat endothelial layer and the few round cells resulting from the initiation of EHT at this stage. C: Representative FACS analysis showing the percentages of LDLhi and LDLint in the mock- (DMSO) treated culture. D: PSM culture treated with SB216763 20µM at D1 and observed at D4. Note the presence of more round hematopoietic cells in the sample treated with SB216763 compared with the mock-treated culture. E: Enlargement of the frame in D showing the flattening of the cells and the presence of fewer round cells. F: representative FACS analysis showing the percentages of LDLhi et LDLint in the SB216763-treated culture. G: Percentage of the LDLhi cells in mock- vs SB216763-treated cells (n=5, p<0.04). H: Percentage of the LDLint cells in mock- vs SB216763-treated cells (n=5, p<0.02). Means are ± SEM. A, D bar= 400µm; B, E, bar= 20µm.

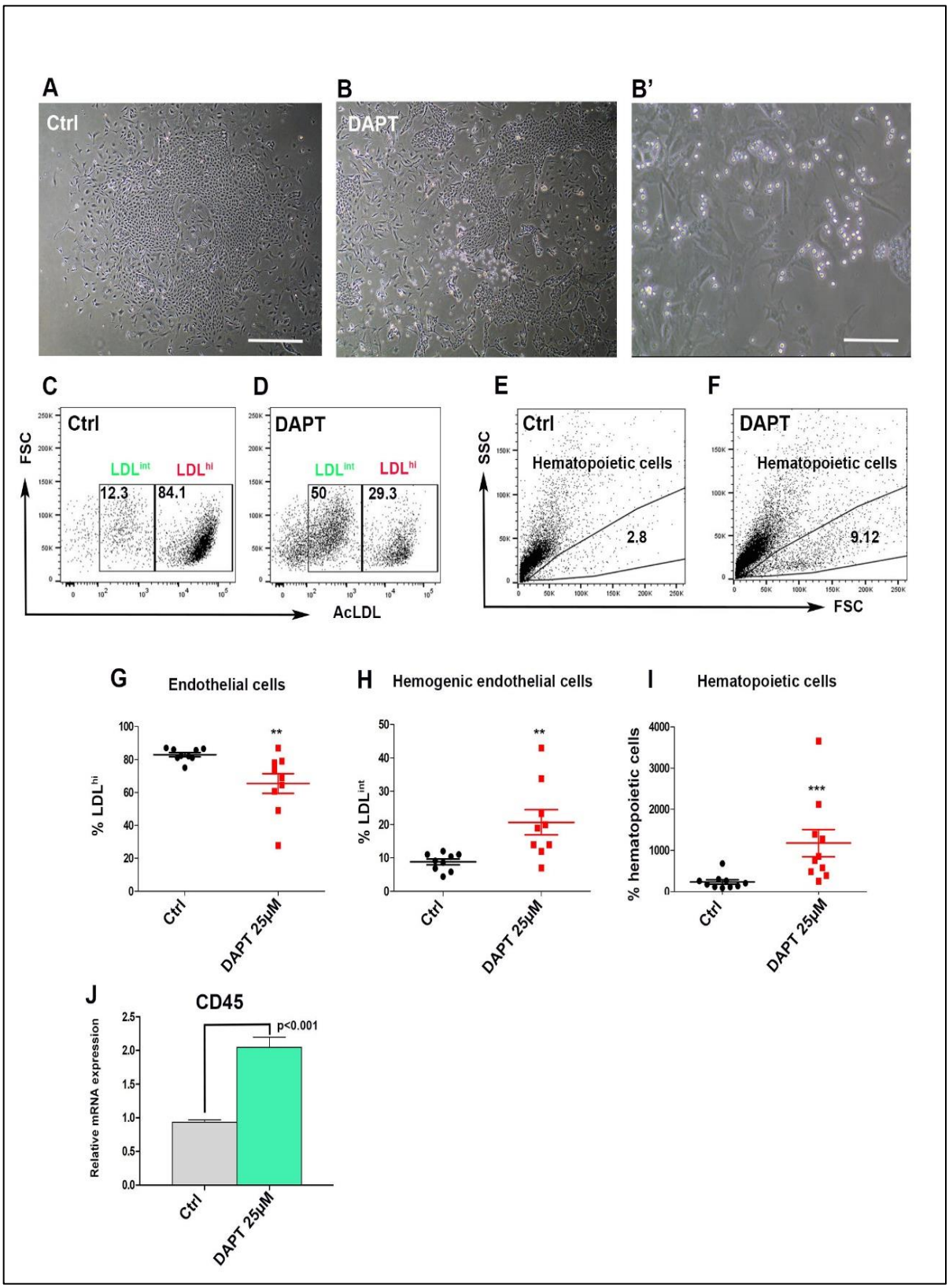

A
Ctrl
B
DAPT
B'
C
Ctrl
FSC
LDL^int
LDL^hi
12.3
84.1
D
DAPT
50
29.3
AcLDL
E
Ctrl
SSC
Hematopoietic cells
2.8
F
DAPT
Hematopoietic cells
9.12
FSC
G
Endothelial cells
% LDL^hi
**
Ctrl
DAPT 25μM
H
Hemogenic endothelial cells
% LDL^int
**
Ctrl
DAPT 25μM
I
Hematopoietic cells
% hematopoietic cells
***
Ctrl
DAPT 25μM
J
CD45
Relative mRNA expression
p<0.001
Ctrl
DAPT 25μM

Figure S4: The gamma secretase inhibitor DAPT enhances EHT

A: Aspect of a control culture at D4. B: aspect of the DAPT-treated culture (25µM) at D4. Note the presence of numerous round cells. B': Enlargement of a significant area of a DAPT-treated culture. Numerous round, bright, floating cells are visible. C: Representative FACS analysis showing the percentages of LDL^{hi} and LDLint in the mock- (DMSO) treated culture.

D: Representative FACS analysis showing the percentages of LDL^{hi} and LDL^{int} in the DAPT-treated culture. E: Representative FACS analysis showing the percentages of floating hematopoietic cells in the mock- (DMSO) treated culture. F: Representative FACS analysis showing the percentages of floating hematopoietic cells in the DAPT- treated culture. A 3.2-fold increase is obtained compared to the DMSO-treated culture. G: LDL^{hi} cells in DMSO- vs DAPT-treated cultures. H: LDL^{int} cells in DMSO- vs DAPT-treated cultures. I: Hematopoietic cells in DMSO- vs DAPT-treated cultures ($p<0.001$). J: q-RT-PCR of CD45 expression in DMSO- vs DAPT-treated cultures showing a more than two-fold increase in CD45 expression. A, B bar= 400µm; B' bar= 20µm.