THE UNIVERSITY
of EDINBURGH

Capturing the First Haematopoietic Stem Cell

-

The Needle in the Haystack

Christeun Sebastiaan Vink

THE UNIVERSITY
of EDINBURGH

Thesis submitted in fulfilment of the
requirement for the degree of

Doctor of Philosophy

(by Research Publications)

The University of Edinburgh

December 2021

Supervisors

Prof Elaine Dzierzak, principal supervisor

Prof Katrin Ottersbach, second supervisor

Viva voce/examination committee

Dr Linus Schumacher, internal examiner

Prof Anna Bigas, external examiner

Prof Thierry Jaffredo, external examiner

Prof Adriano Rossi, non-examining chair

Viva voce date

4th February 2022

Student number

s2215970

To Mom,

for giving up her career in science

when I stepped into this world.

Believe you can and you're halfway there.

– Theodore Roosevelt

- 26th President of the United States -

Nothing is impossible, the word itself says "I'm possible."

– Audrey Hepburn

- British-Dutch actress -

Without a song, each day would be a century.

– Mahalia Jackson

- American gospel singer -

Table of Contents

Declaration

I declare that this thesis has been composed by myself, that the work has not been submitted for any other degree or professional qualification, and that I have made a substantial contribution – clearly indicated – to each of the published papers. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

Signature: ……

Christeun S. Vink
Edinburgh
December 2021

Abstract

The most powerful cell in the blood differentiation hierarchy is the haematopoietic stem cell (HSC). It is the only cell capable of building an entire haematopoietic system from scratch, i.e. long-term (LT) repopulating a multilineage blood system within a lethally irradiated recipient. Though primitive blood is generated at early embryonic stages, it is only from embryonic day (E)10.5 of mouse development (4-5 weeks in human gestation) onwards that the definitive adult haematopoietic system is established with the formation of LT-HSCs.

These first HSCs emerge in small ventral intra-aortic haematopoietic clusters (IAHCs) in the aorta-gonad-mesonephros (AGM) region. Like with many systems, a tight regulation of gene expression by master transcription factors (TF) is what ultimately drives cell fate change. Gata2 is one such TF known to play a crucial role in mouse definitive haematopoiesis. Through the work of de Pater et al. (2013) (*Chapter 1*) it became clear that Gata2 is both required for the generation of the very first HSCs and also, unlike other key TFs like Runx1, the survival of HSCs. The earliest HSCs are generated through an endothelial-to-haematopoietic transition (EHT), a process where flat arterial haemogenic endothelial cells (HECs) round up into IAHCs, containing HSCs.

Solaimani Kartalaei et al. (2015) (*Chapter 4*) discovered that *Gpr56* is one of the most highly upregulated genes during EHT (in mouse) and that it is essential for HSC generation (in zebrafish). *Gpr56* is also a target of the key haematopoietic 'heptad' TFs in both mouse and human blood progenitors. To examine another signalling pathway required in the embryo for HSC generation, we used a *BMP responsive element* (*BRE*)-*GFP* transgenic mouse to show that all AGM HSCs are BMP-activated (Crisan et al., 2015) (*Chapter 6*). At slightly later stages, in the foetal liver and bone marrow, HSC heterogeneity starts to appear with the presence of genetically distinct BMP- and non-BMP activated HSCs.

Given the importance of the 'heptad' TFs in establishing HSC identity, we next developed a novel *Gata2-Venus* (*G2V*) reporter mouse to isolate and examine the dynamics and function of live Gata2-expressing and non-expressing cells (Kaimakis et al., 2016) (*Chapter 2*). *Gata2* is expressed in all HSCs, however haematopoietic progenitors can be generated either Gata2-dependently or -independently, the latter being less potent and genetically distinct from the first. With Eich et al. (2018) (*Chapter 3*) we went further and deciphered the role of Gata2 in haematopoietic fate establishment. By using time-lapse imaging of E10.5 live aortic sections of *G2V* embryos, we discovered rapid pulsatile level changes in Gata2 expression, specifically in those single cells undergoing EHT. This finding indicated the highly unstable genetic state of individual cells undergoing fate change. Furthermore, aortic cells haploinsufficient for *Gata2* show many fewer pulsatile and EHT events, emphasizing the importance of Gata2 levels in this process.

Having defined the specific (medium) levels of Gata2(Venus) protein in the population containing the first HSCs, we used our *G2V* model to isolate a pure population of HSCs as they are generated for the first time in development (Vink et al., 2020) (*Chapter 5*). Through iterations of index-sorting, single-cell transcriptomics and functional analyses we were able to greatly enrich for HSCs (~70x compared to Eich et al. (2018)). Based on refined CD31, cKit and CD27 expression and specific physical parameters we isolated the first cells with an HSC identity and identified their precise gene expression profile. Immunohistochemistry localized these HSCs to only the smallest IAHCs. The combined effort of my research and novel tools generated to facilitate the creation of this body of work/portfolio have now led us for the first time to study the one or two cells per embryo that harbour HSC identity at the developmental time when they are beginning to be made. Now that we have cracked the genetic code of the true HSC, the field is one step closer to translating these findings and applying them to producing human clinically relevant HSCs, which will ultimately improve transplantation therapies and benefit the fight against leukaemia.

Lay summary

Human blood is composed of billions of cells that are all derived from one common precursor called the blood/haematopoietic stem cell (HSC). In adults, HSCs reside in the spongy tissue (marrow) inside the long bones of the body (e.g. femur). Healthy bone marrow is needed throughout life, and therefore many people with blood cancers, such as leukaemia and lymphoma, rely on bone marrow transplants for survival. Unfortunately, these treatments suffer from a big shortage of compatible donor HSCs. Thus far it has been impossible to generate healthy functioning HSCs in the laboratory. To gain a better understanding of what it takes to make an HSC, we study the mouse embryo at the time of development when HSCs are beginning to be made. In both human and mouse embryos those very first HSCs are made within the major blood vessel –the aorta. When a mouse embryo is 10.5 days old, 1-2 HSCs are generated from the wall of the aorta and remain in small clusters of cells until required by the developing embryo. These first HSCs have the power to produce the entire blood system of an adult mouse. Since these first HSCs are so rare, no lab has ever been able to isolate and study them as a pure population. To fully understand what an HSC is, and ultimately how to make one in the laboratory for clinical purposes, we would need to know the 'genetic code/programme' that makes an HSC in the embryo, i.e. what pieces of DNA ('genes') need to be activated for an HSC to be produced and survive. Up until now, whenever this was attempted, many 'contaminating' blood cells were present in the 'HSC population', making it impossible to capture the 'genetic code/programme' exclusive to HSCs. Through my work we can now report the 'genetic code/programme' of these incredibly rare functional HSCs. We discovered new combinations of molecules inside (Gata2) and on the surface (CD31, cKit and CD27) of HSCs and physically isolated the individual cells that make an entire adult blood system, thus allowing us to for the first time 'crack' the genetic programme of HSCs as they are made during development. Our data demonstrate that this 'code' is unique compared to the other closely related blood cells. We also

show that these HSCs can only be found in the smallest aortic clusters of 1-2 cells. All these findings bring us one step closer to producing these clinically important blood stem cells, which will ultimately improve transplantation therapies and benefit the fight against blood cancer.

Acknowledgements

No matter what accomplishments you achieve,

somebody helped you.

– Althea Gibson

- first African American to win a Grand Slam -

Whilst listening to Miles Davis' masterpiece 'Kind of Blue', as I have done so often during this PhD, I would like to reflect on this once-in-a-lifetime experience. It fills me with great pride to have completed my doctoral studies in Scotland and at The University of Edinburgh, and to join a long line of graduates going back as far as 1583. The realisation that this body of work will be in the public domain forever, makes this section extra special. For it is truly the support and help of those acknowledged here that enabled me to do the research which has culminated in me undertaking a PhD.

First and foremost, I'm deeply indebted to Prof Elaine Dzierzak. When joining your lab in September 2009 I could not have imagined moving (the lab) to Scotland 4.5 years later, publishing a first-author paper as a lab manager and deciding to undertake a PhD at the end of 2020. Thank you for creating such a stimulating environment where all of this was possible. It has made me grow personally, intellectually and scientifically. Something I will be forever grateful for. Thank you as well for all the valuable advice you've given prior to and during this PhD, and for sharing so much of your extensive knowledge and experiences, both in and out of the lab. For me personally, putting together my paper and going over the manuscript side by side has taught me so much about scientific writing, the English language and all it takes to produce a high-quality paper. An immensely valuable experience that has helped me a lot during this PhD and will undoubtedly continue to do so in the future. I feel privileged to be able to call such an excellent scientist as yourself my supervisor, boss and mentor. Apart from science, attending

the concerts by Bruce Springsteen and The Rolling Stones were absolute highlights. These, and many other, wonderful memories are part of the reason why the past twelve years have flown by. Thank you for the science journey and for relocating the lab to such a wonderful city and country.

I shall remember to take it as a good sign when, during any next job interview, I'll be asked my opinion about Bob Dylan.

I also had the pleasure of doing two fantastic wet-lab internships as part of my undergraduate studies, prior to starting in the Dzierzak group. I'd like to thank Dr Marieke von Lindern and my direct supervisor Dr Rastislav Horos, and Prof Marella de Bruijn and direct supervisor Dr Andrew Jarratt for making these such positive and formative experiences that have contributed to me being where I am today.

I would like to extend my sincere thanks to Dr Linus Schumacher, Profs Anna Bigas, Thierry Jaffredo and Adriano Rossi for committing their time and agreeing to be part of my examining committee, and to Prof Katrin Ottersbach for being my second supervisor.

I would also like to acknowledge Dr David Munro for making me aware of the possibility of doing a PhD by Research Publications, and Camiel Janssen, perhaps unbeknownst to him, for being one of the people stimulating me to pursue this PhD.

Many thanks as well to all past members of the (ErasmusMC) Dzierzak group, but especially to Drs Emma de Pater, Mihaela Crisan, Christina Eich, Parham Solaimani-Kartalaei, Tomoko Yamada-Inagawa, Polynikis Kaimakis and Parisa Imanirad. I'm grateful for having been involved in your studies and for having been able to make significant contributions to your first-author publications, some of which are now in this thesis. Thank you for your great team spirit and for all the good times we had when the lab was still in Rotterdam. In Mihaela's words: "Life [was and] is beautiful!"

xx

Special thanks to Dr Thomas Clapes for his friendship and all the fun (food and drink) times in Rotterdam, Freiburg, the South of France and Edinburgh. Bring on the next trip! I'm losing my French.

A huge thank you to all past and present members of the (University of Edinburgh) Dzierzak group for creating such a wonderful international atmosphere and team spirit (in that tiny – but cosy – office of ours), and for all the many amazing lab dinners and outings. Coming to work has been such a joy with all of you around. We've always had (and have) so much fun and so many random topics to discuss, but when work and big experiments need(ed) to be done, we work(ed) hard and are/were always there to help each other out. Being abroad, like most of us, creates an extra special bond and I feel very fortunate that bond was and is with people like you. I'm immensely grateful to Drs Samanta Mariani, Antonio Maglitto, Carmen Rodriguez-Seoane and Mari-Liis Lukke. Somehow the stars must've aligned for a Dutch, an Estonian, an Italian, a Sicilian and a Galician to all find their way to Edinburgh and end up in the same research group and, most importantly, have such a fantastic click, both in and out of the lab. Truly unique encounters and friendships I will cherish forever. Samanta, *grazie mille* for your generous friendship, being our Queen of FACS (and curing our FRD), for all your invaluable help and advice in the lab and with my paper, for all things Italian you've taught me (and had me taste, experience, introduced me to [*Italia* itself!]), and for genuinely just being too good for this world. *Ti voglio bene*. Antonio, *grazie mille anche a te! Sei un bravo ragazzo!* (Probably the first Italian you taught me). Thanks for all the unforgettable experiences, trips, parties (The Netherlands, Malta, *Sicilia*, NSJ, Brisbane, Hong Kong etc.), all the wonderful music (more gigs please!), for teaching me how to cook *Pasta alla Norma* (which catalysed my love for cooking *La Cucina Italiana*), and of course also for your great contribution to my paper. Carmen, my long-time back-to-back office partner, *moitas grazas* for all the office and lab fun, the great nights out, for all your invaluable help in keeping the lab up and running and with the experiments for my paper, your southern

temperamento and for promoting Galicia (I can't wait to visit!). Mari-Liis, *tänan sind väga* for your friendship. I'm so glad I made the decision to move to Edinburgh as it meant we got to know each other so much better and became true friends. I remember it like it was yesterday when we were painting our office (and the paint is still on the walls!). I have such fond memories of all the concerts we attended, the club nights (Espionage!), NSJ festival and your beautiful wedding that brought the whole Edinburgh crew together. Of course, also the many fantastic trips together with Drs Marina Gabriel-Salazar and Victor Casanova. *Moltes gràcies* my favourite Catalans! Hopefully we can soon again plan another one! Marina, I'm so glad you took it upon yourself to do an extra lab internship before your PhD and join me and Mari-Liis in Edinburgh. Thank you for all the beautiful and fun moments. It was great as well I got to spend some extra time with you in Tarragona and the Ebro Delta after your amazing wedding this summer. Victor, you were not officially in our group, but you basically were. Thank you for your friendship and for all the fun times. I've been without *Ratafia* for too long and my *Romesco* sauce is nearly finished, so we need to meet up!

Special thanks to Dr Zhuan Li, Dr Fokion Glykofrydis, Dr Xabier Cortes-Lavaud, Eoghan Forde and Karen Colvin as well. Zhuan, I feel fortunate to have learned so much about China from you and I truly hope I get to visit you in Guangzhou one day. *Xièxie* for all the good times in the lab. Karen, you wonderfully crazy Scottish *lass*, it's been so great getting to know you. Thanks for teaching me Scots and for all the *crack*. It's been *well barry*! I would also like to extend my thanks to Anna Popravko and Lorna Mackintosh. Thank you for your enthusiasm, determination and bringing your unique sets of expertise to the group. We might be a small team now, but I'm very happy to see there's again plenty of cake in the office and that you've continued our group's strong festive tradition (I secretly do enjoy it). Anna, many thanks as well for all the bioinformatics analyses you've performed on my data. It's always a miracle to me how you do it and it's so exciting to discover there's so much more in there. It's great to see the story continue into your PhD. Thanks should also go to the QMRI Flow Cytometry and Cell

Sorting Facility, Shonna Johnston and Drs William Ramsay and Mari George, for all their help and excellent cell sorting for my paper, and for all the fun as office neighbours and during all social events.

I also cannot forget about the non-lab members of my Edinburgh family. Many thanks especially to Dr Luca Cassetta for his friendship, all the great events and holidays (in our homelands) over the years and for knowing there'll always be a place to stay at yours. Special thanks as well to Drs Francesca Vacante and Themis Tsarouchas. Truly getting to know you two during the past two years has been wonderful. *Grazie mille* and *σε ευχαριστώ πάρα πολύ* for your generous friendship, enthusiasm, actively staying in touch and organising many fun dinners with a great bunch of people. I feel fortunate to be part of it. Cheers to many more! I also wish to thank Drs Calum Robb and Jamie Scott for their support during the PhD and all the many fun times we've had. Thanks as well to Bas de Wolf and Prof Dies Meijer for bringing a bit of The Netherlands to Scotland. Dies, many thanks for all the nice food and good company over the years. I will also never forget that time we nearly got stuck on the Holy Isle of Lindisfarne. Bas, it's been great getting to know another like-minded Dutchie. Thank you for the adventures (the near avalanche in The Pentlands was quite the highlight, in hindsight), the interesting discussions and for celebrating *Koningsdag* out on a terrace in the freezing cold, due to Covid restrictions. To many more!

Lastly, I'm extremely thankful to my family without whom I'd be nowhere. The completion of this thesis and the beautiful journey it took to get here would not have been possible without the unconditional support and love of my parents, Hans and Joke, my siblings Koen and Jolien and their respective partners Tamara and Camiel, and my cousin (/sister), Lieneke. *Ik houd van jullie*. Mom, dad, I cannot begin to express how grateful I am for all your financial support over the years and for the values with which you've raised us, in such a beautiful warm and stable home, resulting in all three of us doing well in our studies and beyond. Koen, Jolien and Lieneke, thank you

for all the craziness that just always happens whenever we get together. I feel incredibly privileged and lucky to be able to call you family and friends at the same time.

Portfolio of published work

The following six research publications are submitted as part of this thesis and will be discussed in depth in the Critical Review:

1. De Pater, E., Kaimakis, P., **Vink, C.S.**, Yokomizo, T., Yamada-Inagawa, T., Van der Linden, R., Camper, S.A., Speck, N.A., Dzierzak, E. (2013). *Gata2* is required for HSC generation and HSC survival. *Journal of Experimental Medicine*, *210*(13), 2843-50. PMID: 24297996. (***Chapter 1*** of thesis).

2. Kaimakis, P.*, De Pater, E.*, Eich, C.*, Solaimani-Kartalaei, P., Kauts, M-L., **Vink, C.S.**, Van der Linden, R., Jaegle, M., Yokomizo, T., Meijer, D., Dzierzak, E. (2016). Functional and molecular characterization of mouse *Gata2*-independent hematopoietic progenitors. *Blood*, *127*(11), 1426-37. PMID: 26834239. (***Chapter 2*** of thesis).

3. Eich, C., Arlt, J., **Vink, C.S.**, Solaimani-Kartalaei, P., Kaimakis, P., Mariani, S.A., Van der Linden, R., Van Cappellen, W.A., Dzierzak, E. (2018). *In vivo* single-cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. *Journal of Experimental Medicine*, *215*(1), 233-248. PMID: 29217535. (***Chapter 3*** of thesis).

4. Solaimani-Kartalaei, P.*, Yamada-Inagawa,T.*, **Vink, C.S.**, De Pater, E., Van der Linden, R., Marks-Bluth, J., Van der Sloot, A., Van den Hout, M., Yokomizo, T., Van Schaick-Solerno, M.L., Delwel, R., Pimanda, J.E., Van IJcken, W.F., Dzierzak, E. (2015). Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *Journal of Experimental Medicine*, *212*(1), 93-106. PMID: 25547674. (***Chapter 4*** of thesis).

5. **Vink, C.S.**, Calero-Nieto, F.J., Wang, X., Maglitto, A., Mariani, S.A., Jawaid, W., Göttgens, B., Dzierzak, E. (2020). Iterative single-cell analyses define the transcriptome of the first hematopoietic stem cells. *Cell Reports*, *31*(6), 107627.
   PMID: 32402290. (***Chapter 5*** of thesis).

6. Crisan, M., Solaimani-Kartalaei, P., **Vink, C.S.**, Yamada-Inagawa, T., Bollerot, K., Van IJcken, W., Van der Linden, R., De Sousa Lopes, S.M., Monteiro, R., Mummery, C., Dzierzak, E. (2015). BMP signalling differentially regulates distinct haematopoietic stem cell types. *Nature Communications*, *6*(8040)/*6*(8793; corrigendum).
   PMID: 26282601. (***Chapter 6*** of thesis)

* Authors contributed equally.

See Appendix for full publications and supplemental materials.

Critical review

General introduction

Adult haematopoiesis and the discovery and purification of its most potent contributor

The average human (*Homo sapiens*) body contains roughly 5.0 litres of blood, accounting for 7% of body weight (Alberts et al., 2015). Adult mice (*Mus musculus*) have 1.0-2.4 millilitres, which accounts for a similar percentage of their body weight (Wolfensohn and Lloyd, 2013). To maintain this specialised form of connective tissue in both adult mice and men, billions of circulating and migratory haematopoietic cells of more than ten different cell lineages are generated predominantly from the bone marrow (BM) each day, a process referred to as haematopoiesis (reviewed in Nombela-Arrieta and Manz (2017); Seita and Weissman (2010)).

The most superior and at the same time rarest cell from which all haematopoietic lineages originate is the haematopoietic stem cell (HSC), and ever since its first characterisation it has been placed at the source of a branching tree-like haematopoietic hierarchy (Figure 1) (reviewed in Laurenti and Gottgens (2018); Seita and Weissman (2010)). This hierarchy comprises of various lineages, defined as the developmental trajectories of a (population of) cell(s) from birth until terminal division and differentiation into a specific cell type, also known as cell fate. The HSC is experimentally defined as a cell that 1) is able to long-term self-renew, i.e. divide and give rise to one daughter HSC that differentiates and one that retains HSC properties upon serial transplantation into myeloablated adult recipients, and thereby persist throughout an animal's lifespan, 2) has the ability to reconstitute the entire multilineage haematopoietic differentiation hierarchy, and 3) has the capacity to repopulate and home to the adult mammalian BM

(reviewed in Dzierzak and Bigas (2018); Seita and Weissman (2010)). This myriad of different known blood cell types and lineages makes the haematopoietic system the most intricate and probably best-characterised stem cell system in mammals (reviewed in Weissman (2000)). Whereas most tissues can be repaired following injury, blood, a tissue without a known three-dimensional organisation, is unique and can actually be fully restored to its original homeostatic state (reviewed in Schroeder (2008)).

Considering that HSCs do not have clearly distinct morphological characteristics as compared to other haematopoietic cells, the transfusion of (potentially) HSC-containing populations into lethally irradiated recipients and the detection of multilineage reconstitution remains the “gold standard” to show HSC functionality. Such transplantations trigger an HSC towards expansion and multilineage differentiation to ultimately reveal its true potential. When examined in their biological undisturbed (non-transplant) environment, the actual *in vivo* cell fate of these HSCs might not necessarily be the same (reviewed in Haas et al. (2018), Jacobsen and Nerlov (2019)).

Our current understanding of HSCs started to take shape thanks to a big surge in the HSC research field following World War II, the development of the atomic bomb and with the realisation that BM is an extremely radiosensitive tissue (reviewed in Eaves (2015)). Some of the earliest studies used lead shielding or transplantation of haematopoietic tissues (spleen, liver and BM) during or following lethal irradiation and concluded that a powerful ‘humoral factor’ could rescue the damaged haematopoietic system of irradiated recipients or lead-shielded mice (Jacobson et al., 1951; Lorenz et al., 1951). The comprehension of the existence of a powerful transplantable cell with haematopoietic repopulation/chimaeric capacity came from the replacement of the damaged BM of lethally irradiated mice by intravenous transplants of genetically distinct (and thus distinguishable post-transplantation) adult mouse/rat donor BM cells (Ford et al., 1956; Nowell et al., 1956; Vos et al., 1956). In these scenarios the destructive properties of irradiation created ‘space’ in the recipient’s BM for donor BM cells to repopulate it and create haematopoietic chimaeras, hereby for the first time

providing substantial proof for the cellular (and not humoral) basis of blood production. The results of these three independent studies by groups from the USA (Nowell et al., 1956), England (Ford et al., 1956) and The Netherlands (Vos et al., 1956) all simultaneously published in March 1956, would lead E. Donnall Thomas to perform the first allogeneic (genetically different, but same species) intravenous BM transplantation on irradiated human leukaemia patients (Thomas et al., 1957). His transplantation work would later be recognised with a Nobel Prize.

The previous discoveries fuelled research into the quantification of the cells responsible for the repopulation of mice, and specifically to determine the minimum number that would be sufficient to rescue the effects of lethal irradiation (McCulloch and Till, 1960). These first limiting dilution assays also aided the discovery of large macroscopic colonies (‘gross nodules’) on the spleens of irradiated mice 10-11 days after they had received a BM transplant (Till and McCulloch, 1961). Strikingly, the number of BM cells injected correlated linearly with the number of colonies observed, thereby allowing for a direct means of quantifying these highly potent blood-forming cells, that they termed spleen colony-forming units (CFU-S) (Till and McCulloch, 1961). The newly-discovered cells were rare and present at about one CFU-S per 10,000 nucleated BM cells (Till and McCulloch, 1961). A couple years later, by using BM donor cells with at random (and therefore most likely unique) radiation-induced chromosomal differences/markers, the same research group showed these colonies were clonal and derived from one single CFU-S cell (Becker et al., 1963). The multipotency of (at least some) colony-forming ‘HSCs’ was proven by first showing they could give rise to myeloid (erythropoietic, granulocytic and megakaryocytic) lineages/offspring within the spleen colonies (Fowler et al., 1967; Wu et al., 1967) and subsequently that lymphoid cells either descended from CFU-S or a common progenitor (Wu et al., 1968). Conclusive evidence of the existence of a multipotent BM HSC, capable of producing both myeloid and functional B- and T-cells, came later (Abramson et al., 1977).

The other key HSC-defining characteristic (and nowadays deemed that of any other non-haematopoietic stem cell type) came from the evidence that a primary recipient's spleen colonies contain cells that are able to give rise to the same multilineage colonies in secondary transplant recipients, hereby proving the "self-renewal" capacity of these single (CFU-S) HSCs (Siminovitch et al., 1963). The composition of the spleen colonies was found to be very heterogeneous in terms of numbers of differentiated and immature cells, with "few colonies having many colony-forming cells [CFU-S], and many colonies having few." (Fowler et al., 1967; Siminovitch et al., 1963). Till & McCulloch would try to explain this heterogeneity by proposing a "birth-and-death" process with random generation/proliferation of CFU-S ("birth") or differentiation towards certain mature lineages (and thus loss ["death"] of CFU-s potential), a model they called "haemopoiesis engendered randomly," introducing the idea of stochasticity into the field of haematopoiesis for the first time (Till and McCulloch, 1980; Till et al., 1964). The majority of CFU-S were later found not to be actual HSCs, but rather a type of oligopotent (more committed erythromyeloid) progenitor (Na Nakorn et al., 2002). More definitive and direct proof for the existence of single long-lived BM HSCs that can clonally expand, self-renew and produce multilineage progeny, came from experiments that used retroviral integration sites to mark clonality and allowed for fate tracing of each clone's offspring upon primary (minimum of 8 months) and secondary (additional 7 months) transplantation (Keller et al., 1985; Keller and Snodgrass, 1990; Lemischka et al., 1986; Williams et al., 1984), a marking method that proved more efficient and less harmful than the earlier radiation-induced one (Abramson et al., 1977; Becker et al., 1963; Fowler et al., 1967; Wu et al., 1967; Wu et al., 1968). Thus, early BM transplants and CFU-S assays led to the functional characterisation of what we now call an HSC and established and defined the core stem cell principles of self-renewal and multipotency.

For many years, the cellular heterogeneity and diversity of the haematopoietic tissue made it very challenging to obtain large and homogenous enough populations, especially of those most potent stem and

progenitor cells, for proper molecular and functional analyses and characterisation. It was the use of the relatively new fluorescence-activated cell sorting (FACS) technique in combination with isolation based on other physical characteristics that enabled the first great enrichment of the (predominantly erythromyeloid) CFU-S cells (Muller-Sieburg et al., 1986; Visser et al., 1984). Nevertheless, it would take until 1988 before the Weissman group successfully managed to prospectively purify HSCs down to 0.05% of adult mouse BM and with a quantified activity of only 30 of their purified cells needed to rescue half of a group of lethally irradiated recipients (Spangrude et al., 1988). To perform this enrichment, they used multi-colour FACS with monoclonal antibodies against various cell surface antigen molecules/markers, first negatively selecting for mature/differentiated haematopoietic cell lineages (Lin^-; excluding B-, T-, myelomonocytic cells and granulocytes), followed by positive selection for $CD90^{low}$- (*Thy1*) and Sca-1- (Stem cell antigen-1; also known as 'lymphocyte antigen 6 complex, locus A', *Ly6A*) expressing cells. Many further reductionist/enrichment approaches have now shown these cell sorting strategies did not yield a 'pure' population of HSCs: The addition of numerous other markers facilitated the further enrichment. A notable marker still heavily used to this day, is the added positive selection for cKit (*Kit*, also known as CD117) (Ikuta and Weissman, 1992; Okada et al., 1992; Okada et al., 1991), establishing the well-known "LSK" ($Lin^-Sca\text{-}1^+cKit^+$) population. The simplest marker combination yielding the highest purity, was achieved with "LSK" and the additional positive and negative selection for CD150 (*Slamf1*) and CD48, respectively (Kiel et al., 2005). With both being part of the SLAM (Signalling Lymphocyte Activation Molecule) cell surface receptor family, this full combination of markers is now commonly referred to as the "LSK SLAM" HSC purification strategy. At only $0.0058\% \pm 0.0012\%$ of whole BM cells and with one in two cells yielding long-term multilineage reconstitution in irradiated recipients (Kiel et al., 2005), this cell sorting approach still yields the highest level of adult mouse BM HSC enrichment to this date (reviewed in Nombela-Arrieta and Manz (2017)). Interestingly, several markers such as CD34 were shown to be expressed

only on developmental-specific (mouse embryonic) or species-specific (human BM) HSCs (Sanchez et al. (1996); reviewed in Seita and Weissman (2010)). Adult mouse BM HSCs are negative for or express low levels of CD34 and are found at a one-in-five cell frequency within the "$CD34^{low/negative}$ LSK" population (Osawa et al., 1996). Taken together, the development of the FACS technique and discovery of unique HSC markers has greatly helped in the purification, isolation and study of these rare cells.

The heterogeneity of the BM population included cells with other haematopoietic functions and properties. While these might be considered 'contaminating' cells within the HSC populations many were found to be multi-, bi- and mono-lineage-specific haematopoietic progenitor cells (HPC). The existence of a multitude of HPCs with different potencies first became clear from *in vitro* haematopoietic functional colony assays developed by Metcalf and others (Bradley and Metcalf, 1966; Metcalf, 1970; Pluznik and Sachs, 1966). In these assays, haematopoietic cells were cultured in semi-solid agar medium (initially on mouse embryo/spleen feeder layers) supplemented with growth factors like colony stimulating factor (CSF). This allowed for the formation of clusters/aggregates of cells (colony forming unit-culture; CFU-C) which would, by examining the cellular morphology of differentiated progeny within a colony, reveal the potency of the input cells after one to seven days of culture. Based on these initial experiments using whole BM and seminal work using FACS-sorted cells by the Weissman group and others (reviewed in Eaves (2015); Seita and Weissman (2010)), it became evident that heterogeneity exists within the pool of rare HSCs. It was shown that HSCs classically give rise to highly proliferative non-self-renewing multipotent progenitors (MPP), followed by oligopotent common myeloid and lymphoid progenitors (CMP/CLP), lineage restricted (unipotent) progenitors and eventually terminally differentiated/mature cells (Figure 1A,B). All progenitors are progressively restricted (more committed) in their lineage-output potential and their high proliferative state makes for a hierarchical (tree-like) system by which one HSC can give rise to thousands of differentiated offspring. HSCs can therefore be in a dormant/quiescent state

and by asymmetric division still continue to ensure a high rate of haematopoiesis (reviewed in Laurenti and Gottgens (2018)). During the last couple of decades, the haematopoietic hierarchical tree has been revised several times (Figure 1) and due to technological advances in single-cell transcriptomics, the shape of the 'tree' has shifted yet again. A steady and highly grouped hierarchy with a series of fixed intermediate cells (Figure 1A,B) is now molecularly perceived as a continuum with a much more gradual differentiation from one heterogeneous population to the next (Figure 1C) (reviewed in Laurenti and Gottgens (2018)). Still, any of these graphical depictions of the haematopoietic differentiation hierarchy are most-certainly an over-simplification of reality and an underestimation of potential differentiation routes. A more realistic 'tree', including more than one possible differentiation trajectory out of the HSC compartment, has been proposed (Laurenti and Gottgens, 2018) (Figure 1D). In addition to any hierarchical model relying on an HSC compartment at the top, there is growing evidence of long-lived HSC-independently generated haematopoietic cells such as microglia (Ginhoux et al., 2010). The fact that *in vitro* molecular reprogramming of pluripotent or differentiated cells can generate some mature blood lineages but not HSCs, supports a flexible haematopoietic developmental programme that does not always follow one stable differentiation hierarchy (reviewed in Dzierzak and Bigas (2018)).

Altogether, the HSC is the most potent contributor to haematopoiesis, however HPCs also play an important role. The existence of HSC-independently generated HPCs that complement the potent HSC-dependent haematopoietic hierarchy strongly point to more than one programme of generating haematopoietic cells.

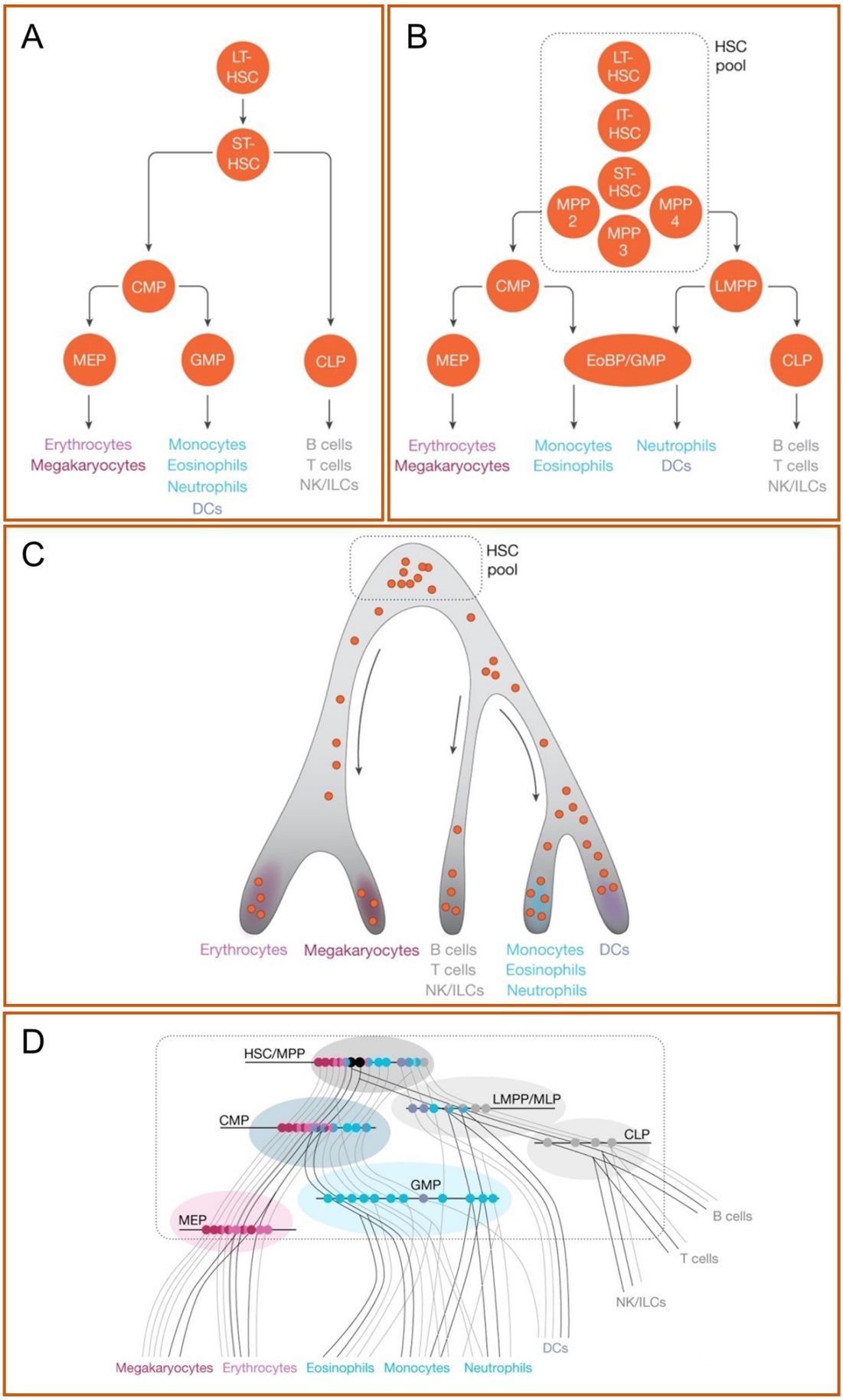

A
LT-HSC
ST-HSC
CMP
MEP
GMP
CLP
Erythrocytes
Megakaryocytes
Monocytes
Eosinophils
Neutrophils
DCs
B cells
T cells
NK/ILCs
B
HSC pool
LT-HSC
IT-HSC
ST-HSC
MPP 2
MPP 3
MPP 4
CMP
LMPP
MEP
EoBP/GMP
CLP
Erythrocytes
Megakaryocytes
Monocytes
Eosinophils
Neutrophils
DCs
B cells
T cells
NK/ILCs
C
HSC pool
Erythrocytes
Megakaryocytes
B cells
T cells
NK/ILCs
Monocytes
Eosinophils
Neutrophils
DCs
D
HSC/MPP
LMPP/MLP
CMP
CLP
GMP
MEP
B cells
T cells
NK/ILCs
DCs
Megakaryocytes
Erythrocytes
Eosinophils
Monocytes
Neutrophils

Figure 1 – The haematopoietic hierarchy over time – Ever since the discovery of the long-term (LT) *in vivo* repopulating ability of a rare bone marrow resident cell, the haematopoietic stem cell (HSC), it has been placed at the top of the blood cell differentiation hierarchy. The hierarchical models **(A-D)** are visualised as tree-like structures with the most potent HSC at the trunk, and terminally differentiated/mature cell lineages at the branched ends. Our views on the number of possible differentiation trajectories and intermediate progenitor stages required to generate specific haematopoietic cell lineages have evolved as technologies for exploration have become more complex. **A)** Hierarchy around the year 2000. All cell populations at transition points appear homogeneous and the main bifurcation is from short-term (ST)-HSC into a common myeloid (CMP) or lymphoid progenitor (CLP) (based on Akashi et al. (2000); Morrison and Weissman (1994)). MEP=megakaryocyte-erythrocyte progenitors; GMP=granulocyte-monocyte progenitors; DCs=dendritic cells; NK=natural killer cells; ILCs=innate lymphoid cells. **B)** Between 2005 and 2015 the model changed and heterogeneity within the HSC pool was incorporated. It would now additionally contain an intermediate-term (IT)-HSC and several multipotent progenitors (MPP). Within the HSC pool (boxed area), cells have different self-renewal (vertical axis) and differentiation (horizontal axis) potentials. The hierarchy now also includes the lymphoid-primed multipotent progenitor (LMPP) and the GMP population shows more heterogeneity/overlap with eosinophil-basophil progenitors (EoBP). Changes to hierarchy supported by Adolfsson et al. (2005); Benveniste et al. (2010); Cabezas-Wallscheid et al. (2014); Doulatov et al. (2010); Pietras et al. (2015); Sanjuan-Pla et al. (2013), amongst others. **C)** From 2016 onwards, advances in single-cell transcriptomics led to a haematopoietic tree that depicts differentiation more like a continuum. Instead of a steady and highly grouped hierarchy with fixed intermediate steps, the differentiation appears much more gradual from one population to the next. Hierarchy supported by Grun et al. (2016); Laurenti and Gottgens (2018); Notta et al. (2016); Paul et al. (2015); Velten et al. (2017), amongst others. **D)** A current more realistic haematopoietic 'tree'

considers the many different possible differentiation trajectories, avoiding over-simplification and an underestimation of potential differentiation routes. Continuous lines indicate differentiation trajectories of single HSCs or MPPs. All trajectories pass through phenotypically defined progenitor compartments (boxed areas). Horizontal lines indicate the potency of different cells/clones within each compartment. Single-colour=unilineage; two-colour=bilineage; three-colour=trilineage; black=multipotent. Most compartments have more uni- than bi- or trilineage cells. **(A-D)** Reprinted from Laurenti and Gottgens (2018), with permission from Springer Nature.

Ontogeny of the haematopoietic system – how the embryo ensures a lifelong supply of blood

In adult mammals, HSCs reside in the BM. This is also where adult haematopoiesis occurs. The BM is not the site in which these powerful HSCs are first *de novo* generated, nor is it the location of the earliest haematopoietic activity in mammals. However, both find their origins in the developing embryo, long before any BM has been established.

Developmentally, the earliest blood formation in vertebrates takes place in the extra-embryonic yolk sac (YS). A little over a century ago, so-called 'blood islands' (BI) were first carefully described as masses of cells which produce haemoglobin and become red blood-corpuscles in the YS of early chick embryos (Sabin, 1920). These red cells would later be termed 'primitive erythrocytes' as they will not reach a final adult enucleated stage. Though the earliest mention of such a model goes back to 1875 (Ranvier, 1875), it was in the chick embryo that the simultaneous emergence of haematopoietic and endothelial cells led P.D.F. Murray to coin the term 'haemangioblast' for a common mesodermal precursor to both cell types (Murray and Wilson (1932); reviewed in Cumano and Godin (2007)). It would take until the end of the 20th and beginning of the 21st century before a functionally similar mammalian cell was identified and could be produced *in vitro* from mouse embryonic stem cells (mESC) (Choi et al., 1998; Fehling et al., 2003). Based on cell surface expression and functional activity a haemangioblast-like cell was later localised *in vivo* to the posterior primitive streak of the embryonic day (E)7.5 gastrulating mouse embryo. It co-expresses the mesodermal marker *brachyury* (*T*) and *Flk1* (*Kdr*, also known as *VEGFR-2*), a marker for lateral plate mesoderm and the earliest differentiation of endothelial and haematopoietic cells (Huber et al., 2004). There haemangioblasts migrate to the YS, where they will rapidly undergo differentiation into more restricted progenitors.

Primitive haematopoiesis

Mesodermal differentiation in the mouse YS observed at E7.25 leads to the formation of haemangioblastic chords, from which primitive-erythrocyte-filled YS BI vessels will form at E7.5 (Figure 2A) (reviewed in Cumano and Godin (2007); McGrath and Palis (2005)). YS BIs appear polyclonal, containing the bipotent haemangioblast, a number of unipotent endothelial and haematopoietic progenitors (Ueno and Weissman, 2006), and other mesodermal precursors initially differentiating into haemogenic endothelium before giving rise to primitive blood (Padron-Barthe et al., 2014; Stefanska et al., 2017). Together, these earliest blood cells produced in the YS BIs form what is called the primitive haematopoietic 'wave' (Figure 2B,C). This transient wave contains primitive (nucleated) erythrocytes expressing embryonic globins, and primitive macrophages and megakaryocytes (Bertrand et al., 2005; Ferkowicz et al., 2003; Palis et al., 1999; Tober et al., 2007). The human YS has been shown to contain similar primitive blood cells around 2-3 weeks post-conception (reviewed in Canu and Ruhrberg (2021)). While the exact role of the early megakaryocytes remains to be determined, primitive erythrocytes are required for oxygen transport as the developing embryo becomes too big to rely on the diffusion of oxygen. Primitive macrophages, that have not gone through a monocyte intermediate like their adult HSC-derived counterparts, play vital roles in tissue remodelling and lymphatic development (reviewed in Dzierzak and Bigas (2018); Kauts et al. (2016)). It is now also well established that some of these YS-derived macrophages persist in the brain until adulthood as microglia (Ginhoux et al., 2010).

Pro-definitive haematopoiesis

Erythroid-myeloid progenitors

Slightly after haemangioblasts move to the YS, mesoderm from the posterior primitive streak migrates to form the extra-embryonic allantois (together with the chorion the prospective placenta) (Figure 2B,C). Thereafter primitive streak mesoderm migrates to what will become the embryo proper, where it forms the lateral plate (blood-forming) mesoderm and other derivatives (reviewed in Cumano and Godin (2007); Dzierzak and Speck (2008)). These mesodermal movements will, roughly one day after the start of the primitive wave, give rise to a second transient 'pro-definitive' haematopoietic wave (Figure 2B,C) that commences just before the onset of circulation at E8.25-8.5. It generates various unipotent (definitive) erythroid and multipotent myeloid (megakaryocytic, mast cell, granulocyte-macrophage) progenitors in the YS (Bertrand et al., 2005; Ferkowicz et al., 2003; McGrath et al., 2015; Palis et al., 1999; Tober et al., 2007). Together, these progenitors are now broadly labelled as EMPs, short for 'erythroid-myeloid progenitors' (Figure 2B,C). They are phenotypically distinct from primitive haematopoietic cells and HSCs by their expression of CD16/32 (*Fcgr3*/*Fcgr2b*) and lack expression of the HSC marker Sca-1 on their surface (McGrath et al., 2015). They are found in increasing numbers in the bloodstream both intra- and extra-embryonically at around E9.5 (reviewed in McGrath and Palis (2005)). These YS-derived EMPs maintain erythropoiesis throughout mouse embryonic development whereas HSCs only start to substantially contribute to the erythroid lineage from around birth (Soares-da-Silva et al., 2021). The pre-circulation allantois and chorion, and later placenta (from E9 onwards) also contain many EMPs and have a strong myeloid potential (reviewed Dzierzak and Speck (2008)).

Lymphoid-primed multipotent progenitors

In addition to the production of EMPs, it is from E9.0-9.5, in the later stage of the second wave, that the YS starts producing various lymphoid-primed multipotent progenitors (LMPPs) that have the potential to generate natural killer (NK), T- and a special subset of B-cells (Boiers et al. (2013); reviewed in Lange et al. (2021); Yamane (2018)) (Figure 2B). These special so-called B-1a (previously 'Ly-1', or $CD5^{+}$) B-cells will eventually reside in the adult peritoneal and pleural cavities and are of particular interest as they are lifelong self-replenishing. In contrast to NK, T- and conventional (B-2) B-cells, they cannot be replaced by HSCs (Kobayashi et al., 2014; Lalor et al., 1989; Yoshimoto et al., 2011) or upon transplantation of adult mouse BM (Ghosn et al., 2012). Later studies leave open the debate whether foetal liver (FL) HSCs could give rise to B-1a cells (Ghosn et al., 2016; Kristiansen et al., 2016). Nonetheless, whether generated HSC-independently or -dependently, it is clear B-1a cells and/or their progenitors find their origin in the embryo.

Multipotent haematopoietic progenitors

The autonomous production of T-cell progenitors by the YS remains controversial as well. Even though YS-derived LMPPs capable of producing these cells have been found (Boiers et al., 2013; Yoshimoto et al., 2012), more recent conflicting results show that all T-cell progenitors are intra-embryonically- (and HSC-)derived (Elsaid et al., 2021). Finally, an additional multipotent YS-derived lymphoid-myeloid progenitor (generated alongside EMPs and LMPPs) has been demonstrated. It reveals HSC function only when transplanted neonatally (Yoder et al., 1997a; Yoder et al., 1997b) or after *in vitro* culturing (Inlay et al., 2014) (Figure 2B). Taken together, the earliest generation of blood cells during embryonic development occurs in the YS with the production of various primitive blood cell types, which is later followed by the sequential (and partly overlapping) generation of EMPs, LMPPs and neonatal/*in vitro*-matured HSCs.

Simultaneous to the YS, the pro-definitive wave also takes place intra-embryonically. The lateral plate mesoderm in the trunk/caudal region of the mouse embryo will give rise to the para-aortic splanchnopleura (pSp; E8.5-10), the earliest intra-embryonic haemogenic site (reviewed in Cumano and Godin (2007)) (Figure 2C, 3). The pSp initially contains the paired dorsal aortae, which are connected to the YS via the vitelline or omphalomesenteric artery (VA) (Figure 2B, 3Ai,E) (Garcia-Porrero et al., 1995; Kaufman, 1992). From E9 the aortae will start to fuse into a single midline dorsal aorta, that will simultaneously become connected to the placenta via the umbilical artery (UA) (Figure 2B, 3E) (Garcia-Porrero et al., 1995; Kaufman, 1992). From about E10-11.5, when the dorsal aorta, gonads and mesonephros have developed within the pSp, this region roughly located between fore- and hindlimbs is known as the AGM, short for 'aorta-gonad-mesonephros' (Figure 3Aii,Aiv,B-D).

The presence of circulation makes it difficult to trace the origins of pro-definitive haematopoietic cells generated from the pSp stage onwards. To overcome this, the E7.5-8 Sp (predecessor of the pSp and prior to the onset of circulation and establishment of the VA) was isolated and taken for *ex vivo* explant culture. This revealed the potential of the Sp to not only *de novo* generate EMPs, like the early pre-circulation YS, but also lymphoid progenitors (Cumano et al., 1996), and even multipotent long-term (low-level) repopulating progenitors (Cumano et al., 2001). Whereas the YS has been shown to produce only B-1a cells, the E8.5-11.5 pSp/AGM contains LMPPs and/or other precursors capable of generating both B-1a and B-2 B-cells autonomously, even though the preferred route of differentiation appears to be towards B-1a cells (Godin et al., 1993; Hadland et al., 2017; Kobayashi et al., 2014; Yoshimoto et al., 2011). From around E8.5-9, the pSp region also contains LMPPs with the potential to form T-cells (Yokota et al., 2006). *In vivo*, the colonization of the thymus by lymphoid progenitors, either fully or partially intra-embryonically derived, starts from around E11 (Figure 2C) (Yokota et al., 2006). Actual T-cell differentiation only starts around E13, with

the first mature T-cells being generated by the thymus just before birth (reviewed in Rothenberg (2021)).

Similar to the YS, the E9-11.5 pSp/AGM also contains a lymphoid-myeloid progenitor that can display HSC function upon *in vitro* culturing (Ganuza et al., 2017; Inlay et al., 2014; Rybtsov et al., 2016) or transplantation into neonates but not adult mice (Boisset et al., 2015; Yoder et al., 1997a) (Figure 2B). Some of these cells could represent predecessors to HSCs ("pre-HSCs") (and/or HPCs ["pre-HSPC"]), or cells in a 'neutral'/plastic state awaiting establishment of another haematopoietic identity. The display of HSC function could be due to *in vitro* culture or embryonic/foetal conditions not necessarily reflective of the *in vivo* adult environment.

Lastly, the pro-definitive wave includes another multipotent (mainly erythromyeloid) progenitor cell, the previously described CFU-S (Figure 2B). E9.5 is the earliest timepoint at which both the YS and embryo proper contain CFU-S activity detectable in the spleen at 8 days post-transplantation. When E10-10.5 YS and AGM are tested for the more primitive/immature CFU-S that can produce spleen colonies 11 days post-transplantation, the AGM was found to vastly exceed the YS in terms of CFU-S progenitors, containing a 20 to 200 times higher frequency, that peaks before any CFU-S can be detected in the FL (Medvinsky et al., 1993). Taken together, from E8.25 to about E10 of mouse development, the pro-definitive wave produces a variety of progenitors of increasing potency at intra- and extra-embryonic anatomical sites (Figure 2B,C). Notably, up until this timepoint all haematopoietic waves and progenitors precede the appearance of HSCs and are thus produced HSC-independently.

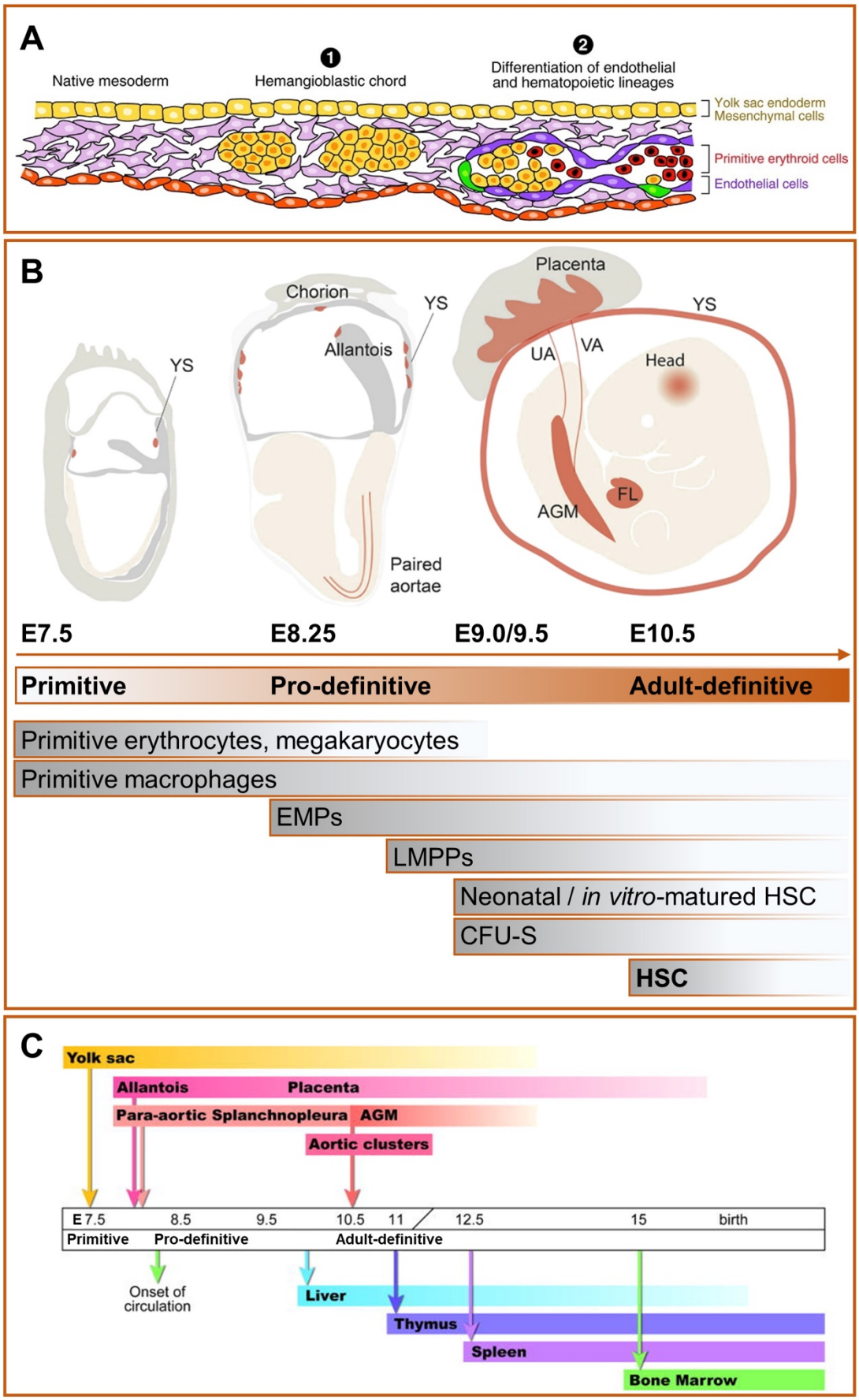

A
Native mesoderm
1
Hemangioblastic chord
2
Differentiation of endothelial and hematopoietic lineages
Yolk sac endoderm
Mesenchymal cells
Primitive erythroid cells
Endothelial cells
B
YS
Chorion
YS
Allantois
Placenta
YS
UA
VA
Head
AGM
FL
Paired aortae
E7.5
E8.25
E9.0/9.5
E10.5
Primitive
Pro-definitive
Adult-definitive
Primitive erythrocytes, megakaryocytes
Primitive macrophages
EMPs
LMPPs
Neonatal / in vitro-matured HSC
CFU-S
HSC
C
Yolk sac
Allantois
Placenta
Para-aortic Splanchnopleura
AGM
Aortic clusters
E7.5
8.5
9.5
10.5
11
12.5
15
birth
Primitive
Pro-definitive
Adult-definitive
Onset of circulation
Liver
Thymus
Spleen
Bone Marrow

Figure 2 – Mouse embryonic haematopoiesis – Mouse embryonic haematopoiesis occurs in several 'waves' and in various organs and tissues. **A)** Schematic representation of yolk sac (YS) blood island (BI) formation. The earliest haematopoietic activity during mouse embryogenesis is observed in the embryonic day (E)7.25 extra-embryonic YS. Mesodermal (native or haemangioblast-derived) differentiation leads to the formation of haemangioblastic chords (1), from which primitive-erythrocyte-filled YS BIs will form around E7.5 (2). Blood formation in the early YS is most-likely reliant on a combination of bipotential haemangioblasts that migrate to the YS from the primitive streak, many unipotent endothelial and haematopoietic progenitors, and haemogenic endothelial cells (green) undergoing transdifferentiation. Reprinted from Cumano and Godin (2007) with permission from Annual Reviews, Inc. **B)** Embryonic haematopoiesis occurs in three waves. The first and 'primitive' wave starts at E7.5 in the YS BIs with the formation of primitive erythrocytes, megakaryocytes and macrophages. Whereas the production of the first two is transient, some of the primitive macrophages persist until adulthood. The second and pro-definitive wave starts just before the onset of circulation at E8.25 with the production of erythroid-myeloid progenitors (EMPs) from the YS. From around E9.0/9.5 these will also be found in increasing numbers in the bloodstream, intra-embryonically and in the chorion, allantois and later the placenta. From around E8.5/9.0, lymphoid potential from lymphoid-primed multipotent progenitors (LMPPs) is detected, slightly earlier intra-embryonically than it is extra-embryonically. At approximately E9.0/9.5, a multipotent myeloid-lymphoid progenitor that displays HSC function only upon *in vitro* culturing or transplantation into neonates, but not adult mice, also emerges both intra- and extra-embryonically. The so-called spleen colony-forming unit (CFU-S) starts appearing at the same time both intra- and extra-embryonically, with the highest activity contained in the E10.0/10.5 aorta-gonad-mesonephros (AGM) region. Finally, at E10.5 the first long-term multilineage adult-repopulation haematopoietic stem cells (HSC) are generated in the intra-embryonic AGM region during the third/adult-definitive wave. Slightly later,

HSCs are also found in the umbilical (UA) and vitelline artery (VA), placenta, YS, and head. They will next migrate to and expand in the foetal liver (FL), together with EMPs. The word ‘definitive’ (third wave) refers both to the lack of primitive erythroid cells and the contributions of this wave to the adult haematopoietic system (reviewed in Dzierzak and Bigas (2018); McGrath and Palis (2005)). Schematic drawings show E7.5, E8.25 and E10.5 embryos with sites of haematopoietic activity coloured in red. Drawings of embryos reprinted from Kauts et al. (2016), with permission from John Wiley and Sons. **C)** Schematic timeline of haematopoietic events at the different sites during mouse embryonic development. Arrows above timeline indicate initiation of haematopoietic cell generation and/or appearance at specific site. Arrows below indicate first time point of these tissues being colonised by haematopoietic cells. Adapted from Dzierzak and Speck (2008).

Adult-definitive haematopoiesis – The HSC

Given the earliest haematopoietic activity in vertebrates occurs in the YS, it was long thought this was the anatomical site where HSCs would first emerge and initiate haematopoiesis. Around 50 years ago, several lines of research supported the notion that HSCs would originate in the mammalian YS, then migrate to the FL, the main haematopoietic organ during development, followed by the thymus, spleen and just before birth the BM (reviewed in Cumano and Godin (2007); Dzierzak and Speck (2008)). Support for this hypothesis came from Moore & Metcalf who, just after developing the *in vitro* CFU-C assay on adult mouse BM, decided to examine the ontogeny of the haematopoietic system in the E7 YS. They indeed showed that the mouse YS is the earliest tissue containing progenitors capable of producing multilineage CFU-Cs (Moore and Metcalf, 1970). Parabiosis (Moore and Owen, 1965) and irradiation (Moore and Owen, 1967) studies on chick embryos already led Moore to conclude that all secondary haematopoietic organs (spleen, BM, thymus, bursa of Fabricius) were dependent on the colonisation by ('blood-borne') HSCs migrated from elsewhere. Altogether, this led them and others to conclude vertebrate HSCs must be of YS origin.

A dogma shift began with the construction of avian inter- and intraspecies chimaeras (chick-chick, chick-quail) in which extra- and intraembryonic (YS and embryo proper) tissues were grafted early in development, before any vascular connections between the two had been established. The species, allogeneic or sex differences allowed for cell fate analyses which would reveal what part/tissue contributes to the adult haematopoietic system and is therefore the site of origin and source of adult blood. This seminal work by Dieterlen-Lièvre and co-workers showed that it is the intraembryonic (destined to give rise to embryo, amnion and allantois) area and not the YS that contains the cells capable of sustaining lifelong haematopoiesis (Dieterlen-Lievre, 1975; Lassila et al., 1978). Later experiments performed in amphibians, namely frog (*Rana*) (Turpen et al.,

1981) and toad (*Xenopus laevis*) (Ciau-Uitz et al., 2000) embryos yielded similar results.

To perform such grafting experiments in mice (mammals) at the time of intra-embryonic blood development would be virtually impossible, given their *in utero* development. The big difference in CFU-S activity contained in the murine YS and AGM did already suggest that the most potent blood-forming cells originate within the intra-embryonic AGM region, as is the case for avian and amphibian species. Definitive proof of the mouse AGM, and not the YS, containing the first long-term self-renewing, multilineage HSCs capable of serially repopulating lethally irradiated recipients came from Muller et al. (1994). By using similarly aged embryos carrying a human B-globin transgene and testing for this and the presence of the Y chromosome marker, donor cells were genetically marked, allowing for the fate of these cells to be tracked in transplant recipients. Different cell fractions from E8-11 embryos were transplanted, but it was only with the E10.5 (34-41 somite pairs, sp) AGM (and not YS, FL, body remnant or circulating cells), that 3 out of 96 primary recipients that had all received 1.16 embryo equivalent (e.e.) of AGM cells showed high-level (>10%) donor chimaerism at >4 months post-transplantation. The same was seen when BM cells from primary repopulated recipients were injected in secondary and tertiary recipients. At E11 (>42sp), the highest HSC activity was found in the AGM, but HSCs were also found in the FL, YS and embryo body remnants, suggesting these had originated in and emigrated from the AGM. Human haematopoietic development seems to occur in a similar spatiotemporal fashion, with a primitive wave that is initiated in the YS, and the first human HSCs generated in the AGM at around 4-5 weeks post-conception (reviewed in Canu and Ruhrberg (2021)).

The discovery of the first long-term high-level self-renewing HSCs in the E10.5 AGM marked the initiation of the third and final adult-definitive wave (Figure 2B,C). The existence of blood flow however allowed for the possibility of pre-HSCs or not fully matured progenitors having migrated from e.g. YS to the AGM, where they would have then matured and/or expanded. Using a novel explant organ culture technique, entire E9-11 AGM, FL and YS

tissues (of male embryos) were cultured *ex vivo* at the air-liquid interface for 2-3 days (Medvinsky and Dzierzak, 1996). By next transplanting these tissues, it was revealed that only the E10/10.5 (32-35sp) AGM region could autonomously generate long-term (>8 months) multilineage HSCs. Furthermore, the *ex vivo* culture system allowed for the number of HSCs (and CFU-S) in the E10-11 AGM to greatly increase over 'fresh' directly assayed AGM (about 15x more high-level repopulated recipients). This could be explained either by increased maturation of HSC precursors or clonal expansion of a few HSCs present at the time of isolation. On the other hand, only minor increases of CFU-S were seen in the other tissues tested. Given that these normally only contain HSCs at later stages, it suggests the AGM colonises these tissues *in vivo* and this *ex vivo* culture system renders HSCs unable to migrate to secondary tissues, hence their accumulation and increase within the AGM. All of this again strongly suggested the source of the adult definitive haematopoietic system is found in the AGM region (Figure 3Aii,Aiv,B-D).

Together, Muller et al. (1994) and Medvinsky and Dzierzak (1996) firmly establish the vital role of the AGM within the adult definitive haematopoietic wave. Subdissection of the E11 AGM into the dorsal aorta and the urogenital ridges (Figure 3Aiv,Bi,D), demonstrated it is the aorta with its surrounding mesenchyme that first contains HSCs (de Bruijn et al., 2000b). All progenitors within the AGM are also confined to the aorta (Godin et al., 1999). Furthermore, the region appears predominantly supportive of haematopoietic stem and progenitor cell (HSPC) generation, rather than haematopoietic differentiation, as shown by a lack of unipotent and an abundance of multipotent progenitors (Godin et al., 1999). In addition to the aorta, both the freshly isolated E11 VA and UA (Figure 2B) also contain HSCs (de Bruijn et al., 2000b). Overall, these data show that it is the major arteries of the midgestation embryo that are the most potent source of adult HSCs and multipotent progenitors.

To further spatially localise the HSCs, a novel transgenic reporter mouse line expressing GFP under the control of the transcriptional regulatory

elements of the well-known BM HSC marker *Ly6A* (Sca-1) was generated (Ma et al., 2002). Whereas adult mouse BM HSCs are both Ly6A-GFP$^+$ and Sca-1$^+$ (Ma et al., 2002; Spangrude et al., 1988), the embryo contains Sca-1$^+$ and Sca-1$^-$ HSCs (de Bruijn et al., 2002). Ly6A-GFP expression however succeeds in marking all functional HSCs in the embryonic aorta-mesenchyme, possibly due to limited Sca-1 surface protein on the first HSCs not detectable by antibody staining, GFP being cytoplasmic and thus not requiring any processing towards the cell surface, and/or the reporter producing a bright fluorescent signal due to eight transgene copies instead of the normal diploid copy number of the endogenous gene (de Bruijn et al., 2002). Following functional assays, immunohistological analyses revealed GFP-expressing cells only within the endothelial lining, and not in the subaortic mesenchyme (de Bruijn et al., 2002). Ly6A-GFP expression in the aorta, VA and UA gradually increased from E9.5 onwards, with expression in the aorta starting on the ventral side. Notably, GFP$^+$ cells were only found in the YS from late E11 (48 sp) onwards. Using a transgenic reporter mouse to visualise *Runx1* expression, a pivotal transcription factor (TF) for the generation of the first HSCs, the spatiotemporal localisation of the first HSCs was confirmed (North et al., 2002). Despite *Runx1* being more widely expressed than Ly6A-GFP in the ventral aspect of the aorta (in both endothelium and mesenchyme), functional HSCs were again only found amongst the cells of the endothelial lining and not the underlying mesenchyme when analysed in wild-type embryos (North et al., 2002). The importance of Runx1 in definitive haematopoiesis is underlined by the fact that upon homozygous germline deletion, primitive haematopoiesis and the embryo's vasculature remain relatively normal, but functional HSCs are not generated and embryonic lethality with/due to severe anaemia occurs at E12.5 (reviewed in Dzierzak and Bigas (2018)).

Functional haematopoietic assays on subdissected ventral and dorsal aspects of live E10.5/11.5 aortae confirmed it is indeed the ventral part of the aorta that predominantly contains HSCs, and also that it is only this part that can autonomously initiate and expand HSC generation upon *ex vivo* explant

culture (Taoudi and Medvinsky, 2007). HPCs/CFU-Cs were found in both parts. When subdissected in three equal parts along the rostral-caudal axis, it was later shown that it is the middle segment of the AGM, near the junction with the VA, that contains all HSCs at E11 (Mascarenhas et al., 2009).

Altogether, these studies provide strong evidence that the first definitive HSCs are found and autonomously generated within the ventral endothelial lining of the middle segment of the midgestation mouse dorsal aorta (de Bruijn et al., 2002; Mascarenhas et al., 2009; North et al., 2002; Taoudi and Medvinsky, 2007).

The number of HSCs generated in the AGM region is very low. Quantitative limiting dilution transplantations revealed that the E10.5/11 AGM contains on average one or two HSCs, which could be increased to about twelve upon *ex vivo* explant culture (Kumaravelu et al., 2002). By E12 the HSC activity within the AGM decreases and the FL has taken over as the main haematopoietic organ.

The murine FL however is haemogenically inactive, i.e. does not generate HSPCs *de novo*, but is instead colonised by HPCs (EMPs) generated in the YS and AGM from late E9 (Houssaint, 1981; Johnson and Moore, 1975) (Figure 2C). From E11 the mouse FL begins to be colonised by AGM-derived HSCs (Medvinsky and Dzierzak, 1996; Muller et al., 1994). From that developmental timepoint onwards, as the haematopoietic activity within the AGM gradually decreases, HSCs can also be found in the mouse placenta (Gekas et al., 2005; Ottersbach and Dzierzak, 2005), YS (Kumaravelu et al., 2002; Medvinsky and Dzierzak, 1996; Moore and Metcalf, 1970; Muller et al., 1994) and head (Li et al., 2012). Given the FL will eventually contain more HSCs than can be generated in this short window of time by the AGM alone, it is highly likely these other HSC-containing/ generating tissues also contribute to the large FL HSC pool, either by facilitating expansion or indeed *de novo* generation. The developing FL itself similarly allows for great expansion of incoming HSCs (Ema and Nakauchi, 2000). The spleen will only be colonised by haematopoietic cells from E12.5

onwards and will contain AGM-derived HSCs from E14.5 (Godin et al., 1999) (Figure 2C).

All in all, both by HSC-independently producing HPCs that give rise to long-lived haematopoietic cells and by *de novo* generating the first adult-definitive HSCs that will ultimately serve as the source of adult blood, the embryo plays an absolutely vital role in ensuring the lifelong supply of blood for all vertebrates.

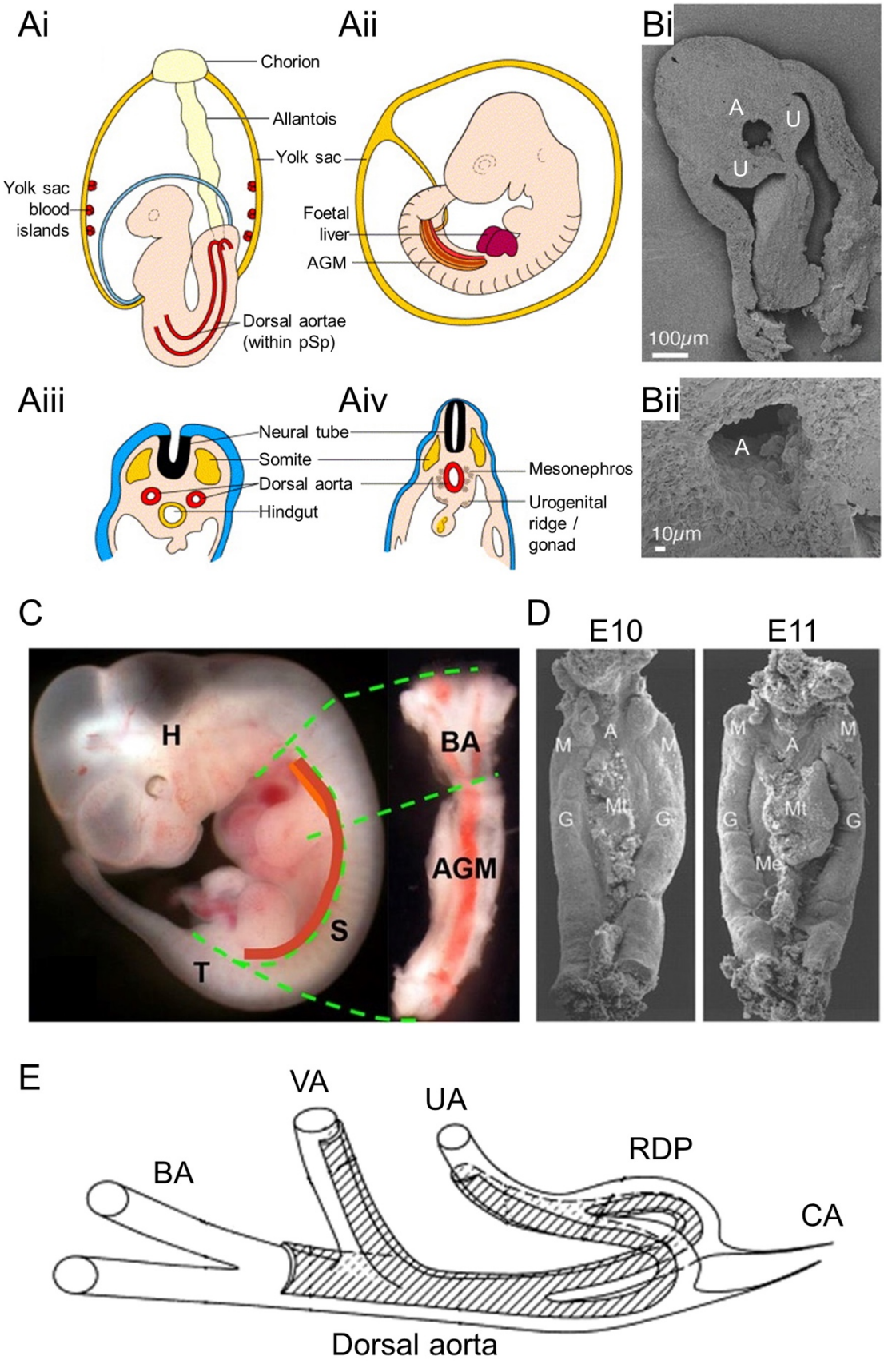

Ai
Aii
Bi
Chorion
Allantois
Yolk sac
Yolk sac blood islands
Foetal liver
AGM
Dorsal aortae (within pSp)
A
U
U
100μm
Aiii
Aiv
Bii
Neural tube
Somite
Dorsal aorta
Hindgut
Mesonephros
Urogenital ridge / gonad
A
10μm
C
D
E10
E11
H
BA
AGM
S
T
M
A
M
G
Mt
G
M
A
M
G
Mt
G
Me
E
VA
UA
RDP
BA
CA
Dorsal aorta

Figure 3 – Anatomy of the mouse embryonic aorta and AGM region – A) schematic representation of **(Ai)** embryonic day (E)8.5/9.0 and **(Aii)** E10.5/11.0 embryo and transverse sections through the **(Aiii)** para-aortic splanchnopleura (pSp) and **(Aiv)** aorta-gonad-mesonephros (AGM) region. Adapted from Dzierzak et al. (1998). **Bi)** Scanning electron microscopy (SEM) image of thick transverse section through the E10 AGM region. A=just outside dorsal aorta; U=urogenital ridge. Top=dorsal; bottom=ventral. **Bii)** zoom-in of **Bi**. A=inside aortic lumen. Adapted from Boisset et al. (2015). **C)** Bright-field image of an unfixed/live E11.5 mouse embryo (left). Thick orange line shows position of dorsal aorta (AGM), including bifurcated rostral part (BA). A dissected AGM (+ BA) region with the ventral side facing up (right). Red line in the centre is blood-filled aorta. H=head; T=tail; S=somites. Reprinted from Rybtsov et al. (2016). **D)** SEM images of a dissected E10 (left) and E11 (right) AGM region. A=dorsal aorta; G=urogenital ridge/gonad; M=pro/mesonephros; Mt=mesentery; Me=mesenchyme. Original magnification=40x. Reprinted from de Bruijn et al. (2000a), with permission from the American Society of Hematology (Elsevier Science & Technology Journals). **E)** Schematic of mouse embryonic dorsal aorta and proximal part of other main arteries. VA=vitelline (also: omphalomesenteric) artery; UA=umbilical artery; CA=caudal arteries; RDP=recurved distal part of dorsal aorta. Adapted from Garcia-Porrero et al. (1998).

The biomechanics of HSC generation

Early observations of intra-aortic clusters

De novo haematopoietic cell generation, i.e. the transition or differentiation from a non-haematopoietic precursor to a haematopoietic cell fate, has been a topic of great interest for many years. Over a century ago, before any functional assays could be performed on embryonic aorta, observational histological studies on early stage embryos of many different vertebrate species (human, mouse, rat, pig, bat, cat, rabbit, mongoose, turtle, and chick, amongst others) revealed the existence of cluster cells tightly associated to the ventral aortic endothelium, frequently near the mouths of vascular branches (Dantschakoff (1907); Emmel (1916); Jordan (1916); Jordan (1917); Maximow (1909); Sabin (1920); reviewed in Adamo and Garcia-Cardena (2012)) (Figure 4Ai,Aii). Given their common appearance during a short period of vertebrate development, the biomechanical process giving rise to these intra-aortic haematopoietic clusters (IAHCs) appears highly conserved (reviewed in Ottersbach (2019)). The IAHCs' fairly regular and gradually increasing spheroidal shape, indicative of intrinsic 'centrifugal growth', and localisation on the ventral aortic wall and nowhere else, implied a unique haemogenic capacity of the subjacent endothelial cells (Figure 4Aii), and made incidental accretions and accumulations of cluster cells from the circulating blood unlikely (Emmel, 1916; Jordan, 1916; Jordan, 1917; Sabin, 1920). Similar intra-arterial masses/clusters were observed in the UA, VA (Figure 4Fiii) and other temporary ventral aortic branches ('rami'), leaving several of these studies to postulate on a causal relationship between IAHC formation and vascular remodelling/atrophy (Emmel, 1916; Jordan, 1916; Jordan, 1917). Indeed, the drastic remodelling of the VA (first being connected to the aorta via the UA, then via a plexus of small ventral aortic branches, to forming a singular arterial connection) was later directly shown to have a strong effect on the intra- and even extravascular cluster/island formation and blood-forming capacity of this artery (Zovein et al., 2010). The

UA was also shown to promote both vascular and blood formation by producing intra-arterial and extravascular clusters/islands, the latter derived from the first (Yzaguirre and Speck, 2016).

The early 20th century observations (Figure 4A), combined with later functional proof of the midgestational arteries' definitive haematopoietic potency (Ciau-Uitz et al., 2000; de Bruijn et al., 2000b; Dieterlen-Lievre, 1975; Lassila et al., 1978; Medvinsky and Dzierzak, 1996; Turpen et al., 1981), strongly suggested a correlation between IAHC formation and HSPC generation. Furthermore, these studies shed light on the unique spatiotemporal relationship that exists between the haematopoietic and (cardio)vascular systems (reviewed in Adamo and Garcia-Cardena (2012); Canu and Ruhrberg (2021)).

The endothelial origin of blood

Like the discovery of the intra-embryonic source of HSCs, the first experimental evidence of endothelial cells capable of changing morphology into IAHCs and producing blood came from the avian model (Jaffredo et al., 1998). Earlier work had proposed the ventral aortic endothelial wall of the chick embryo as the most likely source of the definitive blood system (Dieterlen-Lievre and Martin, 1981). However, it was only by specifically dye-labelling of endothelial cells in chick embryos prior to the onset of haematopoiesis and IAHC formation, that Jaffredo et al. (1998) could show the dye was directly conveyed to newly formed IAHCs within 24 hours. Together with immunohistological analyses showing the (dis)appearance of endothelial and haematopoietic surface markers they would corroborate the endothelial origin of IAHCs in chick. The endothelial cells endowed with haemogenic capacity would be called 'haemogenic endothelium' (HE) (Figure 4Aii,Bi,C), a term used earlier in a study on IAHCs in Mongolian gerbil, pig, mouse and human embryos, where transmission electron microscopy for the first time showed the ultrastructural similarities and tight junctions between

IAHCs and the underlying HE, as well as the tight junctions amongst IAHC cells (Smith and Glomski, 1982) (Figure 4Bi,Bii).

Roughly one decade later, Zovein et al. (2008) would produce the first evidence supporting an endothelial origin of the murine definitive haematopoietic system. By crossing a mouse strain with tamoxifen-inducible Cre-recombinase under the control of the promoter of VE-Cadherin (Vascular Endothelial Cadherin, *Cdh5*; also known as CD144) to a Rosa26 lox-β-galactosidase (LacZ) reporter line, and triggering induction at E9.5, Zovein et al. (2008) could specifically label AGM and not YS endothelium during the time of HSC emergence, and allow for lineage tracing of these cells. The labelled VE-Cadherin-expressing cells gave rise to IAHCs, and offspring was found to colonise the E14.5 FL and subsequently the adult spleen, thymus and BM, where the labelled cells co-expressed HSC and mature haematopoietic cell surface markers (Zovein et al., 2008). Using another Cre line specifically labelling the subaortic mesenchyme, they demonstrated that these cells did not contribute to adult blood. Together, these results showed the ability of murine midgestation aortic endothelial cells, and not underlying mesenchyme, to give rise to HSCs that contribute to the definitive haematopoietic system long-term.

Soon after, the same volume of *Nature* saw the publication of three studies focusing on the same process *in vivo* (Chen et al., 2009) and *in vitro* (Eilken et al., 2009; Lancrin et al., 2009). Chen et al. (2009) used mice with a similar Cre/lox-based system as Zovein et al. (2008) to delete *Runx1* in endothelial (*VE-Cadherin*-expressing) cells during development. They demonstrated both IAHC formation and HSPC generation are abolished upon deletion, and thus confirm both processes have an endothelial origin. Indeed, they show 95% of adult blood is derived from cells that at some point in their life expressed *VE-Cadherin* and therefore that most blood derives from embryonic HE (Chen et al., 2009). Furthermore, they establish Runx1's vital role in establishing haematopoietic fate and, by use of the haematopoietic-specific *Vav-Cre* strain (Stadtfeld and Graf, 2005), how Runx1 is not required in blood cells once they have been generated.

Technological advances in microscopy now also allowed for real-time continuous recording of this process on a single-cell level *in vitro* (reviewed in Schroeder (2008)). Live-imaging was performed using both mESC-derived and freshly isolated primary mesodermal cells as starting point (Eilken et al., 2009). These were differentiated into endothelial cells (ECs)/HE and subsequently blood cells. The entire one-week differentiation, from mesoderm to blood, was imaged continuously. Initially a gradual acquisition of endothelial characteristics (VE-Cadherin expression, functional tight junction formation, morphology, endothelial-specific dye uptake) was observed. Cells then lost these properties and underwent morphological changes into rounded non-adherent cells, started presenting haematopoietic surface markers (first CD41, then CD45 and CD11b) and became highly proliferative in suspension. Some of the non-adherent cells generated were shown to have multilineage CFU-C potential. Haemogenic endothelial cells (HECs) were found to be rare, transiently produced and present at only 1 in 1000 ECs.

Lancrin et al. (2009) came to a similar conclusion. Also using a mESC-derived differentiation model, they would show blood formation is a sequential process, and that differentiation from a bipotential haemangioblast, via an intermediate HE stage, leads to the formation of non-adherent phenotypic blood cells. Additionally, their knockdown experiments showed that the TFs *Scl* (*Tal1*) and *Runx1* are required for the transition from a haemangioblast to HE and HE to definitive haematopoietic cells, respectively. Altogether, these three papers strengthen the proof for existence of HE and show that blood can be formed from (*in vitro*-derived) endothelium as shown by live-imaging and functional *in vivo* experiments.

The advances in time-lapse confocal imaging would ultimately lead to the publication of three elegant studies (simultaneously in *Nature*) visualising the *in vivo* live formation of single HSPCs from the aortic endothelial wall of mouse (Boisset et al., 2010) and zebrafish (*Danio rerio*) (Bertrand et al., 2010; Kissa and Herbomel, 2010) embryos.

Given the location of the mouse aorta deep within the opaque embryo, thick transverse sections were cut from live E10.5 *Ly6A-GFP* embryos (Boisset et al., 2010). As mentioned previously, in this model GFP is expressed in some ECs and IAHC cells (Figure 4Fi) and marks all AGM HSCs (de Bruijn et al., 2002; Ma et al., 2002). Prior to sectioning, intra-aortic blood was removed, and the aortic lumen stained with antibodies against the established endothelial surface marker CD31 (Figure 4D). Embryonic slices were imaged for 15 hours, revealing the live emergence of round GFP^+CD31^+ cells from flat ventral ECs expressing those same markers (Boisset et al., 2010). On average two emerging cells were seen per E10.5 aorta, with each emergence taking roughly one hour. Post-imaging, sections were stained for the IAHC and AGM/adult HSPC marker cKit (Figure 4D), confirming newly emerged cells were $GFP^+CD31^+cKit^+$. The cell surface marker CD41 (*Itga2b*) is generally considered the earliest marker for commitment towards a haematopoietic fate and is expressed before the mature pan-haematopoietic marker CD45 (*Ptprc*) (Eilken et al., 2009; Ferkowicz et al., 2003; Mikkola et al., 2003). Whereas all AGM HSCs express CD41 at intermediate levels (Robin et al., 2011), they do not all express CD45 (North et al., 2002). Boisset et al. (2010) therefore performed live-imaging on *CD41-YFP* knock-in mouse embryos (Zhang et al., 2007) in which yellow fluorescent protein (YFP) marks all *CD41*-expressing cells. Like what was observed using the *Ly6A-GFP* model, emerging cells in the *CD41-YFP* embryonic aorta were $YFP^+CD31^+cKit^+$. Taken together, they showed the *de novo*-generated round cells expressed all phenotypic AGM HSC markers and that the number of emerged cells conforms with previous quantitative HSC studies (Kumaravelu et al., 2002), again stressing the rare nature of HSCs. Unfortunately, technical limitations prevented harvest of these cells to ultimately prove they were functional HSCs capable of repopulating irradiated recipients.

Hemodynamic forces produced by blood flow (shear stress, hydrostatic pressure, and circumferential strain) have been shown to increase *Runx1* expression and to have a positive effect on haematopoiesis

(reviewed in Horton et al. (2021)). Given that blood flow was absent in the Boisset et al. (2010) study, the number of emergence events is possibly higher in intact embryos with a functional heartbeat and circulation. The fact that emergent events still occurred, even without blood flow, shows these biomechanical forces are not absolutely required at the time of emergence.

Using whole live zebrafish embryos, that are transparent and develop *ex utero*, the same process was visualised (Bertrand et al., 2010; Kissa and Herbomel, 2010). In one study, double transgenic reporter fish for haematopoietic (*cmyb*, also known as *myb*) and endothelial (*kdrl*, also known as *flk1* or *vegfr-2*) genes were used to perform confocal time-lapse imaging of the zebrafish AGM equivalent around the time of haematopoietic cell generation (Bertrand et al., 2010). Like the mammalian embryo, flat ventral endothelial (*kdrl*$^{+}$*cmyb*$^{-}$) cells were observed to round up, acquire haematopoietic marker expression (*kdrl*$^{+}$*cmyb*$^{+}$) and bud slightly into the zebrafish aortic lumen. However, contrary to mammals, these cells then migrated extravascularly and entered circulation by intravasation through the dorsal wall of the subjacent caudal vein (Bertrand et al., 2010). FACS of single and double positive cell populations showed a gradual upregulation of *cd41* and later *cd45* when analysed by quantitative (q)PCR (Bertrand et al., 2010), as seen in the mouse embryo (Boisset et al., 2010; Eilken et al., 2009; Mikkola et al., 2003; Yokomizo and Dzierzak, 2010). Lineage tracing with a Cre driven from the transcriptional regulatory elements of *kdrl*, an approach comparable to earlier murine studies using *Cdh5* (Chen et al., 2009; Zovein et al., 2008), showed long-term labelling of most blood cells, suggesting zebrafish haematopoietic tissue derives from HE. The newly emerged cells could also not be directly isolated for any functional tests.

Stunning high-resolution live-imaging on similar transgenic zebrafish embryos was performed by Kissa and Herbomel (2010). In their study they used reporters for endothelial genes *kdrl* and *lmo2* and the early haematopoietic gene *cd41* to capture blood formation *in vivo*. Imaging started at the time of aorta formation, until well past the peak of haematopoietic generation in the AGM. Unlike the mouse imaging study in which the event

appears rare, many *kdrl*$^{+}$ ECs in the zebrafish embryo were shown to contract and bulge into the ventral sub-aortic space (Kissa and Herbomel, 2010). After 1-2 hours a further contraction gradually detached bulging cells from their EC neighbours and allowed round cells to enter circulation via the underlying axial vein. Like Bertrand et al. (2010), ECs (*lmo2*$^{+}$) gradually acquired *cd41* as they budded out of the ventral wall, indicating their haematopoietic commitment (Kissa and Herbomel, 2010). Additionally, at the peak of haematopoietic cell generation in the AGM, the diameter of the zebrafish dorsal aorta temporarily more than doubled in size, implying a role of these remodelling dynamics in nascent zebrafish blood formation (Kissa and Herbomel, 2010). Next, by suppressing *runx1* expression, a massive reduction in endothelial transitions was seen and those cells that did manage to start emerging underwent abortive transitions and disintegrated (Kissa and Herbomel, 2010). Thus, a vital and conserved role for *runx1* in this process occurs in zebrafish, as seen in mouse cells (Chen et al., 2009; Lancrin et al., 2009).

Altogether, both studies showed round blood cells emerged from flat ventral aortic HECs in the zebrafish embryo (Bertrand et al., 2010; Kissa and Herbomel, 2010). The initial stages of this transition appear conserved among several vertebrate species (Bertrand et al., 2010; Boisset et al., 2010; Jaffredo et al., 1998; Kissa and Herbomel, 2010; Zovein et al., 2008). However, the remainder of the process is very different from e.g. mouse, as no IAHC formation is observed in zebrafish, the frequency of transition events seems much higher, and zebrafish haematopoietic cells bud directly into the ventral sub-aortic space as opposed to the aortic lumen, after which they enter circulation via intravasation into the subjacent axial (head end) or caudal (tail end) vein. The chick also has IAHC formation, but unlike mice where these cells remain in the lumen, some of the chick haematopoietic cluster cells have been shown to migrate from the aortic lumen into the underlying mesenchyme ‘para-aortic foci’, a structure unique to chick with a proposed function similar to the mouse FL (Jaffredo et al. (1998); reviewed in Cumano and Godin (2007)).

Kissa and Herbomel (2010) would be the first to describe the event of morphological and cell fate change required for *de novo* definitive blood formation as an 'endothelial-to-haematopoietic transition' (EHT) (Figure 4C,E). Collectively, these studies on mouse, chick and zebrafish embryos establish firmly that the earliest HSCs emerge (directly and/or indirectly) from HE via an EHT event. Though long thought to only occur during early embryonic development, a recent study has observed EHT from BM HE in late foetus/young adult chick and mice (Yvernogeau et al., 2019).

It is remarkable to think that seemingly terminally differentiated endothelial cells have the capability, during a short window of developmental time and/or under unique circumstances, to acquire multipotency, become an HSC, and then differentiate within the haematopoietic hierarchy. Therefore EHT can be defined as a transdifferentiation or lineage switching process, meaning a transformation from one mature somatic cell type (endothelium) into another (blood) without going through a pluripotent state like ESCs (reviewed in Graf and Enver (2009)). This supports the idea that during natural development there is not only unidirectional differentiation towards the terminal cell fate of a certain lineage, but also lineage conversion (transdifferentiation). This type of differentiation and the total number of HSCs as a result thereof, is rare and most vascular ECs will never become endowed with haemogenic potential (reviewed in Lange et al. (2021)). In the end, Emmel (1916) was quite right to conclude "ordinarily passive" embryonic and adult endothelial cells can under certain "abnormal conditions" become proliferative and start contributing to the "free cellular elements" of blood. As cells undergo EHT, they lose an endothelial gene signature, whilst they acquire a haematopoietic phenotype and characteristics. This transition might occur both gradually, with one type of HEC progressively maturing through various intermediate (progenitor) cell fates, and/or spontaneously, with distinct HEC origins for each cell fate (reviewed in Dzierzak and Bigas (2018)). Given its rare occurrence, it is likely that stochasticity also plays a role in EHT (reviewed in Graf and Enver (2009)).

Isolation and characterisation of the first functional HSCs

Most of our current functional and spatiotemporal understanding of early murine HSCs comes from immunohistochemical and FACS analyses of embryonic tissues and is heavily reliant on unique intra- and extracellular molecules/markers (Figure 4E). While the developmental landscapes of HECs, IAHCs and HSPCs are highly similar, these individual populations are phenotypically overlapping and functionally heterogeneous, and therefore challenging to isolate and characterise. Cells undergoing EHT still express varying levels of many endothelial genes and surface markers for some period, while already starting to display haematopoietic features (reviewed in Dzierzak and Speck (2008); Lange et al. (2021)).

Despite their transitional and potentially 'metastable' state, as indicated by co-expression of counteracting/cross-antagonistic (endothelial and haematopoietic) transcriptional programmes (Olsson et al. (2016); reviewed in Graf and Enver (2009)), useful markers of HECs and HSCs/IAHCs exist, and we now know mouse embryos contain around 700 phenotypic IAHC cells at E10.5 (Figure 4G), with only two of these being *bona fide* HSCs at this time of development (Kumaravelu et al., 2002; Yokomizo and Dzierzak, 2010). Both the entire aortic endothelial lining and all IAHCs can be phenotypically characterised and isolated (by FACS) based on their expression of endothelial surface markers like CD31 (*Pecam1*) (Figure 4D,F) and CD34 (Garcia-Porrero et al., 1998; Yokomizo and Dzierzak, 2010). Additionally, all ECs and IAHC cells express VE-Cadherin, Tie2 (*Tek*, also known as CD202B), Endoglin (*Eng*, also known as CD105), and Esam (*Esam*) (Chen et al., 2011; Liakhovitskaia et al., 2009; Roques et al., 2012; Takakura et al., 1998; Yokomizo and Dzierzak, 2010) (Figure 4E). Mutually exclusive expression is seen for Flk1 and CD45, with Flk1 expressed on all ECs and the most basal IAHC cells, and CD45 mainly on the outer/luminal cells of IAHCs (Yokomizo and Dzierzak, 2010) (Figure 4E, Fii), a pattern similar to that observed in chick (Jaffredo et al., 1998). In contrast to these more broadly expressed markers, the surface marker CD41

is only observed on some IAHC cells (presumably those transitioning from Flk1^{+} to CD45^{+}) and no ECs (Figure 4E), whereas the intracellular *Ly6A-GFP* reporter is only expressed in a select number of (H)ECs and IAHC cells, to varying intensity levels (Figure 4Fi) (de Bruijn et al., 2002; Robin et al., 2011; Yokomizo and Dzierzak, 2010). Importantly, when Runx1's non-DNA binding partner "core binding factor β" (CBFβ; also required for HSPC formation) is deleted, HSC but not HPC formation is rescued by expression of CBFβ from the *Ly6A-GFP* regulatory elements, whereas only HPC (EMP) and no HSC generation is rescued when driven from *Tek* regulatory sequences (Chen et al., 2011). This underpins the role of *Ly6A-GFP* as marker for both (HSC-primed) HECs and HSCs, and suggests the existence of several types of HECs or programme plasticity (reviewed in Dzierzak and Bigas (2018); Ottersbach (2019)).

More recently, the hyaluronan receptor CD44 (*Cd44*) was discovered as a functional cell surface marker of HECs at the very moment of EHT when HECs change morphology and express both an endothelial and haematopoietic programme (Figure 4E) (Oatley et al., 2020). Onset of CD44 expression is slightly earlier than CD41 and cKit and starts on HECs that appear more quiescent and metabolically inactive than CD44^{negative} ECs. It continues to be expressed on most IAHC cells (Oatley et al., 2020). Another recent paper reveals wider and in fact pan-endothelial and pan-IAHC expression for CD44 in the midgestation aorta, but also concludes that functionally it is one of the markers of HECs destined to give rise to HSCs (Hou et al., 2020).

Exclusive to all IAHC cells is the expression of stem cell factor (SCF) receptor and well-known adult HSC marker cKit (Yokomizo and Dzierzak, 2010) (Figure 4D,E,Fi,Fiii). Based on its expression, the first IAHCs are observed from E9.5 of murine development and both the total number of cKit^{+} cells within the aorta and the maximum size of individual clusters peaks at E10.5 (± 700 cells, up to 19 cells/IAHC), after which both gradually decrease (Figure 4G) (Yokomizo and Dzierzak, 2010). The IAHCs are present between E9.5-14.5 and found to be of variable size, containing 1-19

and 1-11 cells per IAHC at E10.5 and E11.5 respectively, with a significant portion of those (± 100 out of 500 total cKit$^+$ cells) as single-cell 'clusters' at the latter stage. Based on their shape, IAHCs can generally be classed as either spheroidal, mushroom-like, stacked or single-cell, as shown by scanning electron microscopy of E10 and 11 aortae (Boisset et al., 2015). The fact slightly more functional HSCs are detected at E11.5 versus E10.5, in combination with an increase of the number of single-cell/small clusters at E11.5, might indicate a correlation between IAHC size and HSC functionality (Kumaravelu et al., 2002; Muller et al., 1994; Yokomizo and Dzierzak, 2010). Yokomizo and Dzierzak (2010) also revealed most IAHCs are found in the middle segment of the aorta at E10.5 and 11.5, and that 63% and 37% are localised on the E10.5 ventral and dorsal wall, respectively (Figure 4D). The high density of cKit$^+$ IAHC cells in the ventral middle segment of the aorta therefore matches the location of the earliest *in vivo*-repopulating HSCs (Mascarenhas et al., 2009; Taoudi and Medvinsky, 2007; Yokomizo and Dzierzak, 2010). The clusters observed in both UA and VA (Figure 4Fiii) are phenotypically similar to those within the aorta and also peak in number and size at E10.5, after which both decrease (Yokomizo and Dzierzak, 2010). Compared to aortic clusters, their size is generally much larger and their location does not seem to favour one specific aspect of either of the arteries, except being in closer proximity to the embryo body than the YS or placenta (Yokomizo and Dzierzak, 2010).

Other both functionally and spatiotemporally useful markers for isolation, characterisation and visualisation of HSCs and HECs are the pivotal haematopoietic TFs Gata2 and Runx1 (Figure 4E) (reviewed in Ottersbach (2019)). Reporter expression as driven from *Runx1*'s haematopoietic-specific *+23* enhancer (*23GFP*; 23.5-kb downstream of the ATG in *Runx1*'s exon 1; Nottingham et al. (2007)), is seen in E10.5 aortic ECs (HECs), underlying mesenchymal cells and most IAHCs (Swiers et al., 2013). *23GFP* expression initiates in ECs at E8.5, one day before Runx1 protein and transcript are detected in these cells, revealing early signs of haematopoietic differentiation/competence (and the possibility of prospective

HEC isolation) two full days prior to HSC emergence. From E9.5, the *23GFP*$^{+}$ HECs undergo the first haematopoietic specification, as determined by gene expression. Endogenous *Runx1* expression peaks between E9.5-10.5, after which it rapidly decreases, emphasising the transient nature of HECs (Swiers et al., 2013). When analysed by *in situ* hybridisation (ISH), *Runx1* expression in the AGM (driven from any of its enhancers) is seen in all ECs, IAHCs and in several layers of cells of the subjacent ventral mesenchyme (Bee et al., 2009).

Due to lack of a suitable FACS antibody, the spatiotemporal and functional characterisation of live *Gata2*-expressing cells, with normal levels and function of Gata2, was not possible up until the generation of the *Gata2Venus* reporter mouse (*Chapter 2*; Kaimakis et al. (2016)). Nevertheless, its crucial role in definitive haematopoiesis has long been known (reviewed Crispino and Horwitz (2017)). Even though mouse embryos homozygously deficient for *Gata2* produce (albeit reduced numbers of) primitive haematopoietic cells, they suffer severe anaemia and are embryonic lethal at about E10.5 (67% are dead at E10.5 and none survive beyond E11.5) (Tsai et al., 1994), the time of HSC induction. The embryonic lethality two days prior to the *Runx1* homozygous knockout implies an earlier and upstream role for Gata2 in definitive haematopoiesis. Transgenic analyses and ISH show the earliest expression of *Gata2* in the (intraembryonic) lateral mesoderm next to the primitive streak and in the extraembryonic mesoderm and YS of the E7.5/8 mouse embryo (Minegishi et al., 1999; Minegishi et al., 2003). From E9.5 expression is also detected in the PL and FL (Minegishi et al., 1999; Minegishi et al., 2003). Between E8/9 (dependent on use of knock-in mouse or transgenic reporters with expression driven from different transcriptional regulatory elements) until late E11, ISH and transgenic analyses on mouse aorta and VA localise *Gata2* to a steadily increasing number of (H)ECs and later also to most IAHC cells (Minegishi et al., 1999; Minegishi et al., 2003; Robert-Moreno et al., 2005). *Gata2*'s expression pattern in aortic ECs is like that of (*Runx1*) *23GFP* and *Ly6A-GFP*, however the latter displays a more restricted expression in IAHCs.

Based on functionality and potential, the first definitive HSCs in the AGM can indeed be isolated by use of aforementioned HEC and IAHC cell markers and are therefore at this timepoint phenotypically characterised as: Ly6A-GFP$^+$ (Chen et al., 2011; de Bruijn et al., 2002), CD34$^+$, cKit$^+$ (Sanchez et al., 1996), Gata2$^+$ (Minegishi et al., 1999; Minegishi et al., 2003), CD31$^+$, VE-Cadherin$^+$, Runx1$^+$ (North et al., 2002), Endoglin$^+$ (Roques et al., 2012), Tie2$^+$ (Liakhovitskaia et al., 2009; Takakura et al., 1998), and CD41^{intermediate} (Robin et al., 2011) (Figure 4E). Some markers like CD11b (*Itgam*, also known as Mac-1) (Sanchez et al., 1996), CD45 and Flk1 (North et al., 2002) are only found on some of the first HSCs, with HSC activity detected in both the negative and positive populations for these markers. Given the transitional and heterogeneous nature of HECs and IAHCs, and the non-synchronous expression onset of most markers (Figure 4E), the list is non-exhaustive, and several other endothelial/haematopoietic surface markers and TF reporter lines can be used for purification and/or visualisation of the first HSCs. Our lab has found the use of one pan-endothelial-IAHC marker, in combination with cKit (pan-IAHC), and Ly6A-GFP or one of the pivotal TFs to yield the best AGM HSC enrichment. The necessity for these combinations of embryonic markers is underlined by the fact most adult BM HSC enrichment strategies (e.g. LSK SLAM) are not suitable for isolation of the earliest HSCs. As an example, AGM HSCs see differential expression of adult HSC refinement marker Sca-1 (de Bruijn et al., 2002) and SLAM marker CD150 is only expressed on very few AGM cells, none of which have HSC potential (McKinney-Freeman et al., 2009).

Taken together, it is now well-established that during a narrow window of time in vertebrate development 1) both IAHCs and the first HSCs are *de novo* generated from specialised HECs via a biomechanical and molecular EHT event, 2) their endothelial origin is phenotypically reflected in the unique combination of embryonic/adult endothelial and haematopoietic markers expressed by HSCs/IAHCs and required for their isolation, and 3) the highest density of IAHCs and location of the first functional HSCs co-localise to the ventral middle section of the midgestation dorsal aorta.

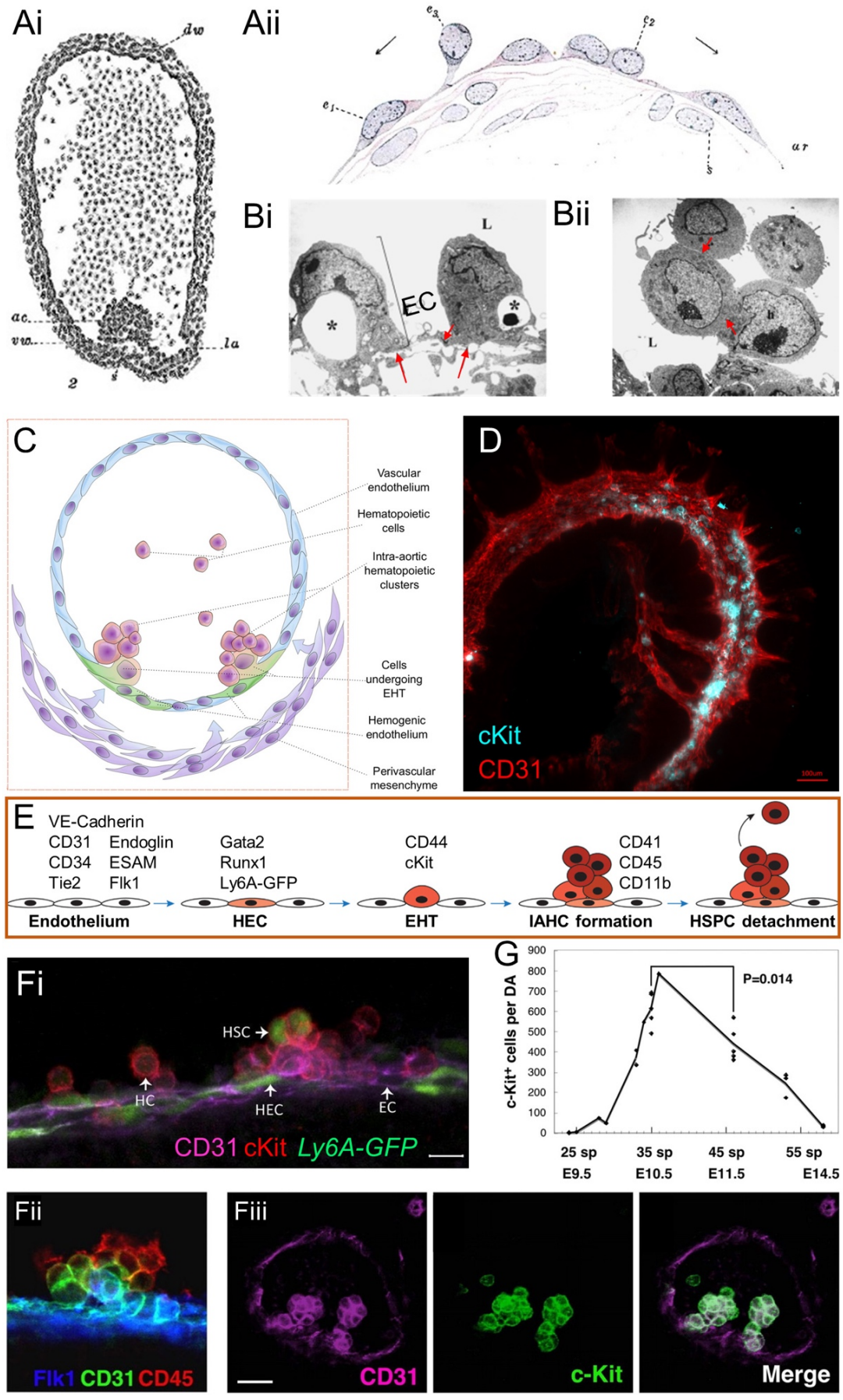

Ai
dw
ac.
vit.
la
2
Aii
e3
e2
e1
ar
s
Bi
L
EC
*
*
Bii
L
C
Vascular endothelium
Hematopoietic cells
Intra-aortic hematopoietic clusters
Cells undergoing EHT
Hemogenic endothelium
Perivascular mesenchyme
D
cKit
CD31
100um
E
VE-Cadherin
CD31 Endoglin
CD34 ESAM
Tie2 Flk1
Gata2
Runx1
Ly6A-GFP
CD44
cKit
CD41
CD45
CD11b
Endothelium
HEC
EHT
IAHC formation
HSPC detachment
Fi
HSC
HC
HEC
EC
CD31 cKit Ly6A-GFP
G
c-Kit+ cells per DA
900
800
700
600
500
400
300
200
100
0
P=0.014
25 sp
E9.5
35 sp
E10.5
45 sp
E11.5
55 sp
E14.5
Fii
Flk1 CD31 CD45
Fiii
CD31
c-Kit
Merge

Figure 4 – The biomechanics of HSC generation – Ai) Early 20th century drawing of a transverse section through the aorta of a 12mm pig embryo. An intra-aortic haematopoietic cluster (IAHC) is attached to the ventral endothelial lining of the aorta. Ac=aortic cluster; vw=ventral wall; dw=dorsal wall; la=lateral aortic ramus (branch); s=mesenchyme. **Aii)** Early drawing of a sagittal section of vascular endothelium at the entrance of an aortic ramus (ar) of a 9mm pig embryo. Morphological changes in (haemogenic) endothelial cells (HECs) are suggestive of IAHC formation (going from endothelial cell e1 to e2 to e3). Right arrow indicates blood flow from aorta into branch. Left arrow indicates direction of circulation along the long axis of the aorta. Both **Ai)** and **Aii)** are effectively some of the first drawings of the endothelial-to-haematopoietic transition (EHT). Reprinted from Emmel (1916), with permission from John Wiley and Sons. **B)** Transmission electron microscopy images of part of the wall of an embryonic day (E)10.5 murine dorsal aorta. Original magnifications 4000x. **Bi)** Endothelial cells (EC) budding out of the ventral aortic wall into the aortic lumen (L) as they undergo EHT. Many of the cells contain large cystic formations (*). Short arrow indicates a tight junction connecting ECs to each other, but not to the underlying mesenchyme. Long arrows point at the basal lamina of extracellular matrix between endothelium and mesenchyme. **Bii)** IAHC of four round haematopoietic cells (h) attached to the endothelial lining. The four cells in the IAHC are interconnected via tight junctions (short arrows). Adapted from Marshall and Thrasher (2001). **C)** Schematic of transverse cross-section of an E10.5/11.5 murine dorsal aorta. Two large ventral IAHCs are being generated through EHT from ventral HECs (green). Reprinted from Lange et al. (2021). **D)** Maximum intensity projection of an E10.5 (35 somite pairs; sp) live whole-mount dorsal aorta stained for endothelial and IAHC marker CD31 (red) and IAHC marker cKit (blue). The highest number of IAHCs is observed on the ventral side near the mouths of several aortic branches. Image by C.S. Vink taken with a Zeiss light-sheet microscope. **E)** Representation of morphological and cell specific marker changes during the gradual transdifferentiation of endothelial cells (HECs) into IAHCs and

haematopoietic stem and progenitor cells (HSPCs) occurring in the midgestation mouse aorta *in vivo*. Vascular endothelial cells express a wide array of pan-endothelial cell surface markers. Through the expression of certain pivotal haematopoietic genes, a sub-fraction of the endothelium will become haemogenic (HECs; light red) and start undergoing EHT, leading to the formation of IAHCs that gain haematopoietic characteristics (dark red) as they advance away from the endothelium. Once required/fully matured the HSPCs will detach from IAHCs and enter circulation. Onset of expression of specific genes and markers is indicated above the schematic of the different EHT stages. Adapted from Ottersbach (2019). **F)** Phenotypic markers of E10.5 IAHCs inside the **(Fi, ii)** dorsal aorta and **(Fiii)** vitelline artery (VA). **Fi)** High magnification 3D confocal image of the vascular wall of a 36sp *Ly6A-GFP* aorta. Vascular ECs and IAHCs are $CD31^{+}$ (magenta). HECs express Ly6A-GFP (green) in addition. IAHCs express both CD31 and cKit (red). Haematopoietic stem cells (HSC) express IAHC markers and Ly6A-GFP. Scale bar=10μm. Adapted from Solaimani Kartalaei et al. (2015). **Fii)** IAHCs show mutually exclusive expression of Flk1 (blue) and CD45 (red). Flk1 is expressed in all ECs and at the base of IAHCs, whereas CD45 is expressed only by cells on the luminal side of IAHCs. **Fiii)** Cell clusters inside the VA (33sp) are phenotypically ($CD31^{+}cKit^{+}$) similar to those within the dorsal aorta. Scale bar=20μm. **G)** Quantification of number of $cKit^{+}$ cells within the murine embryonic dorsal aorta (DA) at different times of development. **Fii, Fiii, G)** Reprinted from Yokomizo and Dzierzak (2010), with permission from Company of Biologists Ltd.

Gata2 at the centre of a haematopoietic web

Timely expression of the right genes is what ultimately drives cell fate decisions. Of great importance is therefore the understanding of how this expression is regulated. Two main classes of players in gene regulation are the *cis*- and *trans*-regulators. Those located in *cis* are non-coding DNA sequences that regulate gene expression intramolecularly, i.e. on the same chromosomal allele or molecule of DNA as the regulated/transcribed gene (Alberts et al., 2015). *Trans*-acting elements/factors encode regulators that bind *cis*-motifs and work both intra- and intermolecularly, i.e. they regulate gene expression on both alleles (Alberts et al., 2015). Critical during development are the relatively short (± 50-500 base pairs; bp) *cis*-acting DNA sequences termed 'promoters' that initiate gene transcription (located proximal to the gene) and 'enhancers' that increase gene transcription. Enhancers typically contain clustered *cis*-motifs and are located intronically or distal to genes (reviewed in Hewitt et al. (2016)). Each *cis*-motif can be bound by one or several of the equally important *trans*-acting transcription factors (TF).

Discovery of Gata transcription factors

Knowledge regarding transcriptional regulation of the haematopoietic system and specifically the erythroid lineage came from a great number of studies focussing on the β-globin locus in the late 1980s (reviewed in Katsumura et al. (2017)). The detection of a consensus *GATA* motif in the enhancers and promoters of most globin and non-globin erythroid genes led to the discovery of the NF-E1 (nuclear factor-erythroid 1) protein. It displayed high-affinity DNA-binding to this short *cis*-motif and regulated erythroid gene expression (reviewed in Orkin (1990)). A further two DNA-binding proteins (NF-E1b and NF-E1c) with developmentally distinct spatiotemporal expression patterns were found to bind the same consensus motif with equally high affinity (Yamamoto et al., 1990). Given their highly similar DNA-

binding and motif selection properties, a new and more straightforward nomenclature emerged at a 1990 Globin Regulation conference, resulting in NF-E1, NF-E1b and NF-E1c henceforth being known as Gata1, Gata2 and Gata3 in mouse (GATA1, GATA2 and GATA3 in human), respectively (reviewed in Orkin (1990)). In both human and mouse the Gata TF family consists of six members (Gata1 to Gata6) all recognising the same naked, i.e. not protected by any proteins or lipids, DNA consensus motif (*A*/*T*)*GATA*(*A*/*G*) ([*C*/*T*]*TATC*[*A*/*T*] in reverse complement) and all having an equally strong preference for binding this motif within enhancers as opposed to promoters (reviewed in Hewitt et al. (2016)). Gata factors display a high degree of motif specificity. However, more than 99% of the *GATA* motifs are not actually occupied, possibly due to inaccessible chromatin and/or absence of other vital *cis*-elements preventing binding (reviewed in Hewitt et al. (2016); Katsumura et al. (2017)). All Gata factors are dual zinc finger TFs that bind the same motif with their two highly conserved and functionally diverse fingers (reviewed in DeVilbiss et al. (2016)). Besides both fingers facilitating protein-protein contact between Gata and other factors, the one nearest the amino-end (N-terminal) also enables chromatin occupancy and augments DNA-binding at palindromic motifs, whereas the finger at the carboxyl-end (C-terminus) is endowed with the sequence-specific DNA-binding (reviewed DeVilbiss et al. (2016)).

Gata1-3 are referred to as the haematopoietic Gata TFs, expressed in and responsible for the development and regulation of various haematopoietic cell types, as well as being expressed in some non-haematopoietic cell types, whereas Gata4-6 have strictly non-haematopoietic functions (reviewed in Katsumura et al. (2017)). Mouse Gata1, Gata2 and Gata3 are located on chromosome X, six and two, respectively (https://www.ncbi.nlm.nih.gov/gene/14460, -61, -62). While human GATA1, GATA2 and GATA3 are located on chromosome X, three and ten, respectively (https://www.ncbi.nlm.nih.gov/gene/2623, -24, -25). Although Gata1 is highly expressed during erythroid differentiation, Gata2 expression precedes that of Gata1 in the earliest erythroid precursors. Overlapping

expression of Gata1 and -2 occurs in mast cells, eosinophils, megakaryocytes and YS-derived primitive erythroblasts (reviewed in Katsumura et al. (2017)). Gata3 fulfils an important role in the regulation of T-cell lymphopoiesis and is expressed in T-cells (reviewed in Katsumura et al. (2017)).

The transition from Gata2-dependence in early erythroid precursors to Gata1-dependence during definitive erythropoiesis, with Gata1 replacing Gata2 at many genomic sites, is referred to as a "Gata-switch" (reviewed in Hewitt et al. (2016)). During both erythropoiesis and megakaryopoiesis (from HSPCs), roughly 30% of Gata2-bound enhancers become Gata1-bound, a switch partly aided by Gata2's shorter half-life (0.5-2 hours) (Lurie et al. (2008); Minegishi et al. (2005); reviewed in Katsumura et al. (2017)). The Gata-switch mechanism highlights the fact different Gata factors (in this case Gata1 and 2) can enhance/initiate unique transcriptional patterns by binding the same genomic site (reviewed in DeVilbiss et al. (2016)). Expression of both Gata1 and Gata2 in YS-derived primitive erythroblasts indicates redundancy in particular circumstances or tissues (reviewed in Hewitt et al. (2016)). Thus, Gata1 and 2 are uniquely expressed in certain cell types but can execute redundant roles as demonstrated by replacing one Gata TF with another on the same *GATA* motif, to regulate establishment of different cell lineages.

Prior to the discovery of the EHT, *Gata2*'s expression pattern in seemingly unrelated cell types like endothelium and multipotent HPCs made its function appear enigmatic (Yamamoto et al. (1990); reviewed in Orkin (1990)). However, knock-out studies firmly established Gata2's vital role in the development of the definitive haematopoietic system and HSPCs, with Gata2 being the only Gata TF family member playing a fundamental role at the very top of the definitive haematopoietic hierarchy (Tsai et al. (1994); reviewed in Katsumura et al. (2017)). Gata2's role in *de novo* HSC generation and expansion is dose-dependent with haploinsufficiency causing a severe reduction in the number of HSCs specifically within the AGM (Ling et al., 2004). Moreover, despite other (later) foetal and adult HSC-harbouring

tissues containing seemingly normal numbers of phenotypic HSCs under haploinsufficient conditions, HSCs produced are qualitatively defective in competitive and serial transplantations, implying a further role for Gata2 in proliferation (Ling et al., 2004). Thus, Gata2 plays two fundamental dose-dependent and distinct roles during HSC generation and proliferation.

Gata2 and gene regulation

The mouse *Gata2* gene consists of seven exons, with only the last five encoding the Gata2 protein (Figure 5A) (Minegishi et al., 1998). The first two exons, termed *1S*(*pecific*) and *1G*(*eneral*) (Figure 5A), both encode distinct (non-protein-coding) 5' untranslated regions that can be transcribed alternatively, whereas the last five are shared by the two transcript variants, resulting in translation of the same Gata2 protein from different m(essenger)RNAs (Minegishi et al., 1998). Expression of the two alternatively spliced transcripts is spatiotemporally regulated by two distinct promoters preceding the two alternative first exons. As their names suggest, initial analyses revealed the proximal 'general' *1G* promoter was active in all tissues and cells expressing *Gata2*, while mRNA transcription initiated from the distal *1S* promoter was only seen in immature BM HPCs, indicating haematopoietic cell-specific *Gata2* expression (Minegishi et al., 1998). More importantly, *1S*-driven *Gata2* expression was later also shown to control HSPC generation in the midgestation pSp/AGM, with specific *Gata2* expression seen in the aortic endothelial lining and IAHCs, and no other mesodermal tissues (Minegishi et al., 1999). A 3.5 kilo bp (kbp) fragment several kbp upstream of the *1S* promoter proved responsible for regulating this tissue-specific expression (Minegishi et al., 1999). Thus, transcription of *Gata2* can be initiated by two different promoters yielding two alternative mRNAs encoding the same Gata2 protein, with one of these promoters uniquely required for *Gata2* expression in the AGM.

Occupation of *cis*-elements by Gata factors in specific genomic locations (in proximity to critical genes) and tissues is what ultimately drives

cell fate change and provides different *cis*-elements with stage-specific roles. The *Gata2* locus is known to contain five highly conserved *cis*-regulatory enhancer elements with *GATA* motifs, indicating at least part of its expression could be controlled by autoregulation. Four of the enhancers are located upstream at position -77, -3.9, -2.8 and -1.8 kbp relative to the *1S* promoter, and one downstream and intronically at position +9.5 kbp (Figure 5A) (reviewed in Katsumura et al. (2017)). Gata2-to-Gata1-switching occurs at all five sites during erythropoiesis, but not all are of equal importance to the control of *Gata2* expression and/or generation of HSPCs. The element at -1.8 kbp functions late during erythropoiesis, deletion of the *-2.8* site only sees a slight reduction of *Gata2* expression, while *-3.9* deletion causes no *Gata2* or haematopoietic changes, additionally, no embryonic lethality is seen upon deletion of any of these three elements making them replaceable during development and haematopoiesis (reviewed in Katsumura et al. (2017)). The remaining two elements are not dispensable and deletion/mutation of the *-77* and *+9.5* elements causes mouse embryonic lethality at E15.5 and E13.5, and is implicated in acute myeloid leukaemia and MonoMAC immunodeficiency in humans, respectively (Figure 5B) (reviewed in DeVilbiss et al. (2016); Hewitt et al. (2016)). In *-77* homozygous knock-out mouse embryos, normal levels of *Gata2* are detected in LSK cells and HSC generation in the AGM occurs normally. However, greatly reduced levels of *Gata2* are seen in erythromyeloid progenitors, resulting in their accumulation and reduced differentiation potential towards certain erythroid and myeloid lineages (Johnson et al., 2015). Thus, the *-77* enhancer seems capable of conferring oligopotentiality onto erythromyeloid HPCs.

The intronic *+9.5* element, located between exon four and five, is of even greater importance and responsible for controlling the *Gata2* expression required for EHT and HSC generation to take place within the AGM region, thus playing a vital role in HECs/IAHCs acquiring multipotentiality (Figure 5B) (Gao et al., 2013). Upon homozygous deletion of this element, virtually no IAHCs are formed within the dorsal aorta and as a result the AGM lacks any functional CFU-Cs and HSCs (Gao et al., 2013). Similar to haploinsufficient

Gata2$^{+/-}$ mice (Ling et al., 2004), two functional *+9.5* enhancer sites are needed for synthesis of the right amount of Gata2 for generation of normal numbers of functional HSCs, with the *+9.5*$^{+/-}$ causing a significant reduction in chimaerism compared to *+9.5*$^{+/+}$ in primary and secondary transplantations (Gao et al., 2013). Thus, the *Gata2* locus contains several *cis*-elements, but only the +9.5 enhancer site is indispensable for HSC induction in the AGM.

Its importance is most likely due to the *+9.5* element not only containing a *GATA* motif, but also a nearby *E-box* motif (*CANNTG*; with N being any nucleotide) that can be bound by basic helix-loop-helix (bHLH) TFs. The exact order of this composite element, with an upstream *E-box*, followed by a 7-9bp spacer and a *GATA* motif (Figure 5A), facilitates maximum assembly of multimeric protein complexes, also known as enhanceosomes (reviewed in Hewitt et al. (2016)). In HSPCs the element is occupied by Gata2 and the pivotal haematopoietic bHLH TF Scl (*Tal1*) (reviewed in Katsumura et al. (2017)). It is the combination of these two DNA-binding proteins that can recruit other DNA- and protein-binding proteins essential for establishment of HSC identity. One such non-DNA-binding protein connecting/bridging Gata2 and Scl at *E-box-GATA* sites is "LIM domain only 2" (Lmo2, *Lmo2*) (Figure 5C), which additionally functions as support for other proteins that join the multimeric complex (reviewed in Hewitt et al. (2016)). Both Scl and Lmo2 are indispensable from early mouse haematopoietic and embryonic development onwards and function in YS and AGM haematopoiesis, whereas *Gata2* and *Runx1* lay downstream and are specifically required for definitive HSC generation in the AGM (reviewed in Dzierzak and Bigas (2018)). Gata2 also activates *cKit* expression by binding an *E-box-GATA* element with a longer spacer than the *+9.5* composite element, at the *cKit* promoter (reviewed in Hewitt et al. (2016)). Adding to the regulatory complexity of composite *E-box-GATA* enhancers are "*E-Twenty Six*" (*ETS*) motifs (*GGAA*) occupied by ETS TFs that are often linked to and found near these composite elements (reviewed in Hewitt et al. (2016)). While the *+9.5* element is located within the *Gata2* locus and is regulating HSC-specific *Gata2* expression through binding of Gata2, Scl and Lmo2

amongst other proteins, equally complex enhancer elements in the *Scl* and *Runx1* gene loci are occupied by multimeric protein complexes containing Gata2 and ETS factors (Fli1 [*Fli1*] and Elf1 [*Elf1*]) (Gottgens et al., 2002), and Gata2, Scl, Runx1 and ETS factors (Fli1, Elf1 and PU.1 [*Spi1*]) (Nottingham et al., 2007), respectively. Indicating Gata2 also regulates *Scl* and *Runx1* expression.

Heptad transcription factors

Fundamental to HSPC specification and generation are the combinatorial interaction and simultaneous co-expression of seven 'heptad' TFs, and particularly the synergistic activity of the core haematopoietic regulators Scl, Gata2 and Runx1, as shown by Wilson et al. (2010) (Figure 5C). The genome-wide analysis of a murine multipotent HSPC cell line (HPC-7) identified unrecognised protein-protein interactions between Runx1 (binding the *TGTGGT* motif) and Gata2, Lmo2, Erg (*Erg*), Fli1, Scl, and Lyl1 (*Lyl1*), the latter being a bHLH TF binding the same *E-box* motif as Scl. Their analyses also confirmed other known binding pairs such as Lmo2 with either Gata2 or Scl, and Gata2 with the ETS factors Erg or Fli1; thus establishing a multimeric heptad complex consisting of Gata2, Scl, Lyl1, Runx1, Lmo2, Erg and Fli1 (Figure 5C) important for HSPC transcriptional control (Wilson et al., 2010). Combinatorial binding of the same heptad factors was later also shown in primary human HSPCs (Beck et al., 2013). The complexity and simultaneous requirement of various TFs in order to establish HSC identity underlines their central role in HSC and in fact stem cell biology, and suggests small level changes of these factors can have serious effects on assembly of multimeric complexes and eventually cell fate.

In summary, HSC-specific enhancers within loci of pivotal haematopoietic genes rely on an intricate combination of conserved (composite) *cis*-elements with a specific order and spacing of individual motifs, to allow for optimal combinatorial levels of proteins to bind, assemble into multimeric complexes and ultimately control gene expression and cell fate establishment.

Developmental onset of expression and upstream regulators

Notch

Whilst binding of the heptad TFs to regulatory elements within key gene loci is critical for EHT, it is the activation and cooperation of several signalling pathways by extrinsic factors/signals that leads to the initial onset of expression of these essential intrinsic regulators. An important pathway directly upstream of *Gata2* is the Notch signalling pathway (Robert-Moreno et al., 2005). It encompasses a family of five transmembrane ligands (Delta-like/DLL1, 3, 4 and Jag[ged]1, 2), four transmembrane receptors (Notch1-4) and various downstream players, all with essential cell-cell contact-dependent roles in various (stem) cell systems (reviewed in Bigas and Porcheri (2018); Bigas et al. (2010)). Signalling is activated through the interaction of a receptor on the surface of one cell with a Notch ligand presented by a neighbouring cell, leading to proteolytic cleavages of the receptor, activation of an intracellular signalling cascade and finally target gene expression (Figure 5Dii) (reviewed in Bigas et al. (2010)). In the developing mouse embryo, Notch signalling is required for the sequential arterial and haemogenic specification of ECs in the dorsal aorta. Despite this crucial role in vascular fate determination, the pathway appears dispensable during the first extraembryonic/YS haematopoietic waves (reviewed in Bigas and Porcheri (2018)). Both *Notch1* and *Notch4*, and ligands *Dll4*, *Jag1* and *Jag2* are expressed in the E9.5-11.5 dorsal aorta (Robert-Moreno et al., 2005; Robert-Moreno et al., 2008). *Notch1* and *Jag1* are detected in most IAHC cells and scattered ECs, with *Notch1* expression mostly restricted to the ventrolateral aortic wall. *Notch4* is highly expressed by all ECs, but much less so in IAHCs, and *Jag2* and *Dll4* showed heterogeneous EC/IAHC expression. Knock-out studies revealed a requirement for Notch1 during HSC generation in the mouse AGM and proved it is the sequential activation by Dll4 and Jag1 that first ensures arterial and then haemogenic/HSC specification, respectively (reviewed in Heck et al. (2020)). Indeed, upon

Jag1 deletion arterial specification itself appears normal, however Notch1 is not present on *Gata2*'s promoter and no *Gata2* expression is seen in the E10.5 aorta, emphasizing the importance of the Notch1-Jag1 axis in controlling *Gata2* expression and initiating HSC generation from arterial (H)ECs (Robert-Moreno et al., 2008). Another downstream target of Notch1 involved in HSC generation in the AGM is the bHLH TF Hes1 (*Hes1*), which blocks the expression of *Gata2* and allows Jag1-activated Notch1 to both directly activate and indirectly block *Gata2* expression to maintain a just-right/physiological dose of Gata2 (Figure 5Dii) (reviewed in Bigas and Porcheri (2018); Katsumura et al. (2017)). While crucial haematopoietic regulators like *Gata2* require Notch1 signalling for activation of their gene expression, EHT and the final maturation into HSCs can only be completed by a timely block/down-regulation of *Notch* as Dll4-induced high Notch strength promotes arterial and suppresses haemogenic specification, and only Jag1-induced low Notch strength is able to induce a haemogenic/HSC programme (reviewed in Ottersbach (2019)). It is therefore that, following the Dll4-Notch1-driven high Notch activity needed to specify aortic/arterial endothelium, ventral ECs will be exposed to Jag1-Notch1-driven low levels of Notch activity to allow for EHT and IAHC formation, whereas the dorsal aspect maintains an arterial programme with high Notch, in accordance with the dorsoventral polarisation of HSC induction in the aorta (Figure 5Di) (reviewed in Bigas and Porcheri (2018)). Taken together, Notch signalling and specifically Notch1 and the sequential interaction with its ligands Dll4 and Jag1 is needed to activate and then suppress an arterial programme, and induce a haematopoietic and ultimately self-renewal programme in the AGM niche.

BMP

Another player said to be directly upstream of *Gata2* in the mouse AGM region is the bone morphogenetic protein (Bmp) signalling pathway and particularly its activation by Bmp4. Bmp4 (*Bmp4*) is a ventralising morphogen

part of the transforming growth factor beta (TGF-β) superfamily and important for embryonic haematopoiesis due to its involvement in dorsoventral mesoderm patterning (reviewed in Sadlon et al. (2004)). All TGF-β members can bind and activate a family of heterodimeric receptors. Dimers of Bmp4 (and other Bmps) are capable of binding either of the two receptor subunits, which then form an activated ligand-receptor complex that will lead to the phosphorylation of members of the downstream Smad family. These will interact with Smad4 (a co-Smad) after which the entire complex is translocated to the nucleus where it regulates target gene expression by interacting with various TFs (reviewed in Sadlon et al. (2004)). Upon Bmp activation it is specifically the receptor-regulated Smads (R-Smad)-1, -5 and -8 that become phosphorylated and get taken to the nucleus. Homozygous deletion of Bmp4 in mice causes embryonic lethality prior to definitive haematopoiesis. This led to the use of lower vertebrates like *Xenopus* for many of these studies and it was in this species that bmp4 was shown to work in a dose-dependent fashion with a gradient from high to low needed for specification of ventral to dorsal mesoderm. Thus establishing it as a strong ventralising morphogen involved in induction of a haematopoietic programme (reviewed in Sadlon et al. (2004)). It was also in *Xenopus* that *Gata2* expression was shown to be induced by (Maeno et al., 1996) and be a direct target of (Friedle and Knochel, 2002) bmp4 signalling. *Xenopus* orthologues of *Runx1*, *Scl* and *Lmo2* are also thought to be directly or indirectly activated by bmp4 (reviewed in Sadlon et al. (2004)). Bmp signalling is required for all stages of *Xenopus* blood development, but specifically induces *Gata2* and *Scl* expression in blood precursors during embryonic development (Walmsley et al., 2002).

In the E11 mouse AGM, Bmp4 appeared to have a similar ventralising role revealing a clear dorsoventral gradient with *Bmp4* mRNA and protein expression mostly detected in the ventral subaortic mesenchyme (Figure 5Ei,Eii,Fi,Gi) (Durand et al., 2007). Bmp4 staining on *Ly6A-GFP* AGM sections revealed cytoplasmic expression directly subjacent to GFP^+ ECs and IAHCs, the latter also presenting foci of Bmp4 expression on their

surface indicative of receptor interaction (Figure 5Gi,Gii) (Durand et al., 2007). When tested on mouse AGM explants, the addition of Bmp4 increased the number of HSCs. When the Bmp antagonist gremlin (and noggin to a lesser extent) was added to explants, no long-term reconstitution and a significant reduction in CFU-S progenitors was seen upon transplantation, strengthening its requirement in HSC induction (Durand et al., 2007). The highly conserved nature of BMP signalling is underscored by the fact BMP4 is expressed with a similar dorsoventral polarity and strong subaortic ventral expression at the time of IAHC formation in the human AGM (Marshall et al., 2000). Prior to IAHC formation virtually no BMP4 staining is observed and once IAHCs have disappeared expression is no longer polarised (Marshall et al., 2000). Finally, like Notch signalling, an active Bmp pathway is only transiently and locally required within the AGM, and during/after the time of IAHC formation the morphogen sonic hedgehog (Shh) activates ventral noggin expression resulting in a block of Bmp signalling (reviewed in Canu and Ruhrberg (2021)). Altogether, Bmp signalling and in particular Bmp4 have a highly conserved role in ventral mesodermal specification and subsequent HSC induction in the AGM region, most likely by directly regulating expression of pivotal haematopoietic TFs like *Gata2*.

In summary, Gata2 finds itself at the centre of a highly conserved haematopoietic web, requiring optimal combinatorial levels of extrinsic pathway-activating signals and intrinsic enhancer-binding TFs, to allow for its own timely transcriptional activation and regulation, which in turn enables it to do the same for other essential haematopoietic TFs down- and upstream. The importance of all factors and signals involved is underlined by the fact dysregulation of any of them is implicated in haematopoietic disorders.

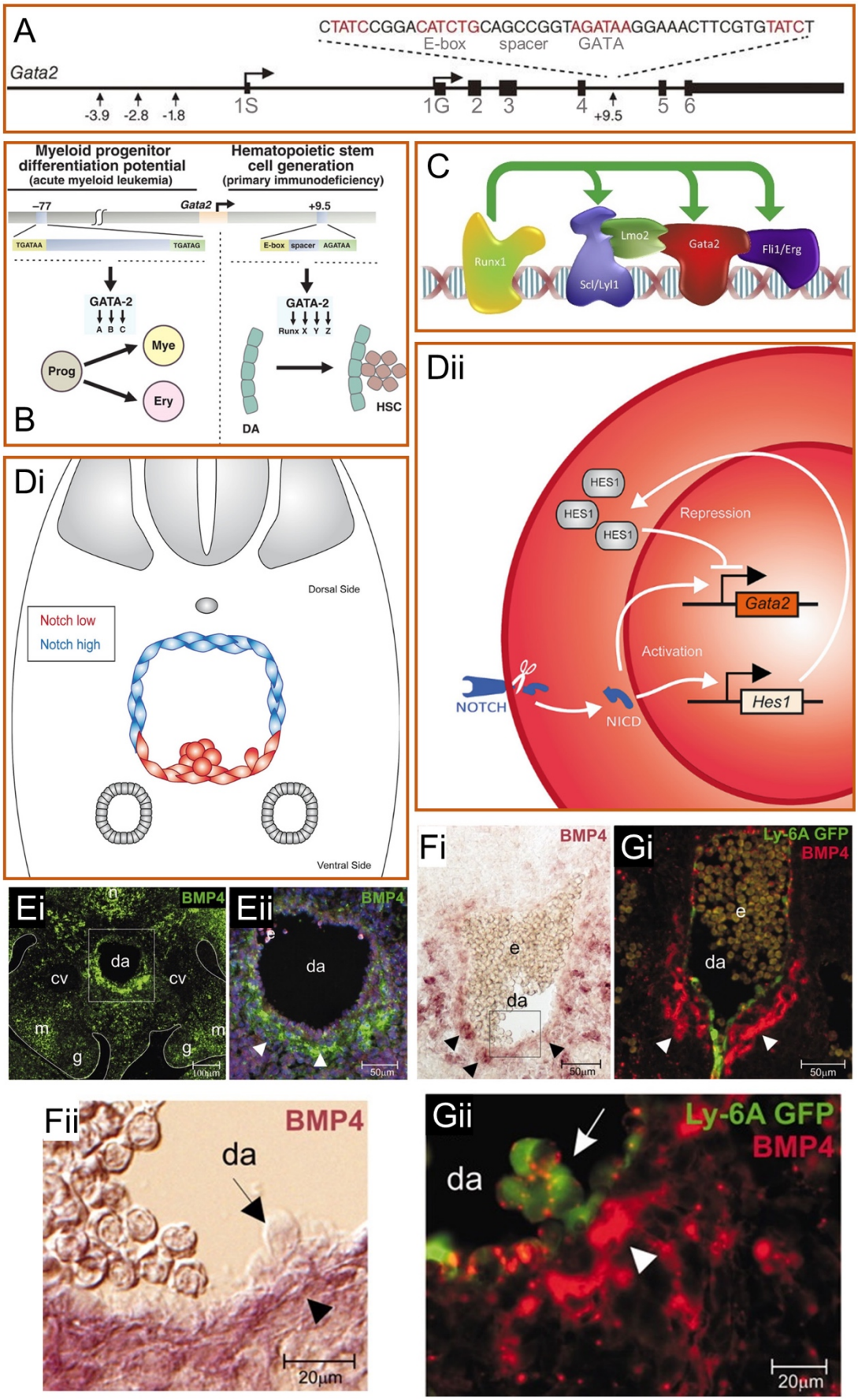

A
CTATCCGGACATCTGCAGCCGGTAGATAAGGAAACTTCGTGTATCT
E-box
spacer
GATA
Gata2
-3.9
-2.8
-1.8
1S
1G
2
3
4
+9.5
5
6
Myeloid progenitor differentiation potential
(acute myeloid leukemia)
Hematopoietic stem cell generation
(primary immunodeficiency)
−77
Gata2
+9.5
TGATAA
TGATAG
E-box
spacer
AGATAA
GATA-2
A B C
Runx X Y Z
Prog
Mye
Ery
DA
HSC
B
C
Runx1
Scl/Lyl1
Lmo2
Gata2
Fli1/Erg
Dii
HES1
HES1
HES1
Repression
Gata2
Activation
NOTCH
NICD
Hes1
Di
Dorsal Side
Notch low
Notch high
Ventral Side
Ei
BMP4
n
da
cv
cv
m
m
g
g
100µm
Eii
BMP4
e
da
50µm
Fi
BMP4
e
da
50µm
Gi
Ly-6A GFP
BMP4
e
da
50µm
Fii
BMP4
da
20µm
Gii
Ly-6A GFP
BMP4
da
20µm

Figure 5 – Gata2 at the centre of a haematopoietic web – A) Schematic of *Gata2* mouse genomic locus with the *-3.9*, *-2.8*, *-1.8* (upstream of exon and promoter *1S*) and *+9.5* (between exon *4* and *5*) *cis*-/enhancer-elements, and all seven exons as indicated. Part of the *+9.5* enhancer with the composite *E-box-spacer-GATA* is indicated above the schematic. A couple of reverse complement GATA motifs (TATC; red) are also present in the enhancer. Adapted from Gao et al. (2013). **B)** *Gata2*'s *-77* and *+9.5* enhancer elements have distinct and essential regulatory roles in haematopoiesis. The *-77* element (left) is needed for successful differentiation of erythromyeloid progenitors (Prog) into certain erythroid (Ery) and myeloid (Mye) lineages. The *+9.5* enhancer (right) is crucial for haematopoietic stem cell (HSC) induction in the mouse midgestation dorsal aorta (DA). Mutational dysregulation of the *-77* and *+9.5* in humans is associated with acute myeloid leukaemia and MonoMAC immunodeficiency, respectively. Reprinted from Johnson et al. (2015). **C)** In both mouse and human the combinatorial interaction and simultaneous co-expression of 'heptad' transcription factors is fundamental to haematopoietic stem and progenitor cell (HSPC) specification and generation. Green arrows indicate novel interactions between Runx1 and Scl/Gata2/Erg. Lmo2 is a non-DNA-binding factor which associates with Scl/Lyl1 and Gata2. The order of proteins in this schematic does not necessarily reflect the most common arrangement. Reprinted from Wilson et al. (2010), with permission from Elsevier. **D)** Notch signalling activity and pathway in mouse aorta-gonad-mesonephros (AGM) region. **Di)** In order for the endothelial-to-haematopoietic transition, and intra-aortic haematopoietic cluster (IAHC) and HSC formation to occur within the mouse AGM region, the ventral aspect of the aorta needs to be exposed to Jag1-Notch1-driven low levels of Notch activity (red cells). Dll4-Notch1 interaction yields high Notch activity that maintains an arterial endothelial programme in the dorsal aspect of the aorta (blue cells). **Dii)** Activation of the Notch receptor leads to proteolytic cleavages (scissor) and the Notch intracellular domain (NICD) migrating to the nucleus where, together with other proteins, it induces expression of target genes *Gata2* and *Hes1*. Hes1 will in turn block *Gata2*

expression. Reprinted from Bigas and Porcheri (2018), with permission from Springer Nature. **E,F,G)** Bmp4 expression in the E11 mouse AGM. **Ei)** Bmp4 protein expression (green) in transverse section through the AGM region is strongest in the ventral subaortic mesenchyme. **Eii)** Higher magnification of boxed area in **Ei** showing a strong dorsoventral gradient with the highest Bmp4 expression seen in the ventral mesenchyme (arrowheads). Blue=DAPI. **Fi)** *In situ* hybridisation for *Bmp4* (purple) shows a similar dorsoventral gradient as seen for protein expression, with the highest expression in the ventral mesenchyme (arrowheads). **Fii)** Higher magnification of boxed area in **Fi** revealing *Bmp4* expression in mesenchymal cells (arrowhead) directly underneath an IAHC (arrow). **Gi,ii)** Bmp4 expression (red) is detected in the ventral mesenchyme (arrowheads) and directly underneath Ly6A-GFP^{+} (green) endothelial cells and **Gii)** IAHCs (arrow). **Gii)** Whereas mesenchymal cells (arrowhead) reveal cytoplasmic Bmp4 expression, IAHCs (arrow) appear to have foci of Bmp4 on their cell surface. Da=dorsal aorta; cv=cardinal vein; m=mesonephros; g=gonad; n=notochord; e=erythrocytes. **E,F,G)** Reprinted from Durand et al. (2007), with permission from The National Academy of Sciences of the USA. Copyright (2007) National Academy of Sciences, U.S.A.

Heterogeneity amongst HSCs and their precursors – Desperately seeking (their) identity

HSCs are rare from the moment they are first generated in the AGM until the time they home to the adult BM, where they will reside for the rest of their life. The long-standing challenge both in the field of developmental and adult haematopoiesis has been to isolate pure populations of immunophenotypic and functional long-term (LT-) multilineage repopulating HSCs. Despite great enrichment (e.g. Kiel et al. (2005)), defined multipotent "HSC" populations continue to show varying heterogeneity on functional, phenotypic and transcriptomic levels both at the time of their first generation (reviewed in Crisan and Dzierzak (2016)) and during adulthood in the BM (reviewed in Haas et al. (2018)). Consequently, the designation "LT-HSC", historically considered a homogenous population, needs re-evaluation.

Functional heterogeneity

From the 1990s onwards functional heterogeneity based on prototypical HSC characteristics was shown in BM HSC populations by limiting dilution transplantations or use of clonally marked HSCs. This revealed the coexistence of various types/states of HSCs with different lineage-bias, self-renewal, repopulating, proliferation, and differentiation potentials (reviewed in Jurecic (2019); Muller-Sieburg et al. (2012)). One study exposing such heterogeneity performed single-cell transplantations and tracked the clonal repopulation kinetics/output of adult BM cells of which >25% had previously been shown to be LT-HSCs (Dykstra et al., 2007). Retrospectively, it was demonstrated that each HSC clone gave rise to one of four (alpha [α], beta [β], gamma [γ], delta [δ]) distinct repopulation patterns. The patterns range from a more balanced (i.e. contributing equally to generation of all mature haematopoietic lineages), to a myeloid- or a lymphoid-biased output, with some cells maintaining these properties upon serial transplantation. Of note, all four types produced multilineage

chimaerism, however their ratios of lymphoid and myeloid cell production differed.

In addition to LT-HSCs, intermediate-term (IT) and short-term (ST) HSCs have been identified in repopulation assays. These BM HSCs can also produce all haematopoietic lineages, however their self-renewal capacity, lifespan (from a couple to >12 months) and lineage-output ratios differ from LT-HSCs: LT are mostly myeloid-biased, whereas IT are balanced or lymphoid/myeloid-biased and ST are mostly lymphoid-biased HSCs (reviewed in Ema et al. (2014)). These classifications, though not entirely overlapping, relate to the retrospectively defined α, β, γ, and δ, and My-(biased), Bala(nced) and Ly-(biased) HSC types, the latter grouping identified by the Muller-Sieburg group who were one of the first to address heterogeneity in the HSC compartment (reviewed in Ema et al. (2014)). Data on My-, Bala- and Ly-subtypes show these HSCs actually produce normal cell numbers of the lineage towards which they are biased, but have a reduced capability to produce other lineages. Furthermore, their potential seems epigenetically fixed and inherited by daughter HSCs, as shown by several rounds of transplantations (reviewed in Muller-Sieburg et al. (2012)).

Recently, an adult megakaryocytic/platelet-biased HSC was discovered near the top of the haematopoietic hierarchy (Rodriguez-Fraticelli et al., 2018; Sanjuan-Pla et al., 2013; Shin et al., 2014). It is within the myeloid-biased fraction, thrombopoietin-dependent and is isolated based on Von Willebrand factor or cKit high level expression (Sanjuan-Pla et al. (2013); Shin et al. (2014); reviewed in Jurecic (2019)). LSK SLAM HSCs with high cKit expression differentiated more easily into megakaryocytes (megakaryocytic-biased), showed increased cell cycling and faster multilineage differentiation, but reduced long-term repopulation and self-renewal. Low cKit expressing cells (within the LSK SLAM population) displayed opposite characteristics (Shin et al., 2014). cKit expression levels also underline a difference between adult BM and AGM HSCs, with the AGM only generating cKit^{high}-expressing HSCs (Sanchez et al., 1996). The adult unipotent megakaryocytic-biased HSC is proposed to directly originate from

multipotent HSCs without going through an intermediate oligopotent state and might even be the major *in vivo* fate of LT-HSCs during normal/unperturbed haematopoiesis, given platelets and erythrocytes make up >90% of the cells produced in the BM each day (Rodriguez-Fraticelli et al. (2018); reviewed in Jurecic (2019)).

Prospective enrichment of lineage-biased HSCs could also be achieved by differential expression of the SLAM marker CD150, with $CD150^{high}$ HSCs predominantly giving a myeloid and $CD150^{low}$ a balanced/lymphoid output (Beerman et al., 2010). Furthermore, they showed myeloid-biased $CD150^{high}$ HSCs start to predominate within the heterogeneous HSC pool as animals age. Indeed, ageing coincides with augmented myelo- and megakaryopoiesis, alongside increased numbers of megakaryocytic- and myeloid-biased HSCs (reviewed in Haas et al. (2018)). Morita et al. (2010) also prospectively enriched for biased BM HSCs based on $CD150^{high/medium/negative}$ expression to yield similar results. They importantly showed that $CD150^{high}$ HSCs had high self-renewal capacity whereas $CD150^{medium/negative}$ had diminished self-renewal potential and a shorter lifespan. Additional heterogeneity existed within the most potent $CD150^{high}$ LT-HSC fraction, with the presence of a possible “latent HSC”, i.e. latent/dormant during primary, but multilineage in secondary transplantations (Morita et al., 2010).

Molecular heterogeneity and HSC signatures

Building on these and other studies that had allowed for adult HSC isolation of up to 40-50% purity (reviewed in Haas et al. (2018)), Wilson et al. (2015) addressed HSC heterogeneity by isolating cells based on five commonly used BM HSC enrichment methods. Using single-cell gene expression analysis, single-cell RNA sequencing (scRNA-seq) and indexomics (FACS index sorting) they discovered vast functional and transcriptomic heterogeneity amongst adult BM HSPCs. Hypothesising there must be an overlapping and functionally homogenous cluster/gene signature

of true HSCs between all five isolation methods, and considering the different published repopulation probabilities, indeed a relatively small fraction of cells molecularly overlapped (MolO) to show homogenous gene expression. ScRNA-seq was performed on the population that contained the highest percentage of durable long-term repopulating cells. Even within this small group of cells substantial heterogeneity was observed by Principal Component Analysis (PCA), with one principal component separating cells. One cluster contained cells primed for proliferation (no molecular overlap; NoMO), the other showed upregulated genes that would negatively regulate proliferation (MolO HSCs). Next, by using a combination of up- and downregulated surface markers found on MolO cells (CD150^{high}, Ly6A^{high}, CD48$^{negative/low}$, and CD201/Procr/Epcr^{high}) they functionally tested and separated MolO cells from 'contaminating' cells within their bulk 'HSC' population. Based on index data of single isolated cells for scRNA-seq, i.e. surface protein expression levels for each cell recorded during FACS, and functional data from single-cell transplants, a positive correlation between the MolO signature and true HSC identity was revealed. Thus, succeeding in further separating transplantable multilineage HSCs from FACS-marker-defined HSCs that lack HSC activity. Altogether, this combination of strategies resulted in over 65% transplantation efficiency by single-cell transplantation experiments (reviewed in Wilson and Gottgens (2018)). The MolO HSC gene signature score was later improved to an "hscScore" (Hamey and Gottgens, 2019).

Metabolic heterogeneity

Additional heterogeneity appears when discriminating between dormant and active HSCs. Dormant HSCs make up roughly 20-30% of BM HSCs, have a very low biosynthetic activity, only divide once every 149-193 days and show a high self-renewal capacity upon transplantation (reviewed in Crisan and Dzierzak (2016); Haas et al. (2018); Jurecic (2019)). These adult HSCs do not express CD34 (LSK SLAM CD34^{negative}). The remaining

70-80% are active HSCs that show reduced reconstitution, divide every 28-36 days, have an increased biosynthetic activity and express CD34 (reviewed in Crisan and Dzierzak (2016); Jurecic (2019)). Importantly, the active and dormant state and accompanying up- and downregulation of CD34 are reversible, thereby keeping HSCs in a state of quiescence and minimising replicative stress when not required to actively respond to e.g. stress (reviewed in Jurecic (2019)).

HSC heterogeneity revealed by unique cell barcoding

Using an alternative approach to address the different clones/types of HSCs present in the bulk HSC compartment, Rodriguez-Fraticelli et al. (2020) used their LARRY (Lineage and RNA RecoverY) lentiviral barcoding of LSK SLAM HSCs to allow for simultaneous analysis of mRNA and lineage information of single HSC clones upon transplantation. Their method allowed them to class roughly 60% of (227) LSK SLAM HSCs as “low output” during “transplantation haematopoiesis”. These cells would self-renew significantly more than differentiate (versus “high output” HSCs that showed the opposite) and therefore be more successful in repopulating secondary recipients than “high output” HSCs. Contrary to previous approaches in which megakaryocyte-biased HSCs were found amongst the (cKit^{high}) LSK SLAM HSCs with reduced self-renewal (Shin et al., 2014), Rodriguez-Fraticelli et al. (2020) find them in their “low output” (highest self-renewal) fraction.

The same study also discovered the TF Tcf15 (*Tcf15*) as an important marker in discriminating, what they call, the most immature and quiescent multipotent HSCs from any other LSK SLAM cell. 38.4% of LSK SLAM cells were Tcf15^{+} and these cells, unsurprisingly, also expressed Procr and Sca-1. Limiting dilution transplants with one and five cells revealed roughly one functional HSC per two Tcf15^{+} LSK SLAM cells.

An RNA barcoding system (*PolyloxExpress*) was used to report both fate (comparing HSC barcodes with offspring) and transcriptome of the same single cell (Pei et al., 2020). Within the LSK SLAM cells examined, four

different types of HSC clones were discovered, namely: 1) differentiation-active multilineage HSCs (barcodes found in all lineages), 2) active erythromyeloid-biased HSCs (barcodes found in mostly myeloid and few lymphoid lineages), 3) active erythromyeloid-restrictive HSCs (barcodes found only in myeloid lineages) and, most surprisingly, 4) differentiation-inactive HSCs (clones had no lineage output). None of these fitted characteristics of previously described categories like LT-, ST- or dormant HSCs. PCA classed the different inactive HSCs and multilineage HSCs as furthest apart from each other. Differentiation-inactive HSCs were also not quiescent based on their gene signature and therefore not dormant. Moreover, both the number of differentiation-inactive and active erythromyeloid-restrictive HSCs did not decrease with age.

In summary, functional and transcriptomic single-cell experiments on the adult BM HSC compartment, even in its most highly enriched form, have revealed many different types of HSCs and large amounts of heterogeneity based on prototypical HSC characteristics like self-renewal, repopulation capacity, high/low/biased lineage output, lifespan, division rate, and biosynthetic activity. Additionally, all these different types of HSCs are incompatible with the classical haematopoietic hierarchy tree (Figure 1A,B) and underline why a hierarchical model with more than one possible differentiation trajectory out of a heterogeneous HSC compartment (Figure 1D) would be more accurate. Still, even within stringent phenotypically enriched subtypes of BM HSCs, heterogeneity of HSCs and contaminating (haematopoietic) subsets remains, hindering us in fully understanding the molecular nature of the most potent HSC.

Developmentally restricted HSCs

The heterogeneity seen in the adult HSC compartment begs the question when and where this originates, and whether the different types of HSCs are specified during embryonic development. Indeed, HSC heterogeneity already exists in the mouse E14.5 FL, with the presence of Ly-,

Bala- and My-HSCs. However, contrary to the BM, balanced HSCs outnumber myeloid-biased HSCs in the FL (reviewed in Crisan and Dzierzak (2016)). FL HSCs are actively cycling, highly proliferative and show superior self-renewal compared to their adult BM counterparts (reviewed in Crisan and Dzierzak (2016); Haas et al. (2018)). These differences with adult HSCs raise the question whether there could be developmentally-related mature cell lineages (not in the classical adult HSC hierarchy) that can only be made from certain types of HSCs during embryonic/foetal stages?

Studies on innate-like B-1a cells suggest further heterogeneity within the FL compartment with the coexistence of so-called developmentally-restricted (B-1a producing) (dr)HSCs and LT-HSCs (reviewed in Ghosn et al. (2019)). Controversy still exists about whether B-1a B-cells are generated HSC-independently, -dependently, or both, and if the drHSC actually fits all criteria of a true HSC (reviewed in Ghosn et al. (2019)). The AGM has been shown capable of producing both B-1a and B-2 cells and E9.5 pre-HSCs have been implied as their common precursor (Hadland et al., 2017). Consequently, it has been suggested that upon maturation these pre-HSCs will turn into the so-called FL drHSCs, whereas the E10.5-11.5 *de novo*-generated AGM HSCs become part of or give rise to the LT-HSCs in the FL (Godin et al. (1993); Hadland et al. (2017); reviewed in Ghosn et al. (2019)).

Molecular characterisation of the earliest HSCs

As shown previously, the number of cells with an HSC identity/ potential at the time when these cells are first generated *in vivo* is extremely low. Unlike the adult BM niche where HSCs are rare but number up to several hundred/thousand per adult mouse, the AGM harbours only 1-2 LT-HSCs per embryo amidst about 700 other cells in the IAHCs (Kumaravelu et al., 2002; Yokomizo and Dzierzak, 2010). Some of these other cells are the so-called “pre-HSCs” (Rybtsov et al., 2011; Taoudi et al., 2008). This “pre-definitive HSC” compartment comprises of a series of intermediate cell types/states (T[ype]1 and T2) defined by a combination of cell surface

markers. The maturation (and proliferation) of these cells into transplantable HSCs requires *in vitro* culture steps and it is therefore uncertain whether all the pre-HSCs that have the power to achieve HSC function *in vitro* will actually achieve HSC function and identity *in vivo*. The sparsity of true AGM HSCs has so far prohibited their precise and comprehensive characterisation e.g. cell cycle status, transcriptome, metabolism etc.

One of the first studies to perform large-scale (bulk) transcriptomics during HSC ontogeny in the mouse embryo was McKinney-Freeman et al. (2012). They isolated (pre-)HSPCs/HSCs from E9 YS, E11.5 AGM (VE-Cadherin^{+}CD45^{+}), E12.5, 13.5 and 14.5 FL, E12.5 placenta, and BM. Following generation of gene expression profiles, hierarchical clustering and PCA revealed a transcriptomic segregation into YS-like, the earliest ("specifying") HSCs (AGM, placenta, some E12.5 FL) and later ("definitive") HSCs (some E12.5, E13.5, E14.5 FL, BM). Few transcriptomic differences were observed between E12, E13 and E14 FL HSCs. However, many more differentially expressed genes (DEG) were seen between their E12 "specifying" and "definitive" FL HSCs, suggesting an important transition point, with some E12 FL HSCs resembling an AGM-like transcriptional signature and some a BM one. Overall, AGM and E12 specifying HSCs are distinct, but transcriptionally related. Whereas some (bulk) populations in this study were highly enriched, others were heterogeneous. The AGM 'HSCs' for example included a nonexclusive mix of pre-HSCs, HPCs/CFU-Cs (Rybtsov et al., 2011; Taoudi et al., 2008) and rare HSCs (North et al., 2002). The transcriptomic signature of the first 1-2 true definitive AGM HSCs would have therefore been lost among the signatures of thousands of other cells. While overall transcriptomic differences are seen between the earliest AGM pre-HSPCs, and FL and BM HSCs, with the biggest difference/transition observed in the FL, this study fails to resolve the transcriptomic signature of the first *de novo*-generated LT-HSCs to address the issue of heterogeneity.

Another study performed single-cell (RNA-seq) analysis on slightly earlier developmental stages of haematopoietic development and examined ECs, T1- and T2 pre-HSCs isolated from E11 AGM (Zhou et al., 2016). The

very rare E11 AGM HSCs were not specifically analysed. Instead, E12 and E14 FL HSCs were included. This study was the first to show the molecular signature of pre-HSCs on a single-cell level. Using the novel phenotypic adult BM HSC marker CD201/Procr/Epcr (specifically CD201^{high}; Wilson et al. (2015)) in addition to the previously described markers (Rybtsov et al., 2011; Taoudi et al., 2008), Zhou et al. (2016) managed to enrich for so-called T1 pre-HSCs at a 1:2.3 frequency (or around 11 cells/E11 AGM). Upon co-culture with OP9-DL1 cells, most of these cells gave multilineage reconstitution in lethally irradiated recipients. T2 pre-HSCs could be enriched for at a 1:2.1 frequency (around 18 cells/E11 AGM). In order to better characterise and further enrich the different populations, several candidates from their scRNA-seq were tested. Roughly half of E10.5 AGM ECs (CD31^{+}) appeared to express CD47 (*Cd47*). When cultured on OP9-DL1 and compared to CD47^{negative} ECs, it was only the positive subset that gave long-term multilineage engraftment and was “HSC-competent.” The same was found true for E11.5 CD47^{+} (cKit^{+}) IAHCs as shown by direct transplantation. Another surface marker, CD168 (*Hmmr*), was only found on actively cycling pre-HSCs, but not on quiescent pre-HSCs or BM HSCs. T1 pre-HSCs were both CD168^{+} and CD168^{-}. Like seen in adult HSCs, cell-cycle status presented heterogeneity within functional T1 pre-HSCs. Those in S/G2/M phases showed the highest repopulation potential and best *in vitro* proliferation. The ones in G0 were slightly less potent. The low proliferative G1 T1 pre-HSCs did not manage to repopulate any animals. As described previously, this is contrary to what is seen in most BM and E14 FL HSCs.

Thus, functional and phenotypic heterogeneity is already seen amongst so-called pre-HSCs of which some have the potential to mature into HSCs, and even though direct proof is lacking, the gradual and asynchronous maturation that some pre-HSCs might undergo to become an HSC could be the origin of heterogeneity witnessed in the FL and BM HSC compartment.

Yet, an even earlier origin of heterogeneity might also be possible. Several publications have now shown the side-by-side existence of various

types of haemogenic and non-haemogenic ECs in the midgestation murine aorta before and during the time of active HSPC generation (Chen et al., 2011; Dignum et al., 2021; Hou et al., 2020; Swiers et al., 2013). The *Runx1 +23* enhancer-mediated reporter marks pan-haematopoietic fate of aortic ECs already two days prior to the first HSC induction (Swiers et al., 2013). These +*23*-labelled HECs lose their endothelial potential at the time of hematopoietic commitment, whilst still part of the endothelial wall.

Hou et al. (2020) used a novel Neurl3 (*Neurl3*) reporter in combination with expression of CD31, cKit, Procr and CD44 (and $CD41^{-}CD43^{-}CD45^{-}$), to achieve further HE enrichment than with the *Runx1 23GFP* reporter, and more specifically managed to highly purify "putative-HSC primed HE." The inclusion of pan-IAHC marker cKit makes that these HECs are probably localised at to the base of bigger IAHCs or within the endothelial layer as single bulging cells/1-to-2-cell clusters. Upon *in vitro* culture, these cells displayed haematopoietic-only, endothelial-only, or rare and never-before-seen haematopoietic-endothelial dual potential, the latter even preserved in so-called T1 pre-HSCs at E11. The number of HSC-primed HECs peaked at E10, half a day before HSC emergence. Novel EC types/states and major bifurcations in aortic EC specification were also discovered. Primitive ECs were shown to first choose between an arterial and venous fate, then mature from primitive arterial into early arterial fate, and at that point either bifurcate in late arterial fate or specify into HE. Notch signalling is involved at the early arterial EC stage and goes down again at the HEC stage.

HPCs/EMPs and HSCs were demonstrated to originate from distinct populations of HECs, with *Ly6A* pivotal in HSC-primed HECs (Chen et al., 2011). More recently it was shown that multipotent progenitors (without HSC potential) also arise from HECs that are phenotypically and transcriptionally unique compared to a rarer population of ($VE\text{-}Cadherin^{+}Procr^{+}CD61^{+}$) CXCR4-marked "HSC-competent HE" (Dignum et al., 2021). According to Dignum et al. (2021), the HE described by Hou et al. (2020) would more likely be capable of early multilineage haematopoiesis rather than specifically mark HSC fate. Altogether, HSC heterogeneity does not appear unique to the

adult and various different types/states of HSCs can already be found in the FL. Given the paucity of LT-HSCs in the AGM, it has been challenging to determine whether functional and/or transcriptomic heterogeneity exists already from the moment HSCs are *de novo* generated, or whether the FL environment is responsible for this. Moreover, heterogeneity amongst so-called 'pre-HSCs' and the existence of several types of HECs might be responsible for heterogeneity amongst HSCs prior to their actual generation. Given the stochastic nature of transcriptional regulation as a result of dose-dependent functioning of the right combinatorial TFs and signals, and their stochastic collision, no two HSCs might ever be the same and generation of different types/states could all be dependent on specific 'just-right' levels of intrinsic and extrinsic factors. Ultimately, single-cell functional and transcriptomic analyses of the first AGM HSCs should allow us to resolve whether heterogeneity plays any part in *de novo* HSC induction.

In the span of a little over a century, we have gone from the discovery of the powerful HSC in the BM to high levels of its purification and establishment of its intraembryonic origin. We now know HSC generation occurs via a conserved EHT event that is ultimately reliant on optimal combinatorial levels of the vital intrinsic (TFs) and extrinsic factors. Further understanding of the functioning and dose regulation of pivotal TFs like Gata2 by other intrinsic and extrinsic factors, and more highly enriched single-cell analyses of HECs and IAHCs involved in EHT and HSC generation, should eventually provide an understanding of what it takes to become and remain an HSC.

The studies presented in this thesis explore these topics and address: 1) the essential involvement of Gata2 during mouse embryonic haematopoiesis, 2) the transcriptomic changes required for EHT, 3) the isolation and transcriptomic signature of the very first *de novo*-generated LT-HSCs by reiterations of single-cell analyses, and 4) BMP signalling and how it is responsible for generation of some of the distinct types of HSCs.

Table 1. Timeline of important scientific discoveries in the field of developmental haematopoiesis related to this thesis

Year(s)	*Discovery*	*Reference(s)*
1907-1920	Intra-aortic haematopoietic clusters (IAHC) observed in many vertebrate embryos.	Dantschakoff (1907); Emmel (1916); Jordan (1916); Jordan (1917); Maximow (1909); Sabin (1920)
1920	Careful description of yolk sac (YS) blood islands; the site of the earliest blood formation during vertebrate development.	Sabin (1920)
1932	Term “haemangioblast” coined for a common mesodermal precursor to both haematopoietic and endothelial cells.	Murray and Wilson (1932)
1951	Following irradiation, the damaged haematopoietic system can be rescued by a ‘humoral factor’ present in haematopoietic tissues.	Jacobson et al. (1951); Lorenz et al. (1951)
1956	Substantial proof for cellular (and not humoral) basis of blood production with the identification of a rare transplantable bone marrow (BM) cell with haematopoietic repopulating ability.	Ford et al. (1956); Nowell et al. (1956); Vos et al. (1956)
1957	E. Donnall Thomas performs first allogeneic BM transplantation.	Thomas et al. (1957)
1961	BM contains blood-forming cells (spleen colony-forming units; CFU-S) that form	Till and McCulloch (1961)

	macroscopic spleen colonies post-transplantation. Their number correlates with the number of injected BM cells.	
1963	Spleen colonies are clonal (derived from one CFU-S) and serially transplantable, giving rise to the principle of "self-renewal" and first valid use of term "haematopoietic stem cell" (HSC).	Becker et al. (1963); Siminovitch et al. (1963)
1964	Heterogeneous composition of spleen colonies leads to a stochastic model for the generation/proliferation of CFU-S (HSC) and haematopoietic differentiation ("haemopoiesis engendered randomly"), and introduces stochasticity into the field.	Till and McCulloch (1980); Till et al. (1964)
1967, 1968	(Some) CFU-S are multipotent (though no direct proof of their ability to generate functional lymphoid cells). Lymphoid cells are either descended from CFU-S or a common progenitor.	(Fowler et al., 1967); Wu et al. (1967); Wu et al. (1968)
1970	Mouse YS is the earliest tissue containing cells with multilineage potential, establishing a YS origin for vertebrate HSCs.	Moore and Metcalf (1970)
1975, 1978	The intraembryonic area (of chick) and not the YS contains the cells capable of sustaining lifelong haematopoiesis.	Dieterlen-Lievre (1975); Lassila et al. (1978)
1975, 1981	Murine foetal liver (FL) is haemogenically inactive and colonised by HSPCs generated elsewhere.	Houssaint (1981); Johnson and Moore (1975)
1977	First conclusive evidence of existence of a multipotent (capable of producing both myeloid and functional B- and T-cells)	Abramson et al. (1977)

	self-renewing BM HSC (by radiation-induced chromosomal marking).	
1984, 1986 1988	Prospective isolation of highly enriched BM HSCs based on physical characteristics and novel surface markers (Sca-1/Ly6A, lineage depletion) by fluorescence-activated cell sorting.	Muller-Sieburg et al. (1986); Spangrude et al. (1988); Visser et al. (1984)
1984, 1985 1986, 1990	Retroviral marking of BM cells (more efficient and less harmful than radiation-induced methods) allows for fate tracing of single HSCs and shows their ability to long-term self-renew, clonally expand and produce multilineage progeny.	Keller et al. (1985); Keller and Snodgrass (1990); Lemischka et al. (1986); Williams et al. (1984)
1990	NF-E1, NF-E1b and NF-E1c become known as Gata1, Gata2 and Gata3.	Orkin (1990)
1991, 1992	cKit established as BM HSC marker, yielding the well-known “LSK” ($Lin^-Sca\text{-}1^+cKit^+$) population.	Ikuta and Weissman (1992); Okada et al. (1992); Okada et al. (1991)
1994	Gata2 plays a fundamental role at the very top of the definitive haematopoietic hierarchy.	Tsai et al. (1994)
1994	BM HSCs are functionally heterogeneous and include long-term (LT-) and short-term (ST-)HSCs and multipotent progenitors (MPP).	Morrison and Weissman (1994)
1994, 1996	The embryonic day 10.5 mouse aorta-gonad-mesonephros (AGM) (and not the YS) contains the first autonomously generated multilineage, long-term self-	Medvinsky and Dzierzak (1996); Muller et al.

	renewing HSCs, that can be isolated based on $CD34^{+}cKit^{+}$ marker expression.	(1994); Sanchez et al. (1996)
1998	Chick IAHCs originate from haemogenic endothelium (HE).	Jaffredo et al. (1998)
1999	*Gata2* expression in aortic endothelial cells in the AGM is driven by its specific *1S* promoter.	Minegishi et al. (1999)
1999, 2000	The midgestation mouse aorta, vitelline and umbilical artery are the most potent source of adult HSCs and HPCs.	de Bruijn et al. (2000b); Godin et al. (1999)
2000	The mouse FL greatly expands the embryonic HSC pool.	Ema and Nakauchi (2000)
2002	Ly6A-GFP expression marks all functional HSCs in the AGM and locates them to the aortic endothelial lining. The AGM contains only one or two HSCs.	de Bruijn et al. (2002); Kumaravelu et al. (2002); North et al. (2002)
2004, 2007	Retrospectively, adult BM HSCs can reveal distinct lineage bias repopulation/ differentiation patterns: balanced, myeloid-biased, or lymphoid-biased (or somewhat overlapping: alpha [α], beta [β], gamma [γ], delta [δ]).	Dykstra et al. (2007); Muller-Sieburg et al. (2004)
2004	*Gata2*'s role in *de novo* HSC generation and expansion is dose-dependent with haploinsufficiency causing a severe reduction of AGM HSCs and those produced are qualitatively defective.	Ling et al. (2004)
2005	"LSK SLAM" BM HSC enrichment strategy (0.0058% ± 0.0012% of whole BM; 1 in 2 cells is LT-HSC) established	Kiel et al. (2005)

	by positive and negative selection for CD150 and CD48, respectively.	
2007, 2009	The first mouse HSCs are autonomously generated within the ventral middle segment of the AGM.	Mascarenhas et al. (2009); Taoudi and Medvinsky (2007)
2008, 2009	Murine midgestation aortic endothelial cells, and not underlying mesenchyme, generate HSCs that give rise to the adult definitive haematopoietic system long-term. *Runx1* is essential for both IAHC and HSPC generation.	Chen et al. (2009); Zovein et al. (2008)
2009	Blood can be directly formed from *in vitro*-derived endothelium.	Eilken et al. (2009); Lancrin et al. (2009)
2010	Visualisation of the *in vivo* live formation of single HSPCs from the aortic endothelial wall of mouse and zebrafish embryos. First mention of the "endothelial-to-haematopoietic transition" (EHT)	Bertrand et al. (2010); Boisset et al. (2010); Kissa and Herbomel (2010)
2010	A multimeric heptad transcription factor complex consisting of Gata2, Scl, Lyl1, Runx1, Lmo2, Erg and Fli1 is important for HSPC transcriptional control.	Wilson et al. (2010)
2013	*Gata2* is required for HSC generation and survival (*Chapter 1*).	de Pater et al. (2013)
2013	*Gata2's +9.5* enhancer is required for EHT and HSC generation in the AGM.	Gao et al. (2013)
2015	*Gpr56* is one of the most highly upregulated genes during EHT (in	Solaimani Kartalaei et al. (2015)

	mouse) and is essential for HSC generation (in zebrafish) **(***Chapter 4*).	
2015	All AGM HSCs are BMP-activated. The FL and BM contain genetically distinct BMP- and non-BMP-activated HSCs (*Chapter 6*).	Crisan et al. (2015)
2015	Establishment of adult BM HSC score/signature (“MolO score”; later “hscScore”; Hamey and Gottgens (2019)).	Wilson et al. (2015)
2016	Generation of *Gata2-Venus* (*G2V*) reporter mouse. *Gata2* is expressed in all HSCs, but progenitors are generated Gata2-dependently or -independently, the latter being less potent and genetically distinct (*Chapter 2*).	Kaimakis et al. (2016)
2016	Molecular signature of mouse embryonic pre-HSCs revealed on a single-cell level.	Zhou et al. (2016)
2018	Rapid pulsatile level changes observed in aortic Gata2 expression, specifically in single cells undergoing EHT. The first HSCs express medium levels of Gata2(Venus). *Gata2* haploinsufficiency causes a severe reduction in pulsatile and EHT events (*Chapter 3*).	Eich et al. (2018)
2020	Single-cell iterations capture the transcriptome of the first one or two functional HSCs in mouse. Specific Gata2, CD31, cKit and CD27 levels define all HSCs in IAHCs (*Chapter 5*).	Vink et al. (2020)

Overall objective of critical review – main research question, hypothesis & specific aims

This thesis consists of a portfolio of six selected publications that together form a coherent body of work. The overall research question, hypothesis and specific aims that have been driving my research in these publications is/are:

Research question

From the wide interest of HSC *de novo* generation for clinical therapies, the field ask, how is the rare and seemingly difficult-to-achieve HSC identity acquisition different from e.g. the more frequent acquisition of other haematopoietic fates?

Hypothesis

Acquisition of HSC identity is stochastic and reliant on reaching optimal combinatorial levels of pivotal intrinsic (Gata2 and other 'heptad' TFs) and extrinsic (BMP, macrophages) intersecting factors within and near single emerging cells and can occur directly through a HE-to-HSC transition.

Specific aims

1. To examine the role of transcription factor Gata2 *in vivo* during embryonic *de novo* HSC development by transplantation, *in vitro* haematopoietic progenitor assays, and whole-mount and vital imaging with loss-of-function and reporter mouse models.
2. To identify the transcriptome of the first functional *de novo*-generated HSCs on the single-cell level by reiterations of *in vivo* and *in vitro* haematopoietic assays and RNA sequencing of highly enriched aortic cells undergoing EHT.
3. To examine effects of extrinsic growth factor BMP on HSC generation and lineage output by using a mouse reporter model of *in vivo* pathway activation.

Chapter 1 – *Gata2* is required for HSC generation and survival

1.1 Summary of results

Gata2 plays a crucial dose-dependent role in definitive haematopoiesis and is expressed in arterial ECs and intra-arterial clusters at the time of definitive HPC and HSC generation in the mouse embryo (Ling et al., 2004; Minegishi et al., 1999; Minegishi et al., 2003; Robert-Moreno et al., 2005; Tsai et al., 1994). To investigate exactly when Gata2 is required *in vivo*, a conditional knock-out (cKO) approach was taken. Using the *VEC-Cre* strain which directs highly restricted expression in aortic ECs from around E8.5 (Chen et al., 2009), one or two floxed alleles of *Gata2* (*Gata2*$^{f/+}$ or *Gata2*$^{f/f}$) were deleted in *VE-Cadherin*-expressing cells prior to the time of HSC induction. *Gata2* was deleted similarly post-HSC generation by use of the haematopoietic-specific *Vav-Cre* strain, which is not expressed in any aortic ECs or IAHCs at E10.5/11.5 (Chen et al., 2009; Stadtfeld and Graf, 2005). Whereas the germline *Gata2*$^{-/-}$ is embryonic lethal at E10.5-11.5 (Tsai et al., 1994), *VEC-Cre:Gata2*$^{f/f}$ were viable until E14, but suffer from FL anaemia. Compared to wild-type (WT), both the E10 *Gata2*$^{+/-}$ and *VEC-Cre:Gata2*$^{f/+}$ AGM showed a similar decrease in HPCs (Figure 6A). However, due to inefficient DNA recombination (only 31% of *VEC-Cre:Gata2*$^{f/f}$ CFU-Cs had deleted two *Gata2* alleles), the number of HPCs in the *VEC-Cre:Gata2*$^{f/f}$ AGM did not decrease as much as the germline *Gata2*$^{-/-}$. The presence of small numbers of HPCs in the homozygous (c)KO also suggested HPCs can be generated *Gata2*-independently. At E11, the cKO AGM showed a more severe drop in HPC numbers than at E10 (Figure 6A). The E10, 11 and 14 *VEC-Cre:Gata2*$^{f/f}$ FL all showed similarly reduced numbers of HPCs (compared to WT). All CFU-C types were reduced in all tissues. Together, this revealed that most AGM HPCs are generated *Gata2*-dependently.

Transplantations of E11 *VEC-Cre:Gata2*$^{f/f \text{ or } f/+}$ AGM cells into lethally irradiated recipients showed these were unable to repopulate any haematopoietic lineage or tissue (Figure 6B), hereby proving the necessity of a WT dose of *Gata2* in *VE-Cadherin*-expressing cells for the formation of functional LT-HSCs. Besides absence of functional AGM HSCs, E14 FL analysis showed complete lack of live (Annexin V^{negative}) LSK SLAM HSCs in the cKO, illustrating *Gata2* is needed for generation of phenotypic HSCs.

IAHC formation in WT and cKO was visually examined by whole-mount immunostainings of E10 (Figure 6C) and E11 aortae and the UAs. A significant reduction in CD31^{+}cKit^{+} cluster cells was seen in the cKO aorta (38±18 at E10 [Figure 6C, top right]; 58 at E11), as compared to WT (634±43 at E10 [Figure 6C, top left]; 411 at E11). The UA showed a similar drop in numbers (Figure 6C; middle panel). Interestingly, the cKit^{+} clusters that did form had a size of 1-2 (morphologically flat) cells (Figure 6C; bottom right). Flow cytometric analysis of ECs (CD31^{+}cKit^{-}) and HECs at the base of emerging IAHCs (Flk^{+}cKit^{+}) revealed no difference in cell number of these populations between WT and cKO. Flk^{+}cKit^{+} cells in the cKO were also not more apoptotic (by Annexin V staining), altogether indicating *Gata2* is crucial during EHT for formation of IAHCs and HSCs.

To examine *Gata2*'s role after AGM HSC induction, *Gata2* was deleted in *Vav*-expressing haematopoietic cells. *Vav-Cre:Gata2*$^{f/f}$ embryos survived past E16, but no pups were born that had both alleles deleted. At E11, the AGM and FL again showed a significant reduction in CFU-Cs (Figure 6D). This decrease was not as severe as with *VEC-Cre* (Figure 6A) and the numbers of phenotypic HSPCs (by FACS) and CD31^{+}cKit^{+} cluster cells (by whole-mount staining) in the E11 *f/f* AGM were not significantly different from WT. At E14, the *Vav-Cre:Gata2*$^{f/+}$ and *Vav-Cre:Gata2*$^{f/f}$ FL both showed a significant decrease in CFU-C numbers (Figure 6D), which was also reflected in a severe reduction of phenotypic FL HSCs in *Vav-Cre:Gata2*$^{f/f}$. None of these HSCs could form CFU-Cs. FL cells from these cKO embryos were also unable to reconstitute irradiated recipients (Figure 6E). Altogether this revealed that *Gata2* is required for the progenitor-forming

and long-term repopulating potential of HSCs after their generation. Since both functional assays rely on extensive HSPC proliferation (and the tested cell to be viable for 10 days to 4 months), cell cycle status and levels of apoptosis were tested. Whereas the cell cycle was similar in E14 WT and *Vav-Cre:Gata2*$^{f/f \text{ or } f/+}$ HSPCs, *Vav-Cre:Gata2*$^{f/f}$ LSK SLAM HSCs were significantly more apoptotic than the other cells, demonstrating *Gata2* is essential for HSC survival in the FL.

1.2 Major findings

- *Gata2* is essential for HSC generation and continues to be required thereafter for HSC survival, and thus functions during and after EHT.
- *Gata2* is responsible for the majority of EHT events and formation of most IAHCs.
- *Gata2* is important for generation of most HPCs in the AGM and FL, even though some can be produced *Gata2*-independently.

1.3 Author contributions

CSV performed all *in vivo* haematopoietic transplantations, including intravenous injections, irradiation of recipients, blood sampling and transgenic donor mouse line maintenance (Figure 6B,E thesis; Figure 2A, 3D publication). CSV performed fast genotyping PCRs on freshly harvested *VEC-/Vav-Cre:Gata2*^{fl} and germline *Gata2 KO* embryos for *in vitro* and *in vivo* functional haematopoietic assays (Figure 1-3 publication). CSV performed whole-mount 3D confocal imaging on mouse embryos, including embryo generation and harvest, dissection, whole-mount staining, genotyping by PCR, confocal imaging and design of figures (Figure 6C thesis; Figure 2D publication).

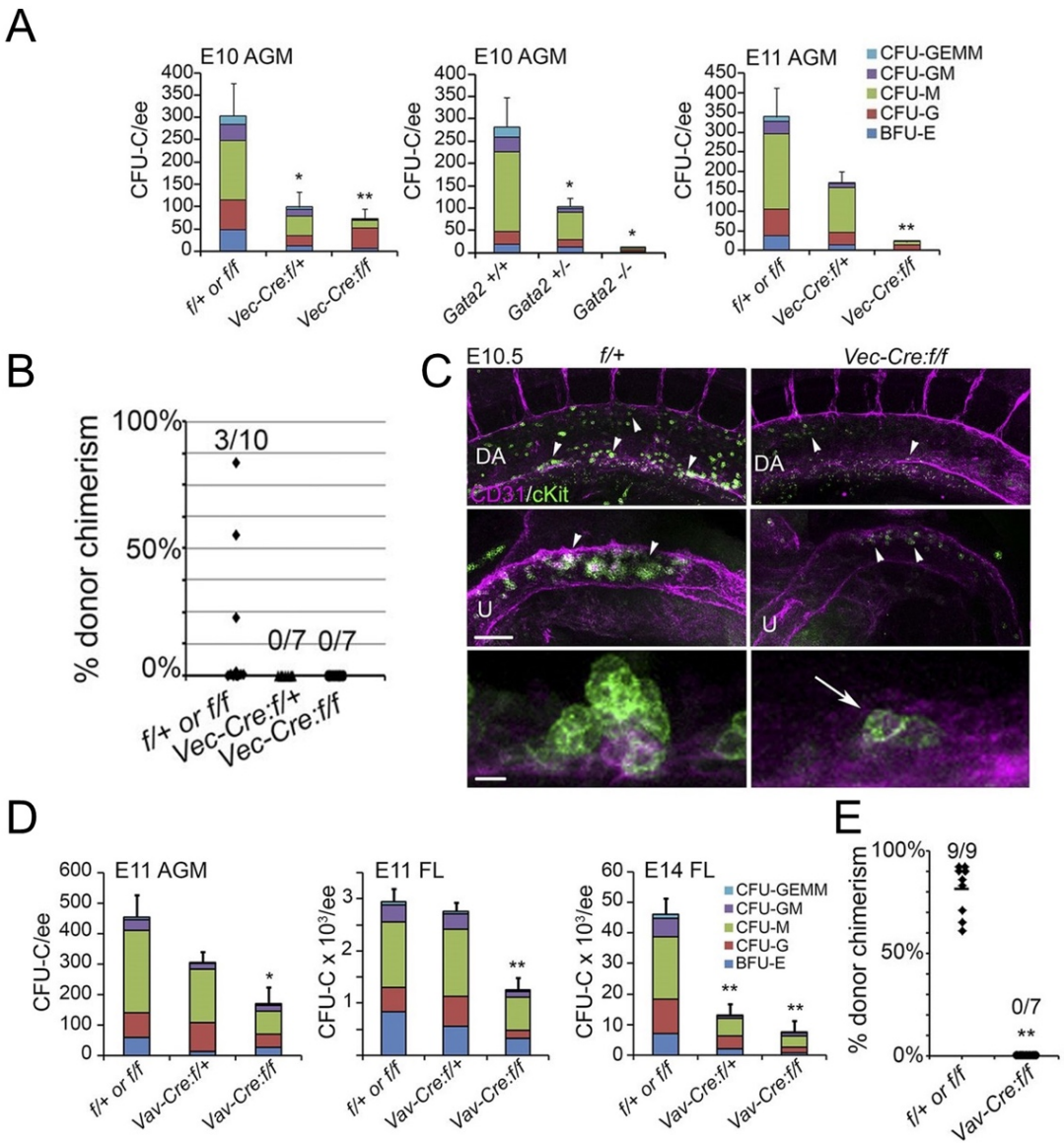


Figure 6 – *Gata2* is required for HSC generation and survival – A) Colony forming unit-culture (CFU-C) numbers per aorta-gonad-mesonephros (AGM) from *Gata2*-deleted embryos. Colony types designated by coloured bars are CFU granulocyte-erythroid-macrophage-megakaryocyte (GEMM); CFU granulocyte-macrophage (GM); CFU macrophage (M); CFU granulocyte (G), and burst forming unit-erythroid (BFU-E). Germline *Gata2*$^{-/-}$ AGMs at embryonic day (E)10 (30-34 somite pairs; sp), n=4. The number of embryos analysed is 8 WT, 11 *Gata2*$^{+/-}$, and 4 *Gata2*$^{-/-}$. For *VEC-Cre*–mediated *Gata2* deletion: WT, *VEC-Cre:Gata2*$^{f/+}$, and *VEC-Cre:Gata2*$^{f/f}$ AGMs at E10 (30-36sp) and E11 (43-48sp). The number of E10 embryos analysed is 5 WT, 6 *VEC-Cre:Gata2*$^{f/+}$, and 6 *VEC-Cre:Gata2*$^{f/f}$; n=3. The number of E11 embryos

analysed is 6 WT, 3 *VEC-Cre:Gata2*$^{f/+}$, and 4 *VEC-Cre:Gata2*$^{f/f}$; n=3. Error bars are mean±SEM of total colony numbers. *P<0.05; **P<0.01, significance determined by Student's t-test. ee=embryo equivalent of specified tissue. **B)** Graph showing the percentage of peripheral blood donor cell chimaerism in adult recipients at 4 months (mo) after transplantation. Recipients were injected with AGM cells from WT (1-2ee; 3 of 10 recipients repopulated; n=3), *VEC-Cre:Gata2*$^{f/+}$ (1-2ee; 0 of 7 recipients repopulated), or *VEC-Cre:Gata2*$^{f/f}$ (1-3ee; 0 of 7 recipients repopulated) E11 embryos. Reconstitution kinetics showed similar outcomes at 1, 2, 3, and 4 mo after transplantation. **C)** Whole-mount immunostaining of E10.5 (36/37sp) WT (*Gata2*$^{f/+}$) and *VEC-Cre:Gata2*$^{f/f}$ embryos showing CD31^{+} vascular endothelial cells (magenta) and cKit^{+} haematopoietic cluster cells (green). The top panels show the dorsal aorta (DA), middle panels the umbilical artery (U), and bottom panels a high-magnification image of a cKit^{+} cluster and aortic endothelium. Arrowheads indicate a few cKit^{+} haematopoietic cluster cells and arrow indicates a flat cKit^{+} cell embedded in the endothelium (n=2; number of embryos analysed is 2 WT and 2 *VEC-Cre:Gata2*$^{f/f}$; significance determined by Student's t test). Bars: 100μm (top and middle); 10μm (bottom). **D)** Graphs showing the CFU-C numbers per WT, *Vav-Cre:Gata2*$^{f/+}$, or *Vav-Cre:Gata2*$^{f/f}$ ee of AGM and FL cells. Error bars indicate mean±SEM. *P<0.05; **P<0.01. The number of E11 AGMs and FLs analysed is 11 WT, 12 *f/+*, and 6 *f/f*; n=4. For E11 AGM, *P<0.05, and E11 FL, **P<0.01. For E14 FL (**P<0.01; n=3), the number of embryos analysed = 10 WT, 5 *f/+*, and 20 *f/f*. Significance was determined by Student's t-test. **E)** Graph showing the percentage of peripheral blood donor cell chimaerism in adult recipients injected with 10^5 WT (n=2; 9 of 9 recipients repopulated) or 10^5 *Vav-Cre:Gata2*$^{f/f}$ E14 FL cells (0 of 7 recipients repopulated, **P<0.01). Diamonds represent individual recipients, with the mean represented by a horizontal bar. Significance was determined by Student's t-test. **(A-E)** Reprinted from de Pater et al. (2013), with permission from Rockefeller University Press.

Chapter 2 – Functional and molecular characterisation of mouse Gata2-independent haematopoietic progenitors

2.1 Summary of results

Several studies have shown the importance of *Gata2* during definitive haematopoiesis and its expression in populations of immature HSCs, HPCs and their precursors (see also *Chapter 1*) (de Pater et al., 2013; Ling et al., 2004; Minegishi et al., 1999; Minegishi et al., 2003; Robert-Moreno et al., 2005; Tsai et al., 1994). During these studies, it had been impossible to prospectively isolate live Gata2-expressing cells during normal haematopoiesis to examine levels and function of Gata2. To overcome this so as to study the function and transcriptome of Gata2-expressing and non-expressing cells, a novel *Gata2Venus* (*G2V*) knock-in reporter mouse was created by recombining the *Venus* fluorochrome gene preceded by an internal ribosome entry site (*IRES*) into *Gata2*'s 3' untranslated region (*UTR*) (Figure 7A). This permitted Gata2 reporting without disturbing Gata2 expression levels or protein function. Indeed, *Venus* and *Gata2* transcripts were only present in Venus-expressing (V^+) BM cells, and equal amounts of Gata2 and LSK cells were found in homozygous $G2^{V/V}$ and WT BM. Competitive limiting dilution transplantations of $G2^{V/V}$ and WT BM cells revealed normal numbers and quality of $G2^{V/V}$ HSC function (Figure 7B). All haematopoietic tissues in E9-11 $G2^{V/+}$ embryos contained V^+ cells as analysed by FACS. Whereas frequencies of V^+ cells in FL, YS and placenta did not change much over time, the AGM showed an increase from 1.8% to 7.9% between E9-11. Immunohistochemistry (IHC) on E9-11 whole-mount embryos and tissue sections confirmed Venus expression in (mostly ventral) aortic endothelial cells and IAHCs, YS BIs and FL (Figure 7C,D). V^+ cells were also detected in the urogenital ridges, neural tube and olfactory bulb. Thus, our *G2V* model is suitable for prospective isolation and

characterisation of Gata2-expressing cells at the time of HSC induction in the embryo.

HPCs/CFU-Cs were found predominantly in the V^+ fraction of E9 and 10 AGM (Figure 7E), YS (Figure 7F), VA+UA, FL and PL. In the AGM, the most primitive (bi- and multipotent) HPCs were exclusively in the V^+ fraction. The other tissues had almost all immature HPCs in the V^+ fraction. The E9 YS contained most Gata2-independently generated (V^-) CFU-Cs. Transplantation of E11 V^+ and V^- AGM cells into lethally irradiated recipients demonstrated all long-term self-renewing multilineage HSCs were confined to the V^+ fraction (Figure 7G). Altogether, all HSCs and immature/primitive HPCs express Gata2, but some HPCs are generated Gata2-independently.

When HPCs were analysed in the context of a *Gata2*$^{-/-}$ or *Gata2*$^{+/-}$ background, E9 AGM (Figure 7H) and YS (Figure 7I), and E10 AGM (Figure 7H), VA+UA and PL all showed significantly reduced numbers of HPCs compared to WT. Like the E9 V^- YS fraction, the E9 *Gata2*$^{-/-}$ YS (Figure 7I) contained most HPCs. The CFU-Cs in any of the *Gata2*$^{-/-}$ tissues were mostly unipotent (granulocyte/macrophage), confirming the *G2V* data.

Differences between Gata2-dependently and -independently generated HPCs were next explored on a transcriptional level. Bulk RNA sequencing (RNA-seq) on $CD31^+cKit^{intermediate}V^-$ and V^+ E10.5 AGM cells (V^- HPCs were mostly $cKit^{intermediate}$) revealed downregulation of Ras signalling pathway genes in V^- cells. Genes in the Notch signalling pathway (e.g. *Notch1* and *Notch4*) were upregulated in the V^+ fraction compared with V^- (Figure 7Ji), confirming *Gata2* as a Notch target (Robert-Moreno et al., 2005). Since most V^- HPCs gave rise to macrophage and granulocytic CFU-Cs, and YS EMPs have been shown to produce macrophages, potential similarities between the two were explored. Flow cytometry revealed most E10 YS and AGM EMPs ($Sca\text{-}1^-cKit^+CD41^+C16/32^+$) were V^-. At E11, overall numbers of EMPs in these tissues decreased and most were now V^+. Several chemokine receptors/ligands known to be transcriptionally lowly (e.g. *Cx3cl1*, *Ccl2*, *Ccl9*) or highly (e.g. *Cxcr4*) expressed by YS EMPs (Kierdorf et

al., 2013), were expressed similarly by E10.5 AGM V^- cells (Figure 7Jii). V^- HPCs therefore transcriptionally resemble EMPs.

Expression of other *Gata* factors was assessed in E10.5 IAHC ($CD31^+cKit^+$) cells and revealed equal levels of *Gata3* in V^+ and V^- cells, but significantly higher *Gata4* expression in V^- cells (Figure 7K). IHC on E10.5 AGM sections confirmed Gata3 expression in aortic endothelial and IAHC cells, some of which also expressed Gata2. Gata3 was also expressed in subaortic mesenchyme. Gata4 was expressed in some endothelial cells. Gata3 and/or 4 might thus be responsible for some HPC function in Gata2-independently generated HPCs.

2.2 Major findings

- Venus levels in the *G2V* model faithfully report Gata2 expression and enable examination of live Gata2-expressing and non-expressing cells.
- Cell isolation based on Venus permits prospective enrichment of all AGM HSCs and the most immature embryonic HPCs.
- Some HPCs do not express Gata2 (V^-)/are generated Gata2-independently and appear functionally and transcriptionally less complex than Gata2-expressing HPCs.
- V^- HPCs are mostly produced in the E9 YS and resemble EMPs.

2.3 Author contributions

CSV performed all IHC for Figure 7C,D (thesis; Fig 2D-F,J-M publication), including embryo generation and harvest, cryosectioning, and confocal microscopy. CSV performed all *in vivo* primary and secondary haematopoietic transplantations (Figure 7B,G thesis; Figure 1D, 3A publication), including intravenous injections, irradiation of recipients and transgenic donor mouse line maintenance. CSV assisted with determination of levels of donor chimaerism (including blood sampling, post-transplant adult

tissue harvests, genomic DNA extraction and semi-quantitative PCR). CSV performed quick genotyping PCRs on freshly harvested embryos for bulk haematopoietic progenitor assays (Figure 7H,I thesis; Figure 4 publication).

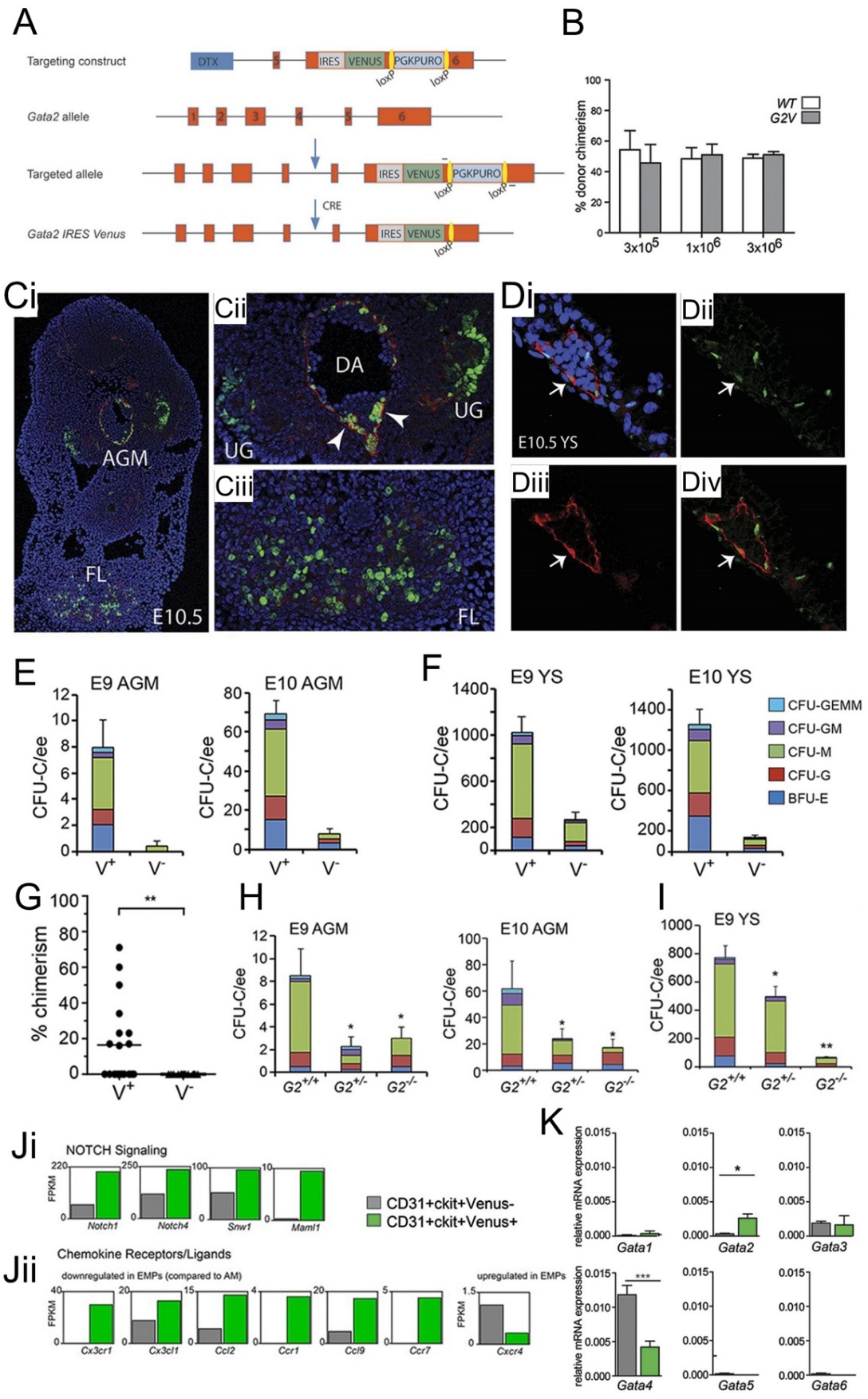

A
Targeting construct
DTX
IRES
VENUS
PGKPURO
loxP
Gata2 allele
Targeted allele
CRE
Gata2 IRES Venus
B
% donor chimerism
WT
G2V
3x10^5
1x10^6
3x10^6
Ci
Cii
Ciii
AGM
FL
E10.5
DA
UG
Di
Dii
Diii
Div
E10.5 YS
E
E9 AGM
E10 AGM
CFU-C/ee
V+
V-
F
E9 YS
E10 YS
CFU-GEMM
CFU-GM
CFU-M
CFU-G
BFU-E
G
% chimerism
**
H
E9 AGM
E10 AGM
G2+/+
G2+/-
G2-/-
I
E9 YS
Ji
NOTCH Signaling
FPKM
Notch1
Notch4
Snw1
Maml1
CD31+ckit+Venus-
CD31+ckit+Venus+
Jii
Chemokine Receptors/Ligands
downregulated in EMPs (compared to AM)
upregulated in EMPs
Cx3cr1
Cx3cl1
Ccl2
Ccr1
Ccl9
Ccr7
Cxcr4
K
relative mRNA expression
Gata1
Gata2
Gata3
Gata4
Gata5
Gata6
***

Figure 7 – Functional and molecular characterisation of mouse *Gata2*-independent haematopoietic progenitors – A) Schematic diagram of the *IRES Venus* reporter selection cassette insertion in the 3' untranslated region (*UTR*) of the mouse *Gata2* locus and Cre-mediated removal of *lox PGK-Puro lox*. Primers used for detection of the targeted and recombined alleles are indicated flanking the *loxP* sites (yellow). **B)** Competitive limiting dilution transplantation strategy used to test the quantity and robustness of $Gata2^{V/V}$ bone marrow (BM) HSCs compared with wild type (WT). Percentage of donor cell chimaerism in adult irradiated recipients co-transplanted with the same number of WT Ly5.1/5.2 and *Gata2Venus* ($G2^{V/V}$) Ly5.2 BM cells. Varying numbers ($1x10^5$, $3x10^5$, $3x10^6$) of BM cells of each genotype were injected, and peripheral blood of recipients was analysed for donor cell engraftment by FACS at 1 and 4 months post-transplantation. N=2 (5 mice per group). **Ci)** Confocal image of a transverse section through the E10.5 aorta-gonad-mesonephros (AGM). DAPI staining (blue), CD31 (red), and Venus fluorescence (green) revealed Gata2 expression in aortic endothelium, haematopoietic cluster cells, urogenital ridges (UG) and the foetal liver (FL). Enlarged images of **Ci)** showing Gata2-expressing cells in **Cii)** dorsal aorta (DA) and UG (arrowheads indicate haematopoietic clusters) and **Ciii)** FL. **Di-iv)** Images of a E10.5 yolk sac (YS) section showing DAPI merged, Venus, CD31, and merged fluorescence. Arrow denotes an endothelial cell expressing Venus and CD31. **E,F)** Haematopoietic progenitor number per tissue in sorted V(enus)$^+$ and V$^-$ cell fractions of **E)** E9 and E10 AGM, **F)** E9 and E10 YS. CFU-C per one embryo equivalent (ee) of tissue is shown. Colony types designated by coloured bars are CFU granulocyte-erythroid-macrophage-megakaryocyte (GEMM); CFU granulocyte-macrophage (GM); CFU macrophage (M); CFU granulocyte (G), and burst forming unit-erythroid (BFU-E). SEM of total CFU-C is shown; 2ee of somite pair–matched tissues were pooled for sorting and yielded 1ee for colony analysis. **G)** Haematopoietic stem cells in sorted V$^+$ and V$^-$ cell fractions of E11 AGM were analysed by transplantation into irradiated adult recipients. Percentage donor cell chimaerism was determined by Venus PCR of peripheral blood DNA at 4

months post-transplantation. Each dot represents one recipient receiving 1.7-6.5ee of AGM cells. Nine of 19 recipients receiving V^+ cells were engrafted (15-71%), whereas none of the 14 V^- recipients showed donor-derived haematopoietic cells; n=7. **P=0.0089. Secondary transplantations: 8 of 12 secondary recipients were repopulated with $3x10^6$ BM cells from primary (V^+) repopulated mice; n=4 **H)** CFU-C numbers per ee found in E9 and E10 AGM, and **I)** E9 YS. *P<0.05; **P<0.01. **J)** Bar graphs of fragments per kilobase million (FPKM) values obtained from RNA sequence analysis of E10.5 AGM *G2V* $CD31^+cKit^{intermediate}V^-$ (gray bar) and $CD31^+cKit^{intermediate}V^+$ (green bar) cells for **Ji)** Notch pathway genes and **Jii)** a selection of chemokine receptor/ligand genes (see Kierdorf et al. (2013); these genes were down-/upregulated in YS EMPs compared with adult microglia [AM]). Gene Expression Omnibus data accession number is GSE76254. **K)** qRT-PCR for expression of *Gata1*, *2*, *3*, *4*, *5*, and *6* transcription factors (normalisation with *Gapdh*) in E11 AGM $CD31^+cKit^+V^+$ (green) and $CD31^+cKit^+V^-$ (gray) cells; n=3. SEM shown with *P<0.05 and ***P<0.001. **(A-K)** Reprinted from Kaimakis et al. (2016), with permission from the American Society of Hematology (Elsevier Science & Technology Journals).

Chapter 3 – *In vivo* single-cell analysis reveals Gata2 dynamics in cells transitioning to haematopoietic fate

3.1 Summary of results

Having established our *G2V* model as a useful tool to study *de novo* HSC generation (*Chapter 2*), and considering that several TFs are known to control crucial cellular functions by dynamic pulsatile expression (reviewed in Levine et al. (2013)), we set out to explore Gata2's *in vivo* expression dynamics during EHT.

First, immunostainings on E10.5 *G2V* fixed whole-mount embryo (Figure 8Ai) and 150 μm-thick vital transverse sections through the AGM (Figure 8Aii) confirmed Gata2 expression in HECs ($CD31^+cKit^-V^+$), endothelial cells bulging/emerging from the aortic wall (BCs; $CD31^+cKit^+V^+$) and IAHC cells ($CD31^+cKit^+V^+$) (Figure 8B). The existence of BCs/IAHCs with medium (V^{med}; 9% of $CD31^+$) and high (V^{high}; 2% of $CD31^+$) Venus levels was shown by flow cytometric analysis on E10.5 AGM (Figure 8C). Bulk RNA-seq confirmed medium and high levels of *Gata2* (Figure 8D) and Western blotting for Venus and Gata2 protein on sorted V^{med} and V^{high} BM cells demonstrated equivalent Gata2:Venus protein ratios (Figure 8E). Thus, the *G2V* model enables precise tracing of *Gata2* transcript and protein expression in live cells.

Next, vital imaging of 49 thick transversal E10 AGM sections (n=15) revealed a total of 13 EHT events (equal to ±20 EHT events/embryo). Analysis of Venus mean fluorescence intensity (MFI) in single cells suggested dynamic pulsatile Gata2 expression in BCs (Figure 8Fi,ii) and, to a lesser degree, in IAHCs (Figure 8Fiii). Quantification of the V^+ EHT cell types (HECs, BCs, IAHCs) revealed similar cell numbers by flow cytometry and microscopy, and showed most V^+ cells were in the aortic endothelium (HECs), followed by BCs and IAHCs. 77% of V^+ cells were on the ventral side of the aorta.

To quantify Gata2 expression dynamics during EHT, and to improve signal-to-noise ratio, confocal images of vital thick embryo truncal sections were deconvoluted and V^+ cells tracked in four dimensions. Single cells were only included if they had been visible for at least ten consecutive frames. Tracing cells in x-, y- and z-dimensions ensured Venus changes were not due to cells moving out of focus. The different EHT cell types revealed various expression patterns/levels. IAHC cells showed significantly higher Venus expression (mean of ten successive time points) than BCs and HECs (Figure 8Gi). Comparing ventral to dorsal cells revealed all ventral subsets expressed significantly higher levels of Venus than their dorsal counterparts (Figure 8Gii). Furthermore, both ventral BCs and IAHCs expressed significantly more Gata2 than ventral HECs, indicating the importance of Gata2 levels during EHT and haematopoietic fate acquisition. To confirm dynamic Venus levels were not due to the imaging procedure, AGM sections were stained with the nuclear dye DRAQ5 that should stably fluoresce during imaging. Indeed, no great intensity changes in DRAQ5 were observed (Figure 8H), confirming Venus level changes were not caused by any imaging noise.

The pulsatile Gata2 expression in single cells was further characterised with custom-made data-processing software. This allowed for calculation of trough-to-peak amplitude (fold change between peak and preceding trough MFI) and pulse/oscillation periodicity (the time between two peaks) (Figure 8I). Single EHT cells would reveal zero, one, two or three peaks of Venus expression. As cells were tracked for different lengths of time, oscillation data were normalised (Figure 8J). This revealed significantly more HECs with no peaks compared with BCs and IAHC cells, and thus that Gata2 expression was more dynamic and pulsatile in the latter two (Figure 8J). Periodicity quantifications showed BCs had significantly more 1-/2-hour periodicities than HECs. On the other hand, HECs had more 3-hour periodicities than BCs and IAHCs. Furthermore, whereas HECs and IAHCs had 2-/3-hour periodicities, most BCs had 1-/2-hour periodicities (Figure 8K). All in all, BCs have the most dynamic Gata2 pulsatile/oscillatory expression

as they undergo EHT. As *Gata2* functions dose-dependently, the maximum Venus MFI was analysed for the different EHT cell types. All ventrally localised cells had higher peak intensities. Strikingly, the trough-to-peak amplitude fold change of ventral BCs was significantly higher than that of ventral HECs and IAHCs (Figure 8L). The high fold change in Gata2 expression in BCs suggests up- and downstream regulators might also (need to) be more variable to acquire specific haematopoietic fates.

Imaging data had revealed different levels of Venus expression in BCs and IAHCs. To functionally explore this, the ($CD31^{+}cKit^{+}$)V^{+} fraction was divided into V^{med} and V^{high} HPCs. Most importantly immature CFU-GEMMs were highly enriched in the V^{med} fraction (Figure 8M). This was also the only fraction that contained LT-HSCs (Figure 8N). Both fractions were further characterised by bulk RNA-seq which revealed upregulation of mature myeloid cell genes in V^{high} cells and genes involved in leukocyte extravasation and epithelial adherence junction pathways in V^{med} cells. Pathway-component analysis revealed that compared with V^{high} cells, V^{med} cells expressed significantly more *Notch1*, a direct upstream regulator of *Gata2* and required for IAHC generation (reviewed in Bigas and Porcheri (2018)), and its target *Hes1*. The haematopoietic heptad TFs were expressed in both V^{med} and V^{high} cells. *Lyl*, *Fli1* and *Erg* were more highly expressed in V^{med} cells, and *Lmo2*, *Runx1* and *Scl* showed higher expression in V^{high} cells. Altogether these data, 1) support functional haematopoietic data to show that the V^{med} fraction contains the most immature cells, 2) confirm a role for Notch during EHT and 3) reveal molecular heterogeneity based on heptad TFs.

Lastly, expression dynamics were assessed in *Gata2* haploinsufficient embryos. $G2^{V/V}$ mice were crossed with $G2^{+/-}$ to yield $G2^{V/-}$ and $G2^{V/+}$ embryos: $G2^{V/-}$ contain one intact and one mutated *Gata2* allele. Time-lapse imaging revealed equal numbers of HECs in $G2^{V/-}$ and $G2^{V/+}$ aortae, however BCs and (particularly) IAHCs were much reduced in the $G2^{V/-}$ aorta. The remaining V^{+} cells were mostly ventrally located (Figure 8O). The Venus MFI (mean of ten successive time points) of HECs in $G2^{V/-}$ and $G2^{V/+}$ was similar, however, contrary to the $G2^{V/+}$ (Figure 8P,Gi), Venus expression in $G2^{V/-}$ was

highest in BCs instead of IAHCs (Figure 8P). *G2*$^{V/-}$ and *G2*$^{V/+}$ embryos also had equal numbers of oscillating HECs, but the number of oscillating BCs seemed to be reduced in the *G2*$^{V/-}$. Despite similar numbers, twice as many *G2*$^{V/-}$ HECs displayed pulse periodicities of ≤2 hours compared with *G2*$^{V/+}$. These HECs also showed higher Venus peak intensities. Like *G2*$^{V/-}$ HECs, the BCs in these embryos showed a trend of reduced pulse periodicities and significantly lower trough-to-peak amplitudes compared with *G2*$^{V/+}$ BCs. Altogether, *Gata2* haploinsufficient embryos show more frequent and intense Gata2 oscillation/pulses at an earlier stage of the EHT and are thus dynamically restricted in their Gata2 expression (Figure 8Q). These data support the notion that threshold levels of Gata2 expression together with high frequency oscillations are required for the transition of HECs to IAHC cells.

3.3 Major findings

- The *G2V* model enables precise tracing of *Gata2* transcript and protein expression in single live cells.
- Roughly 20 EHT events (CD31^{+}cKit^{-}V^{+} into CD31^{+}cKit^{+}V^{+}) are observed in an E10 *G2V* embryo, making Gata2 a broader EHT marker than Ly6A-GFP which only sees 2 EHT events/embryo.
- The HEC-to-BC transition sees an increase in Gata2 expression, followed by BCs displaying highly unstable/dynamic pulsatile Gata2 expression as they undergo EHT and acquire haematopoietic fate. A further increase in Gata2 expression is seen from BC to IAHC, after which levels stabilise and periodicity and amplitude decrease.
- The CD31^{+}cKit^{+}V^{med} AGM fraction is highly enriched for (primitive) HPCs and contains all LT-HSCs.
- *Gata2* haploinsufficient embryos show more frequent and intense Gata2 oscillation/pulses at an earlier stage of the EHT and are thus dynamically restricted in their Gata2 expression.

3.4 Author contributions

CSV performed all haematopoietic functional experiments (Figure 8M,N thesis; Figure 5 publication), i.e. *in vitro* colony assays and *in vivo* transplantation experiments (including embryo generation, fine dissections, cell isolations, data analyses, intravenous injections, post-transplant tissue harvests, irradiation of recipients and transgenic mouse line maintenance). CSV isolated all different cell populations (over several experiments) and extracted RNA (plus performed quality control) for bulk RNA-seq (Figure 8D thesis; Figure 1E, 5 publication), including embryo generation, fine dissections and cell preparations for FACS. CSV performed BM harvests and FACS cell preparation for protein level analysis by Western blot. CSV also performed all Western blot analyses and quantification (Figure 8E thesis; Figure S1 publication). Confocal overnight time-lapse imaging (including embryo generation, fine dissection, sectioning of thick live aortic sections) and DRAQ5 staining was performed by CSV (Figure 8H thesis; Figure S3C,D publication). CSV also assisted with some *G2V* live confocal imaging experiments performed by CE. CSV designed functional studies, interpreted data and made figures.

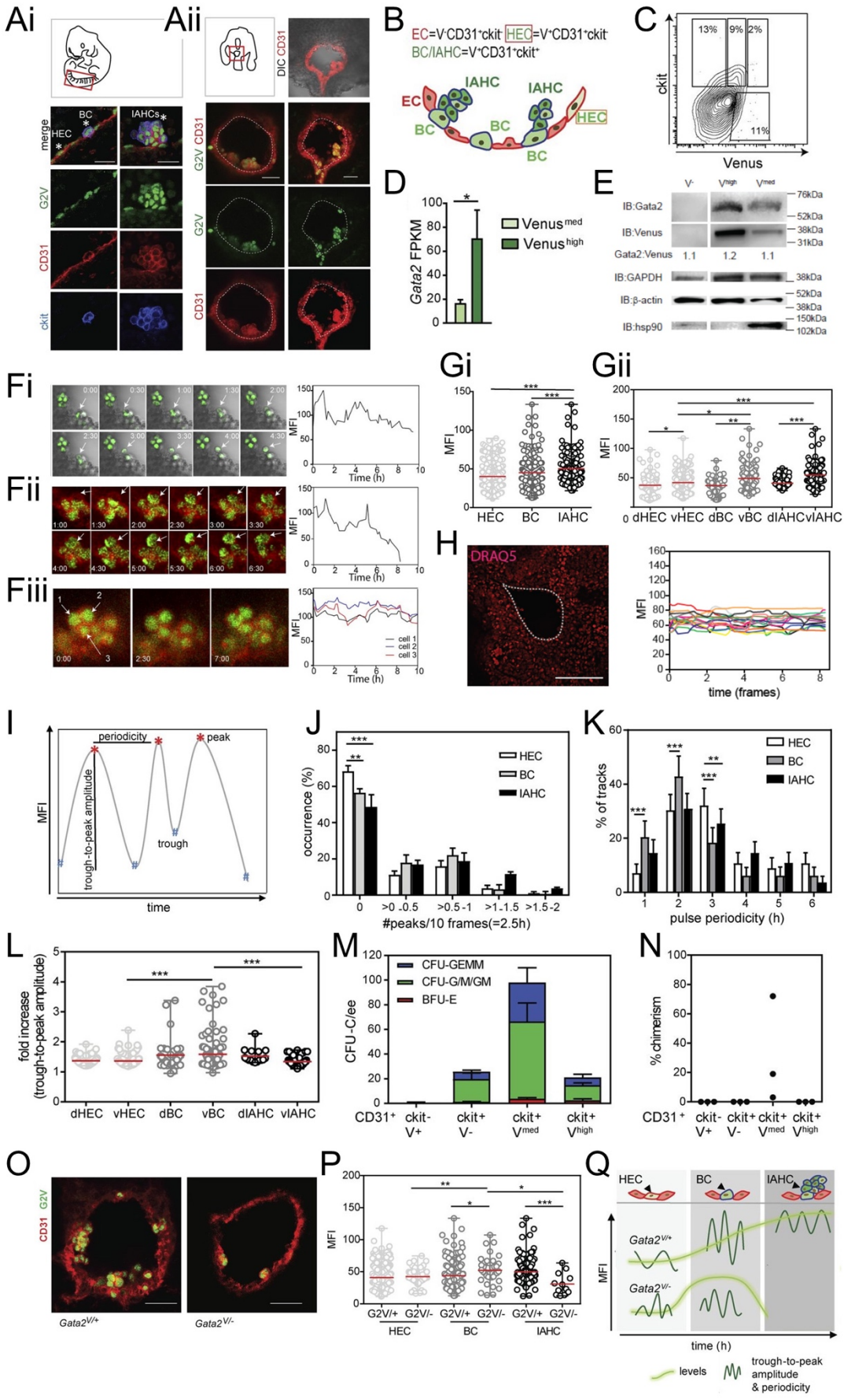

Ai
Aii
B
EC=V-CD31+ckit- HEC=V+CD31+ckit-
BC/IAHC=V+CD31+ckit+
IAHC
IAHC
EC
BC
BC
BC
HEC
C
13%
9%
2%
11%
ckit
Venus
D
Gata2 FPKM
Venus med
Venus high
E
IB:Gata2
IB:Venus
Gata2:Venus
IB:GAPDH
IB:β-actin
IB:hsp90
76kDa
52kDa
38kDa
31kDa
38kDa
52kDa
38kDa
150kDa
102kDa
merge
HEC
BC
IAHCs
G2V
CD31
ckit
DIC CD31
G2V CD31
Fi
Fii
Fiii
MFI
Time (h)
cell 1
cell 2
cell 3
Gi
Gii
HEC
BC
IAHC
dHEC
vHEC
dBC
vBC
dIAHC
vIAHC
H
DRAQ5
time (frames)
I
periodicity
peak
trough
trough-to-peak amplitude
time
J
occurrence (%)
#peaks/10 frames(=2.5h)
K
% of tracks
pulse periodicity (h)
L
fold increase (trough-to-peak amplitude)
M
CFU-GEMM
CFU-G/M/GM
BFU-E
CFU-C/ee
CD31+ ckit- V+
ckit+ V-
ckit+ Vmed
ckit+ Vhigh
N
% chimerism
CD31+ ckit- V+
ckit+ V-
ckit+ Vmed
ckit+ Vhigh
O
Gata2V/+
Gata2V/-
P
G2V/+ G2V/-
HEC
BC
IAHC
Q
HEC
BC
IAHC
Gata2V/+
Gata2V/-
MFI
time (h)
levels
trough-to-peak amplitude & periodicity

Figure 8 – *In vivo* single-cell analysis reveals Gata2 dynamics in cells transitioning to haematopoietic fate – Ai) Diagram of a whole-mount of a 35-somite pair (sp) embryo with the dorsal aorta outlined in red and confocal images of the ventral aspect of the *G2V* dorsal aorta. Haemogenic endothelial cells (HECs) and bulging cells (BCs) (left) and intra-aortic haematopoietic clusters (IAHCs) (right) are indicated by asterisks (CD31, red; cKit, blue; and Venus, green). Bars=20μm. **Aii)** Diagram of a transversal slice of embryo (33-34sp) prepared for vital confocal time-lapse imaging. Confocal images of representative immunostained transversal *G2V* embryo sections (CD31, red; cKit, blue; and Venus, green). Bars=40μm. Ventral side downward. DIC, differential interference contrast. **B)** Schematic representation of Gata2 (green) expression in the different endothelial-to-haematopoietic transition (EHT) cell subsets in the E10.5 mouse dorsal aorta (ventral aspect; ECs, red, red outline; HECs, light green, red outline; BCs, medium green, blue outline; and IAHCs, dark green, blue outline). Cells with blue outline express cKit in addition to CD31. **C)** Flow cytometric contour plot of CD31-gated cells. Percentages of $CD31^{+}cKit^{+}Venus^{med(ium)}$ and $CD31^{+}cKit^{+}Venus^{high}$ expressing cells from E10.5 *G2V* AGMs. **D)** Bar graph of *Gata2* transcription in E10.5 AGM *G2V*-sorted $CD31^{+}cKit^{+}Venus^{med}$ (light green) and $CD31^{+}cKit^{+}Venus^{high}$ (dark green) cells. Y-axis shows FPKM values obtained from RNA-sequencing analysis. The data represent the mean±SEM of three independent experiments and were compared using a Student's t-test (*P=0.0431). **E)** Representative Gata2 and Venus immunoblot (IB) of sorted $Venus^{-}$, $Venus^{med}$, and $Venus^{high}$ cells of $Gata2^{V/V}$ bone marrow (n=2). Although $Venus^{high}$ and $Venus^{med}$-sorted cells show Gata2 protein, Venus-negative sorted cells do not express Gata2 protein. Venus immunoblotting shows that the ratio of Gata2 protein to Venus protein is close to one in all cases, thus indicating a correlation in the protein levels of Gata2 and Venus. β-Actin, GAPDH, and Hsp90 IBs were performed as protein loading controls but show high variability, suggesting that the sorted fractions contain haematopoietic cell types that differ in metabolism, cytoskeleton, etc. **Fi,ii,iii)** Confocal time-lapse imaging of E10 (33-34sp) *G2V*

embryos (Venus, green) immunostained with anti-CD31 (red) antibody. Arrows indicate cells undergoing EHT events **(Fi, ii)** and three IAHC cells **(Fiii)**. **Fi-iii; right)** Quantification of Venus MFI over time (hours) corresponding to the highlighted cells during the process of EHT **(Fi,ii)** and in IAHC **(Fiii)**. Transverse aortic sections were imaged for 10h at 15-min intervals. **Fi,ii** show 30-min snapshots, and **Fiii** shows snapshots at 0, 2.5, and 7h **Gi)** Venus MFI (averaged over frames 3-12) in single EHT subset cells (HECs, BCs, and IAHCs; n=15, 718 cells) of E10 (32-37sp) *G2V* embryos. The data represent the mean±range. Statistical significance was calculated on the pooled data of 15 independent experiments using Mann-Whitney U test (***, P<0.0001). **Gii)** Venus MFI in single EHT subset cells plotted according to their dorsal (d) and ventral (v) location (n=15, 718 cells). The data represent the mean±range. Statistical significance was calculated on the pooled data of 15 independent experiments using Mann-Whitney U test (*P<0.0288; **P=0.0020; ***P<0.0001). **H; left)** Confocal time-lapse imaging of E10 (33-36sp) transverse aortic section stained with DRAQ5 (red). Bar=100μm. Dashed line indicates the inner endothelial boundary. Transverse aortic sections were imaged for 10h at a time interval of 15min. **H; right)** DRAQ5 MFI of individual cells plotted over time. **I)** Schematic representation of the automatic peak detection code. A local MFI maximum is considered a peak if it has at least a 15% higher intensity than its neighbouring minima. The pulse period is the time between two adjacent peaks and the trough-to-peak amplitude the change between peak (highest value) and the preceding trough (lowest value). **J)** Distribution of the occurrence (percentage) of normalized pulse peak numbers in ECs, BCs, and IAHCs tracked over at least 10 consecutive frames (718 cells). To normalise for differences in track length, the data are presented as peaks per 10 frames (2.5h) and represent the mean±SEM (n=15). Statistical significance was calculated using two-way ANOVA with Bonferroni post-test (**P<0.01; ***P<0.001). **K)** Distribution of the pulse periodicities in EHT subset cells showing at least two pulse peaks (n=15, 86 HECs, 80 BCs, and 55 IAHCs). The data represent the mean±SD. Statistical significance was

calculated on the pooled data (n=15) using two-way ANOVA with Bonferroni post-test (**P≤0.01; ***P<0.001). **L)** Trough-to-peak amplitude in the EHT cell subsets, plotted according to their ventral (v) or dorsal (d) location in the aorta (n=13, cells showing at least one peak: 170 HECs, 151 BCs, and 65 IAHCs). The data represent the mean±range. Statistical significance was calculated on the pooled data (n=15) using Mann-Whitney U test (**P=0.0054; ***P<0.0008). **M)** Haematopoietic progenitor numbers in E10.5 (30-36sp) $CD31^{+}cKit^{+}V^{med}$ and $CD31^{+}cKit^{+}V^{high}$ AGM sorted cells. Colony-forming unit-culture per embryo equivalent (CFU-C/ee) is shown, with colony types designated by coloured bars. BFU-E, burst-forming unit erythroid; CFU-G/M/GM, CFU granulocyte, CFU macrophage, and CFU granulocyte-macrophage; CFU-GEMM, CFU granulocyte-erythroid-macrophage-megakaryocyte. The data represent the mean±SEM of four independent experiments. **N)** Percentage donor cell chimaerism in recipient mice injected with $CD31^{+}cKit^{-}V^{+}$, $CD31^{+}cKit^{+}V^{-}$, $CD31^{+}cKit^{+}V^{med}$, or $CD31^{+}cKit^{+}V^{high}$ sorted E11 (41-49sp) AGM cells. Engraftment at 4 months after transplantation was determined by flow cytometric analysis of Ly5.1/Ly5.2 marker expression of peripheral blood cells. Each dot represents one recipient receiving 1.3 to 4.1ee of sorted AGM cells. The data represent the mean±SD. *P≤0.024; **P=0.0085; ***P=0.0003). **O)** Maximum projections of confocal time-lapse images of E10.5 $G(ata)2^{V/+}$ and $G2^{V/-}$ aortae immunostained with anti-CD31 (red; G2V, green). Bars=40µm. $G2^{V/+}$ and $G2^{V/-}$ embryos were harvested from the same litter. Ventral side downward. **P)** Venus MFI (averaged over frames 3-12) in single $G2^{V/-}$ EHT subset cells (n=6; 75 ECs, 37 BCs, and 12 IAHCs). The data were compared with $G2^{V/+}$ EHT subset cells using one-way ANOVA with Bonferroni post-test (mean±SD; *P≤0.024; **P=0.0085; ***P=0.0003). **Q)** Model of Gata2 expression dynamics and pulsatile characteristics during EHT. EHT cell types (top) are shown with accompanying Gata2 dynamic expression changes during EHT directly below. G2V MFI (bright green) and pulse parameters (dark green sinusoids) are shown for $G2^{V/+}$ (middle) and $G2^{V/-}$ (bottom) EHT subset cells. **(A-Q)** Reprinted from Eich et al. (2018).

Chapter 4 – Whole-transcriptome analysis of endothelial-to-HSC transition reveals a requirement for Gpr56 in HSC generation

4.1 Summary of results

Chapter 1-3 of this thesis examined the role of Gata2 in haematopoietic (stem cell) fate acquisition during mouse embryonic development. To capture the whole transcriptome and identify the molecular programme driving the endothelial-to-HSC-transition we next utilised our potent *Ly6A-GFP* HEC-/BC-/HSC-marker mouse model (Boisset et al., 2010; Chen et al., 2011; de Bruijn et al., 2002; Ma et al., 2002). To get a better qualitative and quantitative understanding of Ly6A-GFP expression in the different EHT subtypes at this time of development, IHC was performed on E10 whole-mount (Figure 9Ai,ii) and transversally sectioned embryos (Figure 9Aiii,iv). *Ly6A* expression was found in some aortic ECs (HECs; $CD31^+/CD34^+GFP^+$; both CD31 and CD34 mark ECs and IAHCs) and IAHC cells ($CD31^+/CD34^+cKit^+GFP^+$) (Figure 9A). Based on morphology and localisation, four different GFP^+ cell types could be distinguished: flat (H)ECs, BCs, and IAHC cells either localised juxtaposed or distal from the endothelium (Figure 9Aiv). Between early and late E10, the total number of aortic GFP^+ cells increased more than 5-fold. At each time point most of those cells were HECs, followed by BCs (Figure 9Aiv). GFP^+ juxtaposed and distal IAHC cells, amongst which HSCs reside, were rarest. Thus, the number of Ly6A-GFP^+ cells in the aorta increases over time; however, the number of HECs that undergoes EHT (and gives rise to GFP^+ IAHC cells) remains rare.

Next, haematopoietic function was assessed for enriched EHT subtypes: ECs ($CD31^+cKit^-GFP^-$), HECs ($CD31^+cKit^-GFP^+$), haematopoietic (progenitor) cells (HCs; $CD31^+cKit^+GFP^-$) and HSCs ($CD31^+cKit^+GFP^+$) (Figure 9Aii,Bi). As expected, both HCs and HSCs contained HPCs, with

immature GEMMs highly enriched in the latter (Figure 9Bii). Upon transplantation, as few as 1-5ee of the $CD31^{+}cKit^{+}GFP^{+}$ fraction was sufficient to provide long-term multilineage engraftment of adult irradiated recipients, while higher cell numbers (4-9ee) of any of the other three populations did not yield any repopulation (Figure 9Ci,ii). Thus, the $CD31^{+}cKit^{+}GFP^{+}$ fraction contains most primitive HPCs and all self-renewing LT-HSCs.

To establish an understanding of the molecular basis of EHT, all cell fractions were sorted for bulk RNA-seq analysis. The presence of *CD31*, *cKit* and *Ly6A* transcripts in cell fractions that had been FACS-isolated based on CD31, cKit and GFP protein expression validated the whole-transcriptome analysis (Figure 9D). Comparing EHT datasets, 530 differentially expressed genes (DEGs) were found, of which 340 were differentially expressed between HECs and HSCs (Figure 9E). The DEGs revealed unique HSC- (Figure 9F), HEC-, and EC-patterns when grouped according to their relative expression in each of the other subsets. Gene Ontology (GO) and Gene Set Enrichment Analysis (GSEA) confirmed endothelial characteristics of the EC fraction. HEC genes showed enrichment in cell migration, cell adhesion, hypoxia and Notch pathway gene sets. The HSC genes were linked to pluripotency, haemopoiesis, histone methylation, haematopoietic progenitor and acute myeloid leukaemia gene sets.

Given their crucial role in regulating cell fate change, expression of TFs was explored. HECs showed significant (false discovery rate [FDR] <0.05) upregulation of TFs *Mecom* and *Elk3* compared to ECs. *Notch1* and *Notch4* were also significantly upregulated. Other upregulated TFs (FDR>0.05) included *Gfi1*, *Sox17*, *Sox7*, *Sox18*, and *Lmo2*. Comparing HSCs to HECs revealed significant upregulation (FDR<0.05) of known HSC-related TFs such as: *Myb*, *Gfi1b*, *Hlf* and *Meis1*, and between HSCs and ECs: *Myb*, *Hlf*, *Runx1*, *Myc*, *Dnmt3a*, *Dnmt1* and *Gfi1*. Novel upregulated TFs in HSCs (versus HEC/EC) included zinc-finger proteins *Zfp445*, *Zfp106* and *Zfp763*, and the erythroid/megakaryocytic factor *Nfe2*. Taken together, RNA-

seq of phenotypically enriched EHT populations confirmed a gradual commitment towards a haematopoietic/HSC fate as HECs undergo EHT.

In line with this, expression of the haematopoietic heptad TFs also increased during EHT (Figure 9G). In order to detect new HSC and/or EHT markers, the 530 (EHT) DEGs were compared with 927 previously published heptad targets (Wilson et al., 2010). Fifty-eight overlapped, of which *Gpr56* (*Adgrg1*) was one of the most highly expressed in HSCs (Figure 9Hi). *Gpr56* also ranked at the top of the list of upregulated (DE) receptor genes during the HEC-to-HSC transition (Figure 9Hi,ii). This G-protein-coupled receptor, not previously associated with *de novo* HSC generation or developmental haematopoiesis, contained a heptad composite enhancer element 37kb upstream of its translational start site. Chromatin immunoprecipitation and qPCR on myeloid progenitors confirmed presence of all heptad TFs at the *-37*-enhancer site (Figure 9I). Furthermore, the enhancer showed activation in two progenitor cell lines following transactivation assays. Similar assays also revealed Runx1, Gata2 and Fli1 could activate the *-37*-enhancer site synergistically. In human HSC-enriched cells, all heptad TFs were found to bind a homologous enhancer element 48kb upstream of *GPR56*. In summary, 1) heptad TFs increase their expression during EHT and are directly upstream of *Gpr56*, and 2) *Gpr56* is one of the most highly upregulated and expressed receptor genes during the HEC-to-HSC transition and in AGM HSCs, with a presumptive similar role in human.

Verification of the spatiotemporal expression of *Gpr56* in HSC-containing IAHCs came from ISH on E10.5 AGM sections. *Gpr56* was highly expressed in the majority of IAHC cells and no/few ECs (Figure 9Ji). Co-expression with GFP in the E10.5 *Ly6A-GFP* AGM was observed in some IAHC cells as well (Figure 9Jii). Given its high conservation across vertebrate species, *Gpr56*'s functional involvement in *de novo* HSC generation was assessed in the zebrafish. Upon *gpr56* morpholino oligo (MO) knockdown (KD), the number of *cmyb*$^{+}$ emerging HSPCs was dramatically reduced compared to WT zebrafish embryos (Figure 9Ki). Performing the same KD in *cd41-GFP* embryos also demonstrated a significant reduction in GFP^{+} cells

(Figure 9Kii). Strikingly, the severe reduction in *cmyb*$^{+}$ HSPCs could be completely rescued by injection of zebrafish *gpr56* mRNA or mouse *Gpr56* mRNA, and even triggered some ectopic *cmyb* expression ventral of the aorta (Figure 9Kiii). Thus, *gpr56* plays a crucial role in HSC emergence/ generation and its functional coding sequence is conserved between mouse and zebrafish. Furthermore, the capability to ectopically initiate HSC induction emphasises its significance in this process.

Finally, a role for *GPR56* in stem cell maintenance was shown by stimulating the 32D-CSF3R mouse stem cell line, which had been transduced with WT or constitutively active (MUT) human *GPR56*, with colony-stimulating factor-3 (CSF-3). Whereas the WT line differentiated into neutrophils and lost blast/stem cell characteristics (as expected upon stimulation with CSF-3), the MUT saw a rise in blast and drop in numbers of differentiated cells (Figure 9L).

Altogether, this whole-transcriptome analysis of EHT reveals *Gpr56* as a novel EHT regulator, and demonstrates how these datasets are useful for discovery of other novel EHT/HSC markers.

4.2 Major findings

- The number of Ly6A-GFP^{+} cells in the aorta increases over time; however, the number of HECs that undergoes EHT (and gives rise to GFP^{+} IAHC cells) remains rare.
- Ly6A-GFP expression in (CD31^{+}cKit^{+}) IAHCs marks all LT-HSCs (0.002% of AGM cells) and most primitive HPCs; however, the Ly6A-GFP^{-} IAHC fraction also contains HPCs.
- RNA-seq of enriched EHT populations confirms gradual commitment towards an HSC fate and shows ECs, HECs and HSCs have distinct but closely related transcriptional programmes.
- Heptad TFs increase their expression during EHT and are directly upstream of *Gpr56*.

- *Gpr56* is one of the most highly upregulated and expressed receptor genes during the HEC-to-HSC transition and in AGM HSCs, with a presumptive similar role in human.
- *gpr56* plays a crucial role in zebrafish HSC emergence/generation and its functional coding sequence is conserved between mouse and zebrafish.
- *GPR56* plays a role in stem cell maintenance.

4.3 Author contributions

CSV performed whole-mount 3D confocal imaging on mouse embryos (including embryo generation, dissection, whole-mount staining; Figure 9Ai,ii thesis; Figure 1A,B publication). CSV cryosectioned entire aortae of several embryos, performed IHC, confocal imaging of all stained sections, manually quantified morphologically distinct cells and managed to class these in four novel self-defined categories (including embryo generation, dissection, fixation and embedding; Figure 9Aiii,iv thesis; Figure 1C,D publication). CSV performed all *in vivo* haematopoietic transplantations (including fine embryo dissections for primary transplants, bone marrow harvests for secondary transplants, cell preparations and stainings for FACS, intravenous injections, irradiation of recipients and maintenance of transgenic mouse lines) and determined levels of donor chimaerism (including blood sampling, harvest of adult organs post-transplant, genomic DNA extraction and semi-quantitative PCR) for all recipients (Figure 9Ci,ii thesis; Figure 1G, S1A publication). CSV helped with cell isolations for bulk RNA-seq (including embryo generation, fine embryo dissections and help with cell preparations for FACS). RNA-seq data from these cells were used in Figure 1H,I, 2-4, 5A-C and S2 (of publication). CSV performed RNA expression analysis by ISH on mouse embryo sections, including design, molecular cloning and generation of anti-sense mRNA probe, embryo harvest, paraffin fixation and embedding, microtome sectioning, immunohistological co-staining and all imaging (Figure 9Ji,ii thesis; Figure 6A,B publication). CSV helped in designing functional

experiments, and designed and interpreted imaging experiments and results. CSV was involved in writing parts of manuscript.

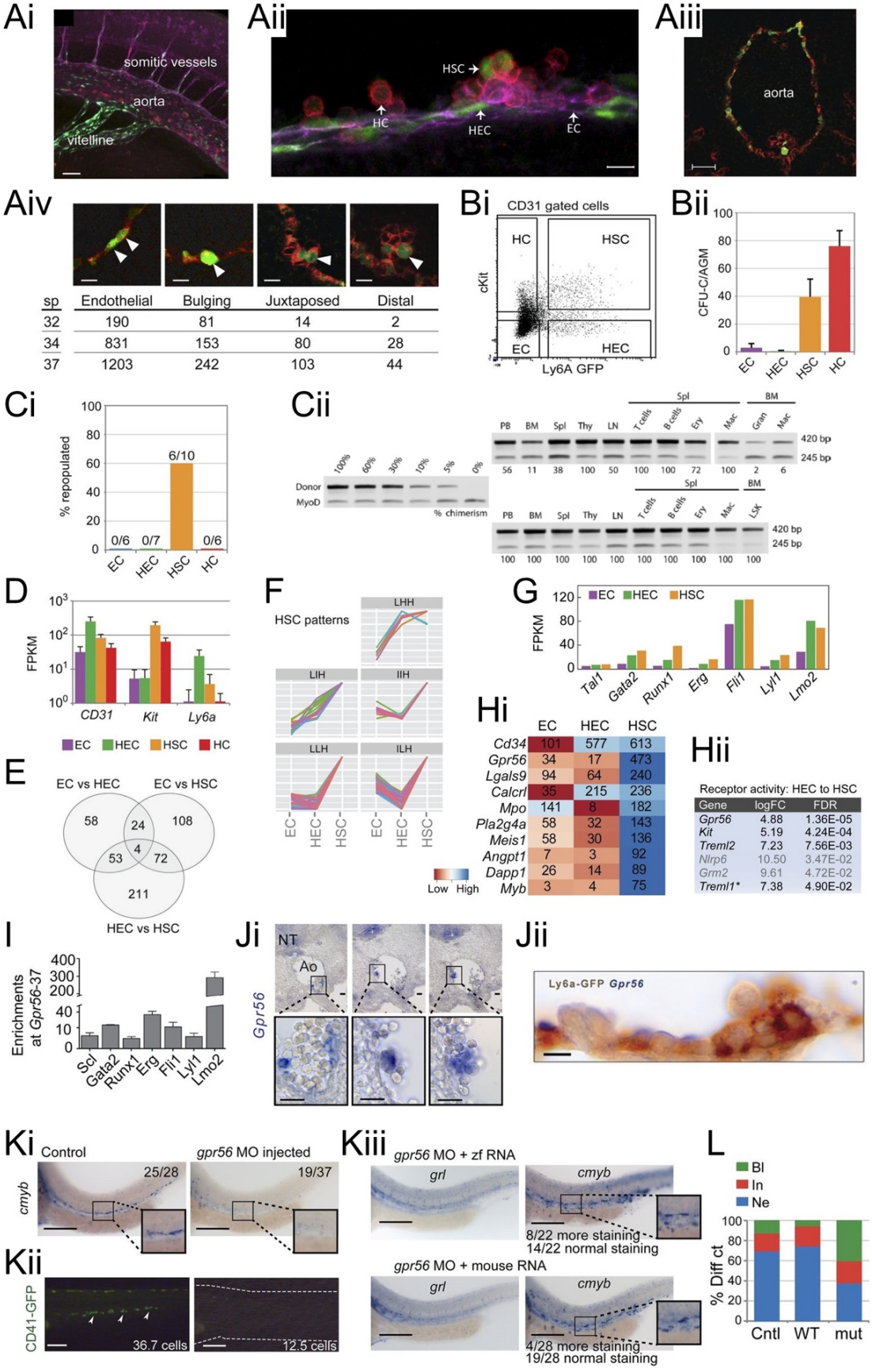

Ai
somitic vessels
aorta
vitelline
Aii
HSC
HC
HEC
EC
Aiii
aorta
Aiv
sp | Endothelial | Bulging | Juxtaposed | Distal
32 | 190 | 81 | 14 | 2
34 | 831 | 153 | 80 | 28
37 | 1203 | 242 | 103 | 44
Bi
CD31 gated cells
HC
HSC
EC
HEC
cKit
Ly6A GFP
Bii
CFU-C/AGM
EC
HEC
HSC
HC
Ci
% repopulated
6/10
0/6
0/7
0/6
EC
HEC
HSC
HC
Cii
Donor
MyoD
% chimerism
PB
BM
Spl
Thy
LN
T cells
B cells
Ery
Mac
Gran
LSK
420 bp
245 bp
D
FPKM
CD31
Kit
Ly6a
EC
HEC
HSC
HC
E
EC vs HEC
EC vs HSC
HEC vs HSC
58
24
108
4
53
72
211
F
HSC patterns
LHH
LIH
IIH
LLH
ILH
EC
HEC
HSC
Low
High
G
FPKM
EC
HEC
HSC
Tal1
Gata2
Runx1
Erg
Fli1
Lyl1
Lmo2
Hi
EC
HEC
HSC
Cd34 101 577 613
Gpr56 34 17 473
Lgals9 94 64 240
Calcrl 35 215 236
Mpo 141 8 182
Pla2g4a 58 32 143
Meis1 58 30 136
Angpt1 7 3 92
Dapp1 26 14 89
Myb 3 4 75
Hii
Receptor activity: HEC to HSC
Gene | logFC | FDR
Gpr56 | 4.88 | 1.36E-05
Kit | 5.19 | 4.24E-04
Treml2 | 7.23 | 7.56E-03
Nlrp6 | 10.50 | 3.47E-02
Grm2 | 9.61 | 4.72E-02
Treml1* | 7.38 | 4.90E-02
I
Enrichments at Gpr56-37
Scl
Gata2
Runx1
Erg
Fli1
Lyl1
Lmo2
Ji
NT
Ao
Gpr56
Jii
Ly6a-GFP Gpr56
Ki
Control
gpr56 MO injected
cmyb
25/28
19/37
Kii
CD41-GFP
36.7 cells
12.5 cells
Kiii
gpr56 MO + zf RNA
grl
cmyb
8/22 more staining
14/22 normal staining
gpr56 MO + mouse RNA
grl
cmyb
4/28 more staining
19/28 normal staining
L
Bl
In
Ne
% Diff ct
Cntl
WT
mut

Figure 9 – Whole-transcriptome analysis of endothelial-to-HSC transition reveals a requirement for Gpr56 in HSC generation – Ai) Whole-mount image of a 34 somite pair (sp) *Ly6A-GFP* embryo showing expression of CD31 (magenta), cKit (red), and GFP (green). The aorta, vitelline artery and somatic vasculature are indicated. **Aii)** Four types of aortic cells during endothelial-to-haematopoietic transition (EHT) in a *Ly6A-GFP* AGM section (36sp) stained with anti-CD31 (magenta) and anti-cKit antibodies (red). Endothelial cells (ECs) are $CD31^{+}cKit^{-}GFP^{-}$, haemogenic ECs (HECs) are $CD31^{+}cKit^{-}GFP^{+}$, haematopoietic stem cells (HSCs) are $CD31^{+}cKit^{+}GFP^{+}$, and other intra-aortic (IA) haematopoietic cluster cells (HCs) are $CD31^{+}cKit^{+}GFP^{-}$. **Aiii)** Transverse section through a 36sp *Ly6A-GFP* embryo showing expression of CD34 (red) and GFP (green). A haematopoietic cluster with some GFP^{+} cells is located ventrally. GFP^{+} ECs are scattered throughout the aorta. **Aiv)** Different GFP^{+} cell types (arrowheads) in an E10.5 *Ly6A-GFP* aorta. (Endothelial) two flat GFP^{+} ECs; (bulging) rounding-up of a GFP^{+} EC (BC); (juxtaposed) round IAHC cells closely adhering to an EC; (distal) round IAHC cell on the distal side of the cluster. The number of cells/aorta is listed below at the 32, 34, and 37sp stages. Bars: **(Ai)** 100μm; **(Aii,iv)** 10μm; **(Aiii)** 50μm. **Bi)** Scatter plot showing the distribution and sorting gates for EHT subsets EC, HEC, HSC, and HC from E10.5/E11 *Ly6A-GFP* AGMs. **Bii)** Haematopoietic progenitor numbers (total CFU-C [CFU-Culture]) per AGM. EHT subsets from E10.5 *Ly6A-GFP* AGMs (34-39sp) were plated in methylcellulose, and colonies were counted at day 12 (SD is shown; *n*=4). **Ci)** HSC long-term repopulating activity in E11 *Ly6A-GFP* AGM EHT subsets (40-49sp). Irradiated adult recipients (*n*=4) were injected with 5-9 embryo equivalents (ee) of ECs, 4-9ee of HECs, 1-5ee of HSCs, and 4-8ee of HCs together with $2x10^{5}$ spleen cells (recipient type). Percentage of high-level repopulated recipients (4 months [mo] after injection) out of total number recipients injected is shown. Indicated above each bar is the number of repopulated recipients/number of recipients injected. **Cii)** Multilineage analysis of two of the adult recipient mice transplanted with the $CD31^{+}cKit^{+}GFP^{+}$ HSC fraction of E11 AGM cells.

At >4 mo after injection, haematopoietic tissues were harvested and cells were sorted by lineage. DNA was prepared from peripheral blood (PB), bone marrow (BM), spleen (Spl), thymus (Thy), and lymph nodes (LNs), and T, B, erythroid (Ery), macrophage (Mac), and granulocytic (Gran) cells. PCR for the donor *Gfp* marker gene and control *MyoD* gene (DNA normalisation) was performed. Percentage of engraftment is indicated under each lane, as determined after *MyoD* normalisation by fragment intensity comparison with percentage of chimaerism standards. Transplantation of BM from primary recipients **(Ci)** showed engraftment of 2/3 and 3/3 secondary recipients (not depicted). **D)** RNA sequencing was performed on each of the EHT subsets (ECs, HECs, HSCs and HCs). N=3. Biological replicate 1 included two 36sp and two 37sp embryos; replicate 2 included four 34sp, two 35sp, three 36sp, and five 37sp embryos; replicate 3 included two 35sp, one 36sp, two 38sp, one 39sp, one 40sp, and one 41sp embryos. Graph shows normalised number of mapped fragments for genes encoding the markers used for sorting EHT fractions. FPKMs (fragments per kilobase exon reads per million fragments mapped) of *CD31*, *cKit*, and *Ly6A-GFP* per fraction are shown (error bars=SD). **E)** Venn diagram showing numbers of differentially expressed genes (DEGs) in comparisons of HECs v(ersu)s ECs, HSCs vs HECs and HSCs vs ECs. Total DEGs was 530. **F)** DEG patterns that are HSC-specific are shown: low-intermediate-high (LIH), LHH, IIH, LLH, and ILH. **G)** Mean FPKM values of heptad factors in EC, HEC, and HSC fractions. **Hi)** Heatmap of top 10 heptad target DEGs based on highest expression in HSCs and with respective mean FPKM values inside heatmap. **Hii)** Receptor-related genes with significant upregulated expression in HEC-to-HSC transition. Gray font, genes with low overall expression levels as defined by edgeR-calculated logCPM of <3 (and higher probability of being false positive); asterisks, genes differentially expressed during EC-to-HEC and HEC-to-HSC transitions; logFC, log fold change. **I)** qPCR for TF enrichment at the *Gpr56 -37*-enhancer as compared with IgG and control in HPC-7 mouse myeloid progenitor cells (*n*=4). **Ji)** *In situ* hybridisation (ISH) of WT mouse E10.5 AGM sections shows specific expression of *Gpr56* in some

IAHCs, a few cells lining the aorta (Ao), and the notochord (NT). The top images show low magnification of AGM cross-section, and the bottom images show high magnification of the boxed areas. Bars=30μm **Jii)** ISH of E10.5 *Ly6A-GFP* AGM shows co-expression of GFP and *Gpr56* in some IAHC cells. Bar=10μm **Ki-iii)** Analysis of WT and *gpr56* morpholino oligo (MO) zebrafish for the presence of HSCs. Numbers in the panels indicate the number of embryos with the depicted phenotype. Bars=100μm **Ki)** ISH with the HSC marker *cmyb* at 30 hours post-fertilisation (hpf). **Kii)** Fluorescent analysis of WT and MO-injected *cd41-GFP* transgenic embryos at 48hpf. Arrowheads (left) indicate *cd41*-expressing HCs in the aorta. The dashed lines (right) indicate the outline of the morphant zebrafish embryo for orientation purposes. **Kiii)** HSC rescue of *gpr56* morphant zebrafish with *gpr56* RNA (zebrafish and mouse) as shown by ISH for *cmyb*. Ectopic *cmyb* expression in the posterior cardinal vein is clearly visible. No vascular abnormalities were observed by *gridlock* ISH. **(Ki,iii)** Insets show boxed areas at higher magnification. **L)** Effect of human *GPR56* activity in neutrophil differentiation of the 32D-CSF3R unipotent stem cell line. 32D-CSF3R cells cultured in medium containing CSF3 efficiently differentiated into neutrophils. Only constitutively active mutant *GPR56* (MUT) could block differentiation. Diff ct, differential count; Cntl, empty vector control; WT, WT human *GPR56* vector; mut, constitutively active human *GPR56* mutant vector; Bl, blast morphology; In, intermediate morphology; Ne, neutrophil morphology. **(A-L)** Reprinted from Solaimani Kartalaei et al. (2015), with permission from Rockefeller University Press.

Chapter 5 – Iterative single-cell analyses define the transcriptome of the first functional HSCs

5.1 Summary of results

Chapter 4 presented the first whole-transcriptome analysis of sorted populations during the endothelial-to-haematopoietic fate/identity transdifferentiation. The utilisation of the *G2V* or *Ly6A-GFP* mouse models (*Chapter 3* and *4*) enabled enrichment of HECs and functional AGM HSCs. However, technical limitations and the low frequency of HSCs compared to other haematopoietic cells, impeded isolation of a pure population of *bona fide* AGM HSCs and the capture of their exclusive transcriptome. To achieve both, *Gata2(Venus)*-mediated enrichment and iterations of single-cell functional and transcriptomic analyses, connecting the inner HSC regulatory programme with function/identity, were performed on IAHC cells.

In *Chapter 3* (Eich et al., 2018), we had identified CD31^{+}cKit^{+}Gata2-medium (G2^{med})-expressing IAHC cells as most potent in terms of HSPC activity. Here, this fraction was subdivided into G2$^{lo(w)}$, G2$^{int(ermediate)}$, cKit^{lo} and cKit$^{hi(gh)}$ (Figure 10Ai). Both at E10 and E11, most HPCs were in the CD31^{+}cKit^{hi}G2^{lo} and -G2^{int} subsets, in which multipotent CFU-GEMMs were also highly enriched or exclusively found (Figure 10Aii). Next, the E11 CD31^{+}cKit^{hi}G2^{int} fraction, that contained most HPCs, was index-sorted to correlate protein level expression (MFI) of enrichment markers with *in vitro* haematopoietic function in single-cell CFU assays (Figure 10Ai,B). G2 expression was more variable in CFU-GMs and -Gs compared to -GEMMs and -Ms, despite mean expression not being significantly different. GEMM-yielding progenitors showed significantly higher cKit and CD31 expression than CFU-Gs (Figure 10Bi,ii). Within the (CD31^{+}cKit^{+}) G2^{hi} control fraction, CFU-GEMMs revealed significantly higher cKit expression than CFU-GMs, -Gs and -Ms (Figure 10Ci). The G2^{hi} CFU-GEMMs also displayed higher CD31 expression than CFU-Gs (Figure 10Cii). Even though G2^{hi} cells had a

higher plating efficiency than $G2^{int}$ (1:2-4 versus 1:5-6 cells producing an HPC), there were relatively more immature (HSC-like) CFU-GEMMs in the $G2^{int}$ fraction (22% versus 9%).

Upon *in vivo* transplantation, E11 $CD31^{+}cKit^{hi}G2^{lo}$ and -$G2^{int}$ cells yielded long-term repopulation of irradiated recipients, indicating that both contained HSCs and that the $G2^{lo}$ or $G2^{int}$ levels cannot be used for additional HSC enrichment. $CD31^{+}cKit^{lo}G2^{med}$ cells did not reconstitute any recipients, and the absence of HSCs in the $G2^{hi}$ subset had previously been demonstrated (Eich et al., 2018). Taken together, 1) the most immature HPCs are $CD31^{hi}$ and $cKit^{hi}$ and 2) further HSC enrichment is achieved based on $cKit^{hi}$ expression.

As all HSC activity was constrained to the $CD31^{+}cKit^{hi}G2^{med}$ IAHC fraction, these were index-sorted for SmartSeq2 single-cell (sc)RNA-seq. In total 968 cells passed quality control (95% of total). *Gata2* and *Venus* were similarly expressed in single cells (correlation coefficient of +0.7) and the different datasets (n=7 biological replicates; 3 RNA-seq processing batches) revealed no batch effect, underlining their reliability and the consistency of cell isolation parameters. The SPRING visualisation was next acquired (Figure 10D) and eleven transcriptomic Louvain clusters (L1.1-1.11) were defined. Five new main clusters emerged based on expression of characteristic EHT genes and hierarchical clustering distances (Figure 10D,E). Surprisingly, within this highly enriched phenotypic and functional HSC population, one of the five clusters (335 cells of 968) revealed an endothelial-like transcriptome (EC) with expression of *Flk1*, *Cdh5* and *AplnR* (Figure 10D,E). Most cells were however in the haematopoietic cluster (HC) and showed expression of genes like *CD41*, *Gpr56* (Figure 10E), *CD45* and *PU.1*. Expression of *Ly6A* and *Cdh5* further separated this cluster into HC1 (405 cells of 968; L1.7 and L1.8) and HC2 (124 cells) (Figure 10D), the first revealing a more immature and the latter a more mature haematopoietic transcriptome. Two additional ‘unknown clusters’ (UC1,2) showed heterogeneous expression of various markers.

Heptad TFs were also examined. Whereas most were highly expressed in the majority of cells, *Gata2* (Figure 10E) and *Tal1* revealed heterogeneous expression in all clusters (and downregulation of *Gata2* in UC1), and *Erg* (Figure 10E) and *Lmo2* were downregulated in some HC2 and UC1,2 cells. Interestingly, *Runx1* expression, expected in all these IAHC cells (Bee et al., 2009), was mostly absent in EC (Figure 10E). Altogether, cKit^{hi}G2^{med}-expressing IAHC cells separated into five transcriptomic clusters revealing a heterogeneous mix of endothelial-like and (im)mature haematopoietic cells within a highly enriched HSC fraction. Heterogeneous and asynchronous expression of the pivotal heptad TFs in these clusters suggests unstable/dynamic expression/regulation and/or a stochastic HSC induction programme, making it altogether impossible to precisely predict what rare cell out of the HEC/'HSC precursor' pool will acquire HSC fate.

A 'hscScore' (Hamey and Gottgens, 2019), indicative of similarity to the adult BM HSC transcriptome, was assigned to each single cell and suggested HC1 contained LT-HSCs (Figure 10F). Our datasets were next compared to published scRNA-seq datasets of AGM endothelium and pre-HSCs, and E12/14 FL HSCs (Zhou et al., 2016). Similarity was seen between their aortic endothelium and our EC cells and revealed that half of HC1 cells (L1.7) had a transcriptomic signature in between pre-HSCs and FL HSCs, whereas the other half (L1.8) mostly had an HSC signature. Thus, HC1 most certainly contains a mix of E11 pre-HSCs/HPCs and HSCs.

To test the haematopoietic potency/function of the five transcriptomic clusters, phenotypic differences that would allow their isolation were explored using the index data obtained during FACS. Cells in HC1 were CD31^{hi} and side-scatter low (SSC^{lo}), whereas EC cells were CD31^{hi} and SSC^{hi}, and UC1+UC2+HC2 were CD31^{lo} (Figure 10Gi). No further separation of the cKit^{hi}G2^{med} population was possible, even though HC1 contained cells expressing the highest levels of cKit (Figure 10Gii). CFU-C assays revealed most HPCs were in the UC1+UC2+HC2 fraction, however HC1 appeared to contain more multipotent (7 vs 5 GEMMs/ee), fewer bipotent (17 vs 25 GMs/ee) and hardly any unipotent HPCs compared to UC1+UC2+HC2

(Figure 10H). EC cells did not reveal any progenitor function (Figure 10H). As predicted, all LT-HSCs were within the HC1 subset as shown by *in vivo* transplantations (Figure 10I), yielding a 10.5-fold HSC enrichment over the Eich et al. (2018) strategy (*Chapter 3*). Single-cell CFU-C assays with index-sorted HC1 cells revealed 92% of CFU-Cs were multi-/bipotent, and GEMMs had significantly higher G2 expression than GMs and Gs, placing them mostly within the former $G2^{int}$ sorting gate (right half of $G2^{med}$). Within this highly enriched HC1 sorting gate, cKit and CD31 no longer provided significantly different expression between CFU-C types. Taken together, all HSCs are contained within the $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}$ (HC1) fraction which is also highly enriched for GEMMs.

To achieve further HSC enrichment, the HC1 transcriptomes were compared to those of cells in the other clusters and revealed high expression of CD27 (*Tnfrsf7*). In fact, 70% of HC1 cells expressed CD27 by flow cytometric analysis (Figure 10J; $CD27^{+W(ide)}$). Transplantation of as few as 442 of the $CD27^{+W}$ HC1 cells/recipient revealed long-term multilineage high-level chimaerism in primary (Figure 10K; open squares) and secondary recipients. Closer examination of CD27 expression on HC1 cells showed that all HSCs were actually amongst the highest/brightest-expressing 19% (Figure 10J; $CD27^{+N(arrow)}$) as demonstrated by transplantations of as few as 238 $CD27^{+N}$ HC1 cells/recipient (Figure 10K; closed squares). No repopulating ability was detected in the remaining negative/low-expressing 81% ($CD27^{-/lo}$ HC1), even upon transplantation of high cell numbers (Figure 10K). Index-sorting and scRNA-seq were next performed on 119 $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{+N}$ cells (from six E11 AGMs). The new dataset merged well with the previous $CD31^{+}cKit^{hi}G2^{med}$ batch and created a new but very similar SPRING plot and Louvain clustering (L2.1-2.12) of 1087 cells. Most of the new cells localised to HC1 (of which 19% to L2.7 and 81% to L2.8) (Figure 10L).

Characterisation of HPC function within the $CD27^{+N}$ HC1 fraction was also explored by index-sorting and single-cell CFU-C assays. CFU-GEMMs revealed significantly lower CD27 and higher G2 expression than CFU-GMs

(Figure 10M). Strikingly, 95% of CFU-GEMMs were found within a very narrow CD27^{med} MFI range (Figure 10M,N, red dotted area). These cells made up 15.5% of HC1 (±20 cells/AGM), yielding a total combined (HC1+CD27^{med}) "HSC/HSC-like HPC"-enrichment of 68.3-fold. Thus, all HSCs and the majority of immature HPCs are highly enriched within the CD27^{+N} / CD27^{med} HC1 fractions, which equal ±20-23 cells/AGM and yield 68.3-fold enrichment over the previous strategy (*Chapter 3*).

The new enrichment allowed for a more accurate capture of the HSC transcriptome. Out of the 119 CD27^{+N} cells isolated for scRNA-seq, 85 localised to HC1. To identify groups of cells and genes with similar transcriptomic behaviour, the 85 cells were examined by Iterative Clustering and Guide-gene Selection (ICGS). This produced a heatmap with three groups of genes and cells (ICGS1-3). The 29 cells in ICGS2 were evenly divided over L2.7 and L2.8, suggesting HSC identity acquisition; while the 54 ICGS1 cells mostly belonged to L2.8 (ICGS3 only contained 2 cells). Correlating this with hscScore and CD27 MFI revealed that ICGS2 cells were most similar to adult BM HSCs and almost all were CD27^{med}, whereas ICGS1 cells had low hscScores and displayed the full CD27^{+N} expression range. Thus, ICGS2 cells are the most immature/HSC-like subset.

The expression of the HSC marker *Ly6A* (de Bruijn et al., 2002) was next added to the enrichment strategy. Out of the 85 CD27^{+N} cells in HC1, 27 expressed *Ly6A*, of which 23 were in the CD27^{med} HSC range (Figure 10O). Eight of the 23 cells, all with high hscScores, belonged to the ICGS2 (HSC) cluster (see Appendix: Vink et al. (2020); Figure 5C and Supp Figure 4,5). Four out of eight localised to L2.7 and the others to L2.8. In summary, through a combination of iterative single-cell functional, transcriptomic and phenotypic (index-sorting) analyses 1) the transcriptomic signature/profile of eight HSCs, originally obtained from six embryos, was identified; 2) the number of HSCs identified (8 per 6 embryos) fits with previous quantifications (Kumaravelu et al., 2002); 3) their allocation in two Louvain clusters indicates HSC acquisition; and 4) the HSC signature is distinct but closely related to immature HPCs.

The signature of the eight cells stands out from the other fifteen to reveal the high expression of genes involved in the generation/regulation of HE/IAHCs and AGM/foetal HSCs, and in the acquisition of self-renewal potential, like Notch signalling pathway genes (*Hes1*, *Hey1*, *Dll4*, *Notch4*) (reviewed in Bigas and Porcheri (2018)) and *Sox17* (Clarke et al., 2013; He et al., 2011; Nobuhisa et al., 2014). Other known HSC genes like *Flk1*, *Flt1*, *Nos3*, *Pdzk1ip1*, *Procr* and *Vwf* were also expressed, altogether strongly supporting the notion that the eight similar gene signatures represent and define the transcriptome of the first functional HSCs.

Lastly, the localisation of the first HSCs was examined based on CD27 expression in *G2V* and *Ly6A-GFP* AGMs. Rare and almost exclusively ventrally located $CD34^{+}G2^{+}CD27^{+}$ cells were observed in small IAHCs (Figure 10P,Q). Some $CD27^{+}$ cells appeared to be bulging/rounding up out of the endothelial layer (Figure 10P, arrow), whilst others were located in 1-/2-cell IAHCs (Figure 10P,Q). $CD27^{+}$ cells were either attached to other $G2^{lo/med}$-expressing cells ($CD27^{+}$ or $CD27^{-}$) (Figure 10Qi,ii), or $G2^{-}CD27^{-}$ cells (Figure 10P,Qiii). The *Ly6A-GFP* AGM also revealed $CD27^{+}$ cells in 1-/2-cell ventral IAHCs (Figure 10Ri,ii) or still closely associated with the endothelial layer (Figure 10Riii). $CD27^{+}$ cells were Ly6A-$GFP^{lo/+}$ or Ly6A-GFP^{-} (Figure 10Ri-iii). Furthermore, macrophages, important for HSPC generation in the AGM (Mariani et al., 2019), were shown to interact with single $CD27^{+}$ cells (Figure 10S). Quantifications showed that indeed 80% of $CD27^{+}$ cell-containing IAHCs had a total size of one to two cells (Figure 10T). Thus, the localisation of single ($G2^{+}$ or $Ly6A^{+}$) $CD27^{+}$ cells in the mouse aorta suggests HSCs emerge directly from HECs into 1-/2-cell clusters.

In summary, iterations of single-cell approaches have allowed capture of the transcriptome, and phenotypic characterisation and localisation of the first functional mouse embryonic HSCs (see graphical abstract on p136).

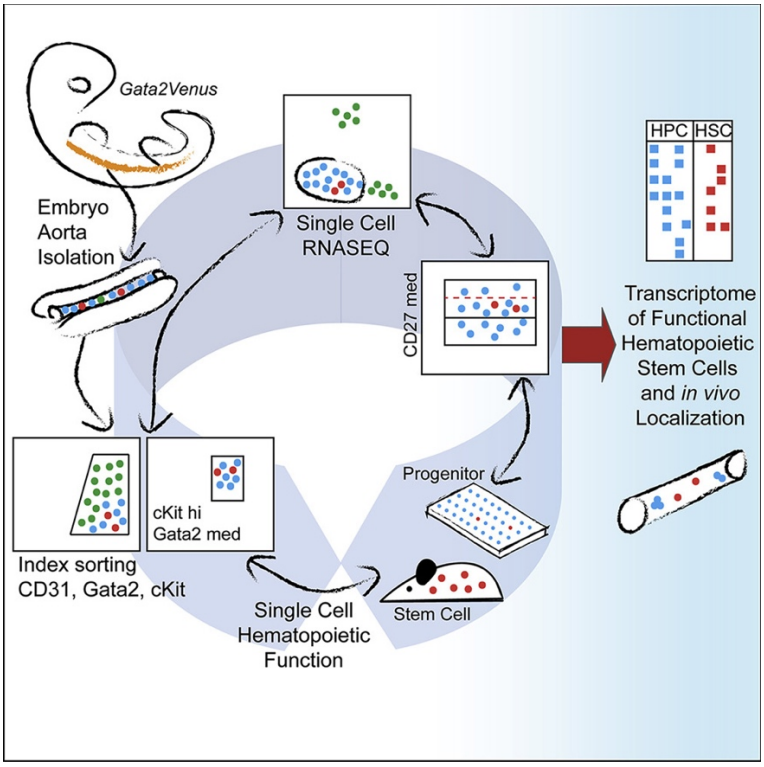
5.2 Major findings

- $CD31^{+}cKit^{hi}G2^{med}$-expressing IAHC cells separate into five transcriptomic clusters and reveal a heterogeneous mix of endothelial-like and (im)mature haematopoietic cells within a highly enriched HSC fraction. Heterogeneous and asynchronous expression of the pivotal heptad TFs in these clusters suggests unstable/dynamic expression/regulation and/or a stochastic HSC induction programme.
- Functional, transcriptomic and phenotypic single-cell iterations capture the transcriptome of the first functional HSCs in mouse (eight HSCs obtained from six embryos) and reveal a signature that is distinctive but closely related to that of immature HPCs, allowing for Gata2 to connect the ‘‘inner’’ HSC regulatory network with single-cell function.
- Specific Gata2, CD31, cKit, and CD27 levels define all functional HSCs (and the majority of immature HPCs) and demonstrate that they are highly enriched within the $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}$ (HC1) $CD27^{+N}$ / $CD27^{med}$ fraction, which equals ±20-23 cells/AGM and yields 68.3-fold enrichment over the previous enrichment strategy (*Chapter 3*).
- HSCs emerge from endothelium as single cells within aortic clusters of 1-2 cells.

5.3 Author contributions

Vink, C.S. et al., 2020. *Cell Reports* (first author; out of eight). In February 2016 CSV initiated work on this project. It was brought to completion and publication in May 2020. During more than four years (at 0.8-0.9 full-time equivalent; FTE) CSV led and coordinated this complex collaborative project, designed the study and interpreted data (together with ED). All wet lab experiments (except for one; by SAM) in this chapter/publication were performed by CSV, i.e. fine embryo dissections (and embryo generation, transgenic mouse line maintenance), cell preparations, flow cytometry, *in vitro* (bulk and single-cell) and *in vivo* haematopoietic

assays (including all animal work, intravenous injections, blood sampling, post-transplant tissue harvests and multilineage analyses, determination of donor chimaerism, writing of experimental animal request forms and providing help with renewal of animal project license), IHC and confocal microscopy (including embryo generation and harvest, and cryosectioning). CSV performed (with AM preparing collection plates) all cell isolations for scRNA-seq to the highest standards (95% of cells passed stringent quality controls). Bioinformatics analyses were done at the University of Cambridge (Gottgens lab; FJCN, XW, WJ, BG). CSV spent a considerable amount of time coordinating these analyses and the collaboration, combining the experimental results and direction of the wet lab research with the required and appropriate bioinformatics analyses. All figures in the main and supplemental text and the graphical abstract were designed and made by CSV (bioinformatics graphs and plots were produced by the Gottgens lab upon request and modified by CSV where needed). The manuscript (including main text, figure legends, title, highlights, eTOC Blurb), response to reviewers and cover letters were written by CSV (together with ED).



Graphical abstract Vink et al. (2020)

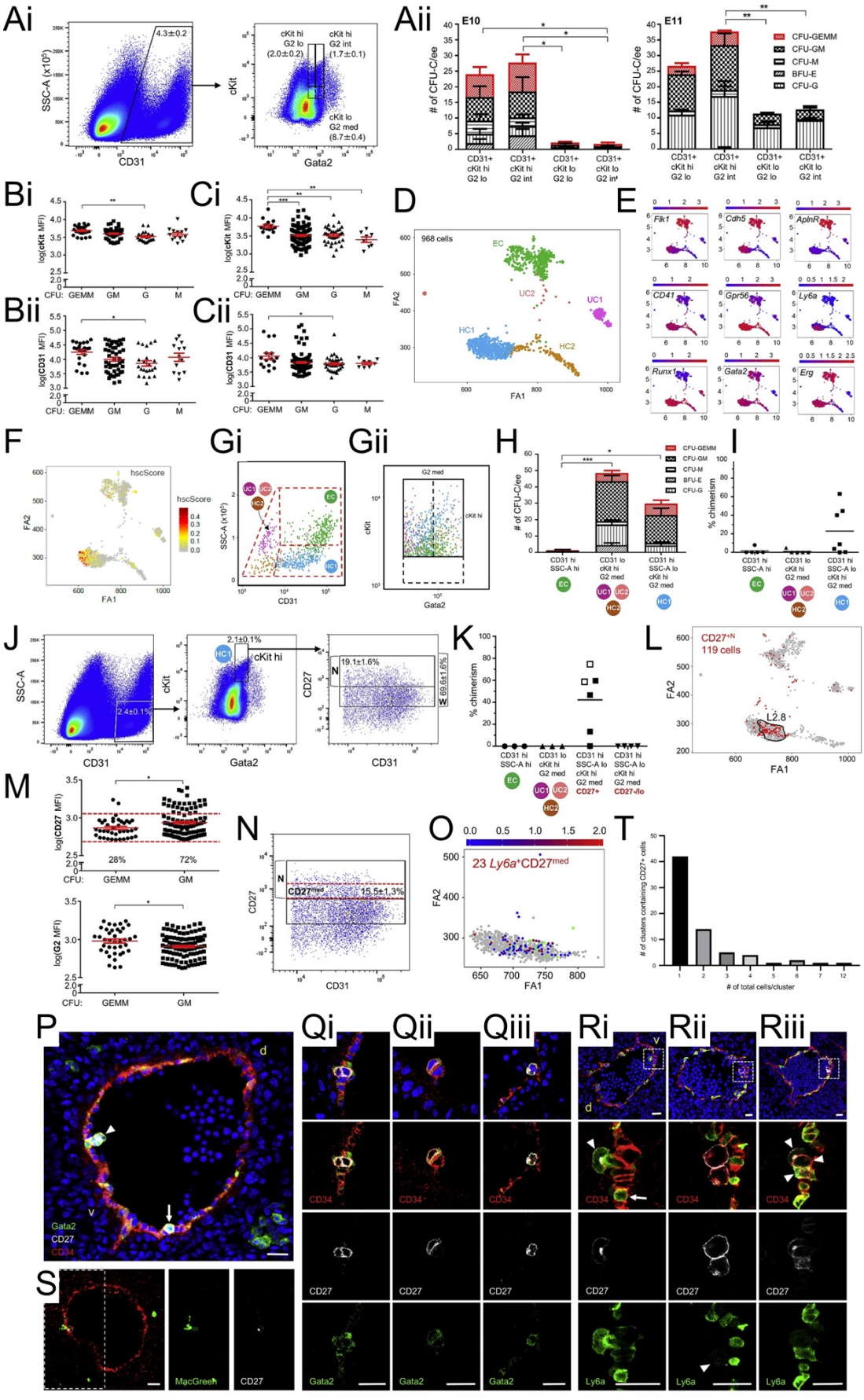

Ai
SSC-A (x10⁵)
CD31
4.3±0.2
cKit
Gata2
cKit hi G2 lo (2.0±0.2)
cKit hi G2 int (1.7±0.1)
cKit lo G2 med (8.7±0.4)
Aii
E10
E11
# of CFU-C/ee
CFU-GEMM
CFU-GM
CFU-M
BFU-E
CFU-G
CD31+ cKit hi G2 lo
CD31+ cKit hi G2 int
CD31+ cKit lo G2 lo
CD31+ cKit lo G2 int
Bi
Ci
Bii
Cii
log(cKit MFI)
log(CD31 MFI)
CFU: GEMM GM G M
D
968 cells
EC
UC1
UC2
HC1
HC2
FA1
FA2
E
Flk1
Cdh5
AplnR
CD41
Gpr56
Ly6a
Runx1
Gata2
Erg
F
hscScore
Gi
Gii
G2 med
cKit hi
H
I
% chimerism
CD31 hi SSC-A hi
CD31 lo cKit hi G2 med
CD31 hi SSC-A lo cKit hi G2 med
J
SSC-A
2.4±0.1%
2.1±0.1%
19.1±1.6%
69.6±1.6%
K
CD27+
CD27-/lo
L
CD27+N
119 cells
L2.8
M
28%
72%
log(G2 MFI)
N
CD27med
15.5±1.3%
O
23 Ly6a+CD27med
T
# of clusters containing CD27+ cells
# of total cells/cluster
P
Gata2
CD27
CD34
Qi
Qii
Qiii
Ri
Rii
Riii
S
MacGreen
Ly6a

Figure 10 – Iterative single-cell analyses define the transcriptome of the first functional HSCs – Ai) Flow-cytometric dot plots of representative E11 *G2V* AGM cells analysed for side scatter (SSC-A) and CD31 expression (left) and cKit and Gata2 (G2) expression of $CD31^{+}$ cells (right). Cell percentage (±SEM) within gated regions. Hi, high; lo, low; int, intermediate; med, medium (lo+int). **Aii)** Haematopoietic progenitor cell (HPC) numbers per lineage per embryo equivalent (ee) of E10 (32-37 somite pairs, sp; n=4; embryos/ experiment: 4, 4, 4, and 4) and E11 (43-48sp; n=3; embryos/experiment: 3, 6, 5, and 7) sorted AGM fractions in the colony-forming unit-culture (CFU-C) assay. GEMM, granulocyte-erythroid-macrophage-megakaryocyte; GM, granulocyte-macrophage; M, macrophage; G, granulocyte; BFU-E, burst-forming unit-erythroid. Error bars represent mean±SEM. Significance for total numbers of CFU-Cs/cell fraction as indicated. For E10 GEMMs: *$cKit^{hi}G2^{int}$ v(ersu)s $cKit^{lo}G2^{lo/int}$. For E11 GEMMs: *$cKit^{hi}G2^{lo}$ vs $cKit^{lo}G2^{lo/int}$, and **$cKit^{hi}G2^{int}$ vs $cKit^{lo}G2^{lo/int}$. For E11 GMs: *$cKit^{hi}G2^{int}$ vs $cKit^{lo}G2^{lo/int}$. **B,C)** E11 *G2V* AGM single cells were sorted for **B)** $CD31^{+}cKit^{hi}G2^{int}$ (11 embryos, 41-47sp) and **C)** $CD31^{+}cKit^{+}G2^{hi}$ cells (20 embryos; 40-49sp) and deposited in methylcellulose. CFU-Cs were counted at days 9-11 (total $CD31^{+}cKit^{hi}G2^{int}$ = 169; total $CD31^{+}cKit^{+}G2^{hi}$ = 105). Each CFU (individual dot) was examined for colony type and plotted against log_{10}-normalised mean fluorescence intensities (MFIs) for expression of **(Bi and Ci**) cKit, and (**Bii and Cii**) CD31. Red lines indicate mean±SEM. **D,E)** SPRING visualisation of merged Louvain clusters corresponding to 968 single $CD31^{+}cKit^{hi}G2^{med}$ IAHC cells (n=6 at E11, 27.6 *G2V* embryos, 41-49sp; n=1 at E10, 2.7 *G2V* embryos, 33-34sp). Five clusters emerge based on the cell-lineage expression characteristics and hierarchical clustering distances. EC, endothelial-like cluster; HC1, haematopoietic cluster 1; HC2, haematopoietic cluster 2; UC1, unknown cluster 1; UC2, unknown cluster 2. **E)** Cells were coloured according to the expression of (top row) endothelial markers, (middle row) haematopoietic markers, and (bottom row) some of the haematopoietic 'heptad' transcription factors. **F)** Force-directed graph (SPRING) visualisation using highly variable genes, the top 90 principal components and the 4

nearest neighbours for all 968 $CD31^{+}cKit^{hi}G2^{med}$ cells, coloured by hscScore (Hamey and Gottgens, 2019). **G)** Flow-cytometric plots showing **(Gi)** side-scatter (SSC-A) and CD31 expression and **(Gii)** cKit and G2 expression of the 968 index-sorted single cells. The colour of each cell indicates presence in specific clusters: EC (green), HC1 (blue), HC2 (ochre), UC1 (magenta), and UC2 (rose). **H)** HPC number and type per ee of bulk-sorted E11 IAHC cells (n=4; 41-50sp; embryos/experiment: 8, 8, 3, 8) corresponding to transcriptomic clusters, as determined by CFU-C assay. For GEMMs: *EC vs HC1; for GMs: *EC vs HC1, and **UC1+UC2+HC2 vs EC. Mean±SEM. **I)** HSC activity in sorted E11 IAHC cells (n=5; 40-49sp; embryos/experiment: 15, 14, 11, 8, and 19) corresponding to transcriptomic clusters by *in vivo* transplantation. Each point indicates percentage of donor-derived cells in peripheral blood of recipients at 4 months post-injection. Cell numbers injected: for HC1, 721-2811, 4.8-8ee; for EC, 5654-14208, 4.7-8.1ee; for UC1+UC2+HC2, 1312-2828, 4.8-8.4ee. Black line indicates mean. **J)** Enrichment sorting parameters for $CD27^{+}$ HC1 cell isolation. SSC-A, CD31, cKit, G2, and CD27 sorting gates and percentage (±SEM) of cells. $CD27^{+}$ cells (%) are indicated within a wide (+W) and a narrow (+N) MFI gate. **K)** Analysis of HSC function within CD27 fractions ($CD27^{+W}$, $CD27^{+N}$, and $CD27^{-/lo}$) of HC1 cells by *in vivo* transplantation. Open squares denote $CD27^{+W}$ HC1 cell recipients (n=2) of 442 cells (2.5ee) and 479 cells (3.3ee) from 15 and 9 (42-50sp) embryos, respectively. Closed squares (n=4; 42-49 sp; embryos/experiment: 25, 11, 22, 28) denote recipients injected with $CD27^{+N}$ HC1 cells (238-957; 3.7-9.8 ee), EC cells (899-2333; 3.9-8.6ee), $CD27^{-/lo}$ HC1 cells (1042-4325; 4.1-9.8ee), and UC1+UC2+HC2 cells (1361-3741; 4.2-9.9ee). Percentage of donor cell chimaerism in peripheral blood of recipients at 4 months post-injection. Black line indicates mean. **L)** Single-cell RNA-seq of 119 $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{+N}$ sorted cells (all passed quality control) from 6 embryos (41-47 sp). These data combined with the previous dataset (gray) resulted in an extended dataset SPRING plot of 1087 cells in total. Overlay of $CD27^{+N}$ cells (red) showing localization to one HC1 subcluster (circled). **M)** Single $CD27^{+N}$ HC1 cells from E11 *G2V* IAHCs (50

embryos; 42-49 sp) deposited into methylcellulose and CFU-C counts at 8-10 days (138 total CFU-Cs). Colony type (individual dot) plotted against log_{10}-normalized MFIs for CD27 and G2 expression. **N)** Representative FACS plot showing newly defined $CD27^{med}$ MFI range (red dotted line) within the $CD27^{+N}$ fraction. Percentage±SEM. **O)** *Ly6A* expression was plotted on the SPRING visualisation of 85 $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{+N}$ cells. 71 are $CD27^{med}$, of which 23 are *Ly6A*$^{+}$ (red) and 58 *Ly6A*$^{-}$ (14 $CD27^{hi}$ cells in green) within HC1 (gray). **(P-S)** Sections through dorsal aorta of E11 (43sp) *Gata2Venus* (green in **P** and **Q**) and *Ly6A-GFP* (green in **R**) and of E10.5 (36sp) *Csf1rGFP* (MacGreen; green in **S**) embryos. Nuclei are indicated in blue (DAPI), endothelial and IAHC cells are indicated in red (CD34) and $CD27^{+}$ cells are indicated in white. v, ventral, d, dorsal. **P)** Arrowhead indicates small cluster containing two $G2^{+}CD27^{+}$ cells attached to one $G2^{+}CD27^{-}$ cell. Arrow indicates a single rounded/bulging $G2^{+}CD27^{+}$ cell. **Q)** In **Qi)**, ventral cluster with two $G2^{med}CD27^{+}$ cells (one bulging out of endothelium). **Qii)** Ventral two-cell cluster ($G2^{med}CD27^{+}$ and $G2^{lo}CD27^{-}$ cells). **Qiii)** Single rounded bulging $G2^{+}CD27^{+}$ cell on dorsal side of aorta. **R)** In **Ri)**, single Ly6A-$GFP^{+}CD27^{+}$ cell (arrowhead) at apex of bulging Ly6A-GFP^{+} cell within ventral endothelium. A neighbouring GFP^{+} cell in the ventral endothelium is $CD27^{-}$ (arrow). **Rii)** Two round $CD27^{+}$ cells (one is Ly6A-GFP^{lo}, indicated by arrowhead) adjacent to each other and attached to interspersed Ly6A-GFP^{+} endothelium. **Riii)** Three $CD27^{+}$ cells (indicated by arrowheads; one is Ly6A-GFP^{+}) in very close association with ventral endothelium. White dashed boxes indicate magnified areas. **S)** One $CD27^{+}CD34^{+}$ cell on dorsal endothelial wall closely associated with a macrophage. White dashed box indicates area for which MacGreen and CD27 expression is shown. Scale bars: 20µm. **T)** Correlation of CD27 expression with IAHC size. A total of 62 transversal AGM sections (10µm) from genetically different embryos (*Gata2Venus*, *Ly6A-GFP* and *Csf1rGFP*/MacGreen) were analysed for CD27 expression (n=4; 35, 36, 42, 43sp). CD27+ cells were quantified and correlated to the size of the cluster in which they were localised. 80% (56 out of 70) of clusters containing $CD27^{+}$

cells had a total size of 1-2 cells. **(A-T)** Statistics: Bartlett's test for homogeneity of variances was performed to ensure variances for all samples were equal. With the assumption of normality, the data were compared for the same means using one-way ANOVA, followed by a pairwise comparison using a Student's t-test, and Bonferroni correction. Significance was determined as $p<0.05$. *$p<0.05$, **$p<0.01$, ***$p<0.001$. FACS percentages, averages (mean) and error bars are ± standard error of the mean (SEM). The number of biological replicates is indicated with 'n'. **(A-T)** Reprinted from Vink et al. (2020).

Chapter 6 – BMP signalling differentially regulates distinct HSC types

6.1 Summary of results

Having examined cell intrinsic factors critical for HSC identity acquisition (*Chapter 1-5*), we next focussed on the extrinsic growth factor BMP (bone morphogenetic protein). BMP is crucial for HSC induction and directly upstream of *Gata2*. Its effect on HSC generation, foetal and adult HSCs, and lineage output was studied by using a *BMP responsive element* (*BRE*)-*GFP* mouse reporting *in vivo* pathway activation. In these mice (Monteiro et al., 2008), GFP expression reports BMP-activated cells. Activation of *BRE*-regulated transcription follows BMP signalling through Smad-1/-5, and thus BMP-Smad pathway activation (Figure 11Ai). In the E10.5 mouse AGM, BMP-activated cells are observed mainly on the ventral side of the aorta (Figure 11Aii). Similar to aortic Ly6A(GFP) and Gata2(Venus) expression, only some endothelial and IAHC cells are BMP-activated (Figure 11Aiii). Transplantation of GFP^+ and GFP^- E11 AGM cells into lethally irradiated recipients revealed that all HSCs were BMP-activated, and thus that BMP signalling is essential during HSC ontogeny (Figure 11B).

Transplantations were next performed with GFP^+ and GFP^- cells from E12-14 FL (Figure 11Ci) and adult BM (Figure 11Cii). Each of the tissues contained long-term multilineage self-renewing HSCs in both fractions (Figure 11Ci,ii), demonstrating the presence and existence of BMP- and non-BMP-activated HSCs. Interestingly, the E14 FL contained significantly more BMP-activated (GFP^+; 73%) than non-BMP-activated (GFP^-) phenotypic LSK SLAM HSCs (Figure 11Di). The adult BM showed a reversed phenotype, with only 8% GFP^+ HSCs (Figure 11Dii). Taken together, FL and BM contain non-BMP- and BMP-activated phenotypic HSCs, there is a developmental progression towards the prevalence of non-BMP-activated HSCs, and the

utilisation of the *BRE-GFP* model allows for prospective isolation of HSCs for molecular characterisation.

Bulk RNA-seq on GFP^+ and GFP^- BM and FL HSCs was used to explore transcriptomic differences between the two types of HSCs. Both fractions (in both tissues) expressed *Smad* genes (Figure 11E), and only the BMP-activated HSCs expressed *Bmpr* genes (Figure 11E) and downstream targets, validating BMP pathway activation. Furthermore, GFP^+ HSCs revealed significant upregulation of *Myc* and *Stat5b* target genes, and enrichment for metabolic and homeostatic genes. GFP^- HSCs on the other hand showed expression of *Nfkb1*, *Sp1* and *Nfe2* targets, as well as genes linked to haematopoietic system development.

To determine HSC frequency/clonal cell number i.e., the number that contained one HSC (calculated based on number of cells transplanted and the resulting frequency of engrafted mice), limiting dilution transplantations were performed with the different fractions and tissues. The highest HSC frequency was found in the GFP^+ fraction of the FL (1:180 GFP^+ cells) (Figure 11Fi), which meant that 81% of FL HSCs (2611/E14 FL) were BMP-activated (Figure 11Fii). The remaining 19% were non-BMP-activated (Figure 11Fii) and were present in the GFP^- FL fraction at a 1:20,545 frequency (Figure 11Fi). The BM contained significantly more non-BMP-activated (1064/BM; 1:17,760 GFP^- cells) than BMP-activated HSCs (110/BM; 1:10,053 GFP^+ cells) (Figure 11Gi). The latter made up only 9% of BM HSCs (Figure 11Gii). For both FL and BM these quantifications corresponded with phenotypic HSC quantifications (Figure 11D). Thus, both FL and BM contain two types of functional HSCs, but the highest frequency and total number of HSCs is found in the BMP-activated FL fraction, with the BM containing mostly non-BMP-activated HSCs at a lower frequency.

To explore whether the HSCs had further differential characteristics, haematopoietic cell lineage-output of the different types of HSCs was tested by clonal transplantations. Four months post-clonal transplant, the recipients' blood was analysed for donor-derived lineage output. The ratio between lymphoid (B- + T-cells) and granulocyte-macrophage (GM) lineages

((B+T)/GM; Muller-Sieburg et al. (2004)) was calculated to determine clonal composition of the different HSC compartments and the type of donor-derived HSC in each recipient. The FL GFP$^+$ fraction contained slightly (though not significantly) more Bala(nced)-HSCs than the GFP$^-$ FL fraction but showed no further differences in clonal composition. In contrast, the BM fractions showed significant lineage-output differences. Bala-HSCs were only found in the BMP-activated BM fraction, while there were significantly more My(eloid-biased)-HSCs in the non-BMP-activated subset (Figure 11H). Similar numbers of Ly(mphoid-biased)-HSCs were found in both BM fractions (Figure 11H). Altogether, BMP activation status can be used to enrich for Bala- and My-HSCs from the adult mouse BM.

Also, the HSC lineage-output analysis revealed that the vast majority (88%; 15/17) of recipients of BMP-activated BM HSCs (Figure 11H) displayed boosted lymphoid differentiation output (Bala+Ly), whereas this was only seen in 53% (8/15) of recipients of the non-BMP-activated BM HSCs. Highlighting this difference was the upregulated expression of key lymphoid regulator genes encoding Ikaros (*Ikzf1*), E2A (*Tcf3*) and Flt3 in GFP$^+$ BM (LSK SLAM) HSCs, as compared to GFP$^-$ HSCs (Figure 11I). The FL HSC compartment, with equal frequencies of Bala- and Ly-HSCs in GFP$^+$ and GFP$^-$ subsets, revealed similar lymphoid gene expression in both FL fractions. Taken together, the BM niche stimulates lymphoid gene expression in BMP-activated HSCs, revealing intrinsic transcriptomic differences between two HSC types differentially regulated by BMP signalling.

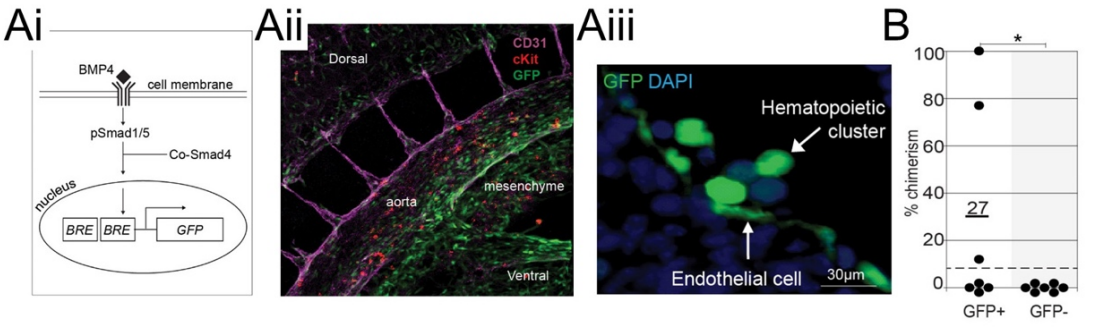
These data propose BMP activation status as one of the developmental origins of HSC heterogeneity.

6.2 Major findings

- Only some E10.5/11 AGM endothelial and IAHC cells are BMP-activated (BRE-GFP^{+}), however all functional AGM HSCs are BMP-activated.
- FL and BM contain non-BMP- and BMP-activated HSCs and utilisation of the *BRE-GFP* model allows for their prospective isolation.
- The highest frequency (1:180 cells) and total number of HSCs (2611/E14 FL; 81% of FL HSCs) is found in the BMP-activated FL fraction, with the BM containing mostly non-BMP-activated HSCs (1064/BM; 91% of BM HSCs) at a lower frequency (1:17,760).
- The BM niche stimulates lymphoid gene expression in BMP-activated HSCs.
- BMP activation status can be used to enrich for Bala- and My-HSCs from the adult mouse BM and might be one of the developmental origins of HSC heterogeneity.

6.3 Author contributions

Most figures/data in this study came from a vast number of *in vivo* haematopoietic transplantation experiments. CSV performed all transplantations (including intravenous injections, irradiation of recipients and maintenance of transgenic donor mouse line) and determined levels of donor chimaerism (including blood sampling, genomic DNA extraction and semi-quantitative PCR) for all recipients. Data produced through these experiments is presented in Figures 1E, 2B,C, 4A,C,E,F, Supp Figure 1C,D, and Supp Table 1-3 (of publication).

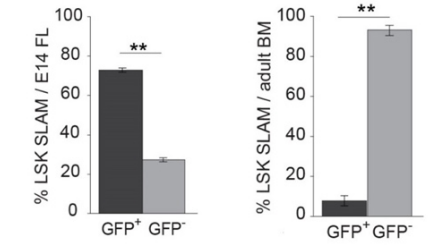


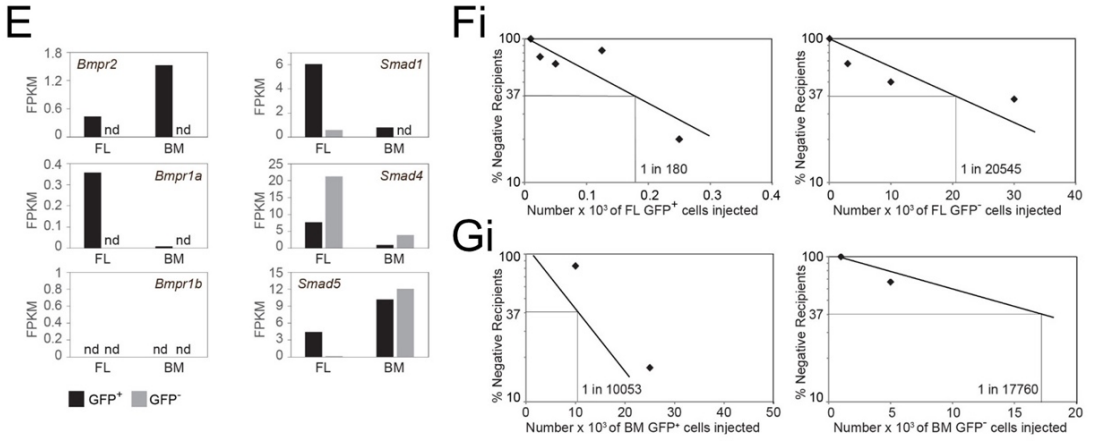
Ci

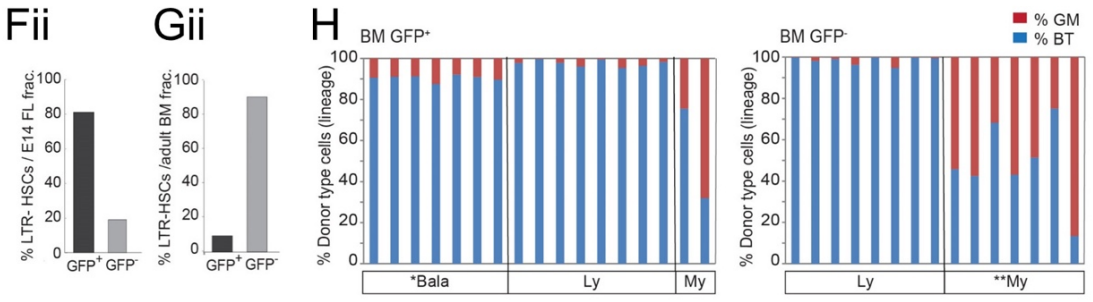
Cells transplanted	FL	Number of injected cells	1º repop/transplanted (% PB chimerism)	2º repop/transplanted (% PB chimerism)
GFP$^+$	E12	$3.3x10^4$	3/3 (100)	-
	E13	$4.8x10^4$	2/2 (100)	3/4 (15,100,100)
	E14	$3.3x10^4$	3/3 (100)	-
GFP$^-$	E12	$1.2x10^6$	3/3 (100)	-
	E13	$2.6x10^6$	2/2 (100)	3/4 (23,85,85)
	E14	$1.8x10^6$	3/3 (100)	-

Cii

Cells transplanted	BM	Number of injected cells	1º repop/transplanted (% PB chimerism)	2º repop/transplanted (% PB chimerism)
GFP$^+$	adult	$5.6x10^4$	5/5 (100)	-
		$3.0x10^4$	2/2 (62,100)	-
		$2.0x10^4$	3/3 (82,86,87)	3/3 (5,6,28)
		$1.7x10^4$	2/2 (10, 92)	-
GFP$^-$	adult	$2.0x10^6$	5/5 (100)	-
		$1.7x10^6$	2/2 (100)	2/3 (100,100)
		$1.0x10^6$	2/2 (100)	1/2 (30)
		$0.5x10^6$	3/3 (82,86,87)	1/3 (100)







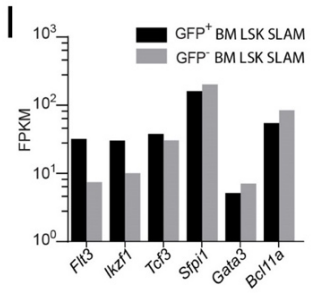


Figure 11 – BMP signalling differentially regulates distinct HSC types – **Ai)** Scheme showing activation of the canonical BMP signalling pathway through phospo-Smad-1/-5 binding of the double *BMP responsive element* (*BRE*) and the resulting transcription of *GFP*. **Aii)** Three-dimensional whole-mount image of an embryonic day (E)10.5 immunostained *BRE-GFP* mouse embryo (37 somite pairs). Anti-CD31 (magenta), cKit (red) and GFP (green) antibody staining shows the predominantly ventral distribution of GFP^+ cells in various cell types within the aorta and underlying mesenchymal cells. **Aiii)** High-magnification image of a transverse section of an E11 *BRE-GFP* dorsal aorta stained with anti-GFP-antibody and DAPI. Intra-aortic haematopoietic clusters contain GFP^+ and GFP^- cells. Some endothelial cells are GFP^+. **B)** Percentage donor cell chimaerism in the peripheral blood of adult irradiated transplant recipients at 4 months after injection of E11 AGM GFP^+ and GFP^- cells (2-4.5 embryo equivalents [ee] of AGM cells injected per mouse; n=3 independent transplantation experiments). *P=0.05 by z-test for proportions. **C)** Transplantation results of adult irradiated recipients injected with GFP^+ and GFP^- cells sorted from **Ci**) E12, E13 or E14 FL (n=3) or **Cii**) adult BM (n=4). Secondary irradiated adult recipients received BM cells, unsorted ($3x10^6$) or GFP^+ ($1.5x10^5$) and GFP^- (0.8, 1.8 and $11.8x10^6$) sorted cells from repopulated primary recipients. Data from all secondary recipients were pooled. Peripheral blood (PB) donor chimaerism is shown for all recipients. **D)** Percentages of LSK SLAM cells that are GFP^+ (black) or GFP^- (grey) as found in **Di)** E14 FL (n=3) and **Dii)** adult BM (n=3). **P<0.001 t-test. **E)** Fragments per kilobase of transcript per million mapped reads (FPKM) values for *BMP receptor* and *Smad* genes expressed by GFP^+ and GFP^- LSK SLAM sorted HSCs from E14 FL and adult BM. ND=not detected. **F-G)** Limiting dilution transplantation experiments show the frequencies of HSCs in GFP^+ and GFP^- cell fractions from **Fi)** E14 FL and **Gi)** adult BM. Bar graphs showing percentages of long-term repopulating (LTR) HSCs that are GFP^+ (black) or GFP^- (grey) in **Fii)** E14 FL (n=5) and **Gii)** adult BM (n=5) as calculated from limiting dilution transplantations. **H)** Summary of lineage-output analysis of repopulated Ly5.1 recipient mice injected with 14-32-week-

old *BRE-GFP* (Ly5.2) GFP^+ and GFP^- BM cells at the clonal HSC level (n=3). Donor chimaerism and lineage output in PB were measured at 4 months post-transplantation. PB was analysed for granulocyte-macrophage (GM) lineages (Gr-1 and CD11b) and B- and T-cell lineages (B220 and CD3) in the donor Ly5.2-cell-marker-expressing cells. Ratio between donor-derived lymphoid and myeloid (L/M = (B+T)/GM ratio) was calculated to reveal the clonal composition of the different HSC compartments and the type of donor-derived HSC in each recipient: My(eloid-biased) (0-4), Bala(nced) (4-16) and Ly(mphoid-biased) (>16). P values were calculated by Fisher exact test between the GFP^+ and the GFP^- fractions (P=0.007) and by two sample test for proportion (X^2) between balanced (*P=0.005), My (**P=0.031) and Ly-HSCs (P=0.727) of the two groups, GFP^+ and GFP^-. **I)** FPKM values for selected lymphoid genes expressed by GFP^+ and GFP^- LSK SLAM-sorted HSCs from adult BM. **(A-I)** Reprinted from Crisan et al. (2015).

Discussion

In the body of work presented here, we focussed on *de novo* generation of the rare but most potent contributor to the haematopoietic differentiation hierarchy: the HSC. The difficult-to-achieve acquisition of HSC identity occurs in the midgestation mouse aorta through EHT, a transdifferentiation process where specialised flat HECs begin to exhibit haematopoietic features as they round up and morph into IAHCs. Using a cKO approach we show that the key haematopoietic TF *Gata2* is required both for the generation of the earliest HSCs as well as their survival, revealing for the first time a requirement of *Gata2* during and after EHT.

We next developed a novel *Gata2-Venus* (*G2V*) reporter mouse that enabled precise tracing of *Gata2* (transcript and protein) expression. This revealed that all functional AGM HSCs (and most HPCs) express medium levels of Gata2. Surprisingly, we found that some HPCs, which are transcriptionally less complex and EMP-resembling, are generated Gata2-independently. We used the same model to decipher Gata2's role during haematopoietic fate establishment and discovered highly dynamic and unstable pulsatile Gata2 expression in single aortic cells undergoing EHT. *Gata2* haploinsufficiency changes this pulsatile behaviour and consequently disrupts EHT, stressing Gata2's crucial role in haematopoietic (stem cell) cell fate acquisition.

The transcriptomic changes driving EHT were explored by bulk RNA-seq on different aortic populations isolated from *Ly6A-GFP* embryos. We discovered that *Gpr56* is one of the most highly upregulated and expressed receptor genes during the HEC-to-HSC transition, making it a useful marker for AGM HSCs. Its importance was demonstrated, supported by our studies revealing an essential role for *gpr56* in HSC generation in the zebrafish embryo and a role for human *GPR56* in stem cell maintenance, as shown by us and others. Next, we captured the transcriptome of a pure population of functional AGM HSCs and connected the "inner" HSC regulatory network with single-cell function by iterations of Gata2-based functional,

transcriptomic and phenotypic analyses. This yielded a novel phenotypic HSC signature as defined by highly specific CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{+N} / CD27^{med} expression levels, enabling the highest HSC enrichment to date.

Lastly, HSC generation/maintenance was explored from an extrinsic regulatory angle and revealed that all AGM HSCs are BMP-activated. However, upon examination of FL and BM HSCs we observed HSC heterogeneity (shifts in lineage-bias) and a developmental progression towards the prevalence of non-BMP-activated HSCs.

Altogether, my research of the first functional HSCs arising during development, supports a model of a stochastic HSC induction programme in single embryonic aortic cells acquiring multipotentiality, as indicated by heterogeneous and asynchronous (and potentially unstable/dynamic) expression of pivotal in-/extrinsic genes/factors, hindering precise prediction of what rare HEC/'HSC precursor' will go on to acquire HSC fate.

Gata2 and its role during HSC ontogeny

Here, utilising loss-of-function and reporter mouse models, we examined the role of Gata2 *in vivo* during embryonic *de novo* HSC generation. Haematopoietic function was examined by *in vivo* transplantations and *in vitro* HPC assays. HSCs were localised *in vivo* by whole-mount imaging and expression dynamics examined by vital imaging (*Chapter 1-3*).

Genetic deletion pre- and post-HSC generation

Gata2 and Runx1 are both part of multimeric 'heptad' TF complexes essential for haematopoietic fate acquisition (Wilson et al., 2010). By using the same cKO approach as for *Runx1* (Chen et al., 2009), the functions of both TFs became directly comparable. Unlike *Runx1*, that is only essential during EHT (Chen et al., 2009), we showed that *Gata2* is required both during and post-EHT for HSC generation and survival. The deletion of *Gata2*

or *Runx1* in aortic ECs from E8.5 (*VEC-Cre:Gata2*$^{f/f}$ or *VEC-Cre:Runx1*$^{f/f}$) revealed that some HPCs and small IAHCs are still generated in the absence of *Gata2* (*Chapter 1*), while neither are made without *Runx1* (Chen et al., 2009). Even though both TFs (or at least activation of the *+23* enhancer in *Runx1*'s case) play an active role in haematopoietic fate determination in aortic ECs from E8.5 (Swiers et al., 2013), and Gata2 has been shown to regulate early haematopoietic *Runx1* expression (Nottingham et al., 2007), the complete absence of IAHCs/HPCs in the *Runx1* cKO (Chen et al., 2009) suggests that it is responsible for pan-haematopoietic cell emergence/fate determination and/or play a slightly earlier/more upstream role than Gata2 in this process. This would be in line with *Runx1*'s recently reported role in regulating the number of so-called pre-HECs that can become HECs by overcoming/passaging through a developmental bottleneck/barrier (Zhu et al., 2020).

Given the stark drop in IAHCs/HPCs upon *Gata2* deletion, Gata2 indisputably plays a key role in HECs undergoing EHT. However, it might not be required for cell emergence per se, but rather/additionally play an important role in establishment and maintenance of an immature HPC/HSC fate, as supported by *Gata2*-independent emergence/generation of only few IAHCs containing unipotent HPCs (but not bi-/multipotent HPCs or HSCs), and its crucial role in maintenance of stemness/long-term repopulating potential (post-HSC generation). The existence of *Gata2*-independent unipotent IAHCs/HPCs and the fact mouse embryos generate dorsal IAHCs that only contain HPCs but no HSCs (Taoudi and Medvinsky, 2007) also tells us that cell emergence/morphological change itself is not enough for establishment of HSC identity.

In summary, our work indicates Gata2 and Runx1 have partially overlapping and distinct functions during haematopoietic (stem/progenitor) fate acquisition and thereafter in HSCs and HPCs.

Generation of a reporter mouse model and isolation of Gata2-dependent and -independent HPCs

To study function and transcriptome of Gata2-expressing and non-expressing cells, without disturbance of Gata2 expression levels or protein function, we generated a reporter line with an *IRESVenus* sequence recombined into the *Gata2 3'UTR*. This produces bicistronic *Gata2IRESVenus* transcripts, resulting in 5'cap-dependent and *IRES*-directed translation into distinct Gata2 and Venus proteins, respectively. Although Venus fluorescence was technically reporting for *Gata2* transcript levels, careful analyses of our novel *Gata2*(*Venus*) reporter revealed a strong correlation between Gata2 and Venus protein and transcript expression levels, as shown by Western blot analysis (*Chapter 3*), bulk quantitative RT-PCR (*Chapter 2*) and bulk (*Chapter 3*) and single-cell (*Chapter 5*) RNA-seq. Moreover, whereas recently major differences in transcript and protein levels were revealed for some haematopoietic master TFs like Runx1, Scl and Gata1 (during erythroid differentiation), Gata2 displayed a high correlation between mRNA and protein levels, suggesting *Gata2* transcript levels are naturally indicative of Gata2 protein levels during haematopoiesis (Gillespie et al., 2020). Thus, our approach enabled precise *in vivo*/vital tracking of Gata2 (protein and transcript expression) levels based on Venus fluorescence. Additionally, normal numbers and quality of BM HSCs were observed in $G2^{V/V}$ adults, demonstrating that the inserted construct does not disrupt normal Gata2 function in the haematopoietic system. Flow cytometric and IHC analyses confirmed Gata2 expression in all expected (Minegishi et al., 1999; Minegishi et al., 2003; Robert-Moreno et al., 2005) haematopoietic sites of the midgestation embryo, including aortic ECs and IAHCs. Altogether, *G2V* is a suitable and reliable model for prospective isolation and characterisation of live Gata2-expressing cells.

Our functional assays revealed that all HSCs and the most immature HPCs in the AGM were exclusively found in the V^+ fraction. However, as observed in $Gata2^{-/-}$ (c)KO embryos (de Pater et al., 2013; Tsai et al., 1994)

(*Chapter 1*, *2*), the V^- fraction also contains a subset of less complex/potent HPCs. Most of these are produced in the E9 ($V^-/G2^{-/-}$) YS. The V^- HPCs show transcriptomic resemblance to EMPs based on chemokine receptor/ligand expression (Kierdorf et al., 2013). However, despite the V^- fraction containing most phenotypic EMPs in the E10 YS and AGM, the majority of definitive HPCs are V^+. Furthermore, some HPCs that emerge without Gata2 (V^-) *in vivo*, can initiate Gata2 expression (become V^+) upon *in vitro* CFU culture. Still, those HPCs that emerge Gata2-independently always remain less potent than those that emerge *in vivo* Gata2-dependently. Thus, our data reveals further heterogeneity within the embryonic HPC compartment.

The existence of the Gata2-independent HPCs raises the possibility that there is more than one way to *de novo* generate an HPC or achieve a specific haematopoietic fate: 1) different progenitor potencies could be linked to distinct regulatory programmes with unique (levels of) key players, 2) all YS HPCs might not necessarily all arise via EHT, and/or 3) different types of HECs could give rise to different HPCs (McGrath et al. (2015); reviewed in Dzierzak and Bigas (2018); McGrath and Palis (2005)). The latter scenario supported by proof of distinct HEC types competent of generating HSCs, EMPs or MPPs (Chen et al., 2011; Dignum et al., 2021; Hou et al., 2020), although this remains controversial due to the rarity and transient nature of the cells. Alternatively, these supposed distinct HEC types could be a result of an unstable programme in HECs that defies verbal classification. This would suggest a model where stochasticity already plays a role at the earliest stages of definitive haematopoietic development.

Redundancy

The existence of Gata2-independent HPCs and expression of *Gata3* and *Gata4* in V^- aortic ECs and HPCs suggests Gata factor redundancy during establishment of certain (embryonic) haematopoietic fates. As observed in other scenarios, when *Gata2* or *Gata3* are placed under the control of *Gata1* transcriptional elements in *Gata1* knockdown/deficient mice,

each is sufficient for rescue of embryonic (but not adult) erythropoiesis and the embryonic lethality associated with the knockdown (Takahashi et al., 2000). On the other hand, endogenous Gata2 and Gata3 (expression levels) are unable to rescue the *Gata1* deficiency phenotype, emphasising the importance of transcriptional control/expression levels and distinct Gata factor functions (reviewed in Zaidan and Ottersbach (2018)). The much longer half-life of Gata1 protein as compared to Gata2 protein (4 to >6 hours versus 0.5 to 2 hours, respectively) might also play a role in such rescue experiments (Lurie et al., 2008; Minegishi et al., 2005).

Gata3 is highly expressed in adult BM HSCs and, by postnatal deletion (with *Mx1-Cre*), was shown to be crucial for their maintenance, regulation of cell cycle entry and (HSC-/cell-autonomous) control of the equilibrium between self-renewal capacity and differentiation towards the more proliferative and less quiescent IT-HSCs (reviewed in Zaidan and Ottersbach (2018)). Conversely, using a *Vav-Cre*-driven cKO approach (like in *Chapter 1*; deletion after E11.5), *Gata3* was deemed not required for HSC expansion/survival and self-renewal (Buza-Vidas et al., 2011). Thus, precise functions of Gata3 in adult haematopoiesis remain unclear.

During mouse development within the region of the AGM, *Gata3* is expressed in cells of the sympathetic nervous system (SNS; located in the ventro-lateral mesenchyme), subaortic mesenchyme and in the aortic endothelial lining (Fitch et al., 2012; Swiers et al., 2013). Our results confirmed these observations and showed *Gata3* expression in V^+ and V^- ECs and IAHCs. Interestingly, a suggested cross-regulation of Gata3 and *Gata2* was revealed in the SNS and between cells of the (*Gata3*$^+$) mesonephric ducts and (*Gata2*$^+$) surrounding tissue within the urogenital ridges (Fitch et al., 2012). Upon *Gata3* germline deletion, SNS and mesonephric duct cells lost *Gata2* expression, however *Gata2* remains expressed within the endothelial lining, placing Gata3 upstream of *Gata2* in the SNS and urogenital ridges, but not in the aorta (Fitch et al., 2012). *Gata3* deletion did cause reduced IAHC numbers and HSC activity, but AGM HSCs themselves did not express *Gata3* (Fitch et al., 2012). In turn, Gata3 was

shown to have a non-cell autonomous effect on HSC ontogeny through regulating catecholamine production in SNS cells (Fitch et al., 2012), thereby revealing a novel link between the SNS and HSC generation within the AGM. The fact we and others have observed *Gata3*-expressing (*Gata2*$^{+}$ and *Gata2*$^{-}$) ECs in the aorta leaves open the possibility that Gata3 might play an additional cell-autonomous role in HECs/EHT and perhaps in emergence of Gata2-independent HPCs. Since deletion of *Gata2* has more severe effects on definitive haematopoiesis, it might be upstream of *Gata3* in V^{+} ECs and/or play a more fundamental role in EHT.

Traditionally not a haematopoietic TF, Gata4 is mainly associated with mammalian cardiac development. In non-mammalian organisms such as *Xenopus* and zebrafish the cardiac mesoderm is found directly neighbouring a population of haemangioblasts, both of which require Gata4, 5 and 6 for initial development (reviewed in Dobrzycki et al. (2020)). The mouse embryo harbours endocardial cells with haemogenic potential which can transiently *de novo* generate definitive erythroid/myeloid progenitors, as demonstrated by pre- and post-circulation heart tube explants (the latter in *Nxc1* KO embryos that lack heartbeat) (Nakano et al., 2013). Another study tracking *Gata4* expression as driven by its allantois/placenta- and lateral mesoderm/AGM-specific *G2* enhancer (active between E7.5-12.5) showed a significant amount of foetal and adult blood cells (and BM HSCs) derived from the embryonic *Gata4*-expressing cells (Canete et al., 2017). At E10.5, *G2*-driven YFP colocalised with Gata4 and Runx1 in some aortic, FL and placental ECs, confirming our observations for the aorta. Interestingly, YFP^{+}/*Gata4*$^{+}$ cells within the placenta did not express Gata2. YFP^{+} AGM, FL and placenta cells all showed CFU-C activity, however upon transplantation of YFP^{+}-sorted cells in newborn mice, only the FL and placenta provided high-level reconstitution, suggesting a placental origin for these (neonatal) HSCs. Altogether, the observations of Canete et al. (2017) of the presence and function of embryonic *Gata4*$^{+}$Gata2^{-} HPCs further suggest and support a role for *Gata4* in the generation of the Gata2-independent HPCs observed by us.

Expression dynamics during EHT

Through (c)KO approaches and examination of populations of *Gata2*(*Venus*)-expressing cells, we have established *Gata2* as an essential TF for multipotent haematopoietic (stem cell) fate acquisition and HSC survival. Our *G2V* model next facilitated careful examination of *Gata2*'s dynamic expression behaviour in HECs undergoing EHT. Gata2 has a short half-life (0.5-2 hours; Lurie et al. (2008); Minegishi et al. (2005)) which necessitated the use of a fluorescent reporter protein with similar characteristics. Instead of the more commonly used GFP reporter, Venus (an improved version of YFP) enabled efficient reporting of *Gata2* expression. The characteristics of Venus included an improved speed of protein folding/maturation, similar half-life to Gata2 (2 hours instead of 9 for GFP), higher fluorescent intensity (hence its name after Earth's second brightest night sky object) and a lower pH sensitivity (Li et al., 1998; Nagai et al., 2002).

Quantitative comparisons of G2V (*Chapter 2,3*) and Ly6A-GFP (*Chapter 4*) expression in cells of the midgestation aorta, revealed that both are expressed in similar numbers of ECs at E10.5, but that Gata2 is already expressed more widely at earlier stages. Gata2 is found in 1076 ECs/AGM (by FACS; 28-36sp; ±11% of ECs) and the number of Ly6A-GFP^{+} ECs increases from 190 (at 32sp) to 1203 (at 37sp; by IHC; 13-19% of aortic ECs by FACS at E10.5). On the other hand, Ly6A-GFP expression in IAHCs is more restricted (97 cells at 32sp to 389 at 37sp; by IHC) compared with Gata2 (680 cells; by FACS; 28-36sp). We also show that Gata2 marks around 20 EHT events/E10 embryo, whereas Ly6A-GFP only two (Boisset et al., 2010). This suggests Gata2 is a broader and earlier EHT and HEC marker than Ly6A-GFP. Further support for this comes from functional haematopoietic assays that reveal only 33% of AGM HPCs express Ly6A-GFP (*Chapter 4*) while around 80% express Gata2 (*Chapter 3*). Since the first AGM HSCs express both Ly6A-GFP and G2V, these results suggest that

Ly6A-GFP marks the most immature EHT events and Gata2-expressing aortic cells within IAHCs.

Expression levels and associated function

Through our work (*Chapter 1-3*, *5*; Ling et al. (2004)) and that of others (Minegishi et al., 2003; Persons et al., 1999; Rodrigues et al., 2008; Rodrigues et al., 2005; Tipping et al., 2009) it has become clear that expression levels, gene/dosage and timely up- and downregulation of *Gata2* are critical for the generation, survival and maintenance of functional HSCs and differentiation of HSPCs into mature blood cells. *Gata2* haploinsufficiency causes a severe reduction in AGM HSCs and results in qualitatively defective BM HSCs that are more quiescent, apoptotic and have reduced repopulating potential in young adults (Ling et al., 2004; Rodrigues et al., 2005). Reduced GMP functionality is also seen in these mice (Rodrigues et al., 2008). Upon ageing, *Gata2*$^{+/-}$ mice (18-20 months old) lose CLPs, undergo a reversal of the earlier GMP and apoptotic HSC effect, and their BM HSCs become more proliferative, with a reduced and myeloid-biased repopulation ability (Abdelfattah et al., 2021).

Enforced high-level expression of *Gata2* in adult BM cells (to levels seen in HSCs) leads to an increase in quiescence, and a block in proliferation and differentiation (Persons et al., 1999; Tipping et al., 2009). Constitutive overexpression of *Gata2* in AGM cells also blocked differentiation and the production of mature CD45^{+} cells (Minegishi et al., 2003). Still, some mature haematopoietic cell types like mast cells and basophils do require/express high levels of Gata2 (reviewed in Sasaki et al. (2016)). Thus, while both *Gata2* haploinsufficiency and high-level overexpression result in increased quiescence, the downregulation of *Gata2* is required for HSPCs to undergo lineage commitment/differentiation and ensure homeostatic haematopoiesis.

Interestingly, low-level *Gata2* upregulation can achieve self-renewal in cells otherwise lacking this potential, and high(er) *Gata2* levels are

associated with myeloid commitment (Nandakumar et al., 2015). In this study, the use of a tamoxifen-inducible *Gata2*-expression vector resulted in constitutive low-level ('leaky') *Gata2* overexpression in the absence of tamoxifen. This particular *Gata2* level (equivalent to that seen in leukaemic blasts) was sufficient to bestow myeloid HPCs with self-renewal capacity and resulted in myeloid expansion and a lymphoid differentiation block upon transplantation (Nandakumar et al., 2015). Our results (*Chapter 3*, *5)* are in agreement with this and indicate that all self-renewing AGM HSCs (and 71% of HPCs) express medium (V^{med}) levels of Gata2 and that this specific expression level is realised by an upregulation in Gata2 expression during the HEC-to-BC transition. This is then followed by unstable/dynamic pulsatile Gata2 expression as cells undergo EHT, after which a further increase in expression is seen before levels stabilise. We also find some IAHC cells that express higher levels of Gata2 (V^{hi}). This fraction contains HPCs and cells with a more mature myeloid/innate immune cell signature.

Taken together, we reveal the importance of a 'just-right' dose of Gata2 at the right time and for the right duration, in order to *de novo* generate and maintain normal numbers of HSCs. This is also reflected by the fact that a too high or low Gata2 dose is associated with myeloid leukaemogenesis (reviewed in Katsumura et al. (2017)). Gata2's extreme dose-sensitivity is also likely to be connected to its autoregulatory nature/properties (reviewed in Crispino and Horwitz (2017)).

Pulsatile expression and its regulators

Several master regulators are known to exhibit (often stochastic) pulsatile expression. Pulsing of the NF-κB, ERK, MAPK and TGF-β signalling pathways has been shown in mammalian cells, and highlights the possibility that several of these pulsatile circuits and TFs cooperate, coregulate and coexist in single cells (reviewed in Levine et al. (2013)). Dynamic pulsing is caused by genetic circuits that repetitively activate and deactivate key regulators to ultimately allow individual cells to (asynchronously) acquire

specific cellular functions and/or fates. These circuits contain (delayed) negative-feedback loops that permit build-up of a signal/pulse and its subsequent termination (reviewed in Levine et al. (2013)).

Our data present for the first time the necessity of dynamic pulsatile Gata2 expression for successful EHT in the AGM region (*Chapter 3*). The pulsatile and unstable expression observed during EHT, as an EC state gets 'disassembled' and a haematopoietic state/fate assembled, is indicative of a metastable cell state and dynamic instability caused by counteracting (endothelial and haematopoietic) transcriptional programmes, a phenomenon also observed during haematopoietic cell-fate differentiation/specification (Olsson et al. (2016); reviewed in Graf and Enver (2009)). The competing endothelial and haematopoietic (stem cell) transcriptional programmes result in, what one might call, an 'intolerable' level of instability that is eventually resolved by dominant establishment of the immature haematopoietic programme in a rare number of HECs. A critical regulator of *Gata2* in this process/transition, located directly upstream, is the Notch signalling pathway and specifically Notch1 (reviewed in Bigas and Porcheri (2018)). Notch1 additionally regulates expression of *Hes1*, which in turn has a negative feedback on/represses *Gata2* expression, resulting in a Notch1-Gata2-Hes1 "type I incoherent feed-forward loop" (reviewed in Alon (2007); Bigas and Porcheri (2018)). Like *Gata2*, *Hes1* has been shown capable of autoregulation (downregulating its own expression) and oscillatory expression behaviour, further supporting the possibility of multiple pulsatile circuits needing to cooperate to achieve certain cell fates (reviewed in Bigas and Porcheri (2018); Levine et al. (2013)). Hes1 has a short (stable) half-life of around 22 minutes, however oscillatory expression of one of its regulators results in a de- or increase of Hes1's half-life, which initiates *Hes1*'s own oscillatory expression (Yoshiura et al., 2007). Our data reveal significantly higher expression of both *Notch1* and *Hes1* in our most immature V^{med} cells compared to the more mature V^{hi} cells (*Chapter 3*). Notch1-induced *Gata2* and *Hes1* expression is therefore likely to result in Hes1 blocking *Gata2* transcription upon reaching a certain threshold, making Hes1 the main driver

of *Gata2*'s pulsatile expression. Additionally, Hes1's relatively short half-life and/or its own pulsatile behaviour might affect both periodicity and amplitude of *Gata2*'s expression. The fact that low *Notch1* and *Hes1* levels are detected in more mature cells that express high levels of *Gata2* (V^{hi}) indicates this *Gata2* expression might no longer be driven by Notch1 (and blocked by Hes1), in line with the downregulation of *Notch1* seen upon HSPC differentiation (reviewed in Ottersbach (2019)). Thus, Gata2 pulsatile expression during the EC-to-HSC transition is likely a result of the combined regulation of *Gata2* by Notch1 and Hes1.

Gata2 haploinsufficiency resulted in more frequent and intense Gata2 pulses at an earlier stage of the EHT. In this model (Tsai et al., 1994), all *Gata2*'s *cis*-regulatory elements are still intact, but one of its zinc-finger domains has been deleted, giving *Gata2*$^{+/-}$ embryos half the amount of functional TF protein. Thus, the change in expression dynamics should not be due to a change in binding of upstream regulators to *Gata2*'s *cis*-elements, but rather the decrease in positive autoregulation by Gata2 itself.

In summary, we show that 1) our *G2V* model is a unique tool for prospective isolation of live (non-)Gata2-expressing cells especially when they undergo EHT and transit through an (otherwise-hard-to-isolate) metastable state, 2) *Gata2* is not required for cell emergence/shape change in the midgestation aorta per se, as indicated by *Gata2*-independent formation of IAHCs, 3) dynamic and pulsatile upregulation of Gata2 expression to optimal V^{med} levels is required for HECs to undergo EHT and (stochastically) acquire and maintain multipotentiality and self-renewal capacity, and 4) Gata2's pulsatile expression is likely driven by the combined actions of Gata2 autoregulation and upstream regulators Notch1 and Hes1.

We therefore suggest that Gata2's unstable/dynamic expression behaviour together with asynchronous fluctuations of other key regulators is critical for the stochastic acquisition of HSC fate, in a similar way the asynchronous combinatorial actions of Gata3, Runx1 and Notch signalling are required for T-cell fate establishment (Kueh et al., 2016).

In search of a transcriptomic and phenotypic signature for the first functional HSCs

In *Chapter 4* and *5* we identify the transcriptome and novel (combinations/expression levels of) markers of the first functional *de novo*-generated HSCs by bulk RNA-seq, reiterations of single cell *in vivo* and *in vitro* haematopoietic assays, and scRNA-seq of highly enriched aortic cells acquiring HSC fate.

RNA-seq of enriched EHT populations

To get a better understanding of transcriptomic changes occurring during/responsible for EHT and HSC generation we used our *Ly6A-GFP* model (*Chapter 4*). Confocal microscopy revealed that the number of Ly6A-GFP$^+$ cells in the aorta increases over time, but that the number of HECs undergoing EHT and giving rise to GFP$^+$ IAHC cells remained rare. Functional haematopoietic assays showed that both ECs and HECs (CD31$^+$Ly6A-GFP$^-$ or GFP$^+$) contained no HPC or HSC potential. Instead, all HPCs were found as part of IAHCs (CD31$^+$cKit$^+$), with 33% of total HPC numbers being Ly6A-GFP$^+$ and 64% GFP$^-$. Ly6A-GFP expression in IAHC cells exclusively marked all LT-HSCs (0.002% of AGM cells) and thus allowed for high-level HSC enrichment.

Here, we isolated small cell numbers of highly enriched EHT populations (ECs, HECs and HSCs) from E10.5 AGM and processed them for RNA-seq using the SMARTer (Switching Mechanism At 5' end of RNA Template) protocol. Earlier transcriptomic (gene expression profiling/ microarray) studies on HSC ontogeny either did not include HECs (Mascarenhas et al., 2009; McKinney-Freeman et al., 2012) or examined HECs and other EHT populations for expression of only limited numbers of genes and E10.5 AGM HSPCs not as highly enriched for HSCs as our Ly6A-GFP$^+$ subset (Swiers et al., 2013). Furthermore, the wide (*Runx1*)*23GFP* expression seen at E8.5 (68% of the conceptus' ECs; 29% of pSp+VA+UA

ECs; Swiers et al. (2013)), when these (pre-)HECs are not ready to undergo EHT (and Ly6A-GFP is not yet expressed in HECs (de Bruijn et al., 2002)), indicates an important early role for *Runx1* and suggests our Ly6A-GFP^{+} HEC subset is more likely to include the HECs actively undergoing EHT at E10.5. Indeed, *Runx1* was recently found to play a critical role in controlling a developmental bottleneck/barrier at the stage of the pre-HEC-to-HEC transition (Zhu et al., 2020).

Additionally, our use of RNA-seq has the advantage that it can detect a much larger range of expression levels (and thus more DEGs) than microarrays and is specifically suitable for detection of low-abundance transcripts. RNA-seq provides superior specificity and sensitivity for detection of the whole transcriptome, including novel genes and transcript isoforms. Our datasets revealed that one *Gpr56* transcript variant was solely expressed in HSCs, whereas another variant was detected in ECs, HECs and HSCs. This latter variant showed higher abundance in HSCs. Thus, our datasets allow for the most thorough and complete analysis of transcriptomic changes during (with some likely to be responsible for) EHT.

We show that there are 139 DEGs between EC and HECs, and 340 between HECs and HSCs. Genes upregulated in HECs showed (GO) enrichment in adhesion and (regulation of) cell migration gene sets. This underlines their potential for (or the fact they are actively undergoing) morphological shape/adhesive property change during the HEC-to-HSC/HPC/IAHC transition to a new cell state that requires adherence to other IAHC cells and migratory potential when taking on haematopoietic function. It also highlights the HECs' intermediate state whilst transitioning from static ECs into motile haematopoietic cells. Our RNA-seq datasets confirm the gradual downregulation/disassembly of an endothelial transcriptional programme and establishment of a haematopoietic one during EHT. We detect similar expression of *Gata2*, *Runx1*, *Meis1*, *Scl*, *Sfpi1*, *Gfi1*, *Myb* and *Gata3* in our cell fractions as in E10.5 EHT subsets examined by single-cell qPCR (Swiers et al., 2013), thereby validating expression of these master regulators in our datasets. Our HEC population also revealed

upregulation of receptors linked to angiogenesis when compared to ECs. Some of the highest upregulated ones were *Ramp2* and the G-protein-coupled receptor *Calcrl* (*CRLR*). Heterodimers of Ramp2 (Receptor activity modifying protein 2) and Calcrl modify the latter to become sensitive/selective to/for adrenomedullin (*Adm*) and mediate its pro-angiogenic effect on ECs (reviewed in Nagaya et al. (2005)). Furthermore, under hypoxic conditions both *Adm* and *Calcrl* are upregulated by hypoxia-inducible factor-1α (*Hif1a*) (reviewed in Nagaya et al. (2005)), which we have previously demonstrated to be an important regulator of HSC and HPC generation in the mouse AGM (Imanirad et al., 2014). Indeed, some ECs and entire IAHCs were found to be hypoxic in the E10 AGM (Imanirad et al., 2014). More recently, both *Adm* and *Ramp2* were found expressed in IAHCs and aortic ECs, and shown to fulfil a conserved role in HSPC generation in mouse, chicken and zebrafish embryos (Yvernogeau et al., 2020). Thus, revealing and confirming that both angiogenic- and hypoxic-induced/-dependent regulators play a role in ECs taking on a haematopoietic fate.

Gpr56

Our RNA-seq analysis revealed that all AGM HSCs express high levels of *Gpr56*, as a result of it being one of the most highly upregulated receptor genes during the HEC-to-HSC transition. We also showed all heptad TFs are located directly upstream of *Gpr56*. This is in line with the severe downregulation of *Gpr56* in E11.5 AGM explants in which *Gata2*'s *+9.5* enhancer, essential for EHT and *de novo* HSC generation, was deleted (Gao et al., 2013). Furthermore, we demonstrate *Gpr56* is expressed in rare ECs and most IAHCs and co-localises with Ly6A-GFP in the midgestation AGM. Through KD experiments, we reveal that *gpr56* is essential for HSC emergence/generation in zebrafish embryos, and that its functional coding sequence is conserved between mouse and zebrafish, with both *Gpr56* and *gpr56* mRNA capable of rescuing HSPC generation in morphant fish. The rescue experiments also resulted in some ectopic HSPC induction.

Altogether, our data suggest an important conserved role for Gpr56 in *de novo* HSC generation.

In contrast, two earlier studies using the same germline *Gpr56* KO mice (deletion of exon 2 and 3) revealed surprisingly mild (conflicting) phenotypes (Rao et al., 2015; Saito et al., 2013). These mice were viable and saw no change in HSPC maintenance and function, but only revealed reduced HSC-repopulating capability upon serial transplantation (Rao et al., 2015). This study also confirmed the high *Gpr56* expression we had observed in WT IAHCs (and some adjacent ECs) and revealed high expression in WT BM LT-HSCs and a gradual downregulation upon differentiation. Saito et al. (2013) discovered a slightly stronger phenotype in *Gpr56* KO mice, with a significant decrease in numbers of BM HSCs and their self-renewal potential. They also revealed reduced cell adhesion and increased migration of BM HSCs in the absence of *Gpr56*. Both would be in line with the fact most adhesion GPCRs, like Gpr56, are known to have a double role in signal transduction and cell adhesion (reviewed in Liebscher et al. (2021)). Given its high expression in IAHCs, as demonstrated here by us (*Chapter 4*) and others (Rao et al., 2015), one of Gpr56's roles could be to regulate/maintain the IAHC microenvironment and, through controlling cell adhesion and migration properties, determine the duration of an HSPC's residency within an IAHC.

The mild and conflicting results obtained with these *Gpr56* KO mice are likely due to one of *Gpr56*'s four splicing variants (S4; start site in exon 4) still being transcribed as a result of a hypomorphic instead of a full KO allele. This would explain the low levels of residual Gpr56 protein (possibly sufficient for some HSPC function) seen in these mice (Rao et al., 2015). Therefore, a complete *Gpr56* (c)KO (of all isoforms) is required to reveal its true function in mouse HSCs.

The lack of a strong phenotype could also be due to redundancy by the highly homologous Gpr97 (*Adgrg3*) and/or Gpr114 (*Adgrg5*), found in the same 5'-to-3' orientation, in relative close proximity down- and upstream of *Gpr56*, respectively (Maglitto et al. (2021); reviewed in Liebscher et al.

(2021)). Indeed, we have recently shown that *Gpr97* can rescue HSPC generation in *gpr56* morphant zebrafish (Maglitto et al., 2021). Furthermore, *Grp97* expression was upregulated in mouse FL HSCs and mESC-derived HPCs upon conditional (*Vav-Cre* and *VEC-Cre*) or permanent deletion of all *Gpr56* isoforms, respectively (Maglitto et al., 2021). These mutant cells revealed a myeloid bias in clonal transplantations and HPC assays. Donor-derived BM HSCs (post-cKO FL transplantation) did not show an upregulation of *Gpr97*. Earlier on in development, the *Gpr56* cKO affected E9 YS haematopoiesis, however expression of *Gpr97* was normal. No strong phenotype (or upregulation of *Gpr97*) was observed in the *Gpr56* cKO E10.5 AGM, and both AGM and FL contained normal numbers of phenotypic HSPCs. Strikingly, differentiated mESCs showed a near complete loss of HPC potential upon deletion of both *Gpr56* and *97* (Maglitto et al., 2021).

Thus, deletion of *Gpr56* revealed that Gpr97 might play a redundant/ compensatory role in FL HSCs and mESC-derived definitive HPCs, but not at the onset of definitive haematopoiesis or in the post-transplant BM. It remains unclear whether both genes undergo similar transcriptional/translational regulation and if the myeloid bias is due to upregulation of *Gpr97* or loss of *Gpr56*. The latter would indicate that at least one of Gpr56's roles is to ensure HSC balanced multipotency/lineage output. HSCs could become endowed with this property the moment they arise via the HEC-to-HSC transition in the AGM, as suggested by our observation of the high upregulation of *Gpr56* during EHT. Since low levels of Gpr56 appear sufficient for relatively normal HSC function (Rao et al., 2015), the same might hold true for Gpr97. Upon cKO of *Gpr56*, *Gpr97* remains expressed to normal/WT levels in the AGM. Its redundancy in the AGM and during HSC ontogeny can therefore not be excluded; especially given *Gpr97*'s potential to rescue zebrafish HSPC generation and the dramatic reduction in mESC-derived HPCs upon deletion of *Gpr56* and *Gpr97*. A double germline (c)KO mouse, if viable at all, will be required to examine the roles of both genes during haematopoiesis. Additionally, *in vitro* and *ex vivo* (explant) manipulations of single or double (*Gpr56*, *97* and *114*) (c)KO cell lines and

AGM tissues, respectively, with specific small-molecule modulatory (partial) agonists for Gpr56 and Gpr114 (Stoveken et al., 2018), or Gpr97 (Gupte et al., 2012), or antagonists for Gpr56 and Gpr114 (Stoveken et al., 2016), could provide further insights in the roles these GPCRs play during HSC generation/maintenance. These compounds are hypothesised to be more efficacious in activating adhesion GPCRs than natural extracellular matrix protein ligands like collagens, as they can bypass the natural multistep/-component activation process of adhesion GPCRs and directly activate them (Stoveken et al., 2016; Stoveken et al., 2018).

Iterative single-cell functional, phenotypic and transcriptomic analyses

Our previous Ly6A-GFP-based HSC enrichment strategy had allowed for isolation of 774-1776 cells from 4-14 E10 embryos for bulk RNA-seq (*Chapter 4*). Whereas this analysis had revealed novel genes highly expressed in an HSC-enriched AGM population, the unique transcriptome of the first 1-2 functional HSCs per embryo was only represented at a low frequency compared to that of other haematopoietic cells.

The earliest HSCs are found in IAHCs on the ventral side of the middle segment of the E10.5-11.5 mouse dorsal aorta, near the VA (Mascarenhas et al., 2009; Taoudi and Medvinsky, 2007). This region also contains the highest density of IAHCs, which vary in size from 1-19 cells to 1-11 at E10.5 and E11.5, respectively (Yokomizo and Dzierzak, 2010). We and others have revealed great phenotypic, functional and transcriptomic heterogeneity within IAHCs (*Chapter 2-6*; Baron et al. (2018); Yokomizo and Dzierzak (2010); Zhou et al. (2016)). The scarcity of HSCs amongst this heterogeneous mix of different HPCs and pre-HSCs suggests the acquisition of HSC identity is more challenging than that of any other haematopoietic fate. Up until now, IAHC heterogeneity and lack of unique (combinations/ levels of) markers have impeded isolation, precise localisation and transcriptomic capture of a pure population of HSCs.

Given the pivotal role of the heptad TFs during HSC ontogeny, here we performed iterations of Gata2-mediated single-cell functional and transcriptomic enrichment analyses in combination with index-sorting, to directly connect the “inner” HSC regulatory programme with single-cell function/identity (*Chapter 5*). Functional single-cell assays revealed multipotent HPCs were highly enriched, and HSCs exclusively found within the $CD31^{+}cKit^{hi}G2^{med}$ IAHC subfraction. From around E9.5-10.5 the AGM becomes colonised by primordial germ cells (PGCs) that migrate via the hindgut into the subaortic (ventrolateral) mesenchyme to eventually end up in the urogenital ridges, where they will differentiate into oo- and spermatogenic lineages (reviewed in Saitou and Yamaji (2012)). Even though PGCs do not have haematopoietic function, they share some phenotypic markers and are $CD31^{+}cKit^{low/-}$ (and $SSEA\text{-}1^{+}$) (Yokomizo and Dzierzak, 2010). Others have tried to overcome any potential mesenchymal contamination by performing intra-aortic (as opposed to whole-tissue digest) staining for cKit and CD31, and mechanical isolation of entire IAHCs (Baron et al., 2018). Although enabling isolation and transcriptomic analysis of ventral and dorsal clusters, this approach impeded the functional haematopoietic analysis of the different transcriptomic/*in silico*-predicted populations. Furthermore, some aortic $cKit^{low}$-expressing cells were recently found to be HSC-/multipotent-primed HECs (Dignum et al., 2021; Hou et al., 2020). Thus, our stringent enrichment approach based on a combination of specific CD31, medium Gata2 and the high cKit expression levels, only exhibited by a fraction of IAHCs and nowhere else within the AGM, allowed for careful functional, transcriptomic and phenotypic analysis of the most immature IAHC cells.

Runx1-negative IAHC cells

We isolated 968 cells from around 30 AGMs (embryos) and performed scRNA-seq using SmartSeq2. Despite high-level HSPC enrichment, Louvain analysis revealed five transcriptomic clusters within our subset, with a majority (529 of 968) of cells exhibiting an (im)mature haematopoietic

transcriptome. Unexpectedly, the other major cluster within our cKit^{hi} IAHC subset consisted of *Runx1*-negative IAHC cells with an endothelial-like signature. Another scRNA-seq study had previously also revealed *Runx1*$^{-}$ cells as part of (cKit^{+}) IAHCs; these cells however did not group in a unique transcriptomic cluster in their analysis (Baron et al., 2018).

Index data obtained during FACS provided us with protein expression levels of all fluorescent enrichment markers/genes for each single cell, and permitted isolation and functional testing of the five transcriptomic clusters. We showed that our *Runx1*$^{-}$ fraction did not contain any directly measurable HPC/CFU-C or adult-repopulating HSC function. While *Runx1* plays an important role in haematopoietic specification prior to EHT (Swiers et al., 2013; Zhu et al., 2020) and is crucial for cell emergence during EHT, it is no longer required for HSPC function thereafter (Chen et al., 2009). Our *Runx1*$^{-}$ subset could therefore contain cells that were specifying towards a haematopoietic fate and previously expressed *Runx1* long enough to enable their emergence into IAHCs, but subsequently downregulated its expression resulting in (or as a result of) an abortive EHT and lack of haematopoietic function. Abortive EHT events have so far only been revealed by live-imaging of *Runx1* MO knockdown zebrafish embryos (Kissa and Herbomel, 2010). In these embryos, the initiation of rare EHT was observed, however HECs were found to burst as soon as they bulged out of the endothelium, establishing a crucial role for *Runx1* in aortic cell emergence in zebrafish. However, EHT in fish is different from mouse, with HEC-derived HSPCs directly budding into the sub-aortic space instead of forming IAHCs (Boisset et al., 2010; Kissa and Herbomel, 2010). Such differences could be due to the extra whole-genome duplication zebrafish underwent during evolution, causing them to currently still have many duplicate gene copies and consequently different gene regulation, including for example two functionally distinct paralogous genes (*gata2a* and *gata2b*) of the haematopoietic master regulator *Gata2* (reviewed in Dobrzycki et al. (2020)). Therefore, whereas abortive EHT events in fish result in the bursting of cells (as IAHCs cannot be formed), in

mice, these could very well result in accumulation of such cells in IAHCs, which could explain our *Runx1*$^{-}$ subset.

Alternatively, given some of our endothelial-like *Runx1*$^{-}$ cells express medium/high CD41 and/or low CD43, these could represent precursors to HSCs ("pro-/pre-HSCs") (Rybtsov et al., 2011; Taoudi et al., 2008) (and/or "pre-HPCs") unable to directly reveal a haematopoietic output. Such cells would only display HSC function upon *in vitro* culturing (Ganuza et al., 2017; Inlay et al., 2014; Rybtsov et al., 2016) or transplantation into neonates but not adult mice (Boisset et al., 2015; Yoder et al., 1997a). A subset of pre-HSCs supposedly also serves as a second embryonic origin of B-1a cells (Hadland et al., 2017). Upon maturation, these pre-HSCs would not give rise to LT-HSCs, but instead turn into developmentally-restricted (B-1a producing) FL HSCs (Godin et al. (1993); Hadland et al. (2017); reviewed in Ghosn et al. (2019)). Since maturation (and proliferation) of any of these precursors into transplantable HSCs requires *in vitro* culture steps, it is uncertain whether all pre-HSCs that can achieve HSC function *in vitro* will actually achieve HSC function and identity *in vivo*.

Besides pre-HSCs, some of our cells could be in a 'neutral'/plastic state awaiting establishment of another haematopoietic identity, possibly the MPP that plays an important role in sustaining unperturbed (non-transplant) steady-state haematopoiesis (Busch et al., 2015; Rodriguez-Fraticelli et al., 2018; Sun et al., 2014).

Taken together, our scRNA-seq dataset separates into five transcriptomic clusters and reveals a heterogeneous mix of (im)mature haematopoietic and *Runx1*$^{-}$ endothelial-like cells within a phenotypically highly enriched HSC fraction.

Heterogeneous heptad TF expression

The notion that all HSCs must mature from HECs via a trajectory of three sequential (intermediate) pro-/pre-HSC 'states' (Rybtsov et al., 2011; Taoudi et al., 2008), reminiscent of the many well-defined progenitor stages

of T-cell development (reviewed in Rothenberg (2021)), seems unlikely. Whereas some HECs might mature via intermediate pre-HSC states, we suggest that others could directly transition into HPCs or HSCs, as indicated by the direct emergence of single HSPCs from flat HECs (Boisset et al., 2010), existence of HE with different potencies (Chen et al., 2011; Dignum et al., 2021; Hou et al., 2020), the stochastic nature of HSC fate acquisition (and/or HECs) as suggested by Gata2's unstable/dynamic expression behaviour (*Chapter 3*), and the heterogeneous and asynchronous expression of heptad TFs within the different transcriptomic clusters and single cells of our highly enriched scRNA-seq dataset (*Chapter 5*).

HSPC specification, generation and transcriptional control is reliant on the simultaneous co-expression and combinatorial interaction of Gata2, Scl, Runx1, Lyl1, Lmo2, Erg and Fli1 with each other and/or the HSPC cistrome, culminating in a multimeric heptad TF complex occupying pivotal *cis*-regulatory elements (Wilson et al. (2010); reviewed in Hewitt et al. (2016)). The consequences of single/double (haploinsufficient) deletion or overexpression of Gata2 and Runx1 on definitive (more complex) haematopoiesis, as observed by us and others, suggest small level changes of any of the heptad TFs might have serious effects on assembly of such a protein complex and ultimately cell fate (Wilson et al. (2010); reviewed in Dzierzak and Bigas (2018)). The dynamic instability during EHT, together with/as a result of pulsatile and unstable Gata2 expression (*Chapter 3*) and variable/heterogeneous (and perhaps stochastic) expression of the other heptad TFs in single cells (*Chapter 5*), implies stochasticity and the need of a combinatorial 'code' of optimal heptad TF expression levels to unlock an HSC identity 'combination lock' in single HECs. Altogether, this indicates the realisation of a stabilised immature HSC programme is stochastic, which would be in line with the extremely low number of functional HSCs in the midgestation mouse embryo and the fact simple (non-dynamic/-pulsatile) overexpression of TFs *in vitro* can only generate some less 'complex' blood lineages but not HSCs (reviewed in Dzierzak and Bigas (2018)).

Capturing the HSC transcriptome and localising the first HSCs

Reiterations of our various single-cell methods have enabled the highest AGM HSC enrichment to date. All HSCs were contained in the HC1 transcriptomic cluster, which our index data had allowed us to isolate by altering our strategy from CD31^{+}cKit^{hi}G2^{med} to CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}, hereby achieving a further 10.5-fold enrichment. Interestingly, in line with characteristics of our E11 (HC1) HSPCs, E9.5-10.5 HSC-primed HECs with haematopoietic-only (and not endothelial-only or dual) potential were also recently found to be SSC^{lo} (Hou et al., 2020), indicating both might have similar cellular granularity/complexity. To achieve additional enrichment, we analysed Louvain clusters (L1.7 and 8) for potential HSC markers. *CD27* (*Tnfrsf7*), a pro-inflammatory tumour necrosis factor receptor family member, was expressed at varying (high) levels by nearly all HC1 cells and thus presented itself as a putative candidate. Its expression had previously been shown on AGM, FL (Li et al., 2017) and adult BM (pre-)HSCs (Vazquez et al., 2015). We found 70% of HC1 cells expressed CD27; however, several rounds of transplantations revealed it was only a narrow FACS gate (CD27^{+N} HC1) with the brightest-expressing 19% of HC1 cells that contained all HSCs, while none were found in the remaining negative/low-expressing 81% of HC1 (CD27$^{-/lo}$ HC1), even with injections of >4 to 18 times more cells. Single-cell CFU-C assays on CD27^{+N} HC1 cells led to discovery of an even higher enriched CD27^{medium} expression level that contained 95% of HSC-like progenitors, yielding an extra 6.5-fold enrichment in addition to the 10.5-fold HC1 enrichment. Notably, no unipotent HPCs were present in the CD27^{+N} HC1 fraction. Thus, all functional HSCs (and most immature HPCs) are found within the CD31^{hi}SSC^{lo}cKit^{hi}G2^{med} (HC1) CD27^{+N} / CD27^{med} IAHC fraction; consisting of ±20-23 cells/AGM.

Other potential enrichment markers were *Selp* (CD62P/P-selectin) and *Mpl* (CD110). P-selectin is expressed on adult BM HSCs and its deletion increases HSC self-renewal, suggesting a role in regulating the fine balance between self-renewal and differentiation (Sullivan et al., 2011). *Mpl* was

previously shown expressed in IAHCs, with deletion causing delayed onset of HSC generation and decreased HSC survival/self-renewal in the FL (Petit-Cocault et al., 2007). Thus, both markers provide further possibilities for isolation of the first 1-2 HSCs.

A scRNA-seq reiteration with $CD27^{+N}$ cells revealed nearly all (85 of 119) merged into the transcriptomic HC1 cluster. The remaining cells mostly located to the endothelial-like/*Runx1*$^{-}$ cluster and could perhaps be some of the pre-HSCs shown to express CD27 (Li et al., 2017). Out of the 85 cells, our index data revealed that 23 expressed both the $CD27^{med}$ HSC range and HSC marker *Ly6A* (see *Chapter 4*; de Bruijn et al. (2002)). The near-equivalent transcriptomic profiles of eight of the 23 cells stood out. These cells displayed the highest hscScores (Hamey and Gottgens, 2019) and were equally distributed between and located at the neighbouring borders of L2.7 and L2.8, suggestive of active acquisition of HSC identity. Their signature revealed high expression of *Hes1*, *Hey1*, *Dll4*, *Sox17* and *Nos3*, amongst others. *Hes1* is downstream of Notch1 and is a negative regulator of *Gata2*, possibly affecting Gata2 pulsatile expression during EHT (*Chapter 3*; reviewed in Bigas and Porcheri (2018)); Dll4-induced Notch1 signalling is required for specification of aortic/arterial endothelium, after which Jag1-Notch1-driven Notch activity allows for EHT and IAHC formation (reviewed in Heck et al. (2020)). Recently, Dll4 was shown to (negatively) regulate IAHC size by blocking recruitment of haemogenic cells, and as a result control HSC production (Porcheri et al., 2020); Sox17, upstream of *Notch1*, also plays a role in generation/regulation of HE/IAHCs and embryonic HSCs, and is important for maintenance of self-renewal (Clarke et al., 2013; He et al., 2011; Nobuhisa et al., 2014); and *Nos3* (*eNOS*) expression, possibly induced by blood flow, is important for IAHC formation and generation of functional HSCs (North et al., 2009). Taken together, the stratified results from our functional, transcriptomic and phenotypic/index-sorting single-cell iterations capture the likely transcriptome and genetic identity of the first functional HSCs in mouse (eight HSCs obtained from six embryos). We reveal a signature that is distinctive but closely related to that of immature HPCs

which provides important clues to further our understanding of HSC fate establishment/acquisition.

Given the extremely rare nature of HSCs and a lack of exclusive markers, their precise location within IAHCs is much debated. Do HSCs emerge as pro-/pre-HSCs that require maturation, or can they undergo a HEC-to-HSC transition and directly acquire definitive HSC fate upon emergence? Do HSCs remain in small 1-cell clusters, or do they require multicellular IAHCs? Or are both scenarios possible with one a natural 'back-up' mechanism for the other?

As mentioned previously, live-imaging in mouse and zebrafish embryos demonstrates direct emergence of single round cells displaying all phenotypic (and functional) HSC criteria (Bertrand et al., 2010; Boisset et al., 2010; Kissa and Herbomel, 2010), and the E11.5 aorta contains around 500 $cKit^{+}$ IAHC cells of which $\pm$100 are present as single-cell 'clusters', as shown by whole-mount 3D confocal imaging (Yokomizo and Dzierzak, 2010). Additionally, zebrafish EHT causes extravasation of HSPCs without the formation of IAHCs. One might propose zebrafish have a simpler blood hierarchy than mice, which allows for direct emergence of HSPCs from HECs without the need for any maturation. Whereas in mouse, with a potentially more complex blood, IAHCs might function for maturation of more complex HPCs/pre-HSPCs/developmentally-restricted HSCs. Given HSPCs can directly emerge from HECs in fish, the direct HEC-to-HSC transition could be the most conserved part of blood formation and might therefore also take place in mice.

Here, we provide further evidence for the latter. Through correlating CD27 with Ly6A-GFP and Gata2 expression by IHC, we demonstrate that the first HSCs localise to small 1- to 2-cell clusters in the E10.5-11.5 AGM, which is in support of a direct HEC-to-HSC transition. We also reveal that macrophages, crucial for HSPC generation in the AGM (Mariani et al., 2019), interact with these single emerged $CD27^{+}$ cells. The high *Dll4* expression observed in our eight HSCs and its absence in the other 15 ($CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{med}$) cells is in line with Porcheri et al. (2020)

showing that Dll4 maintains smaller IAHCs. Taken together these results suggest that Dll4 might play a role in maintenance of HSC quiescence.

In summary, our thorough analysis of the transcriptomic changes driving EHT and defining HSCs reveals a gradual commitment towards an HSC identity, with important early roles for Notch signalling, and angiogenic-, hypoxic- and blood flow-induced regulators. Gpr56 is highly upregulated during EHT and most likely ensures a balanced HSC lineage output in mice. Stratified results from our single-cell methods reveal several transcriptomic clusters and a heterogeneous mix of (im)mature haematopoietic and *Runx1*$^{-}$ endothelial-like cells within a small HSC-containing IAHC fraction. Our stringent Gata2-based single-cell enrichment approach yielded a novel phenotypic/functional HSC signature (CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{med}; ±20-23 cells/AGM) and suggests some of the remaining cells could be products of an abortive EHT event or are unable to reveal direct HSPC output. Our stratified scRNA-seq data also enabled capture of the genetic identity of the rare first functional HSCs (eight cells from six embryos) and reveal a distinct signature (cell state) closely related to that of immature HPCs. Combining these data and those of others with our visual localisation of HSCs supports the notion of a direct HEC-to-HSC transition, with HSCs subsequently residing in small IAHCs until required. Altogether, our work implies that the realisation of a stabilised immature HSC programme is stochastic and reliant on a combinatorial 'code' of unique TF expression levels, in line with the extremely low number of *de novo*-generated functional HSCs.

Extrinsic factors in HSC ontogeny

In *Chapter 6*, we explored HSC generation, maintenance and lineage output from an extrinsic regulatory angle, facilitated by use of a reporter mouse for *in vivo* BMP pathway activation in which all BMP-activated cells express GFP. Rather than relying on surface markers, not always functionally relevant, the *BRE-GFP* model allowed us to directly connect the activation

status of HSCs with lineage-output potential. We demonstrated here that all *de novo*-generated AGM HSCs are BMP-activated *in vivo*. In contrast, upon *ex vivo* explant culture, despite an increase in GFP^+ cells, we found that the AGM is capable of producing (functionally distinct) BMP- and non-BMP-activated HSCs (Crisan et al., 2016), indicative of a shift towards FL conditions, which is perhaps due to lack of adjacent dorsal and ventral tissues. These results underline the importance of the AGM niche and the influence of HEC/HSC-extrinsic factors/tissues on HSC generation. They may also explain why both FL and BM, with dissimilar adjacent tissues, are mostly considered sites of HSC expansion/maintenance rather than *de novo* generation, like the AGM. Nonetheless, EHT has recently been reported of late foetus/young adult BM HECs (Yvernogeau et al., 2019). Interestingly, this study showed that out of all BM (LSK) HSPCs it was only those generated from BM HECs that revealed enrichment for BMP (and Notch) pathway activation genes, underpinning the importance of these pathways and their extrinsic activators for EHT, irrespective of the haematopoietic site.

Though none have facilitated exclusive isolation, various approaches have succeeded in prospectively enriching for functionally distinct/biased BM HSCs by use of molecular phenotypes based on levels of Hoechst dye efflux (Challen et al., 2010) or differential expression of cKit, CD150, or CD229 (reviewed in Crisan and Dzierzak (2016)). As shown here by our results, responsiveness to extrinsic growth factors provided another means of separating distinct HSC types. Both BMPs and TGF-βs are part of the TGF-β superfamily of growth factors. Whereas receptor binding of BMPs leads to phosphorylation of R-Smad-1, -5 and -8, TGF-βs cause R-Smad-2 and -3 phosphorylation, resulting in regulation/control of different downstream targets (reviewed in Blank and Karlsson (2011)). Indeed, upon *in vitro* exposure or *in vivo* administration, TGF-β1 was shown to have a stimulatory effect on proliferation of My-HSCs, while it inhibited Ly-HSCs. However, clonal transplantations had revealed functional heterogeneity in their My-/Ly-HSCs-enriched fractions (Challen et al., 2010). Our data on BMP responsiveness showed that 9% of BM HSCs were BMP-activated and 91%

non-BMP-activated. All Bala-HSCs were exclusively in the BMP-activated fraction, whereas the non-BMP-activated fraction contained significantly more My-HSCs, probably regulated by TGF-β signalling. Given the equal proportions of Ly-HSCs amongst both of our BM fractions, the vast majority (88%) of BMP-activated HSCs displayed increased lymphoid (Bala- and Ly-HSCs) output, compared to only 53% of non-activated cells. In line with this, BM HEC-derived HSPCs (that showed BMP activation) also appeared to preferentially differentiate towards lymphoid lineages (Yvernogeau et al., 2019). In fact, this study showed that around 5% of LSK SLAM HSCs had derived from (young adult) BM HECs, which is similar to the 9% of BMP-activated adult (LSK SLAM) BM HSCs we observed, and suggests that some of these HSCs were *de novo* generated in the BM. Contrary to the BM, 81% of our FL HSCs were BMP-activated and 19% not activated. Proportions of Bala-/Ly-/My-HSCs were not significantly different between the two differentially responsive subsets, but My-HSCs were most infrequent in both subsets. Thus, BMP activation status has allowed us to prospectively and reliably isolate functionally distinct HSCs at a clonal level and our results reveal a developmental progression towards the prevalence of non-BMP-activated HSCs in the adult. BMP signalling/activation may play its biggest role during/just after EHT towards HSC fate, or only in *de novo*-generated HSCs (and not those expanded in the FL), as supported by an increased lymphoid output of BMP-activated BM HSCs and the apparent similar functional output of FL HSCs irrespective of BMP activation status. Our discovery of non-BMP-activated HSCs might also explain the lack of phenotype upon conditional deletion of downstream BMP targets *Smad1* and *5* during FL and BM haematopoiesis (*Smad8* is not expressed in FL/BM HSCs) (Singbrant et al., 2010). Additional studies will be required to examine the extrinsic/niche factors responsible for maintenance/generation of the non-BMP-activated HSCs, and to determine whether HSCs can actively switch between BMP-activated and non-activated states or only transition from activated to non-activated (or remain activated) and not vice versa, as indicated by our explant studies (Crisan et al., 2016).

AGM HSCs are most likely directly activated by the Bmp4 that is produced in high levels in mesenchymal cells immediately surrounding the ventral aspect of the dorsal aorta (Durand et al., 2007). Many other extrinsic factors also differentially influence *de novo* HSPC generation (reviewed in Heck et al. (2020)). Recently we showed that the direct dynamic interaction of HSC-independent pro-inflammatory macrophages with IAHCs in the AGM positively regulates HPC and HSC generation (Mariani et al., 2019). In *Chapter 5*, we show that AGM macrophages interact with single round/emerged $CD27^{+}$ cells in the aortic endothelial wall. It is these, and possibly other, innate immune/primitive myeloid cells and even (H)ECs that are the likely local sources of inflammatory cytokines like TNF-α and IFN-γ, both essential for HSC generation, and that establish the AGM as a (sterile) pro-inflammatory microenvironment (Li et al., 2014; Mariani et al., 2019). Whereas IFN-γ is crucial for HSC generation, IFN-α was later shown important for AGM HSC maturation and expansion. The roles of both cytokines are conserved between mouse and zebrafish (reviewed in Demerdash et al. (2021)). Interestingly, AGM-associated macrophages also showed upregulated expression of *Bmp2*, which binds the same type II receptor subunit (Bmpr2) as Bmp4, providing another possible source of Bmp for the generation of the first HSCs (*Chapter 6*; Mariani et al. (2019); reviewed in Sadlon et al. (2004)). *Bmp2k* (Bmp2-inducible kinase), a downstream target of Bmp2, was upregulated in our HSC signature (*Chapter 5*). Though not much is known about this protein, it might provide a further link between HSCs and macrophages. The importance of IFN-γ for HSC generation (Li et al., 2014) nicely ties in with our *Ly6A-GFP* model (*Chapter 4* and *5*). *Ly6A-GFP* contains a transcriptional regulatory fragment with two hypersensitive sites responsible for high-level IFN-γ-inducible (*Ly6A-*)*GFP* expression (Ma et al., 2001; Sinclair et al., 1996), thereby enabling great enrichment of the earliest *de novo*-generated HSCs (de Bruijn et al., 2002).

Blood flow-associated mechanical (hemodynamic) forces like frictional wall shear stress (parallel to endothelial surface) and the circumferential cyclic stretch/strain (perpendicular to endothelial surface; blood pressure) are

equally critical for establishment of HSC-generating potential in the aortic endothelial lining through nitric oxide (NO) signalling-induced upregulation of *Runx1* expression (reviewed in Heck et al. (2020); Horton et al. (2021)). Our embryonic HSC signature data (*Chapter 5*) reveal the high expression of NO pathway member *Nos3* (*eNOS*), emphasising the importance of this pathway in the first functional HSCs. Interestingly, shear stress has also been shown to (in)directly activate both the Bmp and Notch signalling pathways, ultimately resulting in expression of *Gata2*, amongst other crucial genes (reviewed in Heck et al. (2020)). The combination of (unique levels of) hemodynamic forces (and downstream targets), working specifically on/in the ventral aortic floor in the middle segment of the AGM where blood flow is disrupted at the intersections with the remodelling VA (and its tributaries) (Zovein et al., 2010), might ultimately determine why it is that this particular region at this particular time of development contains the highest density of IAHCs (Yokomizo and Dzierzak, 2010) and all functional HSCs (Mascarenhas et al., 2009; Taoudi and Medvinsky, 2007).

Still, our *ex vivo* live imaging (*Chapter 3*; Boisset et al. (2010)) shows normal/predicted numbers of HSC EHT events in the absence of any blood flow, arguing against the necessity for hemodynamic forces in HSC generation. One possible explanation could be that these forces have already induced a programme of haematopoietic specification in aortic HECs prior to isolation of the embryos.

Taken together, BMP activation status has enabled reliable and prospective isolation of functionally distinct (Bala- and My-) BM HSCs. All AGM HSCs are BMP-activated *in vivo*, whereas examination of FL and BM HSCs revealed a developmental progression towards the prevalence of non-BMP-activated HSCs. Besides the extrinsic growth factor Bmp, the successful generation of HSCs from ventral aortic HECs is reliant upon correct timing, dosing and presence of pro-inflammatory cytokines and cells, growth factors and a unique combination of hemodynamic forces.

Conclusions

As HSCs are the potent stem cells for the blood system, the first (allogeneic) HSC transplantation (Thomas et al., 1957) paved the way for their current use in treatment of a wide range of (acquired and congenital) blood disorders and a growing number of autoimmune diseases (reviewed in Copelan (2006)). Despite the broad applicability of these transplantations, with often no curative alternatives, the lack of suitable donors and high risk of allogeneic complications/death highlights the need for an unlimited source of engraftable HSCs. Though great progress has been made, many years of research have not yet enabled us to engineer customised/patient-specific autologous or 'universal' allogeneic long-term, high-level, self-renewing multilineage HSCs, or efficiently expand *in vivo*-generated HSCs with preservation of all their characteristics (reviewed in Blau and Daley (2019); Dzierzak and Bigas (2018); Lange et al. (2021); Rowe et al. (2016)). The mammalian embryo holds the key to overcome these challenges, given that *de novo* generation of HSCs from aortic endothelium occurs only during a limited time in development. A better understanding of the molecular mechanisms behind the *in vivo* HSC fate/identity acquisition in the embryonic aorta is therefore of vital importance and is what has driven the research presented in this body of work, leading to the following conclusions:

Gata2 is essential for generation and survival of HSCs, and hence functions during and after the EHT. Our novel *G2V* model provided a unique tool for prospective isolation and examination of live Gata2-expressing cells, especially during EHT, when HECs transit through an otherwise-hard-to-isolate unstable/metastable state. We show that *Gata2* is not required for haematopoietic cell emergence per se, but rather for establishment and maintenance of an immature HPC/HSC fate. HECs can only undergo EHT and achieve such a fate (and multipotentiality and self-renewal capacity) by dynamic and pulsatile upregulation of Gata2 expression to optimal V^{med} levels. It is most likely the combined actions of Gata2 autoregulation and its upstream regulators Notch1 and Hes1 that drive this pulsatile expression.

Our data therefore suggest that the combinatorial effects of Gata2's unstable/dynamic expression and asynchronous fluctuations of other key regulators leads to the stochastic acquisition of HSC identity.

Our whole-transcriptome analysis of EHT showed that *Gpr56*, which we found directly downstream of all heptad TFs, is one of the most highly upregulated and expressed genes during EHT and consequently in the first mouse HSCs. In zebrafish, we revealed that *gpr56* is essential for HSC generation. Some of our later work indicates that *Gpr56* most likely ensures a balanced HSC lineage output in mice.

Focussing on a highly enriched phenotypic and functional HSC-containing IAHC fraction, our single-cell approaches revealed transcriptomic heterogeneity and a mix of (im)mature haematopoietic and *Runx1*$^{-}$ endothelial-like cells. Heterogeneous and asynchronous expression of the pivotal heptad TFs was again suggestive of unstable/dynamic expression/regulation and/or a stochastic HSC induction programme. Our stringent Gata2-based single-cell enrichment yielded a novel phenotypic/functional HSC signature (CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{med}) and indicated some of the remaining cells in our dataset could be products of an abortive EHT or unable to reveal direct HSPC output. We also conclude the first HSCs most probably reside in small 1- to 2-cell IAHCs, which supports the notion of a direct HEC-to-HSC transition. Ultimately, our stratified scRNA-seq results enabled capture of the genetic identity of the rare first functional HSCs (eight cells from six embryos). The sum of their gene/TF expression revealed a distinct signature (cell state) closely related to that of immature HPCs. Based on our functional validations of the single-cell transcriptomics clusters, this cell state/signature is likely representative of that of the very first cells that have acquired HSC fate/identity. Our unique HSC dataset and interactive resource website will therefore be valuable for future research aiming to understand and mimic *de novo* generation of HSCs, as indicated by its use in several recent studies (Bresciani et al. (2021); Demirci et al. (2020a); Dignum et al. (2021); Jackson et al. (2021); Maglitto et al. (2021); Tang et al. (2021); Ulloa et al. (2021); reviewed in Demirci et al. (2020b);

Horton et al. (2021); Jaffredo et al. (2021); Koyunlar and de Pater (2020); Liggett and Sankaran (2020); Weijts et al. (2021)).

Our research on the role of the extrinsic morphogen Bmp in HSC generation/maintenance led us to reliably and prospectively isolate functionally distinct (Bala- and My-) BM HSCs based on BMP activation status. During ontogeny, all AGM HSCs are BMP-activated *in vivo*, and we discovered a developmental progression towards the prevalence of non-BMP-activated HSCs by examination of FL and BM HSCs.

In conclusion, the great diversity of indispensable in- and extrinsic HSC-inductive signals highlights the robustness of *de novo* HSC generation where an optimal system is created through timely expression/presence and precise dosage and combination of niche factors (growth factors, pro-inflammatory cytokines, macrophages, hemodynamic forces) and availability of sufficient numbers of specified arterial HSC precursors/HECs (distinct HSC-primed or with programme plasticity). My research supports a model where this (in)direct interaction instigates asynchronous and unstable/ dynamic expression of pivotal (heptad) TFs in these precursors, and their optimal combinatorial levels result in a rare number of unstable/metastable HECs that can stochastically acquire self-renewal capacity and multipotentiality (HSC identity), as they undergo a direct HE-to-HSC transition. Other less complex/potent haematopoietic (progenitor) fates might only require combinatorial expression of some of these TFs or can be generated in more than one way, resulting in an easier fate to achieve and a greater abundance.

The research presented here has brought the field one step closer to the realisation of strategies for producing human clinically relevant HSCs. It will require the integration of single-cell transcriptomics, proteomics, epigenetics and functional analyses of AGM HSCs, together with novel vector systems allowing for controlled, dosed (pulsatile) and timed, expression of pivotal factors, to ultimately achieve the engineering of customised HSCs.

Future work

The conclusions of the studies presented in this thesis work provide the foundations for, and stimulate, the following immediate and future plans:

- **Spatiotemporal localisation, isolation and characterisation of the *Runx1*$^{-}$ (*AplnR*$^{+}$) IAHC population in the AGM.** This previously unexplored population of transcriptomically-similar cells, discovered in our HSC signature scRNA-seq dataset (*Chapter 5*), may have a supportive role, be the result of abortive EHTs or contains cells unable to directly reveal haematopoietic function. SmartSeq2 offers complete transcript coverage for each single cell, including all splice variants. By determining the ratio of spliced to unspliced transcripts for a given gene, RNA velocity (La Manno et al., 2018) can establish changes in gene expression and indicate how actively a gene is expressed at the time of cell isolation, i.e. presence of more unspliced transcripts is indicative of active expression. When done for all transcripts in a single cell (and all cells in the dataset), this is predictive of a cell's future transcriptome, and direction and speed of cell change. Our preliminary RNA velocity results indicate that the HC1 (HSC) cluster does not derive from the *Runx1*$^{-}$ endothelial-like IAHC cells, but rather these are both likely to originate from a population that is missing from our HSC dataset. Given that no good antibodies exist to isolate live AplnR^{+}Runx1^{+} cells, we will initially try to localise these cells in fixed AGM tissue sections. Additionally, we will examine phenotypically-similar cells derived from *in vitro* haematopoietic differentiations of a novel *AplnR-tdTomato* mESC reporter line (Jackson et al., 2021), and attempt to generate a reporter mouse strain for *in vivo* studies. The *Runx1-IRES-GFP* mouse model (Lorsbach et al., 2004) may also be used. Alternatively and in addition, several fluorescent ligands for AplnR will be tested for use in cell isolations for functional assays.

- **Retrospective analysis of the active transcriptome of functionally validated *de novo*-generated HSCs and HPCs, using a novel DCM-ID technology.** This technique links the bacterial methyl transferase DCM to RNA polymerase II subunit b (Rpol2b). Expression of *DCM-Rpol2b* is regulated by a doxycycline-inducible promoter. When the transgene is activated, DCM will localise to and methylate actively transcribing genes. This bacterial methylation is not normally present in mammalian cells. To obtain the active HSC/HPC transcriptome, we will activate DCM gene marking at the time of HSC/HPC generation in the midgestation mouse embryo and analyse the actively transcribed genes in differentiated HSC/HPC offspring by methylated DNA sequencing (MeD-seq; Boers et al. (2018)). For HSCs this would be in differentiated haematopoietic cell offspring post-transplantation of marked AGM cells, whereas HPCs (marked prior to HSC generation) will be examined by CFU-C assay. Altogether, this should yield the active HSC/HPC transcriptome signature and provide direct information about HSC cell fate. This will nicely complement our scRNA-seq HSC signature (*Chapter 5*) which is more indicative of HSC cell state and exact levels of gene expression. Initial experiments will try to establish whether DCM marking (transgene integrated in mouse *Col1A1* locus) can be achieved in cells of the haematopoietic lineage. For this, we will activate marking around the time of EHT and analyse offspring in the FL $CD45^+$ compartment.
- **Develop an embryonic HSC score (E-hscScore)** predictive of a cell's transcriptomic similarity to the first *de novo*-generated HSCs, based on our unique scRNA-seq AGM HSC signature and similar to the adult HSC score (hscScore/MolO score; Hamey and Gottgens (2019); Wilson et al. (2015)).
- **Perform simultaneous multi-colour live imaging on AGM/aorta to visualise and quantify gene expression behaviour of all heptad TFs during EHT.** Whereas FPKMs of the heptad TFs in our eight

HSCs provide information about the gene expression levels at the time of isolation, they do not reveal whether dynamic/pulsatile expression was required to reach this HSC state, as is the case for Gata2. To address this, we would generate fluorescent reporter mouse models carefully reporting for Runx1, Fli1, Scl, Lmo2, Lyl1 and Erg protein expression (like our *Gata2Venus* model). We would start by focussing on Runx1, given its well-known crucial role during EHT, and try to generate a reporter with a fluorescent protein distinct from (Gata2)Venus/GFP to perform simultaneous multi-colour live imaging for both Runx1 and Gata2 (and CD31, cKit and CD27). Instead of using confocal imaging on thick live sections of AGM, we would perform live imaging on entire aortae using light-sheet microscopy. This would enable the capture of all rare EHT events occurring in an intact aorta in one go. Light-sheet microscopy is well-suited for long-term imaging of large/millimetre-sized samples and provides much faster image acquisition and lower photobleaching and -toxicity than point scanning techniques like confocal microscopy (reviewed in Weber et al. (2014)). These less harsh imaging conditions, allowing for longer imaging and maintenance of fluorescence, should help us determine how many (phenotypic) HSCs emerge through a direct HE-to-HSC versus a HE-to-IAHC-to-HSC event (requiring maturation in IAHCs). In addition to live imaging, this technique is suitable for rapid acquisition of static multi-colour/marker 3D images of the entire midgestation aorta (see e.g. Figure 4D). We will therefore also use it for quick testing and visualisation of expression of some of our novel putative HSPC markers. The live aorta is flushed with a solution of the desired directly-conjugated antibodies, dissected out and then directly imaged. This technique thus has the advantage over conventional IHC on fixed tissue sections or fixed whole embryos (and confocal microscopy), that it does not require fixations (possibly affecting antibody binding) or clearing of tissues, and allows for use of directly-conjugated antibodies (instead of primary and secondary antibodies)

and direct visualisation of fluorescent reporter proteins (instead of antibodies against these).

- **Generation of double *Gpr56-Gpr97* germline (c)KO mouse** to examine the roles of both genes during haematopoiesis.

Bibliography

Abdelfattah, A., A. Hughes-Davies, L. Clayfield, J.B. Menendez-Gonzalez, A. Almotiri, B. Alotaibi, A. Tonks, and N.P. Rodrigues. 2021. Gata2 haploinsufficiency promotes proliferation and functional decline of hematopoietic stem cells with myeloid bias during aging. *Blood Adv* 5:4285-4290.

Abramson, S., R.G. Miller, and R.A. Phillips. 1977. The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems. *J Exp Med* 145:1567-1579.

Adamo, L., and G. Garcia-Cardena. 2012. The vascular origin of hematopoietic cells. *Dev Biol* 362:1-10.

Adolfsson, J., R. Mansson, N. Buza-Vidas, A. Hultquist, K. Liuba, C.T. Jensen, D. Bryder, L. Yang, O.J. Borge, L.A. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson, and S.E. Jacobsen. 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* 121:295-306.

Akashi, K., D. Traver, T. Miyamoto, and I.L. Weissman. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 404:193-197.

Alberts, B., A. Johnson, J. Lewis, D. Morgan, M.C. Raff, K. Roberts, P. Walter, J. Wilson, and T. Hunt. 2015. Molecular biology of the cell. Garland Science, Taylor and Francis Group,, New York, NY.

Alon, U. 2007. Network motifs: theory and experimental approaches. *Nat Rev Genet* 8:450-461.

Baron, C.S., L. Kester, A. Klaus, J.C. Boisset, R. Thambyrajah, L. Yvernogeau, V. Kouskoff, G. Lacaud, A. van Oudenaarden, and C. Robin. 2018. Single-cell transcriptomics reveal the dynamic of haematopoietic stem cell production in the aorta. *Nat Commun* 9:2517.

Beck, D., J.A. Thoms, D. Perera, J. Schutte, A. Unnikrishnan, K. Knezevic, S.J. Kinston, N.K. Wilson, T.A. O'Brien, B. Gottgens, J.W. Wong, and J.E. Pimanda. 2013. Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood* 122:e12-22.

Becker, A.J., C.E. Mc, and J.E. Till. 1963. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature* 197:452-454.

Bee, T., K. Liddiard, G. Swiers, S.R. Bickley, C.S. Vink, A. Jarratt, J.R. Hughes, A. Medvinsky, and M.F. de Bruijn. 2009. Alternative Runx1 promoter usage in mouse developmental hematopoiesis. *Blood Cells Mol Dis* 43:35-42.

Beerman, I., D. Bhattacharya, S. Zandi, M. Sigvardsson, I.L. Weissman, D. Bryder, and D.J. Rossi. 2010. Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a

mechanism of clonal expansion. *Proc Natl Acad Sci U S A* 107:5465-5470.
Benveniste, P., C. Frelin, S. Janmohamed, M. Barbara, R. Herrington, D. Hyam, and N.N. Iscove. 2010. Intermediate-term hematopoietic stem cells with extended but time-limited reconstitution potential. *Cell Stem Cell* 6:48-58.
Bertrand, J.Y., N.C. Chi, B. Santoso, S. Teng, D.Y. Stainier, and D. Traver. 2010. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* 464:108-111.
Bertrand, J.Y., A. Jalil, M. Klaine, S. Jung, A. Cumano, and I. Godin. 2005. Three pathways to mature macrophages in the early mouse yolk sac. *Blood* 106:3004-3011.
Bigas, A., and C. Porcheri. 2018. Notch and Stem Cells. *Adv Exp Med Biol* 1066:235-263.
Bigas, A., A. Robert-Moreno, and L. Espinosa. 2010. The Notch pathway in the developing hematopoietic system. *Int J Dev Biol* 54:1175-1188.
Blank, U., and S. Karlsson. 2011. The role of Smad signaling in hematopoiesis and translational hematology. *Leukemia* 25:1379-1388.
Blau, H.M., and G.Q. Daley. 2019. Stem Cells in the Treatment of Disease. *N Engl J Med* 380:1748-1760.
Boers, R., J. Boers, B. de Hoon, C. Kockx, Z. Ozgur, A. Molijn, I.W. van, J. Laven, and J. Gribnau. 2018. Genome-wide DNA methylation profiling using the methylation-dependent restriction enzyme LpnPI. *Genome Res* 28:88-99.
Boiers, C., J. Carrelha, M. Lutteropp, S. Luc, J.C. Green, E. Azzoni, P.S. Woll, A.J. Mead, A. Hultquist, G. Swiers, E.G. Perdiguero, I.C. Macaulay, L. Melchiori, T.C. Luis, S. Kharazi, T. Bouriez-Jones, Q. Deng, A. Ponten, D. Atkinson, C.T. Jensen, E. Sitnicka, F. Geissmann, I. Godin, R. Sandberg, M.F. de Bruijn, and S.E. Jacobsen. 2013. Lymphomyeloid contribution of an immune-restricted progenitor emerging prior to definitive hematopoietic stem cells. *Cell Stem Cell* 13:535-548.
Boisset, J.C., T. Clapes, A. Klaus, N. Papazian, J. Onderwater, M. Mommaas-Kienhuis, T. Cupedo, and C. Robin. 2015. Progressive maturation toward hematopoietic stem cells in the mouse embryo aorta. *Blood* 125:465-469.
Boisset, J.C., W. van Cappellen, C. Andrieu-Soler, N. Galjart, E. Dzierzak, and C. Robin. 2010. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* 464:116-120.
Bradley, T.R., and D. Metcalf. 1966. The growth of mouse bone marrow cells in vitro. *Aust J Exp Biol Med Sci* 44:287-299.
Bresciani, E., B. Carrington, K. Yu, E.M. Kim, T. Zhen, V.S. Guzman, E. Broadbridge, K. Bishop, M. Kirby, U. Harper, S. Wincovitch, S. Dell'Orso, V. Sartorelli, R. Sood, and P. Liu. 2021. Redundant mechanisms driven independently by RUNX1 and GATA2 for hematopoietic development. *Blood Adv* 5:4949-4962.
Busch, K., K. Klapproth, M. Barile, M. Flossdorf, T. Holland-Letz, S.M. Schlenner, M. Reth, T. Hofer, and H.R. Rodewald. 2015. Fundamental

properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* 518:542-546.
Buza-Vidas, N., S. Duarte, S. Luc, T. Bouriez-Jones, P.S. Woll, and S.E. Jacobsen. 2011. GATA3 is redundant for maintenance and self-renewal of hematopoietic stem cells. *Blood* 118:1291-1293.
Cabezas-Wallscheid, N., D. Klimmeck, J. Hansson, D.B. Lipka, A. Reyes, Q. Wang, D. Weichenhan, A. Lier, L. von Paleske, S. Renders, P. Wunsche, P. Zeisberger, D. Brocks, L. Gu, C. Herrmann, S. Haas, M.A.G. Essers, B. Brors, R. Eils, W. Huber, M.D. Milsom, C. Plass, J. Krijgsveld, and A. Trumpp. 2014. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell* 15:507-522.
Canete, A., R. Carmona, L. Ariza, M.J. Sanchez, A. Rojas, and R. Munoz-Chapuli. 2017. A population of hematopoietic stem cells derives from GATA4-expressing progenitors located in the placenta and lateral mesoderm of mice. *Haematologica* 102:647-655.
Canu, G., and C. Ruhrberg. 2021. First blood: the endothelial origins of hematopoietic progenitors. *Angiogenesis*
Challen, G.A., N.C. Boles, S.M. Chambers, and M.A. Goodell. 2010. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* 6:265-278.
Chen, M.J., Y. Li, M.E. De Obaldia, Q. Yang, A.D. Yzaguirre, T. Yamada-Inagawa, C.S. Vink, A. Bhandoola, E. Dzierzak, and N.A. Speck. 2011. Erythroid/myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells. *Cell Stem Cell* 9:541-552.
Chen, M.J., T. Yokomizo, B.M. Zeigler, E. Dzierzak, and N.A. Speck. 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* 457:887-891.
Choi, K., M. Kennedy, A. Kazarov, J.C. Papadimitriou, and G. Keller. 1998. A common precursor for hematopoietic and endothelial cells. *Development* 125:725-732.
Ciau-Uitz, A., M. Walmsley, and R. Patient. 2000. Distinct origins of adult and embryonic blood in Xenopus. *Cell* 102:787-796.
Clarke, R.L., A.D. Yzaguirre, Y. Yashiro-Ohtani, A. Bondue, C. Blanpain, W.S. Pear, N.A. Speck, and G. Keller. 2013. The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol* 15:502-510.
Copelan, E.A. 2006. Hematopoietic stem-cell transplantation. *N Engl J Med* 354:1813-1826.
Crisan, M., and E. Dzierzak. 2016. The many faces of hematopoietic stem cell heterogeneity. *Development* 143:4571-4581.
Crisan, M., P.S. Kartalaei, C.S. Vink, T. Yamada-Inagawa, K. Bollerot, I.W. van, R. van der Linden, S.M.C. de Sousa Lopes, R. Monteiro, C. Mummery, and E. Dzierzak. 2015. Corrigendum: BMP signalling differentially regulates distinct haematopoietic stem cell types. *Nat Commun* 6:8793.

Crisan, M., P. Solaimani Kartalaei, A. Neagu, S. Karkanpouna, T. Yamada-Inagawa, C. Purini, C.S. Vink, R. van der Linden, W. van Ijcken, S.M. Chuva de Sousa Lopes, R. Monteiro, C. Mummery, and E. Dzierzak. 2016. BMP and Hedgehog Regulate Distinct AGM Hematopoietic Stem Cells Ex Vivo. *Stem Cell Reports* 6:383-395.

Crispino, J.D., and M.S. Horwitz. 2017. GATA factor mutations in hematologic disease. *Blood* 129:2103-2110.

Cumano, A., F. Dieterlen-Lievre, and I. Godin. 1996. Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura. *Cell* 86:907-916.

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

Cumano, A., and I. Godin. 2007. Ontogeny of the hematopoietic system. *Annu Rev Immunol* 25:745-785.

Dantschakoff, V. 1907. Uber das erate auftreten der blutelemente in huhnerembryo. *Folia Haematol* 4:159-166.

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

de Bruijn, M.F., M.C. Peeters, T. Luteijn, P. Visser, N.A. Speck, and E. Dzierzak. 2000a. CFU-S(11) activity does not localize solely with the aorta in the aorta-gonad-mesonephros region. *Blood* 96:2902-2904.

de Bruijn, M.F., N.A. Speck, M.C. Peeters, and E. Dzierzak. 2000b. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J* 19:2465-2474.

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

Demerdash, Y., B. Kain, M.A.G. Essers, and K.Y. King. 2021. Yin and Yang: The dual effects of interferons on hematopoiesis. *Exp Hematol* 96:1-12.

Demirci, S., J.J. Haro-Mora, A. Leonard, C. Drysdale, D. Malide, K. Keyvanfar, K. Essawi, R. Vizcardo, N. Tamaoki, N.P. Restifo, J.F. Tisdale, and N. Uchida. 2020a. Definitive hematopoietic stem/progenitor cells from human embryonic stem cells through serum/feeder-free organoid-induced differentiation. *Stem Cell Res Ther* 11:493.

Demirci, S., A. Leonard, and J.F. Tisdale. 2020b. Hematopoietic stem cells from pluripotent stem cells: Clinical potential, challenges, and future perspectives. *Stem Cells Transl Med* 9:1549-1557.

DeVilbiss, A.W., N. Tanimura, S.C. McIver, K.R. Katsumura, K.D. Johnson, and E.H. Bresnick. 2016. Navigating Transcriptional Coregulator Ensembles to Establish Genetic Networks: A GATA Factor Perspective. *Curr Top Dev Biol* 118:205-244.

Dieterlen-Lievre, F. 1975. On the origin of haemopoietic stem cells in the avian embryo: an experimental approach. *J Embryol Exp Morphol* 33:607-619.
Dieterlen-Lievre, F., and C. Martin. 1981. Diffuse intraembryonic hemopoiesis in normal and chimeric avian development. *Dev Biol* 88:180-191.
Dignum, T., B. Varnum-Finney, S.R. Srivatsan, S. Dozono, O. Waltner, A.M. Heck, T. Ishida, C. Nourigat-McKay, D.L. Jackson, S. Rafii, C. Trapnell, I.D. Bernstein, and B. Hadland. 2021. Multipotent progenitors and hematopoietic stem cells arise independently from hemogenic endothelium in the mouse embryo. *Cell Rep* 36:109675.
Dobrzycki, T., M. Lalwani, C. Telfer, R. Monteiro, and R. Patient. 2020. The roles and controls of GATA factors in blood and cardiac development. *IUBMB Life* 72:39-44.
Doulatov, S., F. Notta, K. Eppert, L.T. Nguyen, P.S. Ohashi, and J.E. Dick. 2010. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nat Immunol* 11:585-593.
Durand, C., C. Robin, K. Bollerot, M.H. Baron, K. Ottersbach, and E. Dzierzak. 2007. Embryonic stromal clones reveal developmental regulators of definitive hematopoietic stem cells. *Proc Natl Acad Sci U S A* 104:20838-20843.
Dykstra, B., D. Kent, M. Bowie, L. McCaffrey, M. Hamilton, K. Lyons, S.J. Lee, R. Brinkman, and C. Eaves. 2007. Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell* 1:218-229.
Dzierzak, E., and A. Bigas. 2018. Blood Development: Hematopoietic Stem Cell Dependence and Independence. *Cell Stem Cell* 22:639-651.
Dzierzak, E., A. Medvinsky, and M. de Bruijn. 1998. Qualitative and quantitative aspects of haematopoietic cell development in the mammalian embryo. *Immunol Today* 19:228-236.
Dzierzak, E., and N.A. Speck. 2008. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat Immunol* 9:129-136.
Eaves, C.J. 2015. Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* 125:2605-2613.
Eich, C., J. Arlt, C.S. Vink, P. Solaimani Kartalaei, P. Kaimakis, S.A. Mariani, R. van der Linden, W.A. van Cappellen, and E. Dzierzak. 2018. In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. *J Exp Med* 215:233-248.
Eilken, H.M., S. Nishikawa, and T. Schroeder. 2009. Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* 457:896-900.
Elsaid, R., S. Meunier, O. Burlen-Defranoux, F. Soares-da-Silva, T. Perchet, L. Iturri, L. Freyer, P. Vieira, P. Pereira, R. Golub, A. Bandeira, E.G. Perdiguero, and A. Cumano. 2021. A wave of bipotent T/ILC-restricted progenitors shapes the embryonic thymus microenvironment in a time-dependent manner. *Blood* 137:1024-1036.
Ema, H., Y. Morita, and T. Suda. 2014. Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol* 42:74-82 e72.

Ema, H., and H. Nakauchi. 2000. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood* 95:2284-2288.
Emmel, V.E. 1916. The cell clusters in the dorsal aorta of mammalian embryos. *American Journal of Anatomy* 19:401-421.
Fehling, H.J., G. Lacaud, A. Kubo, M. Kennedy, S. Robertson, G. Keller, and V. Kouskoff. 2003. Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* 130:4217-4227.
Ferkowicz, M.J., M. Starr, X. Xie, W. Li, S.A. Johnson, W.C. Shelley, P.R. Morrison, and M.C. Yoder. 2003. CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. *Development* 130:4393-4403.
Fitch, S.R., G.M. Kimber, N.K. Wilson, A. Parker, B. Mirshekar-Syahkal, B. Gottgens, A. Medvinsky, E. Dzierzak, and K. Ottersbach. 2012. Signaling from the sympathetic nervous system regulates hematopoietic stem cell emergence during embryogenesis. *Cell Stem Cell* 11:554-566.
Ford, C.E., J.L. Hamerton, D.W. Barnes, and J.F. Loutit. 1956. Cytological identification of radiation-chimaeras. *Nature* 177:452-454.
Fowler, J.H., A.M. Wu, J.E. Till, E.A. McCulloch, and L. Siminovitch. 1967. The cellular composition of hemopoietic spleen colonies. *Journal of Cellular Physiology* 69:65-71.
Friedle, H., and W. Knochel. 2002. Cooperative interaction of Xvent-2 and GATA-2 in the activation of the ventral homeobox gene Xvent-1B. *J Biol Chem* 277:23872-23881.
Ganuza, M., B. Hadland, A. Chabot, C. Li, G. Kang, I. Bernstein, and S. McKinney-Freeman. 2017. Murine hemogenic endothelial precursors display heterogeneous hematopoietic potential ex vivo. *Exp Hematol* 51:25-35 e26.
Gao, X., K.D. Johnson, Y.I. Chang, M.E. Boyer, C.N. Dewey, J. Zhang, and E.H. Bresnick. 2013. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med* 210:2833-2842.
Garcia-Porrero, J.A., I.E. Godin, and F. Dieterlen-Lievre. 1995. Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. *Anat Embryol (Berl)* 192:425-435.
Garcia-Porrero, J.A., A. Manaia, J. Jimeno, L.L. Lasky, F. Dieterlen-Lievre, and I.E. Godin. 1998. Antigenic profiles of endothelial and hemopoietic lineages in murine intraembryonic hemogenic sites. *Dev Comp Immunol* 22:303-319.
Gekas, C., F. Dieterlen-Lievre, S.H. Orkin, and H.K. Mikkola. 2005. The placenta is a niche for hematopoietic stem cells. *Dev Cell* 8:365-375.
Ghosn, E., M. Yoshimoto, H. Nakauchi, I.L. Weissman, and L.A. Herzenberg. 2019. Hematopoietic stem cell-independent hematopoiesis and the origins of innate-like B lymphocytes. *Development* 146:
Ghosn, E.E., J. Waters, M. Phillips, R. Yamamoto, B.R. Long, Y. Yang, R. Gerstein, C.A. Stoddart, H. Nakauchi, and L.A. Herzenberg. 2016. Fetal Hematopoietic Stem Cell Transplantation Fails to Fully

Regenerate the B-Lymphocyte Compartment. *Stem Cell Reports* 6:137-149.
Ghosn, E.E., R. Yamamoto, S. Hamanaka, Y. Yang, L.A. Herzenberg, H. Nakauchi, and L.A. Herzenberg. 2012. Distinct B-cell lineage commitment distinguishes adult bone marrow hematopoietic stem cells. *Proc Natl Acad Sci U S A* 109:5394-5398.
Gillespie, M.A., C.G. Palii, D. Sanchez-Taltavull, P. Shannon, W.J.R. Longabaugh, D.J. Downes, K. Sivaraman, H.M. Espinoza, J.R. Hughes, N.D. Price, T.J. Perkins, J.A. Ranish, and M. Brand. 2020. Absolute Quantification of Transcription Factors Reveals Principles of Gene Regulation in Erythropoiesis. *Mol Cell* 78:960-974 e911.
Ginhoux, F., M. Greter, M. Leboeuf, S. Nandi, P. See, S. Gokhan, M.F. Mehler, S.J. Conway, L.G. Ng, E.R. Stanley, I.M. Samokhvalov, and M. Merad. 2010. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* 330:841-845.
Godin, I., J.A. Garcia-Porrero, F. Dieterlen-Lievre, and A. Cumano. 1999. Stem cell emergence and hemopoietic activity are incompatible in mouse intraembryonic sites. *J Exp Med* 190:43-52.
Godin, I.E., J.A. Garcia-Porrero, A. Coutinho, F. Dieterlen-Lievre, and M.A. Marcos. 1993. Para-aortic splanchnopleura from early mouse embryos contains B1a cell progenitors. *Nature* 364:67-70.
Gottgens, B., A. Nastos, S. Kinston, S. Piltz, E.C. Delabesse, M. Stanley, M.J. Sanchez, A. Ciau-Uitz, R. Patient, and A.R. Green. 2002. Establishing the transcriptional programme for blood: the SCL stem cell enhancer is regulated by a multiprotein complex containing Ets and GATA factors. *EMBO J* 21:3039-3050.
Graf, T., and T. Enver. 2009. Forcing cells to change lineages. *Nature* 462:587-594.
Grun, D., M.J. Muraro, J.C. Boisset, K. Wiebrands, A. Lyubimova, G. Dharmadhikari, M. van den Born, J. van Es, E. Jansen, H. Clevers, E.J.P. de Koning, and A. van Oudenaarden. 2016. De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell Stem Cell* 19:266-277.
Gupte, J., G. Swaminath, J. Danao, H. Tian, Y. Li, and X. Wu. 2012. Signaling property study of adhesion G-protein-coupled receptors. *FEBS Lett* 586:1214-1219.
Haas, S., A. Trumpp, and M.D. Milsom. 2018. Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. *Cell Stem Cell* 22:627-638.
Hadland, B.K., B. Varnum-Finney, P.K. Mandal, D.J. Rossi, M.G. Poulos, J.M. Butler, S. Rafii, M.C. Yoder, M. Yoshimoto, and I.D. Bernstein. 2017. A Common Origin for B-1a and B-2 Lymphocytes in Clonal Pre-Hematopoietic Stem Cells. *Stem Cell Reports* 8:1563-1572.
Hamey, F.K., and B. Gottgens. 2019. Machine learning predicts putative hematopoietic stem cells within large single-cell transcriptomics data sets. *Exp Hematol* 78:11-20.
He, S., I. Kim, M.S. Lim, and S.J. Morrison. 2011. Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. *Genes Dev* 25:1613-1627.

Heck, A.M., T. Ishida, and B. Hadland. 2020. Location, Location, Location: How Vascular Specialization Influences Hematopoietic Fates During Development. *Front Cell Dev Biol* 8:602617.
Hewitt, K.J., K.D. Johnson, X. Gao, S. Keles, and E.H. Bresnick. 2016. The Hematopoietic Stem and Progenitor Cell Cistrome: GATA Factor-Dependent cis-Regulatory Mechanisms. *Curr Top Dev Biol* 118:45-76.
Horton, P.D., S.P. Dumbali, K.R. Bhanu, M.F. Diaz, and P.L. Wenzel. 2021. Biomechanical Regulation of Hematopoietic Stem Cells in the Developing Embryo. *Curr Tissue Microenviron Rep* 2:1-15.
Hou, S., Z. Li, X. Zheng, Y. Gao, J. Dong, Y. Ni, X. Wang, Y. Li, X. Ding, Z. Chang, S. Li, Y. Hu, X. Fan, Y. Hou, L. Wen, B. Liu, F. Tang, and Y. Lan. 2020. Embryonic endothelial evolution towards first hematopoietic stem cells revealed by single-cell transcriptomic and functional analyses. *Cell Res* 30:376-392.
Houssaint, E. 1981. Differentiation of the mouse hepatic primordium. II. Extrinsic origin of the haemopoietic cell line. *Cell Differ* 10:243-252.
Huber, T.L., V. Kouskoff, H.J. Fehling, J. Palis, and G. Keller. 2004. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* 432:625-630.
Ikuta, K., and I.L. Weissman. 1992. Evidence that hematopoietic stem cells express mouse c-kit but do not depend on steel factor for their generation. *Proc Natl Acad Sci U S A* 89:1502-1506.
Imanirad, P., P. Solaimani Kartalaei, M. Crisan, C. Vink, T. Yamada-Inagawa, E. de Pater, D. Kurek, P. Kaimakis, R. van der Linden, N. Speck, and E. Dzierzak. 2014. HIF1alpha is a regulator of hematopoietic progenitor and stem cell development in hypoxic sites of the mouse embryo. *Stem Cell Res* 12:24-35.
Inlay, M.A., T. Serwold, A. Mosley, J.W. Fathman, I.K. Dimov, J. Seita, and I.L. Weissman. 2014. Identification of multipotent progenitors that emerge prior to hematopoietic stem cells in embryonic development. *Stem Cell Reports* 2:457-472.
Jackson, M., A. Fidanza, A.H. Taylor, S. Rybtsov, R. Axton, M. Kydonaki, S. Meek, T. Burdon, A. Medvinsky, and L.M. Forrester. 2021. Modulation of APLNR Signaling Is Required during the Development and Maintenance of the Hematopoietic System. *Stem Cell Reports* 16:727-740.
Jacobsen, S.E.W., and C. Nerlov. 2019. Haematopoiesis in the era of advanced single-cell technologies. *Nat Cell Biol* 21:2-8.
Jacobson, L.O., E.L. Simmons, E.K. Marks, and J.H. Eldredge. 1951. Recovery from radiation injury. *Science* 113:510-511.
Jaffredo, T., A. Balduini, A. Bigas, R. Bernardi, D. Bonnet, B. Canque, P. Charbord, A. Cumano, R. Delwel, C. Durand, W. Fibbe, L. Forrester, L. de Franceschi, C. Ghevaert, B. Gjertsen, B. Gottgens, T. Graf, O. Heidenreich, O. Hermine, D. Higgs, M. Kleanthous, H. Klump, V. Kouskoff, D. Krause, G. Lacaud, C.L. Celso, J.H.A. Martens, S. Mendez-Ferrer, P. Menendez, R. Oostendorp, S. Philipsen, B. Porse, M. Raaijmakers, C. Robin, H. Stunnenberg, K. Theilgaard-Monch, I. Touw, W. Vainchenker, J.V. Corrons, L. Yvernogeau, and J.J.

Schuringa. 2021. The EHA Research Roadmap: Normal Hematopoiesis. *Hemasphere* 5:e669.
Jaffredo, T., R. Gautier, A. Eichmann, and F. Dieterlen-Lievre. 1998. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* 125:4575-4583.
Johnson, G.R., and M.A. Moore. 1975. Role of stem cell migration in initiation of mouse foetal liver haemopoiesis. *Nature* 258:726-728.
Johnson, K.D., G. Kong, X. Gao, Y.I. Chang, K.J. Hewitt, R. Sanalkumar, R. Prathibha, E.A. Ranheim, C.N. Dewey, J. Zhang, and E.H. Bresnick. 2015. Cis-regulatory mechanisms governing stem and progenitor cell transitions. *Sci Adv* 1:e1500503.
Jordan, H.E. 1916. Evidence of hemogenic capacity of endothelium. *The Anatomical Record* 10:417-420.
Jordan, H.E. 1917. Aortic Cell Clusters in Vertebrate Embryos. *Proc Natl Acad Sci U S A* 3:149-156.
Jurecic, R. 2019. Hematopoietic Stem Cell Heterogeneity. *Adv Exp Med Biol* 1169:195-211.
Kaimakis, P., E. de Pater, C. Eich, P. Solaimani Kartalaei, M.L. Kauts, C.S. Vink, R. van der Linden, M. Jaegle, T. Yokomizo, D. Meijer, and E. Dzierzak. 2016. Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. *Blood* 127:1426-1437.
Katsumura, K.R., E.H. Bresnick, and G.F.M. Group. 2017. The GATA factor revolution in hematology. *Blood* 129:2092-2102.
Kaufman, M.H. 1992. The atlas of mouse development. Academic Press,
Kauts, M.L., C.S. Vink, and E. Dzierzak. 2016. Hematopoietic (stem) cell development - how divergent are the roads taken? *FEBS Lett* 590:3975-3986.
Keller, G., C. Paige, E. Gilboa, and E.F. Wagner. 1985. Expression of a foreign gene in myeloid and lymphoid cells derived from multipotent haematopoietic precursors. *Nature* 318:149-154.
Keller, G., and R. Snodgrass. 1990. Life span of multipotential hematopoietic stem cells in vivo. *J Exp Med* 171:1407-1418.
Kiel, M.J., O.H. Yilmaz, T. Iwashita, O.H. Yilmaz, C. Terhorst, and S.J. Morrison. 2005. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121:1109-1121.
Kierdorf, K., D. Erny, T. Goldmann, V. Sander, C. Schulz, E.G. Perdiguero, P. Wieghofer, A. Heinrich, P. Riemke, C. Holscher, D.N. Muller, B. Luckow, T. Brocker, K. Debowski, G. Fritz, G. Opdenakker, A. Diefenbach, K. Biber, M. Heikenwalder, F. Geissmann, F. Rosenbauer, and M. Prinz. 2013. Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat Neurosci* 16:273-280.
Kissa, K., and P. Herbomel. 2010. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* 464:112-115.
Kobayashi, M., W.C. Shelley, W. Seo, S. Vemula, Y. Lin, Y. Liu, R. Kapur, I. Taniuchi, and M. Yoshimoto. 2014. Functional B-1 progenitor cells are present in the hematopoietic stem cell-deficient embryo and depend

on Cbfbeta for their development. *Proc Natl Acad Sci U S A* 111:12151-12156.
Koyunlar, C., and E. de Pater. 2020. From Basic Biology to Patient Mutational Spectra of GATA2 Haploinsufficiencies: What Are the Mechanisms, Hurdles, and Prospects of Genome Editing for Treatment. *Front Genome Ed* 2:602182.
Kristiansen, T.A., E. Jaensson Gyllenback, A. Zriwil, T. Bjorklund, J.A. Daniel, E. Sitnicka, S. Soneji, D. Bryder, and J. Yuan. 2016. Cellular Barcoding Links B-1a B Cell Potential to a Fetal Hematopoietic Stem Cell State at the Single-Cell Level. *Immunity* 45:346-357.
Kueh, H.Y., M.A. Yui, K.K. Ng, S.S. Pease, J.A. Zhang, S.S. Damle, G. Freedman, S. Siu, I.D. Bernstein, M.B. Elowitz, and E.V. Rothenberg. 2016. Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. *Nat Immunol* 17:956-965.
Kumaravelu, P., L. Hook, A.M. Morrison, J. Ure, S. Zhao, S. Zuyev, J. Ansell, and A. Medvinsky. 2002. Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development* 129:4891-4899.
La Manno, G., R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber, M.E. Kastriti, P. Lonnerberg, A. Furlan, J. Fan, L.E. Borm, Z. Liu, D. van Bruggen, J. Guo, X. He, R. Barker, E. Sundstrom, G. Castelo-Branco, P. Cramer, I. Adameyko, S. Linnarsson, and P.V. Kharchenko. 2018. RNA velocity of single cells. *Nature* 560:494-498.
Lalor, P.A., L.A. Herzenberg, S. Adams, and A.M. Stall. 1989. Feedback regulation of murine Ly-1 B cell development. *Eur J Immunol* 19:507-513.
Lancrin, C., P. Sroczynska, C. Stephenson, T. Allen, V. Kouskoff, and G. Lacaud. 2009. The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* 457:892-895.
Lange, L., M. Morgan, and A. Schambach. 2021. The hemogenic endothelium: a critical source for the generation of PSC-derived hematopoietic stem and progenitor cells. *Cell Mol Life Sci*
Lassila, O., J. Eskola, P. Toivanen, C. Martin, and F. Dieterlen-Lievre. 1978. The origin of lymphoid stem cells studied in chick yold sac-embryo chimaeras. *Nature* 272:353-354.
Laurenti, E., and B. Gottgens. 2018. From haematopoietic stem cells to complex differentiation landscapes. *Nature* 553:418-426.
Lemischka, I.R., D.H. Raulet, and R.C. Mulligan. 1986. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45:917-927.
Levine, J.H., Y. Lin, and M.B. Elowitz. 2013. Functional roles of pulsing in genetic circuits. *Science* 342:1193-1200.
Li, X., X. Zhao, Y. Fang, X. Jiang, T. Duong, C. Fan, C.C. Huang, and S.R. Kain. 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 273:34970-34975.

Li, Y., V. Esain, L. Teng, J. Xu, W. Kwan, I.M. Frost, A.D. Yzaguirre, X. Cai, M. Cortes, M.W. Maijenburg, J. Tober, E. Dzierzak, S.H. Orkin, K. Tan, T.E. North, and N.A. Speck. 2014. Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev* 28:2597-2612.

Li, Y., L. Gao, B. Hadland, K. Tan, and N.A. Speck. 2017. CD27 marks murine embryonic hematopoietic stem cells and type II prehematopoietic stem cells. *Blood* 130:372-376.

Li, Z., Y. Lan, W. He, D. Chen, J. Wang, F. Zhou, Y. Wang, H. Sun, X. Chen, C. Xu, S. Li, Y. Pang, G. Zhang, L. Yang, L. Zhu, M. Fan, A. Shang, Z. Ju, L. Luo, Y. Ding, W. Guo, W. Yuan, X. Yang, and B. Liu. 2012. Mouse embryonic head as a site for hematopoietic stem cell development. *Cell Stem Cell* 11:663-675.

Liakhovitskaia, A., R. Gribi, E. Stamateris, G. Villain, T. Jaffredo, R. Wilkie, D. Gilchrist, J. Yang, J. Ure, and A. Medvinsky. 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.

Liebscher, I., O. Cevheroglu, C.C. Hsiao, A.F. Maia, H. Schihada, N. Scholz, M. Soave, K. Spiess, K. Trajkovic, M. Kosloff, and S. Promel. 2021. A guide to adhesion GPCR research. *FEBS J*

Liggett, L.A., and V.G. Sankaran. 2020. Unraveling Hematopoiesis through the Lens of Genomics. *Cell* 182:1384-1400.

Ling, K.W., K. Ottersbach, J.P. van Hamburg, A. Oziemlak, F.Y. Tsai, S.H. Orkin, R. Ploemacher, R.W. Hendriks, and E. Dzierzak. 2004. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med* 200:871-882.

Lorenz, E., D. Uphoff, T.R. Reid, and E. Shelton. 1951. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *J Natl Cancer Inst* 12:197-201.

Lorsbach, R.B., J. Moore, S.O. Ang, W. Sun, N. Lenny, and J.R. Downing. 2004. Role of RUNX1 in adult hematopoiesis: analysis of RUNX1-IRES-GFP knock-in mice reveals differential lineage expression. *Blood* 103:2522-2529.

Lurie, L.J., M.E. Boyer, J.A. Grass, and E.H. Bresnick. 2008. Differential GATA factor stabilities: implications for chromatin occupancy by structurally similar transcription factors. *Biochemistry* 47:859-869.

Ma, X., K.W. Ling, and E. Dzierzak. 2001. Cloning of the Ly-6A (Sca-1) gene locus and identification of a 3' distal fragment responsible for high-level gamma-interferon-induced expression in vitro. *Br J Haematol* 114:724-730.

Ma, X., C. Robin, K. Ottersbach, and E. Dzierzak. 2002. The Ly-6A (Sca-1) GFP transgene is expressed in all adult mouse hematopoietic stem cells. *Stem Cells* 20:514-521.

Maeno, M., P.E. Mead, C. Kelley, R.H. Xu, H.F. Kung, A. Suzuki, N. Ueno, and L.I. Zon. 1996. The role of BMP-4 and GATA-2 in the induction and differentiation of hematopoietic mesoderm in Xenopus laevis. *Blood* 88:1965-1972.

Maglitto, A., S.A. Mariani, E. de Pater, C. Rodriguez-Seoane, C.S. Vink, X. Piao, M.L. Lukke, and E. Dzierzak. 2021. Unexpected redundancy of Gpr56 and Gpr97 during hematopoietic cell development and differentiation. *Blood Adv* 5:829-842.

Mariani, S.A., Z. Li, S. Rice, C. Krieg, S. Fragkogianni, M. Robinson, C.S. Vink, J.W. Pollard, and E. Dzierzak. 2019. Pro-inflammatory Aorta-Associated Macrophages Are Involved in Embryonic Development of Hematopoietic Stem Cells. *Immunity* 50:1439-1452 e1435.

Marshall, C.J., C. Kinnon, and A.J. Thrasher. 2000. Polarized expression of bone morphogenetic protein-4 in the human aorta-gonad-mesonephros region. *Blood* 96:1591-1593.

Marshall, C.J., and A.J. Thrasher. 2001. The embryonic origins of human haematopoiesis. *Br J Haematol* 112:838-850.

Mascarenhas, M.I., A. Parker, E. Dzierzak, and K. Ottersbach. 2009. Identification of novel regulators of hematopoietic stem cell development through refinement of stem cell localization and expression profiling. *Blood* 114:4645-4653.

Maximow, A.A. 1909. Untersuchungen uber blut und bindegewebe 1. Die fruhesten entwicklungsstadien der blut- und binde-gewebszellan bein saugetierembryo, bis zum anfang der blutbilding unden leber. *Arch Mikroskop Anat* 73:444-561.

McCulloch, E.A., and J.E. Till. 1960. The radiation sensitivity of normal mouse bone marrow cells, determined by quantitative marrow transplantation into irradiated mice. *Radiat Res* 13:115-125.

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

McGrath, K.E., and J. Palis. 2005. Hematopoiesis in the yolk sac: more than meets the eye. *Exp Hematol* 33:1021-1028.

McKinney-Freeman, S., P. Cahan, H. Li, S.A. Lacadie, H.T. Huang, M. Curran, S. Loewer, O. Naveiras, K.L. Kathrein, M. Konantz, E.M. Langdon, C. Lengerke, L.I. Zon, J.J. Collins, and G.Q. Daley. 2012. The transcriptional landscape of hematopoietic stem cell ontogeny. *Cell Stem Cell* 11:701-714.

McKinney-Freeman, S.L., O. Naveiras, F. Yates, S. Loewer, M. Philitas, M. Curran, P.J. Park, and G.Q. Daley. 2009. Surface antigen phenotypes of hematopoietic stem cells from embryos and murine embryonic stem cells. *Blood* 114:268-278.

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

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

Metcalf, D. 1970. Studies on colony formation in vitro by mouse bone marrow cells. II. Action of colony stimulating factor. *J Cell Physiol* 76:89-99.

Mikkola, H.K., Y. Fujiwara, T.M. Schlaeger, D. Traver, and S.H. Orkin. 2003. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* 101:508-516.

Minegishi, N., J. Ohta, N. Suwabe, H. Nakauchi, H. Ishihara, N. Hayashi, and M. Yamamoto. 1998. Alternative promoters regulate transcription of the mouse GATA-2 gene. *J Biol Chem* 273:3625-3634.

Minegishi, N., J. Ohta, H. Yamagiwa, N. Suzuki, S. Kawauchi, Y. Zhou, S. Takahashi, N. Hayashi, J.D. Engel, and M. Yamamoto. 1999. The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood* 93:4196-4207.

Minegishi, N., N. Suzuki, Y. Kawatani, R. Shimizu, and M. Yamamoto. 2005. Rapid turnover of GATA-2 via ubiquitin-proteasome protein degradation pathway. *Genes Cells* 10:693-704.

Minegishi, N., N. Suzuki, T. Yokomizo, X. Pan, T. Fujimoto, S. Takahashi, T. Hara, A. Miyajima, S. Nishikawa, and M. Yamamoto. 2003. Expression and domain-specific function of GATA-2 during differentiation of the hematopoietic precursor cells in midgestation mouse embryos. *Blood* 102:896-905.

Monteiro, R.M., S.M. de Sousa Lopes, M. Bialecka, S. de Boer, A. Zwijsen, and C.L. Mummery. 2008. Real time monitoring of BMP Smads transcriptional activity during mouse development. *Genesis* 46:335-346.

Moore, M.A., and D. Metcalf. 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.

Moore, M.A., and J.J. Owen. 1965. Chromosome marker studies on the development of the haemopoietic system in the chick embryo. *Nature* 208:956 passim.

Moore, M.A., and J.J. Owen. 1967. Chromosome marker studies in the irradiated chick embryo. *Nature* 215:1081-1082.

Morita, Y., H. Ema, and H. Nakauchi. 2010. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med* 207:1173-1182.

Morrison, S.J., and I.L. Weissman. 1994. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity* 1:661-673.

Muller, A.M., A. Medvinsky, J. Strouboulis, F. Grosveld, and E. Dzierzak. 1994. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* 1:291-301.

Muller-Sieburg, C.E., R.H. Cho, L. Karlsson, J.F. Huang, and H.B. Sieburg. 2004. Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood* 103:4111-4118.

Muller-Sieburg, C.E., H.B. Sieburg, J.M. Bernitz, and G. Cattarossi. 2012. Stem cell heterogeneity: implications for aging and regenerative medicine. *Blood* 119:3900-3907.

Muller-Sieburg, C.E., C.A. Whitlock, and I.L. Weissman. 1986. Isolation of two early B lymphocyte progenitors from mouse marrow: a committed

pre-pre-B cell and a clonogenic Thy-1-lo hematopoietic stem cell. *Cell* 44:653-662.
Murray, P.D.F., and J.T. Wilson. 1932. The development in vitro of the blood of the early chick embryo. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character* 111:497-521.
Na Nakorn, T., D. Traver, I.L. Weissman, and K. Akashi. 2002. Myeloerythroid-restricted progenitors are sufficient to confer radioprotection and provide the majority of day 8 CFU-S. *J Clin Invest* 109:1579-1585.
Nagai, T., K. Ibata, E.S. Park, M. Kubota, K. Mikoshiba, and A. Miyawaki. 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol* 20:87-90.
Nagaya, N., H. Mori, S. Murakami, K. Kangawa, and S. Kitamura. 2005. Adrenomedullin: angiogenesis and gene therapy. *Am J Physiol Regul Integr Comp Physiol* 288:R1432-1437.
Nakano, H., X. Liu, A. Arshi, Y. Nakashima, B. van Handel, R. Sasidharan, A.W. Harmon, J.H. Shin, R.J. Schwartz, S.J. Conway, R.P. Harvey, M. Pashmforoush, H.K. Mikkola, and A. Nakano. 2013. Haemogenic endocardium contributes to transient definitive haematopoiesis. *Nat Commun* 4:1564.
Nandakumar, S.K., K. Johnson, S.L. Throm, T.I. Pestina, G. Neale, and D.A. Persons. 2015. Low-level GATA2 overexpression promotes myeloid progenitor self-renewal and blocks lymphoid differentiation in mice. *Exp Hematol* 43:565-577 e561-510.
Nobuhisa, I., M. Osawa, M. Uemura, Y. Kishikawa, M. Anani, K. Harada, H. Takagi, K. Saito, M. Kanai-Azuma, Y. Kanai, A. Iwama, and T. Taga. 2014. Sox17-mediated maintenance of fetal intra-aortic hematopoietic cell clusters. *Mol Cell Biol* 34:1976-1990.
Nombela-Arrieta, C., and M.G. Manz. 2017. Quantification and three-dimensional microanatomical organization of the bone marrow. *Blood Adv* 1:407-416.
North, T.E., M.F. de Bruijn, T. Stacy, L. Talebian, E. Lind, C. Robin, M. Binder, E. Dzierzak, and N.A. Speck. 2002. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* 16:661-672.
North, T.E., W. Goessling, M. Peeters, P. Li, C. Ceol, A.M. Lord, G.J. Weber, J. Harris, C.C. Cutting, P. Huang, E. Dzierzak, and L.I. Zon. 2009. Hematopoietic stem cell development is dependent on blood flow. *Cell* 137:736-748.
Notta, F., S. Zandi, N. Takayama, S. Dobson, O.I. Gan, G. Wilson, K.B. Kaufmann, J. McLeod, E. Laurenti, C.F. Dunant, J.D. McPherson, L.D. Stein, Y. Dror, and J.E. Dick. 2016. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* 351:aab2116.
Nottingham, W.T., A. Jarratt, M. Burgess, C.L. Speck, J.F. Cheng, S. Prabhakar, E.M. Rubin, P.S. Li, J. Sloane-Stanley, A.S.J. Kong, and M.F. de Bruijn. 2007. Runx1-mediated hematopoietic stem-cell

emergence is controlled by a Gata/Ets/SCL-regulated enhancer. *Blood* 110:4188-4197.
Nowell, P.C., L.J. Cole, J.G. Habermeyer, and P.L. Roan. 1956. Growth and continued function of rat marrow cells in x-radiated mice. *Cancer Res* 16:258-261.
Oatley, M., O.V. Bolukbasi, V. Svensson, M. Shvartsman, K. Ganter, K. Zirngibl, P.V. Pavlovich, V. Milchevskaya, V. Foteva, K.N. Natarajan, B. Baying, V. Benes, K.R. Patil, S.A. Teichmann, and C. Lancrin. 2020. Single-cell transcriptomics identifies CD44 as a marker and regulator of endothelial to haematopoietic transition. *Nat Commun* 11:586.
Okada, S., H. Nakauchi, K. Nagayoshi, S. Nishikawa, Y. Miura, and T. Suda. 1992. In vivo and in vitro stem cell function of c-kit- and Sca-1-positive murine hematopoietic cells. *Blood* 80:3044-3050.
Okada, S., H. Nakauchi, K. Nagayoshi, S. Nishikawa, S. Nishikawa, Y. Miura, and T. Suda. 1991. Enrichment and characterization of murine hematopoietic stem cells that express c-kit molecule. *Blood* 78:1706-1712.
Olsson, A., M. Venkatasubramanian, V.K. Chaudhri, B.J. Aronow, N. Salomonis, H. Singh, and H.L. Grimes. 2016. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* 537:698-702.
Orkin, S.H. 1990. Globin gene regulation and switching: circa 1990. *Cell* 63:665-672.
Osawa, M., K. Hanada, H. Hamada, and H. Nakauchi. 1996. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* 273:242-245.
Ottersbach, K. 2019. Endothelial-to-haematopoietic transition: an update on the process of making blood. *Biochem Soc Trans* 47:591-601.
Ottersbach, K., and E. Dzierzak. 2005. The murine placenta contains hematopoietic stem cells within the vascular labyrinth region. *Dev Cell* 8:377-387.
Padron-Barthe, L., S. Temino, C. Villa del Campo, L. Carramolino, J. Isern, and M. Torres. 2014. Clonal analysis identifies hemogenic endothelium as the source of the blood-endothelial common lineage in the mouse embryo. *Blood* 124:2523-2532.
Palis, J., S. Robertson, M. Kennedy, C. Wall, and G. Keller. 1999. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* 126:5073-5084.
Paul, F., Y. Arkin, A. Giladi, D.A. Jaitin, E. Kenigsberg, H. Keren-Shaul, D. Winter, D. Lara-Astiaso, M. Gury, A. Weiner, E. David, N. Cohen, F.K. Lauridsen, S. Haas, A. Schlitzer, A. Mildner, F. Ginhoux, S. Jung, A. Trumpp, B.T. Porse, A. Tanay, and I. Amit. 2015. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* 163:1663-1677.
Pei, W., F. Shang, X. Wang, A.K. Fanti, A. Greco, K. Busch, K. Klapproth, Q. Zhang, C. Quedenau, S. Sauer, T.B. Feyerabend, T. Hofer, and H.R. Rodewald. 2020. Resolving Fates and Single-Cell Transcriptomes of

Hematopoietic Stem Cell Clones by PolyloxExpress Barcoding. *Cell Stem Cell* 27:383-395 e388.
Persons, D.A., J.A. Allay, E.R. Allay, R.A. Ashmun, D. Orlic, S.M. Jane, J.M. Cunningham, and A.W. Nienhuis. 1999. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood* 93:488-499.
Petit-Cocault, L., C. Volle-Challier, M. Fleury, B. Peault, and M. Souyri. 2007. Dual role of Mpl receptor during the establishment of definitive hematopoiesis. *Development* 134:3031-3040.
Pietras, E.M., D. Reynaud, Y.A. Kang, D. Carlin, F.J. Calero-Nieto, A.D. Leavitt, J.M. Stuart, B. Gottgens, and E. Passegue. 2015. Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell* 17:35-46.
Pluznik, D.H., and L. Sachs. 1966. The induction of clones of normal mast cells by a substance from conditioned medium. *Exp Cell Res* 43:553-563.
Porcheri, C., O. Golan, F.J. Calero-Nieto, R. Thambyrajah, C. Ruiz-Herguido, X. Wang, F. Catto, Y. Guillen, R. Sinha, J. Gonzalez, S.J. Kinston, S.A. Mariani, A. Maglitto, C.S. Vink, E. Dzierzak, P. Charbord, B. Gottgens, L. Espinosa, D. Sprinzak, and A. Bigas. 2020. Notch ligand Dll4 impairs cell recruitment to aortic clusters and limits blood stem cell generation. *EMBO J* 39:e104270.
Ranvier, L. 1875. Traité technique d'histologie. F. Savy, Paris.
Rao, T.N., J. Marks-Bluth, J. Sullivan, M.K. Gupta, V. Chandrakanthan, S.R. Fitch, K. Ottersbach, Y.C. Jang, X. Piao, R.N. Kulkarni, T. Serwold, J.E. Pimanda, and A.J. Wagers. 2015. High-level Gpr56 expression is dispensable for the maintenance and function of hematopoietic stem and progenitor cells in mice. *Stem Cell Res* 14:307-322.
Robert-Moreno, A., L. Espinosa, J.L. de la Pompa, and A. Bigas. 2005. RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development* 132:1117-1126.
Robert-Moreno, A., J. Guiu, C. Ruiz-Herguido, M.E. Lopez, J. Ingles-Esteve, L. Riera, A. Tipping, T. Enver, E. Dzierzak, T. Gridley, L. Espinosa, and A. Bigas. 2008. Impaired embryonic haematopoiesis yet normal arterial development in the absence of the Notch ligand Jagged1. *EMBO J* 27:1886-1895.
Robin, C., K. Ottersbach, J.C. Boisset, A. Oziemlak, and E. Dzierzak. 2011. CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells. *Blood* 117:5088-5091.
Rodrigues, N.P., A.S. Boyd, C. Fugazza, G.E. May, Y. Guo, A.J. Tipping, D.T. Scadden, P. Vyas, and T. Enver. 2008. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood* 112:4862-4873.
Rodrigues, N.P., V. Janzen, R. Forkert, D.M. Dombkowski, A.S. Boyd, S.H. Orkin, T. Enver, P. Vyas, and D.T. Scadden. 2005. Haploinsufficiency

of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood* 106:477-484.
Rodriguez-Fraticelli, A.E., C. Weinreb, S.W. Wang, R.P. Migueles, M. Jankovic, M. Usart, A.M. Klein, S. Lowell, and F.D. Camargo. 2020. Single-cell lineage tracing unveils a role for TCF15 in haematopoiesis. *Nature* 583:585-589.
Rodriguez-Fraticelli, A.E., S.L. Wolock, C.S. Weinreb, R. Panero, S.H. Patel, M. Jankovic, J. Sun, R.A. Calogero, A.M. Klein, and F.D. Camargo. 2018. Clonal analysis of lineage fate in native haematopoiesis. *Nature* 553:212-216.
Roques, M., C. Durand, R. Gautier, P.Y. Canto, L. Petit-Cocault, L. Yvernogeau, D. Dunon, M. Souyri, and T. Jaffredo. 2012. Endoglin expression level discriminates long-term hematopoietic from short-term clonogenic progenitor cells in the aorta. *Haematologica* 97:975-979.
Rothenberg, E.V. 2021. Single-cell insights into the hematopoietic generation of T-lymphocyte precursors in mouse and human. *Exp Hematol*
Rowe, R.G., J. Mandelbaum, L.I. Zon, and G.Q. Daley. 2016. Engineering Hematopoietic Stem Cells: Lessons from Development. *Cell Stem Cell* 18:707-720.
Rybtsov, S., A. Ivanovs, S. Zhao, and A. Medvinsky. 2016. Concealed expansion of immature precursors underpins acute burst of adult HSC activity in foetal liver. *Development* 143:1284-1289.
Rybtsov, S., M. Sobiesiak, S. Taoudi, C. Souilhol, J. Senserrich, A. Liakhovitskaia, A. Ivanovs, J. Frampton, S. Zhao, and A. Medvinsky. 2011. Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. *J Exp Med* 208:1305-1315.
Sabin, F.R. 1920. Studies on the origin of blood-vessels and of red blood-corpuscles as seen in the living blastoderm of chicks during the second day of incubation. *Contributions to embryology.* IX:213-262.
Sadlon, T.J., I.D. Lewis, and R.J. D'Andrea. 2004. BMP4: its role in development of the hematopoietic system and potential as a hematopoietic growth factor. *Stem Cells* 22:457-474.
Saito, Y., K. Kaneda, A. Suekane, E. Ichihara, S. Nakahata, N. Yamakawa, K. Nagai, N. Mizuno, K. Kogawa, I. Miura, H. Itoh, and K. Morishita. 2013. Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. *Leukemia* 27:1637-1649.
Saitou, M., and M. Yamaji. 2012. Primordial germ cells in mice. *Cold Spring Harb Perspect Biol* 4:
Sanchez, M.J., A. Holmes, C. Miles, and E. Dzierzak. 1996. Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. *Immunity* 5:513-525.
Sanjuan-Pla, A., I.C. Macaulay, C.T. Jensen, P.S. Woll, T.C. Luis, A. Mead, S. Moore, C. Carella, S. Matsuoka, T. Bouriez Jones, O. Chowdhury, L. Stenson, M. Lutteropp, J.C. Green, R. Facchini, H. Boukarabila, A. Grover, A. Gambardella, S. Thongjuea, J. Carrelha, P. Tarrant, D. Atkinson, S.A. Clark, C. Nerlov, and S.E. Jacobsen. 2013. Platelet-

biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* 502:232-236.
Sasaki, H., D. Kurotaki, and T. Tamura. 2016. Regulation of basophil and mast cell development by transcription factors. *Allergol Int* 65:127-134.
Schroeder, T. 2008. Imaging stem-cell-driven regeneration in mammals. *Nature* 453:345-351.
Seita, J., and I.L. Weissman. 2010. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med* 2:640-653.
Shin, J.Y., W. Hu, M. Naramura, and C.Y. Park. 2014. High c-Kit expression identifies hematopoietic stem cells with impaired self-renewal and megakaryocytic bias. *J Exp Med* 211:217-231.
Siminovitch, L., E.A. McCulloch, and J.E. Till. 1963. The Distribution of Colony-Forming Cells among Spleen Colonies. *J Cell Comp Physiol* 62:327-336.
Sinclair, A., B. Daly, and E. Dzierzak. 1996. The Ly-6E.1 (Sca-1) gene requires a 3' chromatin-dependent region for high-level gamma-interferon-induced hematopoietic cell expression. *Blood* 87:2750-2761.
Singbrant, S., G. Karlsson, M. Ehinger, K. Olsson, P. Jaako, K. Miharada, M. Stadtfeld, T. Graf, and S. Karlsson. 2010. Canonical BMP signaling is dispensable for hematopoietic stem cell function in both adult and fetal liver hematopoiesis, but essential to preserve colon architecture. *Blood* 115:4689-4698.
Smith, R.A., and C.A. Glomski. 1982. "Hemogenic endothelium" of the embryonic aorta: Does it exist? *Dev Comp Immunol* 6:359-368.
Soares-da-Silva, F., L. Freyer, R. Elsaid, O. Burlen-Defranoux, L. Iturri, O. Sismeiro, O.P. Pinto-do, E. Gomez-Perdiguero, and A. Cumano. 2021. Yolk sac, but not hematopoietic stem cell-derived progenitors, sustain erythropoiesis throughout murine embryonic life. *J Exp Med* 218:
Solaimani Kartalaei, P., T. Yamada-Inagawa, C.S. Vink, E. de Pater, R. van der Linden, J. Marks-Bluth, A. van der Sloot, M. van den Hout, T. Yokomizo, M.L. van Schaick-Solerno, R. Delwel, J.E. Pimanda, I.W.F. van, and E. Dzierzak. 2015. Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *J Exp Med* 212:93-106.
Spangrude, G.J., S. Heimfeld, and I.L. Weissman. 1988. Purification and characterization of mouse hematopoietic stem cells. *Science* 241:58-62.
Stadtfeld, M., and T. Graf. 2005. Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development* 132:203-213.
Stefanska, M., K. Batta, R. Patel, M. Florkowska, V. Kouskoff, and G. Lacaud. 2017. Primitive erythrocytes are generated from hemogenic endothelial cells. *Sci Rep* 7:6401.
Stoveken, H.M., L.L. Bahr, M.W. Anders, A.P. Wojtovich, A.V. Smrcka, and G.G. Tall. 2016. Dihydromunduletone Is a Small-Molecule Selective Adhesion G Protein-Coupled Receptor Antagonist. *Mol Pharmacol* 90:214-224.

Stoveken, H.M., S.D. Larsen, A.V. Smrcka, and G.G. Tall. 2018. Gedunin- and Khivorin-Derivatives Are Small-Molecule Partial Agonists for Adhesion G Protein-Coupled Receptors GPR56/ADGRG1 and GPR114/ADGRG5. *Mol Pharmacol* 93:477-488.

Sullivan, C., Y. Chen, Y. Shan, Y. Hu, C. Peng, H. Zhang, L. Kong, and S. Li. 2011. Functional ramifications for the loss of P-selectin expression on hematopoietic and leukemic stem cells. *PLoS One* 6:e26246.

Sun, J., A. Ramos, B. Chapman, J.B. Johnnidis, L. Le, Y.J. Ho, A. Klein, O. Hofmann, and F.D. Camargo. 2014. Clonal dynamics of native haematopoiesis. *Nature* 514:322-327.

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

Takahashi, S., R. Shimizu, N. Suwabe, T. Kuroha, K. Yoh, J. Ohta, S. Nishimura, K.C. Lim, J.D. Engel, and M. Yamamoto. 2000. GATA factor transgenes under GATA-1 locus control rescue germline GATA-1 mutant deficiencies. *Blood* 96:910-916.

Takakura, N., X.L. Huang, T. Naruse, I. Hamaguchi, D.J. Dumont, G.D. Yancopoulos, and T. Suda. 1998. Critical role of the TIE2 endothelial cell receptor in the development of definitive hematopoiesis. *Immunity* 9:677-686.

Tang, W., J. He, T. Huang, Z. Bai, C. Wang, H. Wang, R. Yang, Y. Ni, J. Hou, J. Wang, J. Zhou, Y. Yao, Y. Gong, S. Hou, B. Liu, and Y. Lan. 2021. Hlf Expression Marks Early Emergence of Hematopoietic Stem Cell Precursors With Adult Repopulating Potential and Fate. *Front Cell Dev Biol* 9:728057.

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

Taoudi, S., and A. Medvinsky. 2007. Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. *Proc Natl Acad Sci U S A* 104:9399-9403.

Thomas, E.D., H.L. Lochte, Jr., W.C. Lu, and J.W. Ferrebee. 1957. Intravenous infusion of bone marrow in patients receiving radiation and chemotherapy. *N Engl J Med* 257:491-496.

Till, J.E., and E.A. McCulloch. 1961. A Direct Measurement of the Radiation Sensitivity of Normal Mouse Bone Marrow Cells. *Radiation Research* 14:213-222.

Till, J.E., and E.A. McCulloch. 1980. Hemopoietic stem cell differentiation. *Biochim Biophys Acta* 605:431-459.

Till, J.E., E.A. McCulloch, and L. Siminovitch. 1964. A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. *Proc Natl Acad Sci U S A* 51:29-36.

Tipping, A.J., C. Pina, A. Castor, D. Hong, N.P. Rodrigues, L. Lazzari, G.E. May, S.E. Jacobsen, and T. Enver. 2009. High GATA-2 expression

inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. *Blood* 113:2661-2672.
Tober, J., A. Koniski, K.E. McGrath, R. Vemishetti, R. Emerson, K.K. de Mesy-Bentley, R. Waugh, and J. Palis. 2007. The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. *Blood* 109:1433-1441.
Tsai, F.Y., G. Keller, F.C. Kuo, M. Weiss, J.Z. Chen, M. Rosenblatt, F.W. Alt, and S.H. Orkin. 1994. An Early Hematopoietic Defect in Mice Lacking the Transcription Factor Gata-2. *Nature* 371:221-226.
Turpen, J.B., C.M. Knudson, and P.S. Hoefen. 1981. The early ontogeny of hematopoietic cells studied by grafting cytogenetically labeled tissue anlagen: localization of a prospective stem cell compartment. *Dev Biol* 85:99-112.
Ueno, H., and I.L. Weissman. 2006. Clonal analysis of mouse development reveals a polyclonal origin for yolk sac blood islands. *Dev Cell* 11:519-533.
Ulloa, B.A., S.S. Habbsa, K.S. Potts, A. Lewis, M. McKinstry, S.G. Payne, J.C. Flores, A. Nizhnik, M. Feliz Norberto, C. Mosimann, and T.V. Bowman. 2021. Definitive hematopoietic stem cells minimally contribute to embryonic hematopoiesis. *Cell Rep* 36:109703.
Vazquez, S.E., M.A. Inlay, and T. Serwold. 2015. CD201 and CD27 identify hematopoietic stem and progenitor cells across multiple murine strains independently of Kit and Sca-1. *Exp Hematol* 43:578-585.
Velten, L., S.F. Haas, S. Raffel, S. Blaszkiewicz, S. Islam, B.P. Hennig, C. Hirche, C. Lutz, E.C. Buss, D. Nowak, T. Boch, W.K. Hofmann, A.D. Ho, W. Huber, A. Trumpp, M.A. Essers, and L.M. Steinmetz. 2017. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol* 19:271-281.
Vink, C.S., F.J. Calero-Nieto, X. Wang, A. Maglitto, S.A. Mariani, W. Jawaid, B. Gottgens, and E. Dzierzak. 2020. Iterative Single-Cell Analyses Define the Transcriptome of the First Functional Hematopoietic Stem Cells. *Cell Rep* 31:107627.
Visser, J.W., J.G. Bauman, A.H. Mulder, J.F. Eliason, and A.M. de Leeuw. 1984. Isolation of murine pluripotent hemopoietic stem cells. *J Exp Med* 159:1576-1590.
Vos, O., J.A. Davids, W.W. Weyzen, and D.W. Van Bekkum. 1956. Evidence for the cellular hypothesis in radiation protection by bone marrow cells. *Acta Physiol Pharmacol Neerl* 4:482-486.
Walmsley, M., A. Ciau-Uitz, and R. Patient. 2002. Adult and embryonic blood and endothelium derive from distinct precursor populations which are differentially programmed by BMP in Xenopus. *Development* 129:5683-5695.
Weber, M., M. Mickoleit, and J. Huisken. 2014. Light sheet microscopy. *Methods Cell Biol* 123:193-215.
Weijts, B., L. Yvernogeau, and C. Robin. 2021. Recent Advances in Developmental Hematopoiesis: Diving Deeper With New Technologies. *Frontiers in Immunology* 12:

Weissman, I.L. 2000. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* 100:157-168.
Williams, D.A., I.R. Lemischka, D.G. Nathan, and R.C. Mulligan. 1984. Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* 310:476-480.
Wilson, N.K., S.D. Foster, X. Wang, K. Knezevic, J. Schutte, P. Kaimakis, P.M. Chilarska, S. Kinston, W.H. Ouwehand, E. Dzierzak, J.E. Pimanda, M.F. de Bruijn, and B. Gottgens. 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* 7:532-544.
Wilson, N.K., and B. Gottgens. 2018. Single-Cell Sequencing in Normal and Malignant Hematopoiesis. *Hemasphere* 2:e34.
Wilson, N.K., D.G. Kent, F. Buettner, M. Shehata, I.C. Macaulay, F.J. Calero-Nieto, M. Sanchez Castillo, C.A. Oedekoven, E. Diamanti, R. Schulte, C.P. Ponting, T. Voet, C. Caldas, J. Stingl, A.R. Green, F.J. Theis, and B. Gottgens. 2015. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* 16:712-724.
Wolfensohn, S., and M. Lloyd. 2013. Handbook of laboratory animal management and welfare. Wiley-Blackwell,
Wu, A.M., J.E. Till, L. Siminovitch, and E.A. McCulloch. 1967. A cytological study of the capacity for differentiation of normal hemopoietic colony-forming cells. *J Cell Physiol* 69:177-184.
Wu, A.M., J.E. Till, L. Siminovitch, and E.A. McCulloch. 1968. Cytological evidence for a relationship between normal hemotopoietic colony-forming cells and cells of the lymphoid system. *J Exp Med* 127:455-464.
Yamamoto, M., L.J. Ko, M.W. Leonard, H. Beug, S.H. Orkin, and J.D. Engel. 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev* 4:1650-1662.
Yamane, T. 2018. Mouse Yolk Sac Hematopoiesis. *Front Cell Dev Biol* 6:80.
Yoder, M.C., K. Hiatt, P. Dutt, P. Mukherjee, D.M. Bodine, and D. Orlic. 1997a. Characterization of definitive lymphohematopoietic stem cells in the day 9 murine yolk sac. *Immunity* 7:335-344.
Yoder, M.C., K. Hiatt, and P. Mukherjee. 1997b. In vivo repopulating hematopoietic stem cells are present in the murine yolk sac at day 9.0 postcoitus. *Proc Natl Acad Sci U S A* 94:6776-6780.
Yokomizo, T., and E. Dzierzak. 2010. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development* 137:3651-3661.
Yokota, T., J. Huang, M. Tavian, Y. Nagai, J. Hirose, J.C. Zuniga-Pflucker, B. Peault, and P.W. Kincade. 2006. Tracing the first waves of lymphopoiesis in mice. *Development* 133:2041-2051.
Yoshimoto, M., E. Montecino-Rodriguez, M.J. Ferkowicz, P. Porayette, W.C. Shelley, S.J. Conway, K. Dorshkind, and M.C. Yoder. 2011. Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone

progenitor lacking B-2 potential. *Proc Natl Acad Sci U S A* 108:1468-1473.
Yoshimoto, M., P. Porayette, N.L. Glosson, S.J. Conway, N. Carlesso, A.A. Cardoso, M.H. Kaplan, and M.C. Yoder. 2012. Autonomous murine T-cell progenitor production in the extra-embryonic yolk sac before HSC emergence. *Blood* 119:5706-5714.
Yoshiura, S., T. Ohtsuka, Y. Takenaka, H. Nagahara, K. Yoshikawa, and R. Kageyama. 2007. Ultradian oscillations of Stat, Smad, and Hes1 expression in response to serum. *Proc Natl Acad Sci U S A* 104:11292-11297.
Yvernogeau, L., R. Gautier, L. Petit, H. Khoury, F. Relaix, V. Ribes, H. Sang, P. Charbord, M. Souyri, C. Robin, and T. Jaffredo. 2019. In vivo generation of haematopoietic stem/progenitor cells from bone marrow-derived haemogenic endothelium. *Nat Cell Biol* 21:1334-1345.
Yvernogeau, L., A. Klaus, J. Maas, I. Morin-Poulard, B. Weijts, S. Schulte-Merker, E. Berezikov, J.P. Junker, and C. Robin. 2020. Multispecies RNA tomography reveals regulators of hematopoietic stem cell birth in the embryonic aorta. *Blood* 136:831-844.
Yzaguirre, A.D., and N.A. Speck. 2016. Extravascular endothelial and hematopoietic islands form through multiple pathways in midgestation mouse embryos. *Dev Biol* 415:111-121.
Zaidan, N., and K. Ottersbach. 2018. The multi-faceted role of Gata3 in developmental haematopoiesis. *Open Biol* 8:
Zhang, J., F. Varas, M. Stadtfeld, S. Heck, N. Faust, and T. Graf. 2007. CD41-YFP mice allow in vivo labeling of megakaryocytic cells and reveal a subset of platelets hyperreactive to thrombin stimulation. *Exp Hematol* 35:490-499.
Zhou, F., X. Li, W. Wang, P. Zhu, J. Zhou, W. He, M. Ding, F. Xiong, X. Zheng, Z. Li, Y. Ni, X. Mu, L. Wen, T. Cheng, Y. Lan, W. Yuan, F. Tang, and B. Liu. 2016. Tracing haematopoietic stem cell formation at single-cell resolution. *Nature* 533:487-492.
Zhu, Q., P. Gao, J. Tober, L. Bennett, C. Chen, Y. Uzun, Y. Li, E.D. Howell, M. Mumau, W. Yu, B. He, N.A. Speck, and K. Tan. 2020. Developmental trajectory of prehematopoietic stem cell formation from endothelium. *Blood* 136:845-856.
Zovein, A.C., J.J. Hofmann, M. Lynch, W.J. French, K.A. Turlo, Y. Yang, M.S. Becker, L. Zanetta, E. Dejana, J.C. Gasson, M.D. Tallquist, and M.L. Iruela-Arispe. 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* 3:625-636.
Zovein, A.C., K.A. Turlo, R.M. Ponec, M.R. Lynch, K.C. Chen, J.J. Hofmann, T.C. Cox, J.C. Gasson, and M.L. Iruela-Arispe. 2010. Vascular remodeling of the vitelline artery initiates extravascular emergence of hematopoietic clusters. *Blood* 116:3435-3444.

Appendices

Copyright permissions

Rockefeller University Press - License Terms and Conditions

This is a License Agreement between Mr Christeun S. Vink ("You") and Rockefeller University Press ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Rockefeller University Press, and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date
09-Mar-2021
Order license ID
1102850-1
ISSN
0022-1007
Type of Use
Republish in a thesis/dissertation
Publisher
ROCKEFELLER UNIVERSITY PRESS,
Portion
Chapter/article

LICENSED CONTENT

Publication Title
The Journal of experimental medicine
Article Title
Gata2 is required for HSC generation and survival.
Author/Editor
ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH., ROCKEFELLER INSTITUTE., ROCKEFELLER UNIVERSITY.
Date
01/01/1896
Language
English
Country
United States of America
Rightsholder
Rockefeller University Press
Publication Type
Journal

Start Page
2843
End Page
2850
Issue
13
Volume
210

REQUEST DETAILS

Portion Type
Chapter/article
Page range(s)
1-8, S1
Total number of pages
9
Format (select all that apply)
Print, Electronic
Who will republish the content?
Academic institution
Duration of Use
Life of current edition
Lifetime Unit Quantity
Up to 499
Rights Requested
Main product
Distribution
Worldwide
Translation
Original language of publication
Copies for the disabled?
No
Minor editing privileges?
No
Incidental promotional use?
No
Currency
GBP

NEW WORK DETAILS

Title
Capturing the first haematopoietic stem cell - The needle in the haystack
Instructor name
Mr Christeun S. Vink
Institution name
The University of Edinburgh

Expected presentation date
2022-01-31

ADDITIONAL DETAILS

Order reference number
N/A
The requesting person / organization to appear on the license
Mr Christeun S. Vink

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)
Page 1-8, S1
Editor of portion(s)
de Pater, Emma; Kaimakis, Polynikis; Vink, Chris S; Yokomizo, Tomomasa; Yamada-Inagawa, Tomoko; van der Linden, Reinier; Kartalaei, Parham Solaimani; Camper, Sally A; Speck, Nancy; Dzierzak, Elaine
Volume of serial or monograph
210
Page or page range of portion
2843-2850
Title of the article/chapter the portion is from
Gata2 is required for HSC generation and survival.
Author of portion(s)
de Pater, Emma; Kaimakis, Polynikis; Vink, Chris S; Yokomizo, Tomomasa; Yamada-Inagawa, Tomoko; van der Linden, Reinier; Kartalaei, Parham Solaimani; Camper, Sally A; Speck, Nancy; Dzierzak, Elaine
Issue, if republishing an article from a serial
13
Publication date of portion
2013-12-16

Rockefeller University Press - License Terms and Conditions

This is a License Agreement between Mr Christeun S. Vink ("You") and Rockefeller University Press ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Rockefeller University Press, and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date
09-Mar-2021
Order license ID
1102854-1
ISSN
0022-1007
Type of Use
Republish in a thesis/dissertation
Publisher
ROCKEFELLER UNIVERSITY PRESS,
Portion
Chapter/article

LICENSED CONTENT

Publication Title
The Journal of experimental medicine
Article Title
Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation.
Author/Editor
ROCKEFELLER INSTITUTE FOR MEDICAL RESEARCH., ROCKEFELLER INSTITUTE., ROCKEFELLER UNIVERSITY.
Date
01/01/1896
Language
English
Country
United States of America
Rightsholder
Rockefeller University Press
Publication Type
Journal
Start Page
93
End Page
106
Issue
1

Volume
212

REQUEST DETAILS

Portion Type
Chapter/article
Page range(s)
93-106, S1-3
Total number of pages
17
Format (select all that apply)
Print, Electronic
Who will republish the content?
Academic institution
Duration of Use
Life of current edition
Lifetime Unit Quantity
Up to 499
Rights Requested
Main product
Distribution
Worldwide
Translation
Original language of publication
Copies for the disabled?
No
Minor editing privileges?
No
Incidental promotional use?
No
Currency
GBP

NEW WORK DETAILS

Title
Capturing the first haematopoietic stem cell - The needle in the haystack
Instructor name
Mr Christeun S. Vink
Institution name
The University of Edinburgh
Expected presentation date
2022-01-31

ADDITIONAL DETAILS

Order reference number

N/A

The requesting person / organization to appear on the license

Mr Christeun S. Vink

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)

Page 93-106, S1-3

Editor of portion(s)

Solaimani Kartalaei Yamada-Inagawa; Vink; de Pater; van der Linden; Marks-Bluth; van der Sloot; van den Hout; Yokomizo; van Schaick-Solernó; Delwel; Pimanda; van IJcken; Dzierzak, Elaine

Volume of serial or monograph

212

Page or page range of portion

93-106

Title of the article/chapter the portion is from

Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation.

Author of portion(s)

Solaimani Kartalaei Yamada-Inagawa; Vink; de Pater; van der Linden; Marks-Bluth; van der Sloot; van den Hout; Yokomizo; van Schaick-Solernó; Delwel; Pimanda; van IJcken; Dzierzak, Elaine

Issue, if republishing an article from a serial

1

Publication date of portion

2015-01-12

Elsevier Science & Technology Journals - License Terms and Conditions

This is a License Agreement between Mr Christeun S. Vink ("You") and Elsevier Science & Technology Journals ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier Science & Technology Journals, and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date
12-Mar-2021
Order license ID
1103747-1
ISSN
1528-0020
Type of Use
Republish in a thesis/dissertation
Publisher
AMERICAN SOCIETY OF HEMATOLOGY
Portion
Chapter/article

LICENSED CONTENT

Publication Title
Blood : journal of the American Society of Hematology
Author/Editor
American Society of Hematology., HighWire Press.
Date
01/01/1946
Language
English
Country
United States of America
Rightsholder
Elsevier Science & Technology Journals
Publication Type
e-Journal
URL
http://www.bloodjournal.org/

REQUEST DETAILS

Portion Type
Chapter/article
Page range(s)
1426-1437, S1-4

Total number of pages
16
Format (select all that apply)
Print, Electronic
Who will republish the content?
Academic institution
Duration of Use
Life of current edition
Lifetime Unit Quantity
Up to 499
Rights Requested
Main product
Distribution
Worldwide
Translation
Original language of publication
Copies for the disabled?
No
Minor editing privileges?
Yes
Incidental promotional use?
No
Currency
GBP

NEW WORK DETAILS

Title
Capturing the first haematopoietic stem cell - The needle in the haystack
Instructor name
Mr Christeun S. Vink
Institution name
The University of Edinburgh
Expected presentation date
2022-01-31

ADDITIONAL DETAILS

Order reference number
N/A
The requesting person / organization to appear on the license
Mr Christeun S. Vink

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)
Page 1426-1437, S1-4
Editor of portion(s)

N/A
Volume of serial or monograph
127
Page or page range of portion
Page 1426-1437, S1-4
Title of the article/chapter the portion is from
Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors
Author of portion(s)
American Society of Hematology.; HighWire Press.
Issue, if republishing an article from a serial
11
Publication date of portion
2016-03-17

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS
Mar 09, 2021

This Agreement between Mr. Christeun Vink ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5023050233681
License date	Mar 06, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature
Licensed Content Title	From haematopoietic stem cells to complex differentiation landscapes
Licensed Content Author	Elisa Laurenti et al
Licensed Content Date	Jan 25, 2018
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	4
High-res required	no
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Capturing the first haematopoietic stem cell - The needle in the haystack
Institution name	The University of Edinburgh
Expected presentation date	Jan 2022
Portions	Figures 1A, 1B, 1C. Figure 2A
Requestor Location	Mr. Christeun Vink 47 Little France Crescent Edinburgh, EH16 4TJ United Kingdom Attn: Mr. Christeun Vink
Customer VAT ID	GB592950700
Total	0.00 GBP

Annual Reviews, Inc. - License Terms and Conditions

This is a License Agreement between Mr Christeun S. Vink ("You") and Annual Reviews, Inc. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Annual Reviews, Inc., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date
08-Mar-2021
Order license ID
1102603-1
ISSN
1545-3278
Type of Use
Republish in a thesis/dissertation
Publisher
ANNUAL REVIEWS
Portion
Image/photo/illustration

LICENSED CONTENT

Publication Title
Annual review of immunology
Date
01/01/1983
Language
English
Country
United States of America
Rightsholder
Annual Reviews, Inc.
Publication Type
e-Journal
URL
http://arjournals.annualreviews.org/loi/immunol

REQUEST DETAILS

Portion Type
Image/photo/illustration
Number of images / photos / illustrations
1
Format (select all that apply)
Print, Electronic
Who will republish the content?
Academic institution

Duration of Use
Life of current edition
Lifetime Unit Quantity
Up to 499
Rights Requested
Main product
Distribution
Worldwide
Translation
Original language of publication
Copies for the disabled?
No
Minor editing privileges?
Yes
Incidental promotional use?
No
Currency
GBP

NEW WORK DETAILS

Title
Capturing the first haematopoietic stem cell - The needle in the haystack
Instructor name
Mr Christeun S Vink
Institution name
The University of Edinburgh
Expected presentation date
2022-01-31

ADDITIONAL DETAILS

Order reference number
N/A
The requesting person / organization to appear on the license
Mr Christeun S. Vink

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)
Figure 1a
Editor of portion(s)
N/A
Volume of serial or monograph
25
Page or page range of portion
748
Title of the article/chapter the portion is from

Ontogeny of the Hematopoietic System
Author of portion(s)
Ana Cumano and Isabelle Godin
Issue, if republishing an article from a serial
N/A
Publication date of portion
2007-04-23

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
Mar 09, 2021

This Agreement between Mr. Christeun Vink ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5024241402600
License date	Mar 08, 2021
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	FEBS Letters
Licensed Content Title	Hematopoietic (stem) cell development — how divergent are the roads taken?
Licensed Content Author	Mari-Liis Kauts, Chris S. Vink, Elaine Dzierzak
Licensed Content Date	Sep 1, 2016
Licensed Content Volume	590
Licensed Content Issue	22
Licensed Content Pages	12
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Will you be translating?	No
Title	Capturing the first haematopoietic stem cell - The needle in the haystack
Institution name	The University of Edinburgh
Expected presentation date	Jan 2022
Portions	Figure 1
Requestor Location	Mr. Christeun Vink 47 Little France Crescent Edinburgh, EH16 4TJ United Kingdom Attn: Mr. Christeun Vink
Publisher Tax ID	EU826007151
Customer VAT ID	GB592950700
Total	0.00 GBP

Elsevier Science & Technology Journals - License Terms and Conditions

This is a License Agreement between Mr Christeun S. Vink ("You") and Elsevier Science & Technology Journals ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier Science & Technology Journals, and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date
10-Mar-2021
Order license ID
1103148-1
ISSN
1528-0020
Type of Use
Republish in a thesis/dissertation
Publisher
AMERICAN SOCIETY OF HEMATOLOGY
Portion
Image/photo/illustration

LICENSED CONTENT

Publication Title
Blood : journal of the American Society of Hematology
Author/Editor
American Society of Hematology., HighWire Press.
Date
01/01/1946
Language
English
Country
United States of America
Rightsholder
Elsevier Science & Technology Journals
Publication Type
e-Journal
URL
http://www.bloodjournal.org/

REQUEST DETAILS

Portion Type
Image/photo/illustration
Number of images / photos / illustrations
1

Format (select all that apply)
Print, Electronic
Who will republish the content?
Academic institution
Duration of Use
Life of current edition
Lifetime Unit Quantity
Up to 499
Rights Requested
Main product
Distribution
Worldwide
Translation
Original language of publication
Copies for the disabled?
No
Minor editing privileges?
Yes
Incidental promotional use?
No
Currency
GBP

NEW WORK DETAILS

Title
Capturing the first haematopoietic stem cell - The needle in the haystack
Instructor name
Mr Christeun S. Vink
Institution name
The University of Edinburgh
Expected presentation date
2022-01-31

ADDITIONAL DETAILS

Order reference number
N/A
The requesting person / organization to appear on the license
Mr Christeun S. Vink

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)
Figure 1A
Editor of portion(s)
N/A
Volume of serial or monograph

96
Page or page range of portion
2903
Title of the article/chapter the portion is from
CFU-S11 activity does not localize solely with the aorta in the aorta-gonad-mesonephros region
Author of portion(s)
American Society of Hematology.; HighWire Press.
Issue, if republishing an article from a serial
8
Publication date of portion
2000-10-15

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS
Jun 14, 2021

This Agreement between Mr. Christeun Vink ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	5087601426535
License date	Jun 14, 2021
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Developmental Dynamics
Licensed Content Title	The cell clusters in the dorsal aorta of mammalian embryos
Licensed Content Author	V. E. Emmel
Licensed Content Date	Jan 23, 2005
Licensed Content Volume	19
Licensed Content Issue	3
Licensed Content Pages	21
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	2
Will you be translating?	No
Title	Capturing the first haematopoietic stem cell - The needle in the haystack
Institution name	The University of Edinburgh
Expected presentation date	Jan 2022
Portions	Figure 2, Figure 8
Requestor Location	Mr. Christeun Vink 47 Little France Crescent Edinburgh, EH16 4TJ United Kingdom Attn: Mr. Christeun Vink
Publisher Tax ID	EU826007151
Customer VAT ID	GB592950700
Total	0.00 GBP

The Company of Biologists Ltd. - License Terms and Conditions

This is a License Agreement between Mr Christeun S. Vink ("You") and The Company of Biologists Ltd. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by The Company of Biologists Ltd., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date
14-Jun-2021
Order License ID
1125768-1
ISSN
1477-9129
Type of Use
Republish in a thesis/dissertation
Publisher
COMPANY OF BIOLOGISTS,
Portion
Image/photo/illustration

LICENSED CONTENT

Publication Title
Development
Article Title
Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos.
Author/Editor
Company of Biologists.
Date
01/01/1987
Language
English
Country
United Kingdom of Great Britain and Northern Ireland
Rightsholder
The Company of Biologists Ltd.
Publication Type
e-Journal
Start Page
3651
End Page
3661
Issue
21

Volume
137
URL
http://dev.biologists.org/

REQUEST DETAILS

Portion Type
Image/photo/illustration
Number of images / photos / illustrations
5
Format (select all that apply)
Print, Electronic
Who will republish the content?
Academic institution
Duration of Use
Life of current edition
Lifetime Unit Quantity
Up to 499
Rights Requested
Main product
Distribution
Worldwide
Translation
Original language of publication
Copies for the disabled?
No
Minor editing privileges?
Yes
Incidental promotional use?
No
Currency
GBP

NEW WORK DETAILS

Title
Capturing the first haematopoietic stem cell - The needle in the haystack
Instructor name
Mr Christeun S. Vink
Institution name
The University of Edinburgh
Expected presentation date
2022-01-31

ADDITIONAL DETAILS

Order reference number

N/A
The requesting person / organization to appear on the license
Mr Christeun S. Vink

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)
Figure 3A, 4A, 4B, 4C, 7F
Editor of portion(s)
Yokomizo, Tomomasa; Dzierzak, Elaine
Volume of serial or monograph
137
Page or page range of portion
3651-3661
Title of the article/chapter the portion is from
Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos.
Author of portion(s)
Yokomizo, Tomomasa; Dzierzak, Elaine
Issue, if republishing an article from a serial
21
Publication date of portion
2010-11-01

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS
Jul 09, 2021

This Agreement between Mr. Christeun Vink ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5104801167843
License date	Jul 09, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	Notch and Stem Cells
Licensed Content Author	Anna Bigas, Cristina Porcheri
Licensed Content Date	Jan 1, 2018
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation/distribution	1 - 29
Author of this Springer Nature content	no
Title	Capturing the first haematopoietic stem cell - The needle in the haystack
Institution name	The University of Edinburgh
Expected presentation date	Jan 2022
Portions	Figure 4
Requestor Location	Mr. Christeun Vink 47 Little France Crescent Edinburgh, EH16 4TJ United Kingdom Attn: Mr. Christeun Vink
Customer VAT ID	GB592950700
Total	0.00 GBP

ELSEVIER LICENSE
TERMS AND CONDITIONS
Jul 09, 2021

This Agreement between Mr. Christeun Vink ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5104800873974
License date	Jul 09, 2021
Licensed Content Publisher	Elsevier
Licensed Content Publication	Cell Stem Cell
Licensed Content Title	Combinatorial Transcriptional Control In Blood Stem/Progenitor Cells: Genome-wide Analysis of Ten Major Transcriptional Regulators
Licensed Content Author	Nicola K. Wilson,Samuel D. Foster,Xiaonan Wang,Kathy Knezevic,Judith Schütte,Polynikis Kaimakis,Paulina M. Chilarska,Sarah Kinston,Willem H. Ouwehand,Elaine Dzierzak,John E. Pimanda,Marella F.T.R. de Bruijn,Berthold Göttgens
Licensed Content Date	Oct 8, 2010
Licensed Content Volume	7
Licensed Content Issue	4
Licensed Content Pages	13
Start Page	532
End Page	544
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	Capturing the first haematopoietic stem cell - The needle in the haystack
Institution name	The University of Edinburgh
Expected presentation date	Jan 2022
Portions	Figure 6B
Requestor Location	Mr. Christeun Vink 47 Little France Crescent

	Edinburgh, EH16 4TJ United Kingdom Attn: Mr. Christeun Vink
Publisher Tax ID	GB 494 6272 12
Customer VAT ID	GB592950700
Total	0.00 GBP

From: PNAS Permissions <PNASPermissions@nas.edu>
Sent: 09 July 2021 16:00
To: VINK Chris <C.S.Vink@ed.ac.uk>
Subject: RE: Durand et al

Dear Mr. Vink,

Thank you for your message. Permission is granted for your use of the material as described in your request. Please include a complete citation for the original PNAS article when reusing the material, and include "Copyright (2007) National Academy of Sciences, U.S.A." as a copyright note. Because this material published between 1993 and 2008, a copyright note is needed. There is no charge for this material, either. Let us know if you have any questions.

Best regards,
Delaney Cruickshank for
Diane Sullenberger
PNAS Executive Editor

From: VINK Chris <C.S.Vink@ed.ac.uk>
Sent: Friday, July 9, 2021 10:15 AM
To: PNAS Permissions <PNASPermissions@nas.edu>
Subject: Durand et al

Dear PNAS,

Please could I request permission to reuse the following material (figure) in my PhD thesis:

Your full name, affiliation, and title
Mr Christeun S Vink, PhD student/laboratory manager, Bsc

Your complete mailing address, phone number, and email
Queen's Medical Research Institute, 47 Little France Crescent, EH16 4TJ, Edinburgh, Scotland, UK; c.s.vink@ed.ac.uk, 0044 131 242 9161,

PNAS volume number, issue number, and issue date
Volume 104, issue 52, 26th December 2007

PNAS article title
Embryonic stromal clones reveal developmental regulators of definitive hematopoietic stem cells

PNAS authors' names
Charles Durand, Catherine Robin, Karine Bollerot, Margaret H. Baron, Katrin Ottersbach, and Elaine Dzierzak

Page numbers of items to be reprinted
20841

Figure/table number or portion of text to be reprinted
Figure 4

Requests must also include the following information about the intended use of the material:

Title of work in which PNAS material will appear
Thesis: Capturing the first haematopoietic stem cell - The needle in the haystack

Authors/editors of work
Mr Christeun S vink

Publisher of work
The University of Edinburgh

Retail price of work
£0

Number of copies of work to be produced
Max 30

Intended audience
Thesis committee

Whether work is for non-profit or commercial use
Non-profit

Full publication list

- Maglitto, A.*, Mariani, S.A.*, De Pater, E., Rodriguez-Seoane, C., **Vink, C.S.**, Piao, X., Lukke, M.-L., Dzierzak E. 2021. Unexpected redundancy of Gpr56 and Gpr97 during hematopoietic cell development and differentiation. **Blood Advances** 5(3):829-842. PMID: 33560396.
- Munro, D.A.D., Bradford, B.M., Mariani, S.A., Hampton, D.W., **Vink, C.S.**, Chandran, S., Hume, D.A., Pridans, C., Priller, J. 2020. CNS macrophages differentially rely on an intronic Csf1r enhancer for their development. **Development** 147(23):dev194449. PMID: 33323375.
- **Vink, C.S.**, Calero-Nieto, F.J., Wang, X., Maglitto, A., Mariani, S.A., Jawaid, W., Göttgens, B., Dzierzak, E. 2020. Iterative single cell analyses define the transcriptome of the first hematopoietic stem cells. **Cell Reports** 31(6):107627. PMID: 32402290.
- Porcheri, C., Golan, O., Calero-Nieto, F.J., Thambyrajah, R., Ruiz-Herguido, C., Wang, X., Catto, F., Guillen, Y., Sinha, R., Gonzalez, J., Kinston, S.J., Mariani, S.A., Maglitto, A., **Vink, C.S.**, Dzierzak, E., Charbord, P., Göttgens, B., Espinosa, L., Sprinzak, D., Bigas, A. 2020. Notch ligand Dll4 impairs cell recruitment to aortic clusters and limits blood stem cell generation. **EMBO Journal** 39(8):e104270. PMID: 32149421.
- Li, Z., Mariani, S.A., Rodriguez-Seoane, C., He, W., Ning, X., Liu, B. **Vink, C.S.**, Dzierzak, E. 2019. A role for macrophages in hematopoiesis in the embryonic head. **Blood** 134(22):1929-1940. PMID: 31697805.
- Mariani, S.A., Li, Z., Rice, S., Krieg, C., Fragkogianni, S., Robinson, M., **Vink, C.S.**, Pollard, J.W., Dzierzak, E. 2019. Pro-inflammatory Aorta-Gonad-Mesonephros-associated macrophages are involved in embryonic development of hematopoietic stem cells. **Immunity** 50(6):1439-1452. PMID: 31178352.
- Munro, D.A.D., Wineberg, Y., Tarnick, J., **Vink, C.S.**, Li, Z., Pridans, C., Dzierzak, E., Kalisky, T., Hohenstein, P., Davies, J.A. 2019. Macrophages restrict the nephrogenic field and promote endothelial connections during kidney development. **eLife** 8. pii: e43271. PMID: 30758286.
- Eich, C., Arlt, J., **Vink, C.S.**, Solaimani Kartalaei, P., Kaimakis, P., Mariani, S.A., Van der Linden, R., Van Cappellen, W.A., Dzierzak, E. 2018. In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. **Journal of Exp. Medicine** 215(1):233-248. PMID: 29217535.
- Nasrallah, R., Fast, E.M., Solaimani, P., Knezevic, K., Eliades, A., Patel, R., Thambyrajah, R., Unnikrishnan, A., Thoms, J., Beck, D., **Vink, C.S.**, Smith, A., Wong, J., Shepherd, M., Kent, D., Roychoudhuri, R., Paul, F., Klippert, J., Hammes, A., Willnow, T., Göttgens, B., Dzierzak, E., Zon, L.I., Lacaud, G., Kouskoff, V.,

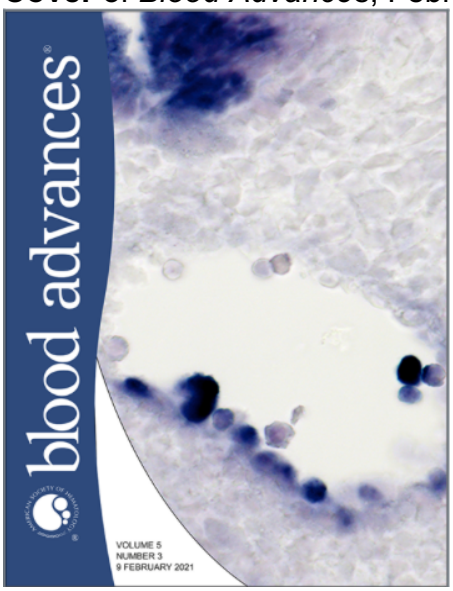
Pimanda, J.E. 2016. Identification of novel regulators of developmental hematopoiesis using endoglin regulatory elements as molecular probes. **Blood** 128(15):1928-1939. PMID: 27554085.

- Kauts, M.L., **Vink, C.S.**, Dzierzak, E. 2016. Hematopoietic (stem) cell development-how divergent are the roads taken? **FEBS Letters** 590(22):3975-3986. PMID: 27543859.
- Li, Z., **Vink, C.S.**, Mariani, S.A., Dzierzak, E. 2016. Subregional localization and characterization of Ly6aGFP-expressing hematopoietic cells in the mouse embryonic head. **Developmental Biology** 416(1):34-41. PMID: 27235813.
- Crisan, M., Solaimani-Kartalaei, P., Neagu, A., Karkanpouna, S., Yamada-Inagawa, T., Purini, C., **Vink, C.S.**, Van der Linden, R., Van IJcken, W., Chuva de Sousa Lopes, S.M., Monteiro, R., Mummery, C., Dzierzak, E. 2016. BMP and Hedgehog regulate distinct AGM hematopoietic stem cells ex vivo. **Stem Cell Reports** 6(3):383-95. PMID: 26923823.
- Kaimakis, P.*, De Pater, E.*, Eich, C.*, Solaimani-Kartalaei, P., Kauts, M-L., **Vink, C.S.**, Van der Linden, R., Jaegle, M., Yokomizo, T., Meijer, D., Dzierzak, E. 2016. Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. **Blood** 127(11):1426-37. PMID: 26834239.
- Crisan, M., Solaimani-Kartalaei, P., **Vink, C.S.**, Yamada-Inagawa, T., Bollerot, K., Van IJcken, W., Van der Linden, R., De Sousa Lopes, S.M., Monteiro, R., Mummery, C., Dzierzak, E. 2015. BMP signalling differentially regulates distinct haematopoietic stem cell types. **Nature Communications** 6:8040/6:8793 (corrigendum). PMID: 26282601.
- Solaimani-Kartalaei, P.*, Yamada-Inagawa,T.*, **Vink, C.S.**, De Pater, E., Van der Linden, R., Marks-Bluth, J., Van der Sloot, A., Van den Hout, M., Yokomizo, T., Van Schaick-Solerno, M.L., Delwel, R., Pimanda, J.E., Van IJcken, W.F., Dzierzak, E. 2015. Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. **Journal of Exp. Medicine** 212(1):93-106. PMID: 25547674.
- Imanirad, P.*, Solaimani-Kartalaei, P.*, Crisan, M.*, **Vink, C.S.**, Yamada-Inagawa,T., De Pater, E., Kurek, D., Kaimakis, P., Van der Linden, R., Speck, N.A., Dzierzak, E. 2014. HIF1α is a regulator of hematopoietic progenitor and stem cell development in hypoxic sites of the mouse embryo. **Stem Cell Research** 12(1):24-35. PMID: 24141110.
- De Pater, E., Kaimakis, P., **Vink, C.S.**, Yokomizo, T., Yamada-Inagawa, T., Van der Linden, R., Camper, S.A., Speck, N.A., Dzierzak, E. 2013. Gata2 is required for HSC generation and HSC survival. **Journal of Exp. Medicine** 210(13):2843-50. PMID: 24297996.
- Chen, M.J., Li, Y., De Obaldia, M.E., Yang, Q., Yzaguirre, A.D., Yamada-Inagawa, T., **Vink, C.S.**, Bhandoola, A., Dzierzak, E., Speck, N.A. 2011. Erythroid/Myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells. **Cell Stem Cell** 9(6):541-552. PMID: 22136929.

- Bee, T., Liddiard, K., Swiers, G., Bickley, S.R.B., **Vink, C.S.**, Jarratt, A., Hughes, J.R., Medvinsky A., de Bruijn, M.F. 2009. Alternative Runx1 promoter usage in mouse developmental hematopoiesis. **Blood Cells, Molecules and Diseases** 43(1):35-42. PMID: 19464215.

- **Cover** of *Blood*, April 13th 2017, vol. 129, issue 15: 2039-2205.

- **Cover** of *Blood Advances*, February 9th 2021, vol. 5, issue 3: 613-925.



Publications submitted for PhD by Research Publications

Publications

1. De Pater, E., Kaimakis, P., **Vink, C.S.** *et al.* (2013). Gata2 is required for HSC generation and HSC survival. *Journal of Experimental Medicine*, *210*(13), 2843-50. (***Chapter 1***)

2. Kaimakis, P.*, De Pater, E.*, Eich, C.*, Solaimani-Kartalaei, P., Kauts, M-L., **Vink, C.S.** *et al.* (2016). Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. *Blood*, *127*(11), 1426-37. (***Chapter 2***)

3. Eich, C., Arlt, J., **Vink, C.S.** *et al.* (2018). In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. *Journal of Experimental Medicine*, *215*(1), 233-248. (***Chapter 3***)

4. Solaimani-Kartalaei, P.*, Yamada-Inagawa,T.*, **Vink, C.S.** *et al.* (2015). Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *Journal of Experimental Medicine*, *212*(1), 93-106. (***Chapter 4***)

5. **Vink, C.S.** *et al.* (2020). Iterative single cell analyses define the transcriptome of the first hematopoietic stem cells. *Cell Reports*, *31*(6), 107627. PMID: 32402290. (***Chapter 5***)

6. Crisan, M., Solaimani-Kartalaei, P., **Vink, C.S.** *et al.* (2015). BMP signalling differentially regulates distinct haematopoietic stem cell types. *Nature Communications*, *6*(8040)/*6*(8793; corrigendum). PMID: 26282601. (***Chapter 6***)

* Authors contributed equally.

Gata2 is required for HSC generation and survival

Emma de Pater,[1] Polynikis Kaimakis,[1] Chris S. Vink,[1] Tomomasa Yokomizo,[1] Tomoko Yamada-Inagawa,[1] Reinier van der Linden,[1] Parham Solaimani Kartalaei,[1] Sally A. Camper,[2] Nancy Speck,[3,4] and Elaine Dzierzak[1]

[1]Erasmus Medical Center, Erasmus Stem Cell Institute, 3000 CA Rotterdam, Netherlands
[2]Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109
[3]Abramson Family Cancer Research Institute and [4]Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, PA 19104

Knowledge of the key transcription factors that drive hematopoietic stem cell (HSC) generation is of particular importance for current hematopoietic regenerative approaches and reprogramming strategies. Whereas GATA2 has long been implicated as a hematopoietic transcription factor and its dysregulated expression is associated with human immunodeficiency syndromes and vascular integrity, it is as yet unknown how GATA2 functions in the generation of HSCs. HSCs are generated from endothelial cells of the major embryonic vasculature (aorta, vitelline, and umbilical arteries) and are found in intra-aortic hematopoietic clusters. In this study, we find that GATA2 function is essential for the generation of HSCs during the stage of endothelial-to-hematopoietic cell transition. Specific deletion of *Gata2* in *Vec* (*Vascular Endothelial Cadherin*)-expressing endothelial cells results in a deficiency of long-term repopulating HSCs and intra-aortic cluster cells. By specific deletion of *Gata2* in *Vav*-expressing hematopoietic cells (after HSC generation), we further show that GATA2 is essential for HSC survival. This is in contrast to the known activity of the RUNX1 transcription factor, which functions only in the generation of HSCs, and highlights the unique requirement for GATA2 function in HSCs throughout all developmental stages.

CORRESPONDENCE
E. Dzierzak:
e.dzierzak@erasmusmc.nl

Abbreviations used: AGM, aorta-gonad-mesonephros; CFU-C, colony-forming unit culture; DA, dorsal aorta; EHT, endothelial-to-hematopoietic transition; FL, fetal liver; HPC, hematopoietic progenitor cell; HSC, hematopoietic stem cell; sp, somite pair; YS, yolk sac.

The permanent adult hematopoietic system initiates with the formation of hematopoietic stem cells (HSCs) (Dzierzak and Speck, 2008). Expansion of HSCs in culture and reprogramming somatic cells into HSCs are as yet not possible and, thus, necessitate an understanding of the molecular programs directing the generation of HSCs. At the time of HSC generation in the mouse aorta-gonad-mesonephros (AGM) region (embryonic day [E] 10.5), clusters of hematopoietic cells are found closely associated with the ventral wall of the dorsal aorta and along the other major arteries (vitelline and umbilical; de Bruijn et al., 2000, 2002; North et al., 2002; Taoudi and Medvinsky, 2007; Zovein et al., 2008; Boisset et al., 2010; Yokomizo and Dzierzak, 2010). The *Gata2* transcription factor is expressed in the mouse embryo in a pattern consistent with a role in hematopoietic cell development (Minegishi et al., 1999; Robert-Moreno et al., 2005). It is first expressed at E7.5 in the primitive streak and the endothelial cells of the paired dorsal aorta. Later, *Gata2* is expressed in endothelial cells lining the dorsal aorta, vitelline, and umbilical arteries and in the intra-arterial cluster cells at the time of definitive hematopoietic progenitor cell (HPC) and HSC formation. *Gata2* is also expressed in hematopoietic cells of the yolk sac (YS), fetal liver (FL), and placenta (PL) and in adult BM HSCs (Ng et al., 1994; Orlic et al., 1995; Minegishi et al., 1999; Nardelli et al., 1999; Robert-Moreno et al., 2005).

Germline $Gata2^{-/-}$ embryos suffer from FL anemia and die at E10, just before the appearance of the first HSCs, and $Gata2^{-/-}$ ES cells do not

T. Yokomizo's present address is Department of Hematology, Juntendo University School of Medicine, Bunkyo-ku, Tokyo 113-8421, Japan.

The Rockefeller University Press $30.00
J. Exp. Med. 2013 Vol. 210 No. 13 2843-2850
www.jem.org/cgi/doi/10.1084/jem.20130751

contribute to definitive hematopoiesis in mouse chimeras (Ng et al., 1994; Tsai et al., 1994; Orlic et al., 1995; Tsai and Orkin, 1997; Minegishi et al., 1999; Nardelli et al., 1999; Robert-Moreno et al., 2005). Haploinsufficient *Gata2*$^{+/-}$ mice are severely affected in the production of early progenitors (Tsai et al., 1994), and *Gata2*$^{+/-}$ BM HSCs are qualitatively defective in serial transplantation assays (Ling et al., 2004; Rodrigues et al., 2008). When crossed with *Ly6A-GFP* mice as a marker for emerging hematopoietic cells, GFP^{+} hematopoietic cells are decreased in the E11 AGM region of *Gata2*$^{+/-}$ embryos (Ling et al., 2004). Recent studies of the *Gata2* intron 4 enhancer show that FL HSCs are affected when *Gata2* is dysregulated (Johnson et al., 2012; Lim et al., 2012). Yet, it is unknown whether GATA2 is required in the endothelial compartment for the formation of intra-arterial clusters and AGM HSCs, and what the function of GATA2 is thereafter. Furthermore, GATA factors are described to act through combinatory interactions with other key regulators, including RUNX1 (Wilson et al., 2010; van Riel et al., 2012). Previously, RUNX1 was shown to be required for the generation of HSCs during the endothelial-to-hematopoietic transition (EHT). Although RUNX1 continues to be expressed, it is not required after HSCs are made (Chen et al., 2009).

Taking a conditional KO (cKO) approach to examine when GATA2 is required, we deleted a floxed allele of *Gata2* with *Vec-Cre* in *VE-Cadherin*–expressing endothelial cells before the generation of HSCs and with *Vav-Cre* in hematopoietic cells after the generation of HSCs (Stadtfeld and Graf, 2005; Charles et al., 2006; Chen et al., 2009). Using the same approach as was used for the conditional deletion of *Runx1*, we conclude that the developmental requirements for GATA2 and RUNX1 differ during the generation of HPCs, and although both transcription factors are required for the generation of HSCs, only GATA2 is further required for HSC survival.

RESULTS AND DISCUSSION

Unlike germline *Gata2*$^{-/-}$ embryos, which die at E10, conditional *Vec-Cre:Gata2*$^{f/f}$ embryos survive until E14 and show FL anemia (Fig. 1 A). This phenotype is similar to embryos with *Gata2* deficiencies directed by the mutation/deletion of the *Gata2* intron 4 enhancer (Johnson et al., 2012; Lim et al., 2012). To examine whether HPC generation and/or function is affected, cells from the AGM and FL of *Vec* cKO embryos were tested in the colony-forming unit culture (CFU-C) assay. Already at E10, HPC numbers in the AGM (including vitelline and umbilical arteries) of *Vec-Cre:Gata2*$^{f/+}$ and *Vec-Cre:Gata2*$^{f/f}$ embryos were decreased by 3-fold ($P < 0.05$) and 3.5-fold ($P < 0.01$), respectively, as compared with WT. The decrease in HPCs in the heterozygous cKO AGM corresponds to the decrease observed in germline *Gata2*$^{+/-}$ AGMs (Fig. 1 B and Table 1). However, homozygous cKO HPC numbers were not decreased to the level observed in germline *Gata2*$^{-/-}$ AGMs (25-fold). Individual colonies were picked from methylcellulose plates and DNA was examined for recombination by PCR. Colonies showed a frequency of 92% recombination for one *Gata2* floxed allele. Surprisingly, 31% of colonies

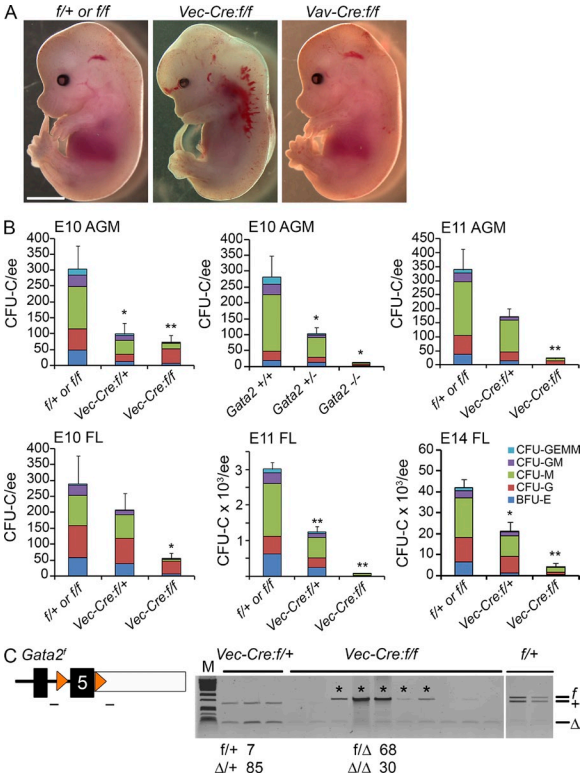


Figure 1. GATA2 is required in VE-Cadherin–expressing cells for hematopoietic progenitor formation. (A) Bright field images of E14 WT, *Vec-Cre:Gata2*$^{f/f}$, and *Vav-Cre:Gata2*$^{f/f}$ embryos. Bar, 2 mm. (B) CFU-C numbers per AGM and FL from *Gata2*-deleted embryos. Germline *Gata2*$^{-/-}$ AGMs at E10 (30–34 sp), $n = 4$. The number of embryos analyzed is 8 WT, 11 *Gata2*$^{+/-}$, and 4 *Gata2*$^{-/-}$. For *Vec-Cre*–mediated *Gata2* deletion: WT, *Vec-Cre:Gata2*$^{f/+}$, and *Vec-Cre:*Gata2$^{f/f}$ AGMs and FLs at E10 (30–36 sp), E11 (43–48 sp), and FL E14, $n = 2$. The number of E10 embryos analyzed is 5 WT, 6 *Vec-Cre:Gata2*$^{f/+}$, and 6 *Vec-Cre:Gata2*$^{f/f}$; $n = 3$. The number of E11 embryos analyzed is 6 WT, 3 *Vec-Cre:Gata2*$^{f/+}$, and 4 *Vec-Cre:Gata2*$^{f/f}$; E14 FL $n = 3$. The number of embryos analyzed is 11 WT, 3 *Vec-Cre:Gata2*$^{f/+}$, and 5 *Vec-Cre:Gata2*$^{f/f}$. Error bars are mean ± SEM of total colony numbers. *, $P < 0.05$; **, $P < 0.01$, significance determined by Student's *t* test. ee, embryo equivalent of specified tissue. (C) Schematic representation of the *Gata2* locus with *Loxp* sites and genotyping PCR primer location. *Loxp* sites are indicated by orange triangles. Gel showing the *Gata2* flox band of 844 bp (f) deleted (Δ) band of 181 bp and WT *Gata2* band (+) of 717 bp for individual CFU-C. The number of identified genotypes of individual colonies is shown beneath the gel.

deleted the second allele (Fig. 1 C), indicating that some progenitors are *Gata2* independent. This is underscored by the finding that the germline *Gata2*$^{-/-}$ AGMs still contain some CFU-C activity (Fig. 1 B and Table 1). At E11, *Vec-Cre:Gata2*$^{f/f}$ AGMs showed a more profound 15-fold ($P < 0.01$) reduction of CFU-C compared with WT. Thus, most AGM HPCs require GATA2 in *VE-Cadherin*–expressing cells during the time of hematopoietic cluster generation (80% at E10 and 95% at E11).

Similar decreases in HPCs were found in the *Vec-Cre:Gata2*$^{f/f}$ FL. At E10, E11, and E14, FL HPCs were significantly reduced by a factor of 5, 45, and 10, respectively, compared

Table 1. CFU-C numbers in *Gata2*-deficient embryos

Tissue	Stage	Genotype		
		Gata2$^{+/+}$	*Gata2*$^{+/-}$	*Gata2*$^{-/-}$
AGM	E10 (30-34sp)	282.2 ± 64.4	103.6 ± 19.1 (P < 0.05)	11.2 ± 2.7 (P < 0.05)
YS	E10 (30-34sp)	918.6 ± 147.9	584.5 ± 72.4	25.7 ± 5.3 (P < 0.01)
		Gata2$^{f/+}$ or *Gata2*$^{f/f}$	*Vec-Cre:Gata2*$^{f/+}$	*Vec-Cre:Gata2*$^{f/f}$
AGM	E10 (30-36sp)	336.0 ± 72.3	99.0 ± 33.5 (P < 0.05)	72.8 ± 21.5 (P < 0.01)
FL	E10 (30-36sp)	1,15.8 ± 35.0	83.0 ± 20.8	22.5 ± 22.2 (P < 0.05)
YS	E10 (30-36sp)	1,132.0 ± 67.2	388.8 ± 93.0(P < 0.01)	70.8 ± 16.0 (P < 0.01)
AGM	E11 (43-48sp)	340.0 ± 73.0	172.0 ± 28.9	22.5 ± 3.4 (P < 0.01)
FL	E11 (43-48sp)	3,015.0 ± 186.1	1,234.0 ± 156.1 (P < 0.01)	66.8 ± 12.6 (P < 0.01)
FL	E14	42,155.3 ± 3,813.5	21,320.8 ± 4,359.2 (P < 0.05)	4,068.1 ± 2,068.3 (P < 0.01)
		Gata2$^{f/+}$ or *Gata2*$^{f/f}$	*Vav-Cre:Gata2*$^{f/+}$	*Vav-Cre:Gata2*$^{f/f}$
AGM	E11 (41-49sp)	456.0 ± 73.1	305.6 ± 35.3	170.5 ± 54.3 (P < 0.05)
FL	E11 (41-49sp)	2,950.0 ± 244.3	2,754.2 ± 173.9	1,249.1 ± 231.4 (P < 0.01)
FL	E14	43,320.2 ± 5,359.6	13,048.1 ± 3,647.0 (P < 0.01)	7,495.1 ± 3,847.1 (P < 0.01)

with WT. The FL of E14 *Vec-Cre:Gata2*$^{f/f}$ embryos contained 3.4× fewer cells ($14.9 \pm 4.3 \times 10^6$ WT and $4.3 \pm 1.6 \times 10^6$ *Vec-Cre:Gata2*$^{f/f}$ [P = 0.06]). Although all CFU-C types were decreased in number, the CFU-M and CFU-G colonies formed from *Vec-Cre:Gata2*$^{f/f}$ tissues were smaller and more compact than WT colonies (unpublished data), indicating a role for *Gata2* after the generation of macrophage and granulocyte progenitors, most likely affecting their differentiation and proliferation. These data are consistent with defects found in macrophage and granulocyte progenitors in germline haploinsufficient *Gata2*$^{+/-}$ adult BM (Rodrigues et al., 2008). The growth, size, and differentiation of the other hematopoietic colony types in *Vec-Cre:Gata2*$^{f/f}$ tissues was unaffected.

To test whether loss of *Gata2* in *VE-Cadherin*–expressing cells affects HSC generation, in vivo transplantation experiments were performed. E11 AGM cells were injected into irradiated adult recipients and assayed at 1, 2, 3, and 4 mo after transplantation. In contrast to WT controls, no recipients receiving *Vec-Cre:Gata2*$^{f/f}$ (or *Vec-Cre:Gata2*$^{f/+}$) cells showed donor repopulation in any hematopoietic tissue or lineage (Fig. 2 A), indicating that GATA2 is required in VE-Cadherin–expressing cells for the generation of functional adult repopulating HSCs.

Next, flow cytometric and immunohistochemical analyses were performed for the specific detection of HPC/HSCs in embryonic tissues and the vasculature; the presence of such phenotypic HSCs in the *Vec-Cre:Gata2*$^{f/f}$ embryos would indicate that GATA2 is essential for HPC/HSC function but not for their generation. It has been shown previously in WT embryos that E14 FL HSCs are Lin$^-$Sca1$^+$cKit$^+$CD48$^-$CD150$^+$ (or LSK SLAM; Kim et al., 2006). Flow cytometric analysis of *Vec-Cre:Gata2*$^{f/f}$ E14 FL showed a severe reduction of LSK cells and a complete absence of viable LSK SLAM cells as compared with WT FL (Fig. 2 B; $n = 4$, $P < 0.05$). To visually examine whether vascular hematopoietic cluster cells were affected, embryo immunostaining was performed with anti-cKit antibody alone (Fig. 2 C) or with anti-cKit and anti-CD31 antibodies (Fig. 2 D). Imaging of an E10 *Gata2*$^{-/-}$ embryo (30–somite pairs [sp] stage) revealed only 3 cKit$^+$ cells in the aorta, as compared with 25 in a *Gata2*$^{+/-}$ and 213 in a WT aorta. At the peak period of hematopoietic cluster formation in the embryo (E10.5), 38 ± 18 cKit$^+$ cluster cells in the *Vec-Cre:Gata2*$^{f/f}$ aorta were observed, as compared with 634 ± 43 cKit$^+$ cells in the WT aorta (Fig. 2 D, top right and left; $n = 2$, $P < 0.01$). Similar decreases in cluster numbers were found in the umbilical artery (Fig. 2 D, middle). Interestingly, cluster size was reduced to 1 or 2 cells, with the remaining cKit$^+$ cells of the *Vec-Cre:Gata2*$^{f/f}$ aorta showing a flat morphology, embedded within the endothelium (Fig. 2 D, bottom right). Also at E11, *Vec-Cre:Gata2*$^{f/f}$ AGMs contained far fewer cKit$^+$ cells in the dorsal aorta than WT embryos (58 in *Vec-Cre:Gata2*$^{f/f}$ versus 411 in WT; unpublished data). These data suggest that either hematopoietic cluster cells are not generated or they are apoptotic in *Gata2* cKO embryos.

To distinguish between these possibilities, we stained E11 AGMs with Annexin V to measure apoptotic status, and with anti-Flk1 and anti-cKit antibodies to identify endothelial cells and the cells at the base of emerging hematopoietic clusters (Yokomizo and Dzierzak, 2010; Fig. 2 E and Table 2). Flk1$^+$cKit$^+$ cells showed no increase in Annexin V staining compared with WT. Normal numbers of endothelial cells (cKit$^-$CD31$^+$ or cKit$^-$Flk1$^+$) were found in the AGM region of *Vec-Cre:Gata2*$^{f/f}$ embryos (not depicted and Table 2). Because we do not find a difference in viability of Flk1$^+$cKit$^+$ cells in *Vec-Cre:Gata2*$^{f/f}$ aortas as compared with WT, we conclude that GATA2 is essential in the EHT for the generation of hematopoietic cluster cells and HSCs.

The fact that Vav is expressed in HSCs and their progeny (Ogilvy et al., 1999), but not in endothelial cells, allowed Chen et al. (2009) to demonstrate that the RUNX1 transcription factor is required during, but not after, the generation of HSCs. The effects of *Gata2* deletion on HPCs and HSCs after their generation were tested using this *Vav-Cre* model (Stadtfeld and Graf, 2005). In contrast to *Vec-Cre:Gata2*$^{f/f}$ embryos, *Vav-Cre:Gata2*$^{f/f}$ embryos showed no FL anemia at E14 (Fig. 1 A;

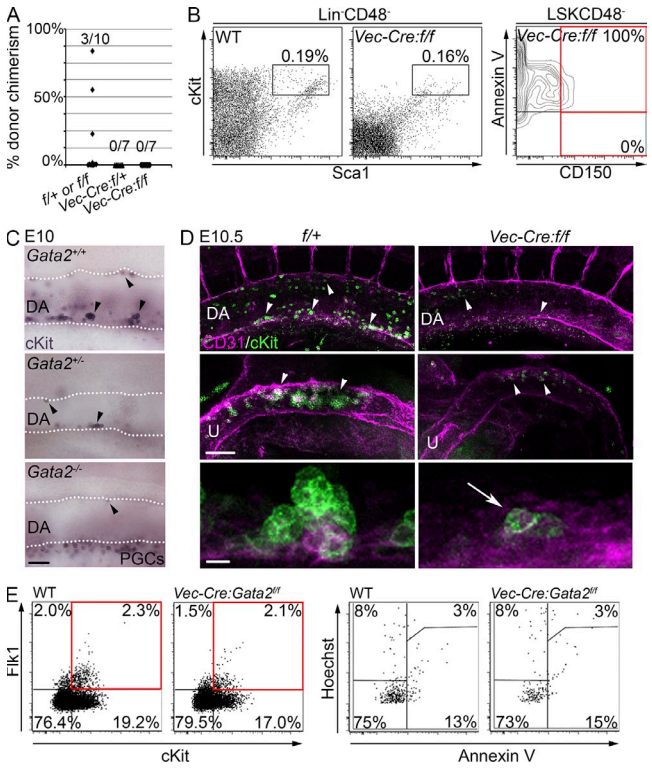


Figure 2. GATA2 is required in VE-Cadherin–expressing cells for the generation of HSCs and vascular hematopoietic cluster cells. (A) Graph showing the percentage of peripheral blood donor cell chimerism in adult recipients at 4 mo after transplantation. Recipients were injected with AGM cells from WT (1–2 ee; 3 of 10 recipients repopulated, $n = 3$), *Vec-Cre:Gata2*$^{f/+}$ (1–2 ee; 0 of 7 recipients repopulated), or *Vec-Cre:Gata2*$^{f/f}$ (1–3 ee; 0 of 7 recipients repopulated) E11 embryos. Reconstitution kinetics showed similar outcomes at 1, 2, 3, and 4 mo after transplantation. (B) FACS analysis of Lin$^-$CD48$^-$ gated, cKit- and Sca1-stained E14 FL cells from WT and *Vec-Cre:Gata2*$^{f/f}$ embryos. Percentage of cKit$^+$Sca1$^+$ cells within the Lin$^-$CD48$^-$ population is shown for WT (0.019 ± 0.01%) and *Vec-Cre:Gata2*$^{f/f}$ (0.16 ± 0.04%, $P < 0.05$). The right panel shows FACS analysis of LSKCD48$^-$ gated, CD150- and Annexin V–stained *Vec-Cre:Gata2*$^{f/f}$ E14 FLs. Phenotypic HSCs are indicated within the red quadrants. 100% of LSK SLAM cells are Annexin V–positive ($n = 2$, number of embryos analyzed = 7 WT, 5 *Vec-Cre:Gata2*$^{f/f}$; significance determined by Student's *t* test). (For WT, see Fig. 3 E.) (C) Immunostaining of E10 (30/31 sp) embryos showing cKit$^+$ hematopoietic cluster cells in germline *Gata2*$^{+/+}$, *Gata2*$^{+/-}$, *and Gata2*$^{-/-}$ aortae. Arrowheads indicate some of the hematopoietic cluster cells along the aortic wall. Bar, 50 µm. DA = dorsal aorta, PGC = primordial germ cells ($n = 1$). (D) Whole mount immunostaining of E10.5 (36/37 sp) *Gata2*$^{f/+}$ and *Vec-Cre:Gata2*$^{f/f}$ embryos showing CD31$^+$ vascular endothelial cells (magenta) and cKit$^+$ hematopoietic cluster cells (green). The top panels show the dorsal aorta (DA), middle panels the umbilical artery (U), and bottom panels a high-magnification image of a cKit$^+$ cluster and aortic endothelium. Arrowheads indicate a few cKit$^+$ hematopoietic cluster cells and arrow indicates a flat cKit$^+$ cell embedded in the endothelium ($n = 2$, number of embryos analyzed is 2 WT and 2 *Vec-Cre:Gata2*$^{f/f}$; significance determined by Student's *t* test). Bars: (top and middle) 100 µm; (bottom) 10 µm. (E) FACS analysis showing Flk1- and cKit-stained WT and *Vec-Cre:Gata2*$^{f/f}$ E11 AGMs. Flk1$^+$ cKit$^+$ cells (red quadrant) and gating strategy for the viability (right) are shown. Hoechst$^-$Annexin V$^-$ cells are viable, Hoechst$^-$Annexin V$^+$ are early apoptotic cells, Hoechst$^+$Annexin V$^+$ are late apoptotic cells, and Hoechst$^+$Annexin V$^-$ cells are dead ($n = 3$; number of embryos analyzed = 7 WT and 6 *Vec-Cre:Gata2*$^{f/f}$; significance determined by Student's *t* test).

15.8 ± 3.5 × 10^6 WT and 13.2 ± 1.6 × 10^6 [$P = 0.4$]) and survived past E16. Only 2 *Vav-Cre:Gata2*$^{f/f}$ offspring were born out of 36 pups in 6 litters (6%). Because both pups showed an incomplete deletion of the floxed *Gata2* alleles (unpublished data), these results suggest that *Vav-Cre*–mediated deletion of *Gata2* affects HPC/HSCs after their generation.

E11 *Vav-Cre:Gata2*$^{f/f}$ AGMs and FLs were significantly decreased in CFU-C numbers (2.7- and 2.4-fold), as compared with WT tissues (Fig. 3 A and Table 1). The observed decreases were not as severe as those in the *Vec-Cre* cKO tissues, and this is underscored by FACS analysis. The number of phenotypically defined CD31$^+$cKit^{high}CD41^{int} HPC/HSCs in E11 *Vav-Cre:Gata2*$^{f/f}$ AGMs was not significantly altered from WT and corresponded to previously published numbers (Sánchez et al., 1996; Yokomizo and Dzierzak, 2010; Robin et al., 2011; Boisset et al., 2013). Moreover, cluster numbers in *Vav-Cre* cKO vessels were normal at E11, as compared with WT (unpublished data). Interestingly, whereas no significant differences in total cell numbers of E14 FLs were found, both E14 *Vav-Cre:Gata2*$^{f/+}$ and *Vav-Cre:Gata2*$^{f/f}$ FLs contained significantly fewer CFU-Cs than WT FLs (3.3- and 5.8-fold reduction, respectively; Fig. 3 A and Table 1). Indeed, flow cytometric analysis of *Vav-Cre:Gata2*$^{f/f}$ E14 FLs showed that phenotypic HSCs, defined by LSK SLAM staining, were reduced by a factor of two, as compared with WT (Fig. 3 B and Table 3). In most *Vav-Cre:Gata2*$^{f/f}$ E14 FLs, Lin$^-$cKit^{high} cell numbers are reduced (not depicted) and LSK SLAM cells were almost completely absent (Table 3). Additionally, whereas 58.4 ± 8.4% of WT LSK SLAM cells were capable of forming a CFU-C (7.9 ± 2.1% CFU-GEMM), only 31.6 ± 6.7% of *Vav-Cre:Gata2*$^{f/+}$ LSK SLAM cells were capable of forming CFU-C (2.3 ± 1.5% CFU-GEMM), and none of the *Vav-Cre:Gata2*$^{f/f}$ LSK SLAM cells resulted in CFU-Cs ($P < 0.01$; Fig. 3 C). These results indicate that GATA2 is essential for the hematopoietic progenitor activity of phenotypic HSCs. Furthermore, E14 *Vav-Cre:Gata2*$^{f/f}$ FL cells did not long-term reconstitute the hematopoietic system of irradiated adult recipients (Fig. 3 D, $P < 0.01$). Two recipients were found with *Vav-Cre:Gata2*$^{f/f}$ donor cell chimerism, but these donor cells retained at least one unrecombined floxed allele (as detected in the recipient peripheral blood and the donor FL cells; unpublished data). Thus, GATA2 is essential in HSCs after their formation.

Transplantation and CFU assays do not provide insight into the processes affected by deletion of *Gata2* after formation of HSCs because both require the cells to extensively proliferate and to be viable for 10 d to 4 mo. Therefore, we postulated that the reduction in the number and function of *Vav-Cre:Gata2*$^{f/f}$ HPC/HSCs in the FL may be due to defective cell amplification and/or lack of cell maintenance/survival. To check whether the loss of *Gata2* affects the cell cycle status of HPC/HSCs, Ki67 staining was performed on E14 FL (total and Lin$^-$cKit^{high}) cells. As expected, more cells in the S/G_2/M

Table 2. Apoptosis analysis of *Gata2* E11 AGM and FL cells

Genotype/apoptotic status	AGM			FL	
	Flk1⁺cKit⁻ endothelial cells	Flk1⁺cKit⁺ cluster base cells	Flk1⁻cKit⁺ hematopoietic cells	CD31⁺cKit⁺ HPCs/HSCs	CD31⁻cKit⁺ hematopoietic cells
	%	%	%	%	%
Gata2$^{f/+}$ or *Gata2*$^{f/f}$ (WT)	Total: 2.0 ± 0.6	Total: 2.3 ± 0.6	Total: 19.2 ± 3.4	Total: 60.2 ± 4.2	Total: 32.4 ± 4.2
Viable	76	76	77	75	2
Early apoptotic	9	13	5	5	1
Late apoptotic	1	3	6	7	21
Dead	14	8	12	11	76
Vec-Cre:Gata2$^{f/f}$	Total: 1.5 ± 0.3	Total: 2.1 ± 0.7	Total: 17.0 ± 2.3	Total: 22.1 ± 0.4**	Total: 68.6 ± 1.4**
Viable	76	74	74	30**	1
Early apoptotic	8	15	6	15	1
Late apoptotic	2	3	7	20	18**
Dead	14	8	12	35	80**

Results of flow cytometric analyses with various cell subset specific markers (Flk1, cKit, and CD31) show the percentage of cells within total populations of E11 AGM or E11 FL (Total). Apoptotic status of cells within these gated populations (cell subsets) was tested by Annexin V and Hoechst staining and is shown below the percentage of the total population. Viable cells are Annexin V⁻Hoechst⁻, early apoptotic are Annexin V⁺Hoechst⁻, late apoptotic are Annexin V⁺Hoechst⁺, and dead cells are Annexin V⁻Hoechst⁺. $n = 3$. Numbers of embryos analysed = 7 WT and 7 *Vec-Cre:Gata2*$^{f/f}$. Significance was determined by Student's *t* test. **, $P < 0.01$.

phase were found in the Lin⁻cKit^{high} enriched FL HPC/HSC population than in the total FL population. However, no alteration in cell cycle could be detected between WT and *Vav-Cre:Gata2*$^{f/f}$ FL cells (Table 3). We also tested the survival of phenotypic HSCs in the E14 FL by Annexin V staining. A significant threefold increase in apoptotic cells was found in *Vav-Cre:Gata2*$^{f/f}$ E14 FL LSK SLAM cells as compared with WT LSK SLAM FL ($P < 0.05$), demonstrating that GATA2 is required in HSCs during the FL stage for their survival (Fig. 3 E). This result is underscored by the finding of Linnemann et al.

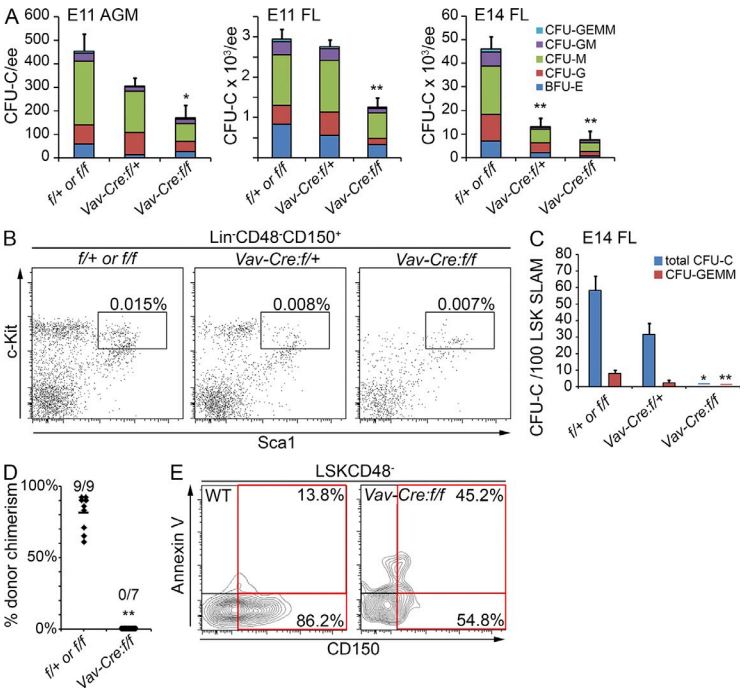


Figure 3. *Gata2* is required in HPCs and in HSCs for survival after their generation. (A) Graphs showing the CFU-C numbers per WT, *Vav-Cre:Gata2*$^{f/+}$, or *Vav-Cre:Gata2*$^{f/f}$ embryo equivalent (ee) of AGM and FL cells. Error bars indicate mean ± SEM. *, $P < 0.05$; **, $P < 0.01$. The number of E11 AGMs and FLs analyzed is 11 WT, 12 *f/+*, and 6 *f/f*; $n = 4$. For E11 AGM, $P < 0.05$, and E11 FL, $P < 0.01$. For E14 FL ($P < 0.01$; $n = 3$), the number of embryos analyzed = 10 WT, 5 *f/+*, and 20 *f/f*. Significance was determined by Student's *t* test. (B) FACS analysis of WT, *Vav-Cre:Gata2*$^{f/+}$, and *Vav-Cre:Gata2*$^{f/f}$ E14 FL for LSK SLAM markers. Lin⁻CD48⁻CD150⁺ gated cells were analyzed for cKit and Sca1. Percentages of LSK SLAM cells (gated region) are shown. (C) Graph of the percentage of total CFU-C and CFU-GEMM per 100 LSK SLAM cells of each genotype ($n = 3$). The number of embryos analyzed = 3 WT, 3 *f/+*, and 3 *f/f*. Error bars are mean ± SEM of total colony numbers and of CFU-GEMM only. *, $P < 0.05$; **, $P < 0.01$; significance determined by Student's *t* test. (D) Graph showing the percentage of peripheral blood donor cell chimerism in adult recipients injected with 10^5 WT ($n = 2$; 9 of 9 recipients repopulated) or 10^5 *Vav-Cre:Gata2*$^{f/f}$ E14 FL cells (0 of 7 recipients repopulated, $P < 0.01$). Diamonds represent individual recipients, with the mean represented by a horizontal bar. Significance was determined by Student's *t* test. (E) Annexin V FACS analysis of WT and *Vav-Cre:Gata2*$^{f/f}$ E16 FL LSK SLAM cells. LSKCD48⁻ gated cells were analyzed for Annexin V and CD150 ($n = 2$). The number of embryos analyzed = 3 WT and 6 *f/f*. Phenotypic HSCs (red quadrant) and the percentages of apoptotic and viable cells are indicated. (Not depicted: for E14 FL $n = 3$; the number of embryos analyzed = 8 WT, 10 *f/+*, and 9 *f/f*.)

Table 3. Analysis of Gata2 conditional knockout E14 FL

E14 FL cell subset/cell type	*Gata2$^{f/+}$ or Gata2$^{f/f}$*	*Vec-Cre:Gata2$^{f/+}$*	*Vec-Cre:Gata2$^{f/f}$*
	%	%	%
Differentiated cells			
T cells	5.1 ± 1.6	4.5 ± 1.5	3.3 ± 1.1
B cells	6.5 ± 1.4	7.5 ± 2.5	5.9 ± 1.5
Erythroblasts	61.7 ± 3.0	62.6 ± 1.3	61.9 ± 3.1
Erythrocytes	28.7 ± 3.1	30.3 ± 1.4	33.2 ± 3.3
LSK-SLAM	0.0150 ± 0.0021	0.0080 ± 0.0011 (P < 0.05)	0.0066 ± 0.0017 (P < 0.01)
Total			
G_0	0.6 ± 0.2	0.6 ± 0.2	0.6 ± 0.3
G_1	44.3 ± 1.8	50.8 ± 2.7	44.8 ± 2.0
$S/G_2/M$	54.5 ± 3.2	48.1 ± 2.6	54.1 ± 2.1
Lin$^-$cKit^{high}			
G_0	0.1 ± 0.03	0.2 ± 0.1	0.2 ± 0.05
G_1	31.5 ± 1.12	36.4 ± 3.9	35.5 ± 2.4
$S/G_2/M$	67.9 ± 1.10	63.1 ± 3.8	63.9 ± 2.3
Annexin$^+$	13.8 ± 5.2	19.1 ± 5.1	45.2 ± 16.0 (P < 0.05)

Values for differentiated cells and LSK-SLAM are percentage of viable cells. Values for Total and Lin^{2}cKit^{high} are percentage of single cells. Values for Annexin$^+$ are percentage of viable cells including pre-apoptotic. For the analysis of differentiated cells, $n = 2$ and number of embryos analyzed = 7, 4, and 8, respectively by genotype; LSK-SLAM, $n = 5$ and number of embryos analyzed = 12, 10, and 13, respectively; total and Lin$^-$cKit^{high}, $n = 2$ and number of embryos analyzed = 7, 3, and 6; and LSK-SLAM Annexin$^+$, $n = 3$ and number of embryos analyzed = 8, 10, and 9. Significance was determined by Student's *t* test. ND = not done.

(2011) that the anti-apoptotic gene *BirC3* was down-regulated upon knockdown of *Gata2* in human endothelial cells.

Our results show for the first time that GATA2 is required for HSC generation in the AGM and continues to be required for HSC survival. GATA2 is thought to act in combination with other transcription factors, such as RUNX1, to promote/enhance transcription of genes relevant to hematopoietic cell development and growth (Wilson et al., 2010; van Riel et al., 2012). Because we used the *Vec-Cre* and *Vav-Cre* cKO approach used previously for *Runx1* deletion (Chen et al., 2009), the requirements for these two transcription factors can be directly compared. Whereas the RUNX1 transcription factor functions only during the EHT and not thereafter (Chen et al., 2009), our study shows that GATA2 functions in HSCs both during and after the EHT stage. At the EHT stage, we observed a severe reduction but not a complete absence of CFU-Cs in *Vec-Cre:Gata2$^{f/f}$* AGMs. 31% of these HPCs have both alleles recombined, indicating that some HPCs are GATA2 independent. This is in contrast to *Vec-Cre:Runx1$^{f/f}$* AGMs, in which no HPCs with both recombined alleles were found (Chen et al., 2009). In *Runx1* cKO embryos, no cKit$^+$ aortic hematopoietic clusters or cells are found, unlike the *Gata2* cKO. The difference between the phenotypes of the *Gata2* and the *Runx1 Vec-Cre* cKOs suggests that there are distinct subsets of HPCs. Collectively, our results are supportive of the concurrent function of GATA2 and RUNX1 in EHT, and they reveal important new information that GATA2 and RUNX1 also act separately to provide unique functions in HPCs and HSCs at different developmental stages.

MATERIALS AND METHODS

Conditional deletion by Vec-Cre or Vav-Cre. WT, *Vec-Cre:Gata2$^{f/+}$*, and *Vec-Cre:Gata2$^{f/f}$* conceptuses were generated by crossing *Vec-Cre:Gata2$^{f/+}$* males or females with *Gata2$^{f/f}$* females or males. Similar breeding strategies were used for Vav-Cre deletion. WT and germline *Gata2$^{+/-}$* and *Gata2$^{-/-}$* conceptuses were generated by crossing *Gata2$^{+/-}$* males and females. Genotyping was performed by PCR using the primers in Table S1. All experiments have been conducted according to Dutch law and have been approved by the animal experiments committee (Stichting DEC consult).

Microscopic and histological analyses. Conceptuses were suspended in phosphate-buffered saline with 10% FCS and visualized with a stereomicroscope. To analyze intra-aortic clusters, embryos were fixed in 2% PFA and stained with anti-cKit or anti-cKit and anti-Pecam1 (CD31) antibodies, and imaged in a 1:2 mix of benzyl alcohol and benzyl benzoate. Samples were analyzed with a laser-scanning confocal microscope (SP5; Leica), as previously described (Yokomizo et al., 2012), or analyzed with a brightfield microscope. Three-dimensional reconstructions were generated from Z-stacks (50–150 optical sections) using LasAF software.

Hematopoietic assays. Methylcellulose colony-forming assays were performed as described previously (Medvinsky et al., 2008). Embryo equivalents (ee; 1/10 and 9/10 ee of E10 and E11 AGMs and E10 FLs; 1/30, 1/10, and 26/30 ee of E11 FLs; and 1/1,000 and 1/10 ee or 10^5 cells of E14 FLs) were seeded per 1 ml M3434 medium and colonies counted after 10–12 d. Colonies were individually isolated from methylcellulose medium and genotyped to determine clonal recombination efficiency. AGMs and FL cells were transplanted as described previously (Medvinsky et al., 2008). After isolation of Ly5.2 cKO and WT conceptuses, YS material was genotyped using a fast genotyping kit (KAPA) and 1–3 AGMs of the same genotype were pooled or 10^5 FL cells were injected into lethally irradiated Ly5.1 recipients. Donor chimerism was determined through flow cytometric analysis of anti-Ly5.1 and anti-Ly5.2 antibody stained peripheral blood (PB) cells of recipients at 1, 2, 3, and 4 mo after transplantation. Multilineage organ chimerism analysis was similarly performed at 4 mo after transplantation. Transplanted recipients were scored positive if PB donor chimerism was at least 5% at 4 mo after transplantation

and multi-organ and myeloid and lymphoid reconstitution was at least 1% at the time of sacrifice (4 mo after transplantation).

FACS analyses. Single cell suspensions were made by 45-min collagenase type I treatment of dissected AGM + vitelline artery + umbilical artery and subsequent dissociation by pipetting or direct dissociation of FLs by pipetting. LSK SLAM analyses were performed according to Kim et al. (2006). Lineage cocktail included CD3e, Ly6C, CD45R, Ter119, and CD48. Annexin V staining was performed simultaneously with LSK SLAM staining in Annexin V binding buffer (BD) in PBS containing 10% FCS. Ki67/cell cycle analysis; Lin^-CD48^- sorted cells were stained with anti-Sca1, anti-cKit, and anti-CD150 antibodies, fixed with 2% paraformaldehyde, and permeabilized with Triton X-100. Ki67 staining was combined with Hoechst staining (BD). Other antibodies include anti-CD31 and anti-FLK1 (BD) and anti-CD41 (Santa Cruz Biotechnology, Inc.). Fluorochrome conjugates used were: FITC, phycoerytherin, phycoerytherin-Cy7, allophycocyanin-, or allophycocyanin–efluor 780. Stained cells were analyzed on a FACSAria III or SORP-FACSAria II flow cytometer (BD). Dead cells were excluded with Hoechst.

Statistical analysis. Statistical significance throughout this study was determined by Student's *t* test.

Online supplemental material. Table S1 shows genotyping primers and conditions. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20130751/DC1.

The authors thank Drs. Catherine Robin, Derk ten Berge, Michael Chen, Gerben Schaaf, and all laboratory members for lively discussions and the Erasmus Dier (Animal) Center for animal care.

This work was supported by National Institutes of Health RO37DKD54077, FES Netherlands Institute for Regenerative Medicine, Erasmus Grant, Landsteiner Society for Blood Research 1109, and ZonMW 911-09-036.

The authors state that there are no competing financial interests.

Submitted: 11 April 2013
Accepted: 12 November 2013

REFERENCES

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

Boisset, J.C., T. Clapes, R. Van Der Linden, E. Dzierzak, and C. Robin. 2013. Integrin αIIb (CD41) plays a role in the maintenance of hematopoietic stem cell activity in the mouse embryonic aorta. *Biol. Open*. 2:525–532. http://dx.doi.org/10.1242/bio.20133715

Charles, M.A., T.L. Saunders, W.M. Wood, K. Owens, A.F. Parlow, S.A. Camper, E.C. Ridgway, and D.F. Gordon. 2006. Pituitary-specific Gata2 knockout: effects on gonadotrope and thyrotrope function. *Mol. Endocrinol.* 20:1366–1377. http://dx.doi.org/10.1210/me.2005-0378

Chen, M.J., T. Yokomizo, B.M. Zeigler, E. Dzierzak, and N.A. Speck. 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature*. 457:887–891. http://dx.doi.org/10.1038/nature07619

de Bruijn, M.F., N.A. Speck, M.C. Peeters, and E. Dzierzak. 2000. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J.* 19:2465–2474. http://dx.doi.org/10.1093/emboj/19.11.2465

de Bruijn, M.F., X. Ma, C. Robin, K. Ottersbach, M.J. Sanchez, and E. Dzierzak. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity*. 16:673–683. http://dx.doi.org/10.1016/S1074-7613(02)00313-8

Dzierzak, E., and N.A. Speck. 2008. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat. Immunol.* 9:129–136. http://dx.doi.org/10.1038/ni1560

Johnson, K.D., A.P. Hsu, M.J. Ryu, J. Wang, X. Gao, M.E. Boyer, Y. Liu, Y. Lee, K.R. Calvo, S. Keles, et al. 2012. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J. Clin. Invest.* 122:3692–3704. http://dx.doi.org/10.1172/JCI61623

Kim, I., S. He, O.H. Yilmaz, M.J. Kiel, and S.J. Morrison. 2006. Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. *Blood*. 108:737–744. http://dx.doi.org/10.1182/blood-2005-10-4135

Lim, K.C., T. Hosoya, W. Brandt, C.J. Ku, S. Hosoya-Ohmura, S.A. Camper, M. Yamamoto, and J.D. Engel. 2012. Conditional Gata2 inactivation results in HSC loss and lymphatic mispatterning. *J. Clin. Invest.* 122:3705–3717. http://dx.doi.org/10.1172/JCI61619

Ling, K.W., K. Ottersbach, J.P. van Hamburg, A. Oziemlak, F.Y. Tsai, S.H. Orkin, R. Ploemacher, R.W. Hendriks, and E. Dzierzak. 2004. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J. Exp. Med.* 200:871–882. http://dx.doi.org/10.1084/jem.20031556

Linnemann, A.K., H. O'Geen, S. Keles, P.J. Farnham, and E.H. Bresnick. 2011. Genetic framework for GATA factor function in vascular biology. *Proc. Natl. Acad. Sci. USA*. 108:13641–13646. http://dx.doi.org/10.1073/pnas.1108440108

Medvinsky, A., S. Taoudi, S. Mendes, and E. Dzierzak. 2008. Analysis and manipulation of hematopoietic progenitor and stem cells from murine embryonic tissues. *Curr. Protoc. Stem Cell Biol.* Chapter 2:6.

Minegishi, N., J. Ohta, H. Yamagiwa, N. Suzuki, S. Kawauchi, Y. Zhou, S. Takahashi, N. Hayashi, J.D. Engel, and M. Yamamoto. 1999. The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood*. 93:4196–4207.

Nardelli, J., D. Thiesson, Y. Fujiwara, F.Y. Tsai, and S.H. Orkin. 1999. Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. *Dev. Biol.* 210:305–321. http://dx.doi.org/10.1006/dbio.1999.9278

Ng, Y.K., K.M. George, J.D. Engel, and D.I. Linzer. 1994. GATA factor activity is required for the trophoblast-specific transcriptional regulation of the mouse placental lactogen I gene. *Development*. 120:3257–3266.

North, T.E., M.F. de Bruijn, T. Stacy, L. Talebian, E. Lind, C. Robin, M. Binder, E. Dzierzak, and N.A. Speck. 2002. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity*. 16:661–672. http://dx.doi.org/10.1016/S1074-7613(02)00296-0

Ogilvy, S., D. Metcalf, L. Gibson, M.L. Bath, A.W. Harris, and J.M. Adams. 1999. Promoter elements of vav drive transgene expression in vivo throughout the hematopoietic compartment. *Blood*. 94:1855–1863.

Orlic, D., S. Anderson, L.G. Biesecker, B.P. Sorrentino, and D.M. Bodine. 1995. Pluripotent hematopoietic stem cells contain high levels of mRNA for c-kit, GATA-2, p45 NF-E2, and c-myb and low levels or no mRNA for c-fms and the receptors for granulocyte colony-stimulating factor and interleukins 5 and 7. *Proc. Natl. Acad. Sci. USA*. 92:4601–4605. http://dx.doi.org/10.1073/pnas.92.10.4601

Robert-Moreno, A., L. Espinosa, J.L. de la Pompa, and A. Bigas. 2005. RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development*. 132:1117–1126. http://dx.doi.org/10.1242/dev.01660

Robin, C., K. Ottersbach, J.C. Boisset, A. Oziemlak, and E. Dzierzak. 2011. CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells. *Blood*. 117:5088–5091. http://dx.doi.org/10.1182/blood-2011-01-329516

Rodrigues, N.P., A.S. Boyd, C. Fugazza, G.E. May, Y. Guo, A.J. Tipping, D.T. Scadden, P. Vyas, and T. Enver. 2008. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood*. 112:4862–4873. http://dx.doi.org/10.1182/blood-2008-01-136564

Sánchez, M.J., A. Holmes, C. Miles, and E. Dzierzak. 1996. Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. *Immunity*. 5:513–525. http://dx.doi.org/10.1016/S1074-7613(00)80267-8

Stadtfeld, M., and T. Graf. 2005. Assessing the role of hematopoietic plasticity for endothelial and hepatocyte development by non-invasive lineage tracing. *Development*. 132:203–213. http://dx.doi.org/10.1242/dev.01558

Taoudi, S., and A. Medvinsky. 2007. Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. *Proc. Natl. Acad. Sci. USA*. 104:9399–9403. http://dx.doi.org/10.1073/pnas.0700984104

Tsai, F.Y., and S.H. Orkin. 1997. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood*. 89:3636–3643.

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

van Riel, B., T. Pakozdi, R. Brouwer, R. Monteiro, K. Tuladhar, V. Franke, J.C. Bryne, R. Jorna, E.J. Rijkers, W. van Ijcken, et al. 2012. A novel complex, RUNX1-MYEF2, represses hematopoietic genes in erythroid cells. *Mol. Cell. Biol.* 32:3814–3822. http://dx.doi.org/10.1128/MCB.05938-11

Wilson, N.K., S.D. Foster, X. Wang, K. Knezevic, J. Schütte, P. Kaimakis, P.M. Chilarska, S. Kinston, W.H. Ouwehand, E. Dzierzak, et al. 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell*. 7:532–544. http://dx.doi.org/10.1016/j.stem.2010.07.016

Yokomizo, T., and E. Dzierzak. 2010. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development*. 137:3651–3661. http://dx.doi.org/10.1242/dev.051094

Yokomizo, T., T. Yamada-Inagawa, A.D. Yzaguirre, M.J. Chen, N.A. Speck, and E. Dzierzak. 2012. Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nat. Protoc.* 7:421–431. http://dx.doi.org/10.1038/nprot.2011.441

Zovein, A.C., J.J. Hofmann, M. Lynch, W.J. French, K.A. Turlo, Y. Yang, M.S. Becker, L. Zanetta, E. Dejana, J.C. Gasson, et al. 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell*. 3:625–636. http://dx.doi.org/10.1016/j.stem.2008.09.018

SUPPLEMENTAL MATERIAL

de Pater et al., http://www.jem.org/cgi/content/full/jem.20130751/DC1

Table S1. Genotyping primers and conditions

Primer	Primer sequence (5′–3′)	Tm	Size PCR product	Strain:
Vec-F-219	CCCAGGCTGACCAAGCTGAG	58	300 bp	VE-Cadherin Cre
Vav-F23	GGCGACAGTTACAGTCACAGAAGAGG	58	500 bp	Vav-Cre
Cre-R107	GCCTGGCGATCCCTGAACATG	58		VE-Cadherin Cre, Vav-Cre
Gata2 fl F	TCCGTGGGACCTGTTTCCTTAC	58	fl = 844bp, WT = 717bp,	Gata2 fl
Gata2 fl R	GCCTGCGTCCTCCAACACCTCTAA	58	delta = 181bp	Gata2 fl
mGATA 39	GGAACGCCAACGGGGAC	60	mutant = 900bp, WT = 600bp	Gata2 KO
mGATA 208	GCTGGACATCTTCCGATTCCGGGT	60		Gata2 KO
Neo 503	GATCTCCTGTCATCTCACCTTGCT	60		Gata2 KO

HEMATOPOIESIS AND STEM CELLS

Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors

Polynikis Kaimakis,[1,*] Emma de Pater,[1,2,*] Christina Eich,[1,*] Parham Solaimani Kartalaei,[1] Mari-Liis Kauts,[1,3] Chris S. Vink,[1,3] Reinier van der Linden,[1] Martine Jaegle,[1] Tomomasa Yokomizo,[1] Dies Meijer,[1,4] and Elaine Dzierzak[1,3]

[1]Erasmus Medical Center Stem Cell Institute, Departments of Cell Biology and Genetics, and [2]Department of Hematology, Erasmus Medical Center, Rotterdam, The Netherlands; and [3]Centre for Inflammation Research, Queens Medical Research Institute, and [4]Centre for NeuroRegeneration, University of Edinburgh, Edinburgh, United Kingdom

Key Points

- A new Gata2 reporter indicates that all HSCs express Gata2 and corroborates findings that Gata2 is not required for generation of all HPCs.
- Isolatable non–Gata2-expressing HPCs show less potency and a distinct genetic program, thus having implications for reprogramming strategies.

The Gata2 transcription factor is a pivotal regulator of hematopoietic cell development and maintenance, highlighted by the fact that *Gata2* haploinsufficiency has been identified as the cause of some familial cases of acute myelogenous leukemia/myelodysplastic syndrome and in MonoMac syndrome. Genetic deletion in mice has shown that Gata2 is pivotal to the embryonic generation of hematopoietic stem cells (HSCs) and hematopoietic progenitor cells (HPCs). It functions in the embryo during endothelial cell to hematopoietic cell transition to affect hematopoietic cluster, HPC, and HSC formation. *Gata2* conditional deletion and overexpression studies show the importance of Gata2 levels in hematopoiesis, during all developmental stages. Although previous studies of cell populations phenotypically enriched in HPCs and HSCs show expression of *Gata2*, there has been no direct study of Gata2 expressing cells during normal hematopoiesis. In this study, we generate a *Gata2Venus* reporter mouse model with unperturbed Gata2 expression to examine the hematopoietic function and transcriptome of Gata2 expressing and nonexpressing cells. We show that all the HSCs are Gata2 expressing. However, not all HPCs in the aorta, vitelline and umbilical arteries, and fetal liver require or express Gata2. These Gata2-independent HPCs exhibit a different functional output and genetic program, including Ras and cyclic AMP response element-binding protein pathways and other Gata factors, compared with Gata2-dependent HPCs. Our results, indicating that Gata2 is of major importance in programming toward HSC fate but not in all cells with HPC fate, have implications for current reprogramming strategies. (*Blood*. 2016;127(11):1426-1437)

Introduction

Gata2 is one of the "heptad" transcription factors that acts on regulatory regions of hematopoietic genes.[1] It is upregulated in vivo in Ly6aGFP$^+$ cells undergoing endothelial-to-hematopoietic cell transition (EHT), a process by which definitive hematopoietic progenitors (HPCs) and hematopoietic stem cells (HSCs) are generated in the embryo.[2,3] As one of the major regulators of HPC and HSC generation, germline deficiency of *Gata2* results in embryonic lethality between embryonic day (E)10 and E10.5 and an anemic phenotype, with a decreased number of primitive and definitive HPCs in the yolk sac (YS) and in *Gata2*$^{-/-}$ embryonic stem (ES) cell hematopoietic differentiation cultures.[4-6] Chimeric embryo generation with *Gata2*$^{-/-}$ ES cells revealed defective production of all hematopoietic lineages.[5] The E10.5 lethality of *Gata2*$^{-/-}$ embryos precludes the study of HSC generation in the aorta-gonad-mesonephros (AGM) region, the first site of de novo HSC production. *Gata2*$^{+/-}$ embryos contain greatly reduced number of HSCs in the AGM region.[7,8] Gata2 haploinsufficiency perturbs adult HSC homeostasis in mice[9] and, in humans, leads to MonoMac syndrome,[10] which is associated with sporadic myelodysplasia and myeloid leukemia. Also, rearrangement of the remote *Gata2* enhancer drives acute myeloid leukemogenesis by activating *Evi1* expression.[11,12] Overexpression studies also reveal that levels of Gata2 expression are important for its hematopoietic function.[13-15] In situ hybridization studies localize *Gata2* expression to aortic endothelial cells, intra-aortic hematopoietic cluster cells, placenta (PL), and fetal liver (FL) in the midgestation mouse.[16-18] Conditional knockout of *Gata2* or *Gata2* regulatory elements in vascular endothelial cells indicates that Gata2 is essential for hematopoietic cluster formation and HSC generation.[7,19,20] Gata2 plays a role in the emergence of cKit-expressing hematopoietic cells from the endothelium.[7] Later, as shown in *VavCre* conditional knockout mice, *Gata2* is essential for HSC maintenance,[7] thus demonstrating a role for Gata2 as previously recognized in bone marrow LSK HSCs.[21]

Submitted October 15, 2015; accepted January 20, 2016. Prepublished online as *Blood* First Edition paper, February 1, 2016; DOI 10.1182/blood-2015-10-673749.

*P.K., E.d.P., and C.E. contributed equally to this work.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

To date, the correlation between Gata2 and hematopoietic cell generation in the embryo has been made in the absence of prospective isolation of viable Gata2-expressing cells.[16] Although some hematopoietic cells remain in the embryo in the absence of Gata2,[5-8] the identity of these cells is unknown. In this study, to further understand the requirement for Gata2 in normal hematopoietic development, we create and use a mouse model in which a fluorescent reporter for Gata2 (*IRES-Venus* knock-in gene) does not affect the normal level or function of Gata2. We demonstrate that all long-term repopulating HSCs and a large percentage of HPCs in the midgestation mouse embryo are Venus$^+$. We isolate and characterize a Venus$^-$ HPC population that corresponds to the HPCs found in *Gata2*-null embryos. Gata2-independent hematopoietic progenitors are functionally less complex and do not follow the same genetic program as Gata2-dependent HPCs.

Materials and methods

Gata2Venus ES cells and mice

Generation of the Gata2-Venus mouse model is described in the supplemental Methods, available on the *Blood* Web site. In short, an *IRES-Venus* fragment and a *loxP-PGK-Puro-loxP* fragment were inserted in the *Gata2* 3′ untranslated region (UTR). IB10 ES cells were transfected and puromycin selected, and 360 clones were polymerase chain reaction (PCR) screened for *Gata2Venus* (right arm junction, 2292 bp). Correct integration was verified by Southern blot (left arm) for 2 clones with normal karyotype. Founders were identified by *Venus* PCR. First-generation *G2V* offspring were crossed with *CAG-Cre* mice[22] and backcrossed (>10 generations) with C57BL/6.

Mice and embryo production

Gata2$^{+/-}$ mice,[5] Ly5.1 (6-8 weeks) and C57BL/6 mice were obtained/maintained (Harlan or locally) and genotyped by PCR (supplemental Methods). Day of plug discovery is E0. Embryos were staged by somite pair (sp): E9.5 = 16 to 28 sp, E10 = 28 to 40 sp, early E10 = 28 to 34 sp, E10.5 = 35 to 40 sp, and E11 = 40 to 50 sp. All mouse experimentation was performed under the UK Animals Scientific Procedures Act 1986 Project License 70/8076 and NL Ethics Committee approval and performed in compliance with Standards for Care and Use of Laboratory Animals.

Immunostaining

Whole-mount conceptuses were stained and imaged[23]; cryosections and cells for flow cytometry were stained[8] using anti-CD34-biotin (1:50; BD), anti-Gata3 (1:10 KT122, 111207H09; Absea), anti-Gata4 (1:50 H-112, sc-9053; SantaCruz), and anti-green fluorescent protein antibodies. For flow cytometry,[8] cells were stained with anti-CD31 (390; BD), anti-CD34 (RAM34; BD), anti-cKit (2B8; BD), anti-CD41 (MWReg3; SantaCruz), anti-Sca1 (D7; Ebiosciences), and anti-CD16/32 (2.4G2; BD) antibodies and Hoechst 33258 (BD) and analyzed (FACSARIAIII; SORP).

Hematopoietic assays

Venus-sorted E9 to E11 AGM, vitelline+umbilical arteries (VA+UA), PL, and YS and E10 FL or *Gata2*-deficient E9 to E10 AGM, VA+UA, and PL were seeded in 3.6 mL methylcellulose (1 mL per dish; M3434; Stem Cell Tech) for 10 to 12 days.[24] Colonies were counted, isolated, and washed, and Venus expression was examined (FACSARIAIII; Fortessa). Sorted *G2V* E11 AGM (Ly5.1/Ly5.2) cells were transplanted[24] into 9.5-Gy irradiated (C57BL/6X129) F1 recipients (Ly5.2/Ly5.2) together with 2×10^5 spleen cells from the recipient strain. Peripheral blood (PB) donor chimerism was determined by *Venus* PCR and/or fluorescence-activated cell sorter (FACS) at 4 months after transplantation and scored positive if PB donor chimerism was >10%.

RNA analyses

Detailed RNA procedures are provided in supplemental Methods. The Gene Expression Omnibus data accession number is GSE76254. Briefly, RNA was isolated from E10.5 AGM CD31 and cKit sorted cells with the mirVana miRNA Kit (Ambion), and quality/quantity was assessed using the 2100 Bioanalyzer (Agilent; RNA Nano/Pico chip). For RNA sequencing, samples were prepared with the SMARTER protocol and analyzed on an Illumina HiSeq200 system.[25-27] For quantitative reverse transcription-PCR (qRT-PCR), SuperScriptII/III ReverseTranscriptase (LifeTechnologies) was used for first-strand cDNA synthesis. Primers are specified in supplemental Table 1.

Results

Generation and validation of a novel *Gata2* reporter mouse model

Previously, analysis of *Gata2*-expressing cells has been limited to a reporter mouse model that results in Gata2 haploinsufficiency.[16] Our approach allows for the expression of the reporter within the *Gata2* genomic locus without affecting the levels of Gata2 expression or protein function. This is particularly important because Gata2 haploinsufficiency greatly reduces the number of HS/PCs generated during development.[7,8,19,20,28] Briefly, an internal ribosome entry site sequence (*IRES*) followed by the *Venus* fluorochrome gene was recombined into the *Gata2* 3′UTR (Figure 1A) in ES cells. The resulting *Gata2Venus* (*G2V*) mice bred normally and showed no overt growth or hematopoietic defects.

To determine whether *Venus* reporter expression parallels that of *Gata2*, *G2V* bone marrow (BM) cells were sorted into Venus-expressing (Venus$^+$) and nonexpressing (Venus$^-$) fractions (Figure 1B). qRT-PCR for *Gata2* and *Venus* transcripts (*Gata2*$^{V/+}$ BM) demonstrated that only Venus$^+$ cells express *Venus* and *Gata2* mRNA (Figure 1C). Western blot analysis revealed that equivalent amounts of Gata2 protein were present in *Gata2*$^{+/+}$ and *Gata2*$^{V/V}$ adult BM cells (data not shown). FACS analysis showed that *Gata2*$^{V/y}$ BM LSK frequency (388/10^5 cells) are comparable to wild-type (WT) BM LSK frequency (378/10^5 cells). Importantly, the results of competitive limiting dilution transplantation analyses of *Gata2*$^{V/V}$ and WT BM cells demonstrate that HSCs are qualitatively and quantitatively normal in this mouse model (Figure 1D). Thus, Venus expression correctly reports Gata2 expression without interfering with its normal expression levels or function.

Gata2 is expressed in emerging aortic hematopoietic cluster cells and other embryonic hematopoietic tissues

Venus expression was examined in midgestation *G2V* hematopoietic tissues. Flow cytometry revealed that E9 to E11 AGM, YS, PL, and FL contained Venus$^+$ cells (Figure 1E). At E9, 6.28 ± 0.47% of viable YS cells and 1.82 ± 0.31% of AGM cells are Venus$^+$. At E10.5 (when the first HSCs are generated), 3.27 ± 0.52% of AGM cells are Venus$^+$, and this increases to 7.86 ± 1.1% at E11. Table 1 shows the frequencies of Gata2-expressing cells.[29,30]

Whole-mount images of E10 and E11 *G2V* embryos immunostained with anti-CD31 antibody (marks all endothelial cells and hematopoietic cluster cells) shows Venus$^+$ cells along the aorta (DA). Venus$^+$ cells are also observed in cells of the neural tube (NT), olfactory bulb (OB), and FL (Figure 2A-C). In the E10.5 AGM region (4,6 diamidino-2-phenylindole [DAPI] and CD31 stained [blue and red, respectively]), Venus expression is found in endothelial and hematopoietic cluster cells mainly on the ventral side of the DA and in

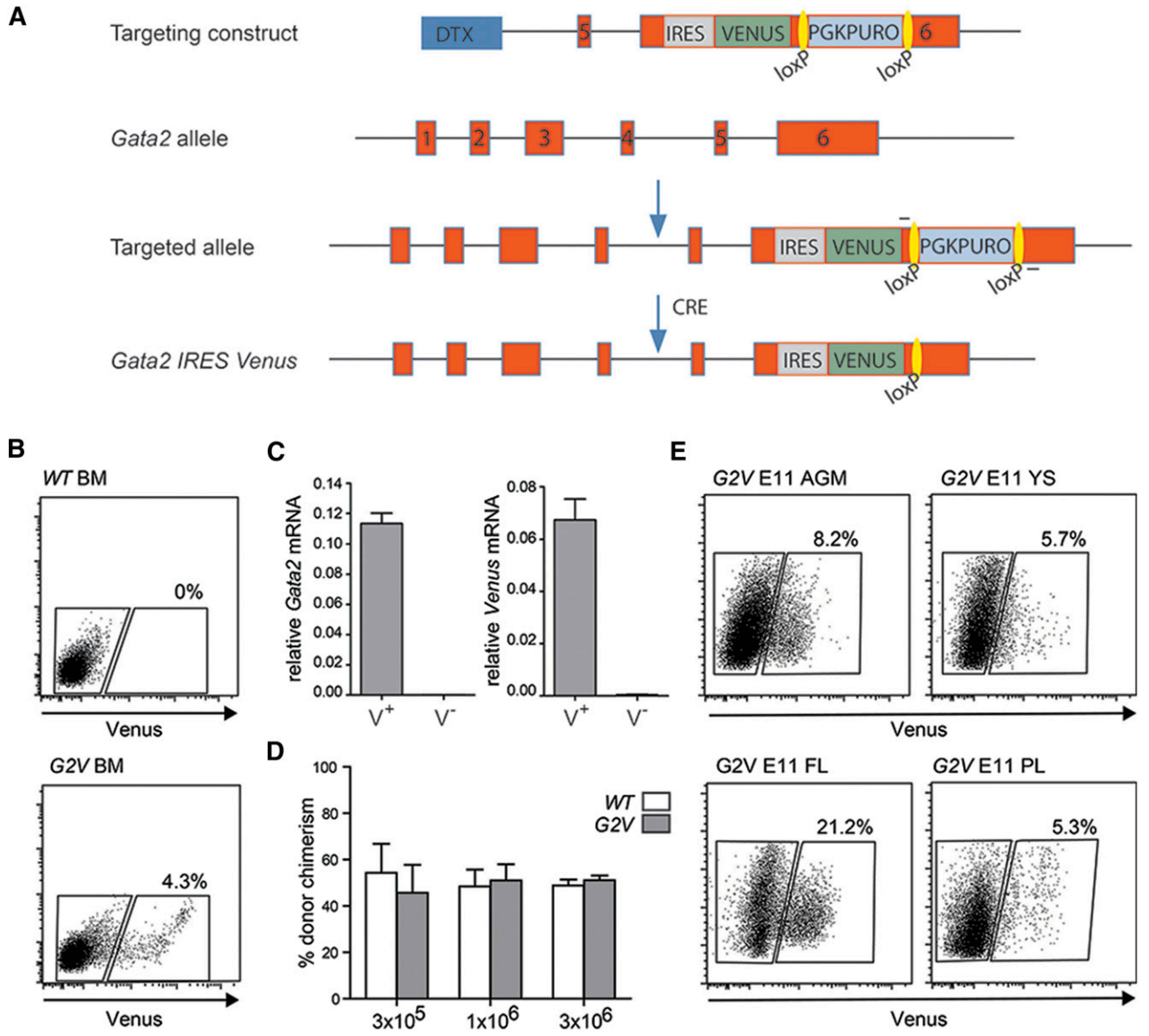


Figure 1. Gata2 Venus reporter construction and validation. (A) Schematic diagram of the *IRES Venus* reporter selection cassette insertion in the 3′UTR of the mouse *Gata2* locus and Cre-mediated removal of *lox PGK-Puro lox*. Primers used for detection of the targeted and recombined alleles are indicated flanking the loxP sites (yellow). (B) Representative flow cytometric analysis and sorting plot of Venus-expressing cells in the BM of adult *Gata2 Venus* (*G2V*) mice. Gated regions show percentage positive and negative viable cells. (C) Relative levels of *Gata2* and *Venus* mRNA in sorted Venus$^+$ and Venus$^-$ *Gata2*$^{V/+}$ BM cells as determined by qRT-PCR. *Gata2* transcripts in Venus$^+$ cells = 0.11337 ± 0.00681 and Venus$^-$ cells = 0.00012 ± 0.00003 (P = .000076). *Venus* transcripts in Venus$^+$ cells = 0.06722 ± 0.00799 and Venus$^-$ cells = 0.00036 ± 0.00010 (P = .00112). Mean ± standard error of the mean (SEM), n = 3. (D) Competitive limiting dilution transplantation strategy used to test the quantity and robustness of *Gata2*$^{V/V}$ BM HSCs compared with wild type. Percentage of donor cell chimerism in adult irradiated recipients cotransplanted with the same number of wild-type (WT) Ly5.1/5.2 and *Gata2 Venus* (*G2*$^{V/V}$) Ly5.2 BM cells. Varying numbers (1×10^5, 3×10^5, 3×10^6) of BM cells of each genotype were injected, and peripheral blood of recipients was analyzed for donor cell engraftment by FACS at 1 and 4 months after transplantation. n = 2 (5 mice per group). (E) Representative FACS plots demonstrating frequency of Venus expressing cells in E11 AGM, YS, PL, and FL. Gates indicate Venus$^-$ and Venus$^+$ cell fractions. Percentages represent the frequency of Venus$^+$ cells within the viable cell fraction (Table 1).

the urogenital (UG) region (Figure 2D-E). In the FL, Venus-expressing cells are found in a punctate distribution pattern (Figure 2D,F). At E9, Venus is expressed in some of the CD34$^+$ (red) endothelial cells of the paired aorta (Figure 2G, arrowhead) and also in some of the endothelial and hematopoietic cluster cells of the UA (Figure 2H). Venus continues to be expressed at E11 in some aortic endothelial cells and emerging/other hematopoietic cluster cells (Figure 2I, arrowhead). The E10.5 YS shows Venus expression in some of the CD31$^+$ (red) endothelial cells. Overall, Venus expression is similar to what has been previously documented for *Gata2* in situ hybridization analysis.[31,32] Thus, our model allows for the prospective identification, isolation, and characterization of Gata2-expressing cells during normal development.

All HSCs, but not all HPCs, express Gata2

To test for HSC activity, E11 AGM Venus$^+$ and Venus$^-$ cells were transplanted into irradiated adult recipients. All long-term repopulating HSCs were found in the Gata2Venus-expressing fraction (Figure 3A). Nine of 19 recipients receiving Venus$^+$ cells were engrafted (15-71%), whereas none of the 14 Venus$^-$ recipients showed donor-derived hematopoietic cells. These HSCs were multilineage repopulating (supplemental Figure 1) and self-renewing (8 repopulated of 12 transplanted with 3×10^6 BM cells from primary repopulated mice; n = 4).

The relationship between Gata2 expression and HPC function was also examined. E9 and E10 AGM Venus$^+$ and Venus$^-$ cells were

Table 1. Frequency of Venus$^+$ cells in embryonic tissues of *G2V* embryos

Tissue	Stage	Number of experiments, embryos analyzed	% Venus$^+$ cells/tissue
AGM	E9, 16-25 sp	n = 4, 22	1.82 ± 0.31
	E10, 28-36 sp	n = 4, 29	3.27 ± 0.52
	E11, 43-49 sp	n = 4, 25	7.86 ± 1.1
FL	E9	nd	nd
	E10, 28-36 sp	n = 4, 10	13.89 ± 0.7
	E11, 43-49 sp	n = 4, 19	19.27 ± 2.14
YS	E9, 16-25 sp	n = 5, 8	6.28 ± 0.47
	E10, 28-36 sp	n = 6, 9	6.17 ± 0.86
	E11, 42-46 sp	n = 1, 3	5.25 ± 0.59
PL	E9, 17-23 sp	n = 3, 4	10.91 ± 0.49
	E10, 28-35 sp	n = 3, 6	10.8 ± 1.92
	E11, 42-46 sp	n = 1, 3	10.01 ± 4.64

The frequency of Venus$^+$ cells within the viable cell fraction of embryonic tissues per embryo is presented. FACS analysis of single cell suspensions of dissected embryonic tissues was performed to define the percentage of cells expressing Venus. AGM contains part of the vitelline and umbilical arteries. Our data for the total number of cells in each tissue (data not shown) correlated with published data for YS and E9.5 AGM[29] and for FL.[30] The data represent mean ± SEM of 3 to 6 independent experiments, with the exception of E11 YS and PL data (data represent the mean ± standard deviation of 1 experiment. E, embryonic day; n, number of independent experiments, number of individual embryos analyzed; nd, not done; sp, somite pairs.

plated in the colony forming unit–culture (CFU-C) assay. High enrichment of HPCs was found in the Venus$^+$ fractions (Table 2). Surprisingly, HPCs were found also in the Venus$^-$ fraction, although there were very few. At E9, the Venus$^+$ and Venus$^-$ fractions, respectively, yielded 8.0 ± 2.1 and 0.4 ± 0.4 CFU-C per AGM. CFU-C numbers increased in the Venus$^+$ cell fraction at E10 (69.0 ± 7.1) and the Venus$^-$ fraction increased to 8.0 ± 2.3 CFU per AGM. However, bi- and multipotent progenitors were found only in the Venus$^+$ fraction (Figure 3B).

HPC activity was also examined in the Venus$^+$ and Venus$^-$ fractions of other hematopoietic tissues (Table 2). E9 and E10 VA+UA (Figure 3C), YS (Figure 3D), and PL (Figure 3E), and E10 FL (Figure 3F) contained progenitors in both fractions. Most HPCs were Venus$^+$. The greatest number of CFU-C arising from Venus$^-$ cells was found in the E9 YS (270.0 ± 69.8 CFU-C/YS). These data indicate that some HPCs are not expressing Gata2. BFU-E, CFU-G, and CFU-M were the predominant colony types in both fractions, and in contrast to the Venus$^+$ fractions, the Venus$^-$ fractions of VA+UA, YS, FL, and PL yielded few or no CFU-GEMM. Thus, all AGM HSCs express Gata2, Gata2 expression is associated with immature HPCs, but not all HPCs are Gata2 expressing.

Some HPCs and vascular cluster cells are formed in the absence of *Gata2*

Because the Venus$^-$ fractions of midgestation *G2V* hematopoietic tissues contain CFU-C, we tested whether such hematopoietic progenitors are present in *Gata2*-deleted embryos. CFU-Cs were detected in the *Gata2*$^{-/-}$ E9 AGM, E10 AGM (Figure 4A), and E10 VA+UA (Figure 4B), although significantly fewer compared with WT (Table 3). *Gata2*$^{+/-}$ tissues also contained fewer CFU-Cs compared with WT. The E9 *Gata2*$^{-/-}$ YS contained the most CFU-C (64.4 ± 12.2; Figure 4C). In *VEC-Cre:Gata*$^{f/f}$ embryos, E10 PL showed significantly decreased CFU-C numbers (Figure 4D), as did E10 AGM and YS[7] compared with WT. The CFU-C remaining in *Gata2*$^{-/-}$ embryos are predominantly CFU-G and CFU-M. Very few *Gata2*$^{-/-}$ CFU-GM and no CFU-GEMM were observed. These data support and validate our findings in *G2V* embryos that not all HPCs are Gata2-expressing, Gata2-independent progenitors exist in each of the early hematopoietic tissues, and the Gata2-expressing cell fraction is more enriched in multipotent progenitors.

Because hematopoietic clusters appear in the VA and UA prior to appearance in the aorta, and are larger than in the AGM,[33] we further examined these vessels. Whole-mount microscopic analysis demonstrates that clusters form in the absence of Gata2. The number and size of cKit$^+$ hematopoietic clusters in early E10 *Gata2*$^{+/-}$ and *Gata2*$^{-/-}$ VA+UA are decreased compared with WT (Figure 4E). The number of cKit$^+$ cells decreases 20-fold in the E10 *Gata2*$^{-/-}$ VA+UA (Figure 4F) in correspondence to the decrease in VA+UA CFU-C (Figure 4B), suggesting that these emerging cKit$^+$ hematopoietic cluster cells are part of the cohort of Gata2-independent HPCs.

Alternative genetic program is expressed in Venus$^-$ hematopoietic cells

The molecular basis for the functional differences observed in Gata2-dependent and -independent HPCs was examined by RNA sequencing. As most CD31$^+$Venus$^-$ HPCs showed cKit intermediate expression, we compared this population to CD31$^+$Venus$^+$cKit^{int} HPCs (Figure 5A). Gene set enrichment analysis on genes sorted by log ratio of Venus$^+$ vs Venus$^-$ FPKMs revealed that genes in the Ras signaling pathway were significantly enriched in the Venus$^+$ compared with the Venus$^-$ fraction (Figure 5B). Genes upregulated by Ras were enriched in the Venus$^+$ fraction, and highly upregulated genes included *Kras, Grb2* (Ras adaptor), and *Sos1* and *Sos2* (RasGEF activators) (Figure 5C). Genes downregulated by Ras were enriched in the Venus$^-$ fraction. RasGAP gene (renders Ras inactive) *Rasa2* was highly upregulated in the Venus$^-$ fraction, whereas *Rasa1* and *Rasa3* were highly upregulated in the Venus$^+$ fraction. *Rasa4* and *NF1* were expressed to similar levels. Also, Venus$^+$ HPCs showed increased levels of *CREB* and *CBP* expression compared with Venus$^-$ HPCs and express protein kinase A catalytic subunit genes, suggesting that Venus$^+$ HPCs have the potential to activate CREB target genes. *Gata2* has *CREB response element* consensus sites (−3 kb, −300 bp upstream transcription start site), suggesting that it is a downstream target.[34,35] As *Gata2* is a Notch target,[18] a two- to fourfold higher expression of *Notch1* and *Notch4* was found in the Venus$^+$ fraction (Figure 5D). Moreover, *Snw1* and *Maml1* (transcriptional coactivators in the Notch pathway that interact with Notch) were upregulated (2- and 30-fold, respectively) in Gata2-expressing HPCs.

Because Venus$^-$ HPCs are mainly restricted in their differentiation potential to the macrophage and granulocytic lineages, we evaluated their similarity to YS-derived erythromyeloid progenitors (EMPs) that give rise to tissue-resident macrophages. Flow cytometric analysis for EMP markers[36] showed that 3.89% of E10 YS and 0.72% of E10 AGM cells were EMPs (Sca1$^-$cKit$^+$CD41$^+$CD16/32$^+$). The majority of EMPs were Venus$^-$ (74% in YS, 85% in AGM; Figure 5F). At E11, the frequency of EMPs in the E11 YS and AGM decreased to 2.11% and 0.11%, respectively (Figure 5F), with 60% of YS and 77% of AGM EMPs now being Venus$^+$. Published transcriptome data on mouse YS EMPs show the low expression of several chemokine receptors/ligands (*Cx3cr1*, *Cx3cl1*, *Ccl2*, *Ccr1*, *Ccl9*, and *Ccr7*).[37] The expression of these genes was low or absent in Venus$^-$ AGM cells compared with Venus$^+$ cells (Figure 5F). Also, *Cxcr4* (highly expressed in EMPs) was highly expressed in Venus$^-$ AGM cells compared with Venus$^+$ cells. These results suggest that the Venus$^-$ population shares similarities to EMPs at the transcription level.

Analysis of FPKMs for heptad transcription factors previously described as expressed in AGM HSCs and HPCs[1-3] showed expression in both the Venus$^+$ and Venus$^-$ AGM fractions (data not shown). Also, other Gata factors were expressed in both fractions. In the mouse, *Gata1, 2,* and *3* are hematopoietic transcription factors, whereas the *Gata4, 5,* and *6* factors are not directly related to hematopoiesis.

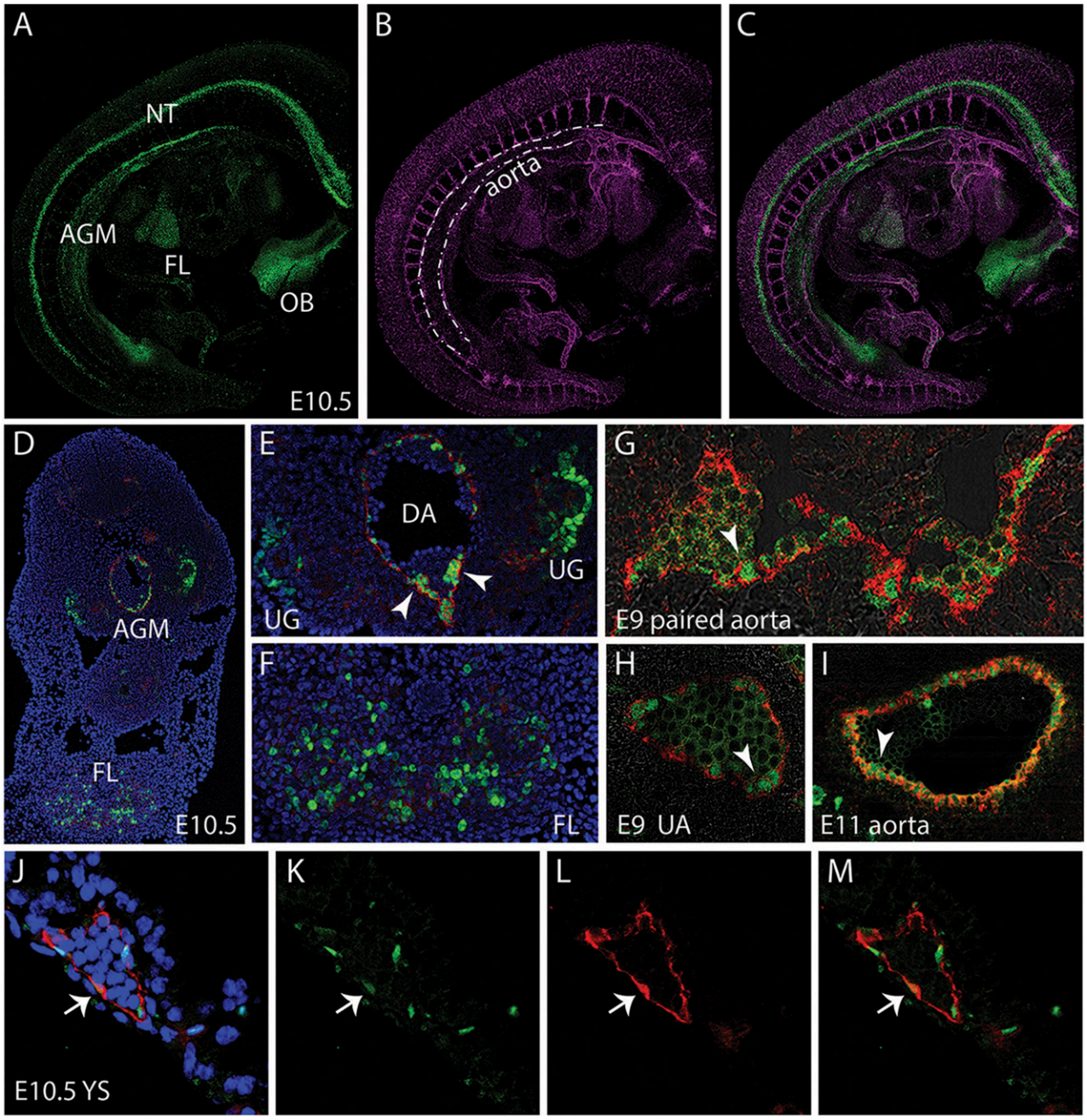


Figure 2. Localization of Gata2Venus-expressing cells in embryonic hematopoietic sites. Confocal images of a whole mount immunostained E10.5 Gata2Venus embryo showing (A) Venus (green), (B) CD31 (magenta), and (C) merged expression. Venus-expressing cells are detected in the AGM along the wall of the dorsal aorta (dotted lines), the FL, NT, and OB. (D) Confocal image of a transverse section through the E10.5 AGM. DAPI staining (blue), CD31 (red), and Venus fluorescence (green) revealed Gata2-expressing aortic endothelial and hematopoietic cluster cells and UG and FL cells. Enlarged images of D showing Gata2-expressing cells in (E) AGM (DA, dorsal aorta; UG, urogenital ridges; arrowheads indicate hematopoietic cluster) and (F) FL. Venus (green) and CD34 (red) fluorescence showing endothelial and hematopoietic cluster cells in (G) E9 paired aorta, (H) umbilical artery (UA) at E9, and (I) E11 aorta. Arrowheads indicate hematopoietic cluster. (J-M) Images of E10.5 YS section showing DAPI merged, Venus, CD31, and merged fluorescence. Arrow denotes an endothelial cell expressing Venus and CD31.

qRT-PCR performed on E10.5 AGM CD31$^+$cKit$^+$ cells (Figure 6A) confirmed *Gata3* expression by both the Venus$^+$ and Venus$^-$ fractions, and *Gata4* was significantly higher in the Venus$^-$ fraction. *Gata1, 5,* and *6* transcripts were low/not detected. Immunostaining of E10.5 AGM (WT) showed Gata3-expressing cells in the mesenchyme underlying the ventral aspect of the aorta, aortic endothelial cells, and some cells emerging from the aortic wall (Figure 6B) in agreement with Fitch.[38] *G2V* E10.5 AGM confirmed that some aortic endothelial cells coexpress Gata3 and Gata2 (Figure 6C). Gata4 expression was also found in aortic endothelial (CD34$^+$) cells, but it did not overlap with Gata2 expression (Figure 6D). Together, these results suggest that Gata3 and/or Gata4 may provide some function in Gata2-independent hematopoietic cells.

Discussion

In this study, we prospectively enriched and characterized Gata2-dependent and -independent HPC subsets from our novel *Gata2Venus* reporter mouse. Molecular analyses, together with the fact that some

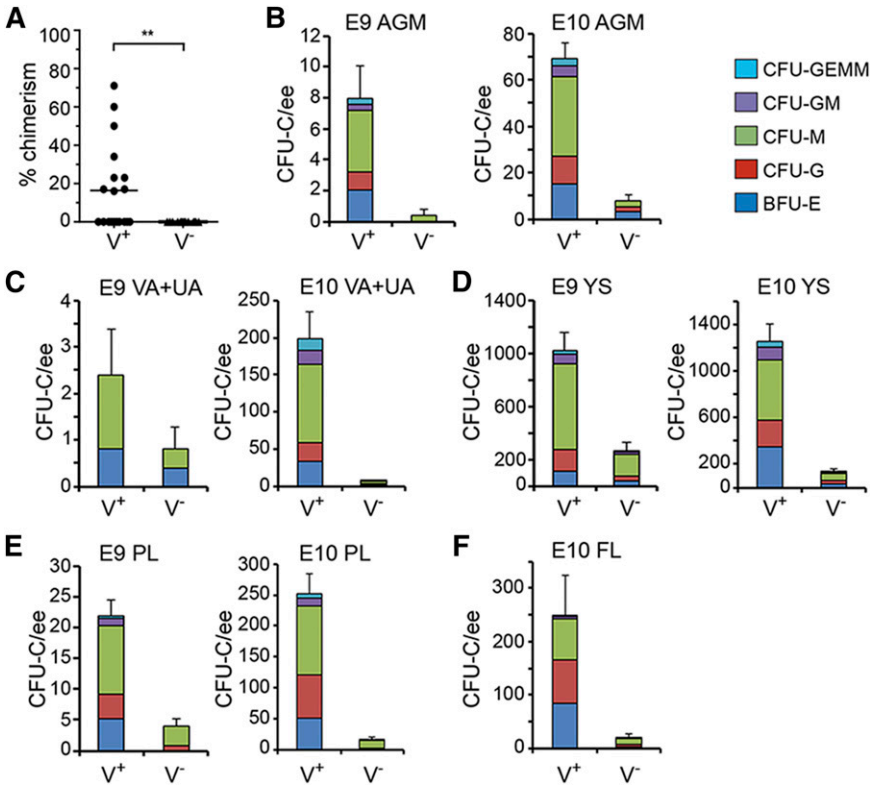


Figure 3. Quantitation of functional HSCs and HPCs in *G2V* embryonic hematopoietic tissues. HSCs in sorted Venus$^+$ and Venus$^-$ cell fractions of E11 AGM were analyzed by transplantation into irradiated adult recipients. (A) Percentage donor cell chimerism was determined by Venus PCR of peripheral blood DNA at 4 months after transplantation. Each dot represents 1 recipient receiving 1.7 to 6.5 embryo equivalent (ee) of AGM cells. n = 7. **P = .0089. (B-F) Hematopoietic progenitor number per tissue in sorted Venus$^+$ and Venus$^-$ cell fractions of (B) E9 and E10 AGM, (C) E9 and E10 VA+UA, (D) E9 and E10 YS, (E) E9 and E10 PL, and (F) E10 FL. CFU-C per 1 ee of tissue is shown. Colony types designated by colored bars are CFU-granulocyte, erythroid, monocyte, megakaryocyte (GEMM); CFU-granulocyte, macrophage (GM); CFU-macrophage (M); CFU-granulocyte (G), and burst forming unit-erythroid (BFU-E). SEM of total CFU-C is shown; 2 ee of somite pair–matched tissues were pooled for sorting and yielded 1 ee for colony analysis.

vascular hematopoietic cluster cells and HPCs persist in the absence of Gata2 expression, suggest that an alternative genetic program exists for the production of HPCs. The transcriptome differences observed between Venus$^+$ and Venus$^-$ HPCs may offer possibilities for pathway modifications to achieve the programming complexities necessary for the generation/function of normal definitive HPCs and provide insights into the factors involved in myeloid leukemogenesis.

Gata2 expression in the developing hematopoietic system

We showed the temporal and quantitatively coordinate transcription of *Venus* and *Gata2* in our *G2V* mouse model. The strategy used[39] eliminates expression level and protein alterations that affect HP/SC development. In *G2V* embryos, we showed that the cells with the most robust and complex hematopoietic potential (all HSCs and most HPCs) are Gata2 expressing. Imaging and FACS analyses of *G2V* embryos confirm that Gata2 is expressed in all hematopoietic sites during midgestation and that the numbers of Gata2-expressing cells reflect the developmental and temporal hematopoietic changes occurring in each site. At E9, Gata2-expressing cells are found predominantly in the YS, which at this time produces the highest numbers of the hematopoietic progenitors (EMP) in the conceptus. Slightly later as hematopoiesis begins in the AGM and FL, the numbers of Gata2-expressing cells also increase. The highest numbers of CD31$^+$cKit$^+$ cluster cells are found in the aorta, VA, and UA at E10.5, as quantitated by whole-mount embryo imaging.[33] Most, but not all, hematopoietic cluster cells express Gata2, and *Gata2* expression may be downregulated as HPCs differentiate. However, we found some hematopoietic cluster cells and HPCs in the E10 *Gata2*$^{-/-}$ vasculature, confirming the existence of Gata2-independent HPCs.[5]

Importantly, Gata2 is expressed in the endothelial cells of the DA. Already at E8.5, endothelial cells lining the paired dorsal aortae express Gata2, and it continues to be expressed in the E10.5 aorta when HSCs are generated, thus highlighting an involvement of Gata2 in the hemogenic program of endothelial cells. Data in *VE-cadherin* conditional *Gata2*-deficient mice and other models[7,19,28,40] strongly support the notion that Gata2 is required in hemogenic endothelium

Table 2. CFU-C number in Venus$^+$ and Venus$^-$ cell fractions of *G2V* embryonic tissues

Tissue	Stage	Number of experiments, embryos analyzed	CFU-C/tissue/sorted cell fraction	
			Venus$^-$	Venus$^+$
AGM	E9, 20-23sp	n = 2, 5	0.4 ± 0.4	8.0 ± 2.1
	E10, 32-35sp	n = 2, 6	8.0 ± 2.3	69.0 ± 7.1
VA+UA	E9, 20-23sp	n = 2, 5	0.8 ± 0.5	2.4 ± 1.0
	E10, 32-35sp	n = 2, 6	7.0 ± 1.5	197.7 ± 36.9
YS	E9, 20-23sp	n = 2, 5	270.0 ± 69.8	1020 ± 137.3
	E10, 32-35sp	n = 2, 5	130.0 ± 25.5	1252.0 ± 156.0
PL	E9, 20-23sp	n = 2, 5	4.0 ± 1.3	22.0 ± 2.5
	E10, 32-35sp	n = 2, 6	14.5 ± 5.3	251.7 ± 32.4
FL	E10, 32-35sp	n = 2, 6	20.5 ± 7.6	248.7 ± 75.2

Number of total CFU-C (mean ± SEM) per tissue per *G2V* embryo for the sorted Venus$^-$ and Venus$^+$ cell fractions at E9 and E10.

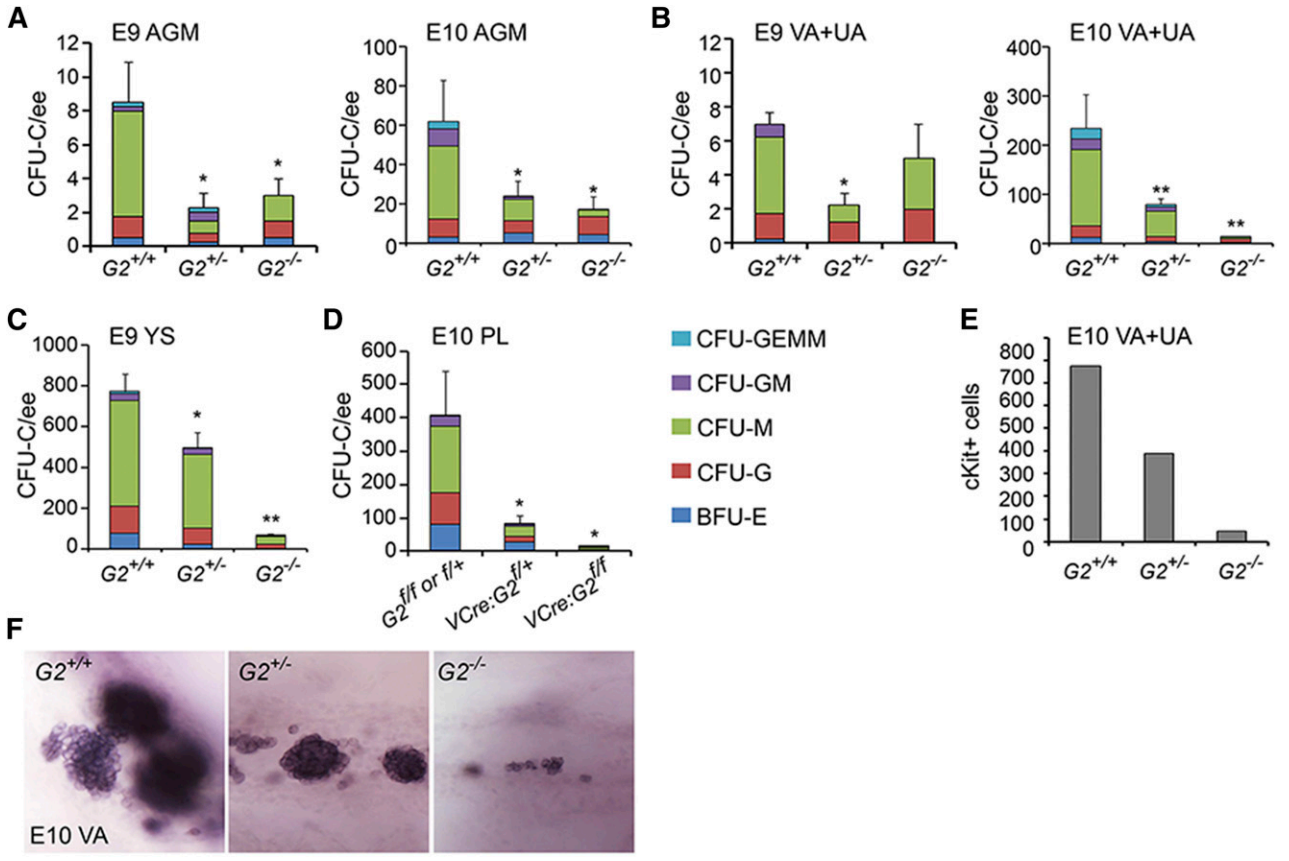


Figure 4. CFU-C numbers and vascular hematopoietic clusters in Gata2$^{-/-}$ embryos. CFU-C numbers per ee found in (A) E9 and E10 AGM, (B) E9 and E10 VA+UA, (C) E9 and E10 YS, and (D) E10 placenta. *$P < .05$; **$P < .01$. (E) Quantitation of cKit$^+$ hematopoietic cluster cells in VA+UA of E10 *Gata2*$^{+/+}$, *Gata2*$^{+/-}$, and *Gata2*$^{-/-}$ embryos. (F) Representative whole mount images of hematopoietic cluster cells in the VA of E10 *Gata2*$^{+/+}$ (30 sp), *Gata2*$^{+/-}$ (31 sp), and *Gata2*$^{-/-}$ (30 sp) embryos stained for cKit expression.

for the emergence of HSCs, as does the morpholino knockdown of *Gata2b* in zebrafish.[41]

Gata2 and the relationship with hematopoietic function

Prospective isolation and in vivo transplantation showed that all HSCs are Gata2 expressing. In contrast, some HPCs are present in the Venus$^-$ cell fractions of *G2V* hematopoietic tissues and *Gata2*$^{-/-}$ hematopoietic tissues. In both cases, the HPCs are restricted in their differentiation potential to predominantly the macrophage and granulocytic lineages. Currently, the EMP population is of high interest as a novel hematopoietic cell subset providing tissue resident macrophages.[37,42-44] Our FACS data revealed that EMPs are mainly in the Venus$^-$ cell population of E10 YS and AGM and increased in the Venus$^+$ population at E11. Chemokine receptor/ligand gene sets obtained from a study on EMP/microglia transcriptome comparisons allowed us to find similarities in chemokine receptor/ligand expression between EMPs and the Venus$^-$ HPC fraction.

Despite prevalence of EMPs in the Venus$^-$ cell population in E10 YS and AGM, definitive progenitors are largely Venus$^+$. The coexistence of these HPC subsets highlights the fact that there is more diversity in the types of progenitors generated in the embryo than was

Table 3. CFU-C number per E9-E10 *Gata2*-deleted hematopoietic tissues

		Genotype					
Tissue	**Stage**	**WT**		***Gata2*$^{+/-}$**		***Gata2*$^{-/-}$**	
AGM	E9, 20-23sp	8.5 ± 2.3	n = 1, 4	2.25 ± 0.9*	n = 1, 4	3.0 ± 1.0	n = 1, 2
	E10, 28-34sp	62.4 ± 20.8	n = 3, 7	24.0 ± 7.4*	n = 3, 13	16.7 ± 6.8*	n = 3, 9
VA+UA	E9, 20-23sp	7.0 ± 0.6	n = 1, 4	2.25 ± 0.6*	n = 1, 4	5.0 ± 2.0	n = 1, 2
	E10, 28-34sp	241.1 ± 67.8	n = 3, 7	78.7 ± 12.3**	n = 3, 13	12.0 ± 2.5**	n = 3, 9
YS	E9, 20-23sp	772.5 ± 85.3	n = 1, 4	500.0 ± 71.5*	n = 1, 4	64.4 ± 12.2**	n = 1, 2
	E10, 28-34sp	918.6 ± 147.9	n = 3, 7	584.6 ± 89.0**	n = 3, 9	25.7 ± 8.0**	n = 2, 3
		***Gata2*$^{f/+}$ or *Gata2*$^{f/f}$**		***VEC-Cre:Gata2*$^{f/+}$**		***VEC-Cre:Gata2*$^{-/-}$**	
FL	E10, 30-34sp	115.8 ± 35.0	n = 2, 5	83.0 ± 20.8	n = 2, 6	22.2 ± 6.7*	n = 2, 6
PL	E10, 30-34sp	406.0 ± 134.0	n = 2, 5	81.0 ± 25.0*	n = 2, 6	12.0 ± 4.0*	n = 2, 6

Number of total CFU-C (mean ± SEM) per tissue shown for WT, *Gata2* germline, and conditional knockout embryos at E9 and E10.
*$P < .05$.
**$P < .01$.

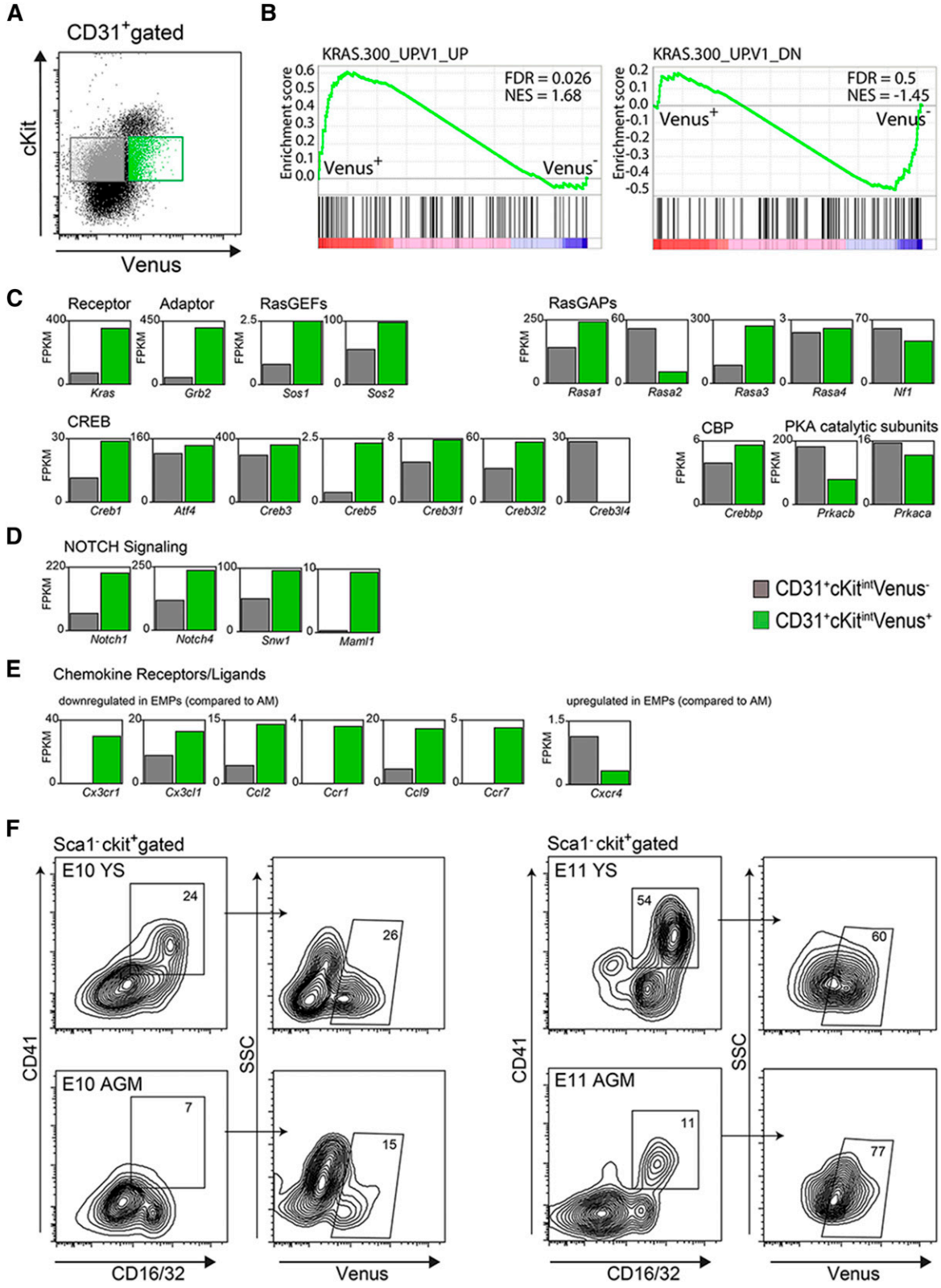


Figure 5. Differential expression of signaling pathway modulators in Gata2-dependent and -independent HPCs. (A) Flow cytometric sorting gates for isolation of E10.5 AGM *G2V* CD31^{+}cKit^{int}Venus^{-} (gray) and CD31^{+}cKit^{int}Venus^{-} (green) HPCs used for RNA sequence analysis. Gene Expression Omnibus data accession number is GSE76254. (B) Gene enrichment analysis for Ras signaling pathway genes. Bar graphs of fragments per kilobase million (FPKM) values obtained from RNA sequence analysis of CD31^{+}cKit^{int}Venus^{-} (gray bar) and CD31^{+}cKit^{int}Venus^{+} (green bar) AGM cells for (C) Ras pathway and cyclic AMP response element-binding protein (CREB) and CREB-binding protein (CBP) transcription factor genes and (D) Notch pathway genes. (E) Bar graphs of FPKM values obtained from RNA sequence analysis for a selection of chemokine receptor/ligand genes (see Kierdorf et al[37]; these genes were down-/upregulated in YS EMPs compared with adult microglia [AM]). (F) Representative FACS plots demonstrating frequency of EMPs in the Venus^{+} fraction, as defined as Sca1^{-}cKit^{+}CD41^{+}CD16/32^{+}, in YS and AGM of E10 (left) and E11 (right) *G2V* embryos. Numbers indicate the percentages of gated cells within the parental cell population.

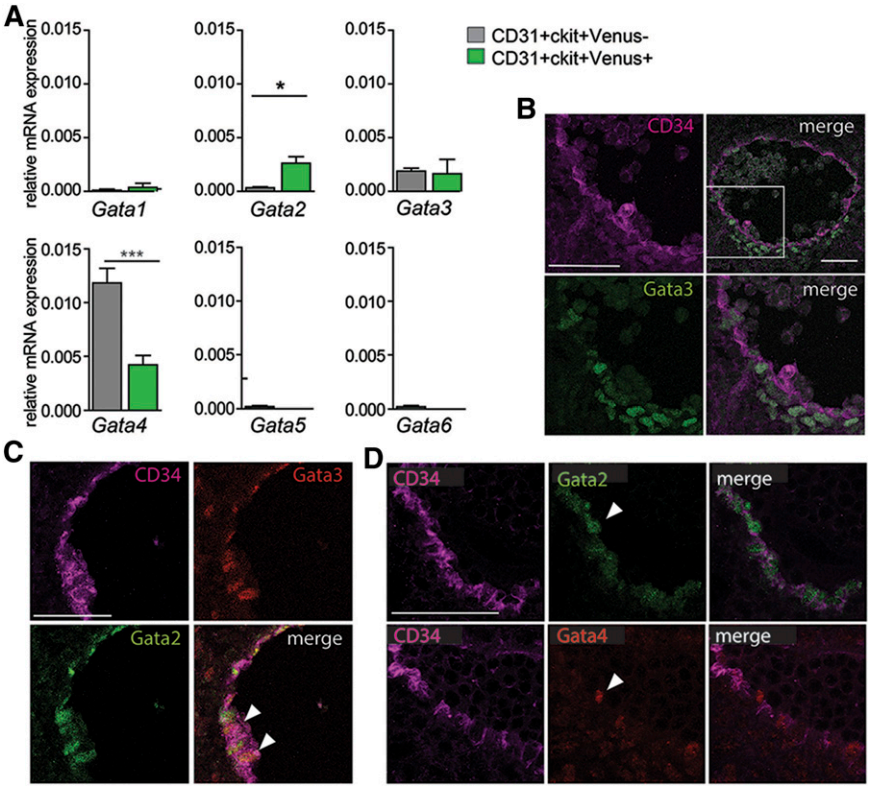


Figure 6. Gata family gene expression in AGM Gata2-dependent and -independent HPCs. (A) qRT-PCR for expression of *Gata1, 2, 3, 4, 5,* and *6* transcription factors (normalization with *Gapdh*) in E11 AGM CD31$^+$cKit$^+$Venus$^+$ and CD31$^+$cKit$^+$Venus$^-$ cells. n = 3. SEM shown with **P* = .05 and ****P* = .001. (B) Transverse section of WT E10.5 AGM immunostained for CD34 (magenta) and Gata3 (green) showing expression of Gata3 in the aortic endothelial cells and some emerging hematopoietic cells and ventral mesenchymal cells directly under the aorta. (C) Transverse section of *G2V* E10.5 AGM immunostained for CD34 (magenta), Gata2 (green), and Gata3 (red) showing some overlapping expression of Gata2 and Gata3 in aortic endothelial cells (arrowheads). (D) Transverse consecutive sections of E11 *G2V* AGM immunostained for CD34 (magenta) and Venus (green) in the top panels and for CD34 (magenta) and Gata4 (red) in the bottom panels. Gata4 expression is observed in some ventral aortic endothelial cells and emerging hematopoietic cells (arrow).

previously appreciated. In support of this are recent data from ES cell hematopoietic differentiation cultures suggesting that there are 2 different hemogenic endothelial cell subsets[45] and the fact that, in vivo, the AGM, VA/UA, YS, PL, and head are all hemogenic tissues.[36,46-49]

The highest number of Venus$^-$ HPCs was found in E9 and E10 YS (270.0 ± 69.8 and 130.0 ± 25.5 CFU-C, respectively) compared with other tissues (PL, AGM, VA+UA). It is clear that Gata2 has an important role in EHT in the hemogenic endothelial cell compartment before or during the generation/emergence of hematopoietic cells, as evidenced by the decrease (but not absence) in the hematopoietic cluster cells in *Gata2*$^{-/-}$ aorta, VA, and UA. However, it is as yet unclear at what frequency EHT occurs in the YS, thus raising the possibility that Gata2-independent HPCs arise differently than Gata2-dependent HPCs (perhaps directly from hemangioblasts[50]).

We found differences in the number of CFUs from E9 YS Venus$^-$ cells (270.0 ± 69) and Gata2$^{-/-}$ cells (64.4 ± 12.2; Figures 3D and 4C; Tables 2 and 3). The fourfold lower CFU number is likely related to observations (ours and others) that colonies from *Gata2*$^{-/-}$ embryos, YS explants, and ES cell differentiations were smaller/less proliferative than WT colonies, due to the complete absence of Gata2.[6,7] Venus$^-$ cells are not defective for Gata2, and the resulting colonies are normal in size. Whereas at the time of sorting they did not express Gata2, Gata2 expression could initiate after seeding Venus$^-$ HPCs in methylcellulose, and cells thus undergo normal proliferation/differentiation. To test whether Venus$^-$ HPCs can convert to Venus$^+$ cells, we analyzed Venus expression in colonies derived from sorted YS fractions after 10 days of differentiation (Figure 7). Venus expression was found in colonies derived from both fractions, indicating that a portion of Venus$^-$ cells start to express Gata2 during formation of a hematopoietic colony. Interestingly, colonies derived from Venus$^+$ cells showed a Gr1$^+$ and Mac1$^+$ phenotype, whereas Venus$^-$-derived colony cells were only Mac1$^+$. This demonstrates that Gata2 is not necessary for a subset of HPCs and that Gata2 promotes more complex hematopoietic function in other progenitor subsets.

Gata3/Gata4 redundancy in Gata2-independent progenitors

The expression of Gata3 and Gata4 in Venus$^-$ AGM HPCs and aortic endothelial cells is intriguing and highlights the potential redundancy of Gata transcription factors in hematopoietic cell generation. *Gata2* and *Gata3* can partially rescue the erythroid phenotype in *Gata1*-deficient mice,[51,52] and although recently it was suggested that *Gata3* is redundant in HSCs,[53] others clearly show that it regulates HSC cell cycle entry[54] and self-renewal.[55] The fact that *Gata3* and *Gata4* are expressed in *Gata2*-nonexpressing enriched HPCs suggests that they may function in this early progenitor subset. *Gata3*-deficient embryos show decreased numbers of FL HP/SCs.[56] Gata3 affects HSC development non-cell autonomously by activating the expression of *Th* (tyrosine hydroxylase) and hence, catecholamine production in the ventro-lateral cells of the sympathetic nervous system underlying the embryonic aorta.[38] HSC production was rescued when catecholamines were administered to the pregnant dams. These investigators also found that some aortic endothelial cells express a *Gata3LacZ* reporter, leaving open the possibility of a direct and overlapping role for Gata3 in some HPCs.

Much less is known concerning Gata4 in hematopoietic development. In zebrafish, there is a close relationship between anterior hemangioblasts and cardiac precursors.[57] Together with *Gata5* and *Gata6, Gata4* specifies these 2 anterior mesoderm derivatives. In mice, *Gata4* is a key component of the cardiac developmental program, with close associations between cardiac, vascular, and hematopoietic lineages.[58-60] Moreover, a subset of mouse endocardial and YS endothelial cells express cardiac markers, possess hemogenic potential, and give rise to transient definitive erythroid/myeloid progenitors.[61] Our results suggest the Gata4 aortic endothelial cells and Venus$^-$

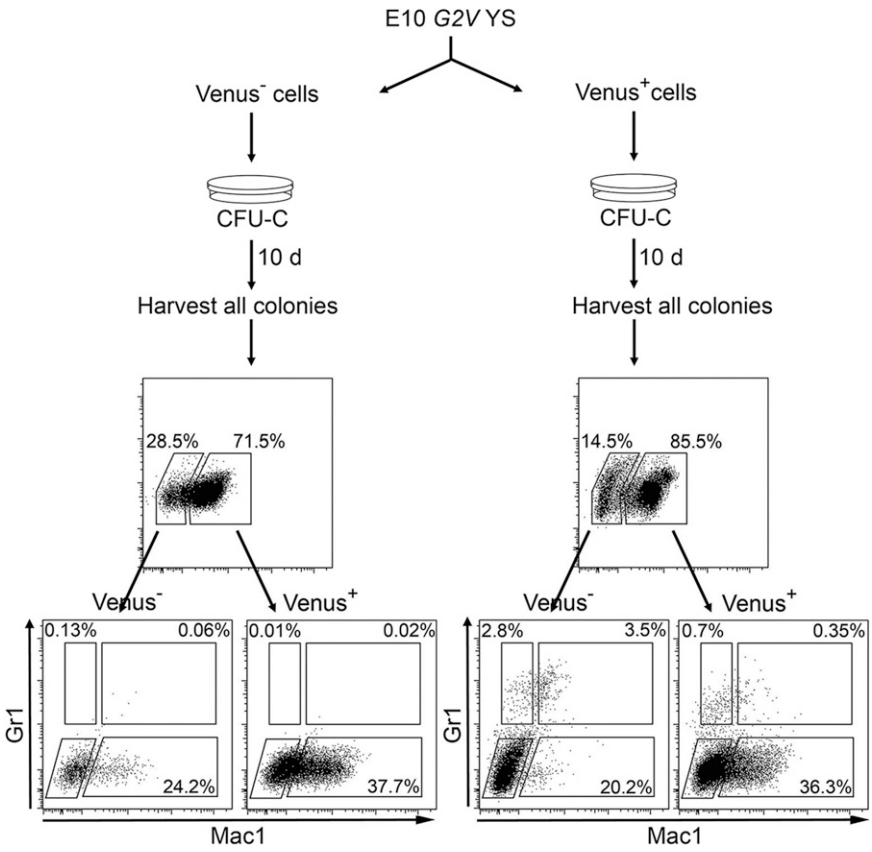


Figure 7. Gata2 is expressed by Venus$^-$ cells after culture. Schematic diagram showing method and FACS analysis by which Gata2 expression was found in the progeny of sorted Venus$^-$ HPCs. *G2V* YS tissue was FACS sorted into Venus$^-$ and Venus$^+$ fractions. Cells were subsequently seeded in methylcellulose, and colonies were analyzed after 10 days of culture. Colonies were harvested from the dish, cells were washed and stained (with anti-Gr1 and anti-Mac1 antibodies), and Venus, Gr1, and Mac1 expression was analyzed by FACS. FACS plots (top) indicate Venus expression in cells harvested from Venus$^-$ (left) and Venus$^+$ (right) CFU-C experiments. Note that both FACS analyses indicate Venus expression in both cultures. FACS plots (bottom) show Gr1 and Mac1 expression in Venus$^-$ and Venus$^+$ populations in both cultures and that cells harvested from the Venus$^+$ culture show a more immature phenotype.

Gata4-expressing HPCs may be derivatives of mesodermal cells with a genetic program that retains cardiac-vascular-hematopoietic potential and can produce HPCs. Further examination of double reporter and deficient mice should reveal the overlapping and/or redundant roles of these Gata factors.

Gata2 as a pivotal regulator of complex hematopoietic function

RNA sequence comparisons of the functionally distinct Venus$^+$ and Venus$^-$ HPC subsets revealed a strong upregulation of *Kras* and *Ras* pathway genes in Venus$^+$ HPCs. This pathway is particularly important in cell differentiation, acting as a molecular switch to relay extracellular growth signals.[62] *Kras* mutations confer a competitive-repopulating advantage to BM HSCs in transplantations and initiate leukemia in mice.[63] In humans, *Kras* mutations (together with other cooperating gene mutations) are prevalent in patients with various forms of myelomonocytic and myeloid leukemia.[64] Interactions between oncogenic Ras and Gata2 have been proposed.[65] The normal function of Kras has not yet been explored fully. However, conditional deletion of *Kras* by *Vav-Cre* or *Mx1-Cre* does not affect HSCs or the adult hematopoietic system.[66] However, chimeric mice, produced by *Kras*$^{-/-}$ ES cell blastocyst injection, show no contribution of *Kras*$^{-/-}$ cells to the hematopoietic system, suggesting that Kras may be important during the embryonic development of the hematopoietic system but not after its generation.

The low expression of *Notch1*, *Notch4*, and coactivators in the Venus$^-$ compared with the Venus$^+$ HPC fraction supports the fact that early hematopoietic cells are generated independent of this signaling pathway or implies that these are differentiated cells that have turned off Notch signaling.[67,68] Others have shown that *Notch1* deletion impairs the development of HSCs and angiogenesis,[69] but not YS primitive or definitive hematopoiesis. Moreover, *Gata2* expression in the aortic endothelium is lost when *Jagged1* (ligand) is deleted.[70] Our data demonstrate a direct relationship for Notch and Gata2 expression, strongly supporting a pivotal role for this pathway in the generation of functionally complex hematopoietic cells. In the absence of Notch signaling, less complex HPCs emerge in the AGM or are immigrants from the YS.[18,69,70] In addition, our observed upregulated expression of some CREB genes in Venus$^+$ HPCs supports the involvement of these regulators in definitive hematopoietic cell generation.[71] Our *Gata2Venus* model, in combination with recently reported *Gata2* distal enhancer-*Evi1* mouse model, will allow for a direct examination of the cells relevant to leukemogenesis.[12]

In conclusion, we enriched, localized, and characterized Gata2-dependent and -independent subsets of hematopoietic progenitors in *Gata2Venus* embryos. The combination of this reporter with other reporter and knockout models will lead to a better understanding of the role of Gata2 (and other factors) in the development and function of multipotential HP/SCs in health, leukemogenesis, and reprogramming.

Acknowledgments

The authors thank Prof Jim Palis for critical comments on this manuscript; Dr Dorota Kurek for providing anti-Gata3 and Gata4 antibodies; Drs Dorota Kurek and Mihaela Crisan for immunostaining support; Dr Siska Driegen for ES cell culture support; and Dr Derk ten Berge for providing Wnt for ES cell cultures.

This work was supported by ZonMw (Dutch Medical Research Council Grant 911-09-036), FES Netherlands Institute for Regenerative Medicine (101675), the National Institutes of Health,

National Institute of Diabetes and Digestive and Kidney Diseases (RO37 DK54077), ZonMw TOP (91211068), European Research Council Advanced Grant (341096), and Landsteiner Foundation for Blood Transfusion Research (1344).

Authorship

Contribution: P.K., E.d.P., C.E., M.-L.K., M.J., C.S.V., and T.Y. performed research; P.S.K. analyzed RNAseq data; R.v.d.L. performed/analyzed flow cytometric data; D.M. provided reagents; P.K., E.d.P., C.E., D.M., and E.D. designed experiments and analyzed and interpreted the data; and P.K., E.D., E.d.P., and C.E. wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Elaine Dzierzak, University of Edinburgh, Centre for Inflammation Research, Queens Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ, United Kingdom; e-mail: e.dzierzak@erasmusmc.nl or elaine.dzierzak@ed.ac.uk.

References

1. Wilson NK, Foster SD, Wang X, et al. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell.* 2010;7(4):532-544.
2. Solaimani Kartalaei P, Yamada-Inagawa T, Vink CS, et al. Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *J Exp Med.* 2015;212(1):93-106.
3. Swiers G, Baumann C, O'Rourke J, et al. Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat Commun.* 2013;4:2924.
4. Fujiwara Y, Chang AN, Williams AM, Orkin SH. Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. *Blood.* 2004;103(2):583-585.
5. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature.* 1994; 371(6494):221-226.
6. Tsai FY, Orkin SH. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood.* 1997;89(10):3636-3643.
7. de Pater E, Kaimakis P, Vink CS, et al. Gata2 is required for HSC generation and survival. *J Exp Med.* 2013;210(13):2843-2850.
8. Ling KW, Ottersbach K, van Hamburg JP, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med.* 2004;200(7):871-882.
9. Rodrigues NP, Janzen V, Forkert R, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood.* 2005;106(2):477-484.
10. Hsu AP, Johnson KD, Falcone EL, et al. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. *Blood.* 2013;121(19):S3830-S3837.
11. Gröschel S, Sanders MA, Hoogenboezem R, et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell.* 2014;157(2): 369-381.
12. Yamazaki H, Suzuki M, Otsuki A, et al. A remote GATA2 hematopoietic enhancer drives leukemogenesis in inv(3)(q21;q26) by activating EVI1 expression. *Cancer Cell.* 2014;25(4): 415-427.
13. Heyworth C, Gale K, Dexter M, May G, Enver T. A GATA-2/estrogen receptor chimera functions as a ligand-dependent negative regulator of self-renewal. *Genes Dev.* 1999;13(14):1847-1860.
14. Persons DA, Allay JA, Allay ER, et al. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood.* 1999;93(2): 488-499.
15. Nandakumar SK, Johnson K, Throm SL, Pestina TI, Neale G, Persons DA. Low-level GATA2 overexpression promotes myeloid progenitor self-renewal and blocks lymphoid differentiation in mice. *Exp Hematol.* 2015;43(7):565-577.
16. Minegishi N, Suzuki N, Yokomizo T, et al. Expression and domain-specific function of GATA-2 during differentiation of the hematopoietic precursor cells in midgestation mouse embryos. *Blood.* 2003;102(3):896-905.
17. Ottersbach K, Dzierzak E. Analysis of the mouse placenta as a hematopoietic stem cell niche. *Methods Mol Biol.* 2009;538:335-346.
18. Robert-Moreno A, Espinosa L, de la Pompa JL, Bigas A. RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development.* 2005;132(5):1117-1126.
19. Gao X, Johnson KD, Chang YI, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med.* 2013;210(13):2833-2842.
20. Lim KC, Hosoya T, Brandt W, et al. Conditional Gata2 inactivation results in HSC loss and lymphatic mispatterning. *J Clin Invest.* 2012; 122(10):3705-3717.
21. Orlic D, Anderson S, Biesecker LG, Sorrentino BP, Bodine DM. Pluripotent hematopoietic stem cells contain high levels of mRNA for c-kit, GATA-2, p45 NF-E2, and c-myb and low levels or no mRNA for c-fms and the receptors for granulocyte colony-stimulating factor and interleukins 5 and 7. *Proc Natl Acad Sci USA.* 1995;92(10):4601-4605.
22. Sakai K, Miyazaki J. A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochem Biophys Res Commun.* 1997;237(2):318-324.
23. Yokomizo T, Yamada-Inagawa T, Yzaguirre AD, Chen MJ, Speck NA, Dzierzak E. Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nat Protoc.* 2012;7(3):421-431.
24. Medvinsky A, Taoudi S, Mendes S, Dzierzak E. Analysis and manipulation of hematopoietic progenitor and stem cells from murine embryonic tissues. *Curr Protoc Stem Cell Biol.* 2008;Chapter 2:Unit 2A.6.
25. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol.* 2013;31(1):46-53.
26. Mootha VK, Bunkenborg J, Olsen JV, et al. Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell.* 2003;115(5):629-640.
27. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545-15550.
28. Johnson KD, Hsu AP, Ryu MJ, et al. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J Clin Invest.* 2012;122(10):3692-3704.
29. Dzierzak E, de Bruijn M. Isolation and analysis of hematopoietic stem cells from mouse embryos. *Methods Mol Med.* 2002;63:1-14.
30. Gekas C, Dieterlen-Lievre F, Orkin SH, Mikkola HK. The placenta is a niche for hematopoietic stem cells. *Dev Cell.* 2005;8(3):365-375.
31. Nardelli J, Thiesson D, Fujiwara Y, Tsai FY, Orkin SH. Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. *Dev Biol.* 1999;210(2):305-321.
32. Pimanda JE, Ottersbach K, Knezevic K, et al. Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc Natl Acad Sci USA.* 2007; 104(45):17692-17697.
33. Yokomizo T, Dzierzak E. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development.* 2010;137(21):3651-3661.
34. Impey S, McCorkle SR, Cha-Molstad H, et al. Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell.* 2004;119(7):1041-1054.
35. Zhang X, Odom DT, Koo SH, et al. Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proc Natl Acad Sci USA.* 2005;102(12):4459-4464.
36. McGrath KE, Frame JM, Fegan KH, et al. Distinct sources of hematopoietic progenitors emerge before HSCs and provide functional blood cells in the mammalian embryo. *Cell Reports.* 2015; 11(12):1892-1904.
37. Kierdorf K, Erny D, Goldmann T, et al. Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat Neurosci.* 2013;16(3):273-280.
38. Fitch SR, Kimber GM, Wilson NK, et al. Signaling from the sympathetic nervous system regulates hematopoietic stem cell emergence during embryogenesis. *Cell Stem Cell.* 2012;11(4): 554-566.
39. Nutt SL, Metcalf D, D'Amico A, Polli M, Wu L. Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *J Exp Med.* 2005;201(2):221-231.
40. Khandekar M, Brandt W, Zhou Y, et al. A Gata2 intronic enhancer confers its pan-endothelia-specific regulation. *Development.* 2007;134(9): 1703-1712.
41. Butko E, Distel M, Pouget C, et al. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development.* 2015;142(6):1050-1061.

42. Ginhoux F, Greter M, Leboeuf M, et al. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science.* 2010;330(6005):841-845.

43. Gomez Perdiguero E, Klapproth K, Schulz C, et al. Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature.* 2015;518(7540):547-551.

44. Herbomel P, Thisse B, Thisse C. Zebrafish early macrophages colonize cephalic mesenchyme and developing brain, retina, and epidermis through a M-CSF receptor-dependent invasive process. *Dev Biol.* 2001;238(2):274-288.

45. Ditadi A, Sturgeon CM, Tober J, et al. Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat Cell Biol.* 2015;17(5):580-591.

46. Palis J, Yoder MC. Yolk-sac hematopoiesis: the first blood cells of mouse and man. *Exp Hematol.* 2001;29(8):927-936.

47. Frame JM, McGrath KE, Palis J. Erythro-myeloid progenitors: "definitive" hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells Mol Dis.* 2013;51(4):220-225.

48. Lux CT, Yoshimoto M, McGrath K, Conway SJ, Palis J, Yoder MC. All primitive and definitive hematopoietic progenitor cells emerging before E10 in the mouse embryo are products of the yolk sac. *Blood.* 2008;111(7):3435-3438.

49. Rhodes KE, Gekas C, Wang Y, et al. The emergence of hematopoietic stem cells is initiated in the placental vasculature in the absence of circulation. *Cell Stem Cell.* 2008;2(3):252-263.

50. Jaffredo T, Nottingham W, Liddiard K, Bollerot K, Pouget C, de Bruijn M. From hemangioblast to hematopoietic stem cell: an endothelial connection? *Exp Hematol.* 2005;33(9):1029-1040.

51. Takahashi S, Shimizu R, Suwabe N, et al. GATA factor transgenes under GATA-1 locus control rescue germline GATA-1 mutant deficiencies. *Blood.* 2000;96(3):910-916.

52. Tsai FY, Browne CP, Orkin SH. Knock-in mutation of transcription factor GATA-3 into the GATA-1 locus: partial rescue of GATA-1 loss of function in erythroid cells. *Dev Biol.* 1998;196(2):218-227.

53. Buza-Vidas N, Duarte S, Luc S, Bouriez-Jones T, Woll PS, Jacobsen SE. GATA3 is redundant for maintenance and self-renewal of hematopoietic stem cells. *Blood.* 2011;118(5):1291-1293.

54. Ku CJ, Hosoya T, Maillard I, Engel JD. GATA-3 regulates hematopoietic stem cell maintenance and cell-cycle entry. *Blood.* 2012;119(10):2242-2251.

55. Frelin C, Herrington R, Janmohamed S, et al. GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. *Nat Immunol.* 2013;14(10):1037-1044.

56. Pandolfi PP, Roth ME, Karis A, et al. Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. *Nat Genet.* 1995;11(1):40-44.

57. Peterkin T, Gibson A, Patient R. Common genetic control of haemangioblast and cardiac development in zebrafish. *Development.* 2009;136(9):1465-1474.

58. Fehling HJ, Lacaud G, Kubo A, et al. Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development.* 2003;130(17):4217-4227.

59. Huber TL, Kouskoff V, Fehling HJ, Palis J, Keller G. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature.* 2004;432(7017):625-630.

60. Van Handel B, Montel-Hagen A, Sasidharan R, et al. Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell.* 2012;150(3):590-605.

61. Nakano H, Liu X, Arshi A, et al. Haemogenic endocardium contributes to transient definitive haematopoiesis. *Nat Commun.* 2013;4:1564.

62. Pierre S, Bats AS, Chevallier A, et al. Induction of the Ras activator Son of Sevenless 1 by environmental pollutants mediates their effects on cellular proliferation. *Biochem Pharmacol.* 2011;81(2):304-313.

63. Sabnis AJ, Cheung LS, Dail M, et al. Oncogenic Kras initiates leukemia in hematopoietic stem cells. *PLoS Biol.* 2009;7(3):e59.

64. Chang YI, You X, Kong G, et al. Loss of Dnmt3a and endogenous Kras(G12D/+) cooperate to regulate hematopoietic stem and progenitor cell functions in leukemogenesis. *Leukemia.* 2015;29(9):1847-1856.

65. Katsumura KR, Yang C, Boyer ME, Li L, Bresnick EH. Molecular basis of crosstalk between oncogenic Ras and the master regulator of hematopoiesis GATA-2. *EMBO Rep.* 2014;15(9):938-947.

66. Damnernsawad A, Kong G, Liu Y, Chang Y-I, et al. Kras plays an important role in generating differentiated blood cells [abstract]. *Blood.* 2013;122(21). Abstract 2451.

67. Richard C, Drevon C, Canto PY, et al. Endothelio-mesenchymal interaction controls runx1 expression and modulates the notch pathway to initiate aortic hematopoiesis. *Dev Cell.* 2013;24(6):600-611.

68. Bertrand JY, Cisson JL, Stachura DL, Traver D. Notch signaling distinguishes 2 waves of definitive hematopoiesis in the zebrafish embryo. *Blood.* 2010;115(14):2777-2783.

69. Kumano K, Chiba S, Kunisato A, et al. Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. *Immunity.* 2003;18(5):699-711.

70. Robert-Moreno A, Guiu J, Ruiz-Herguido C, et al. Impaired embryonic haematopoiesis yet normal arterial development in the absence of the Notch ligand Jagged1. *EMBO J.* 2008;27(13):1886-1895.

71. Kim PG, Nakano H, Das PP, et al. Flow-induced protein kinase A-CREB pathway acts via BMP signaling to promote HSC emergence. *J Exp Med.* 2015;212(5):633-648.

Online Supplemental Methods

Gata2Venus ESC and mice. Gata2 homology arms were amplified with primers as specified in Supplemental Table 1 and cloned into TOPO (Invitrogen). An *IRES-Venus* fragment was inserted between left and middle, and *loxP-PGK-Puro-loxP* between middle and right arms in pSP72. IB10 ESCs were transfected with linearized vector (20μg), puromycin-selected and 360 clones PCR screened for *Gata2 Venus* (right arm junction – primers in Table1; 2292bp). Correct integration was verified by Southern blot (left arm) for 2 clones with normal karyotype. ESC+MEF were cultured in DMEM, 15%FBS, 1%Glutamax, 1%non-essential amino acids, 1%sodium pyruvate, 50μM βmercaptoethanol, 1000U/ml LIF at 37°C, 5%CO_2. Founders were identified by *Venus* PCR. First generation *G2V* offspring were crossed with *CAG-Cre* mice [22] and backcrossed (>10 generations) with C57BL/6.

Mice and embryo production. Gata2$^{+/-}$mice [5], Ly5.1 (6-8 week) and C57BL/6 mice were obtained/maintained (Harlan or locally) and genotyped by PCR. For PCR primers see Supplemental Table1. Day of plug discovery is embryonic day (E)0. Embryos were individually staged by somite pair (sp): E9=16-28sp, E10=28-40sp, early E10=28-34sp, E10.5=35-40sp, E11=40-50sp.

Immunostaining. Whole-mount conceptuses [23] were suspended in PBS/10%FCS, fixed in 2%PFA, and 100%MeOH. The trunk was dissected, anti-cKit and anti-CD31 stained, made transparent in 1:2BABB, and analysed (LeicaSP5) with Z-stack reconstructions (50–150μm) by LasAFsoftware. For cryosectioning [8] embryos were fixed in 2%PFA (20min), suspended in 20%sucrose O/N, snap-frozen in TissueTec, sectioned (HM630 cryostat) and fixed (acetone). Anti-CD34-biotin (1:50, BD), anti-Gata3 (1:10 clone KT122, 111207H09, Absea) and anti-Gata4 (1:50, sc-9053, SantaCruz) and anti-GFP antibodies O/N and secondary antibody incubation at RT, DAPI counterstained, mounted (Vectashield) and imaged (LeicaSP5). For flow cytometry, cell suspensions were made by collagenase treatment or direct dissociation and filtration [8], stained with anti-CD31 (390, BD), anti-CD34 (RAM34, BD), anti-cKit (2B8, BD), anti-CD41 (MWReg3, SantaCruz), anti-Sca1 (D7, Ebiosciences) and anti-CD16/32 (2.4G2, BD) antibodies, Hoechst 33258 (BD) and analysed on a FACSARIAIII/SORP.

Hematopoietic assays. Venus-sorted E9, E10, E11AGM, V+U, PL and YS and E10FL or *Gata2*-deficient E9-E10 AGM, V+U and PL (1 embryo equivalent (ee) of E9, E10 and E11 AGM, 0.1 ee V+U V^+, 0.1 ee PL V^+,0.1 ee YS V^-, 0.03 ee YS V^+ and 0.1 ee E10 FLs) were seeded in 3.6ml Methylcellulose (1ml per dish; M3434, StemCellTech)(or 0.9 and 0.1 ee of

Gata2$^{-/-}$ V+U and *Gata2*$^{-/-}$ YS 0.1 and 0.03 ee were seeded in 1.2 ml and distributed in a single dish) and cultured for 10-12 days[24]. Colonies were isolated, washed and Venus expression examined by FACSARIAIII/Fortessa. Sorted *G2V* E11AGM (Ly5.1/Ly5.2) cells were transplanted as described previously[24]. In short, AGMs were dissected and collagen II treated for 45 minutes. Subsequently dissociated and washed. When appropriate, the cells were stained and subsequently FACS sorted and collected in 50% FCS/ 50% PBS. Sorted *G2V* E11 AGM (Ly5.1/Ly5.2) cells were washed to PBS and injected into 9.5 Gy-irradiated (C57BL/6 x 129)F1 recipients (Ly5.2/Ly5.2) together with 2x10^5 spleen cells from the recipient strain. Peripheral blood (PB) donor chimerism was determined by *Venus* PCR or FACS at 1 and 4mo post-transplantation and scored positive if PB donor chimerism was >10% and multiorgan/lineage chimerism was >1%. For secondary transplantations, 3x10^6 BM cells from primary reconstituted recipients (4mo post-transplantation) were injected into 9.5 Gy-irradiated recipients and repopulation analysed at 4mo post-transplantation.

RNA analyses. E10.5 AGM cells were stained and sorted (CD31-PE-Cy7, 390, Biolegend; ckit-APC, 288, BD), washed, resuspended in RNAlater (LifeTechnologies) and stored at −20°C. RNA preparation was with mirVana miRNA Kit (Ambion) and quality/quantity were assessed on a 2100 bioanalyzer (Agilent Technologies) with a RNA Nano/Pico chip. RNAsequencing used SMARTER protocol for Illumina HiSeq2000. Sequences were mapped to mouse (version 10) and FPKMs calculated using TopHat (v2.0.13) and Cufflinks (v2.2.1) with fragment-bias and multi-read corrections[25]. Expression difference threshold was >2-fold change and higher or lower expressing genes were applied to Enrichr tool. Enrichr output (and calculated p-values from Chi2 tests) were imported into R and corrected for multiple testing (FDR) and a threshold of FDR <0.05 was used. For GSEA, ratios were calculated for AGM CD31$^+$c Kit$^+$Venus$^-$ versus CD31$^+$cKit$^+$Venus$^+$ FPKMs(+1) and used with GSEA (version2.2.0) preranked method using default options[26,27]. GEO Data Accession Number is GSE76254.

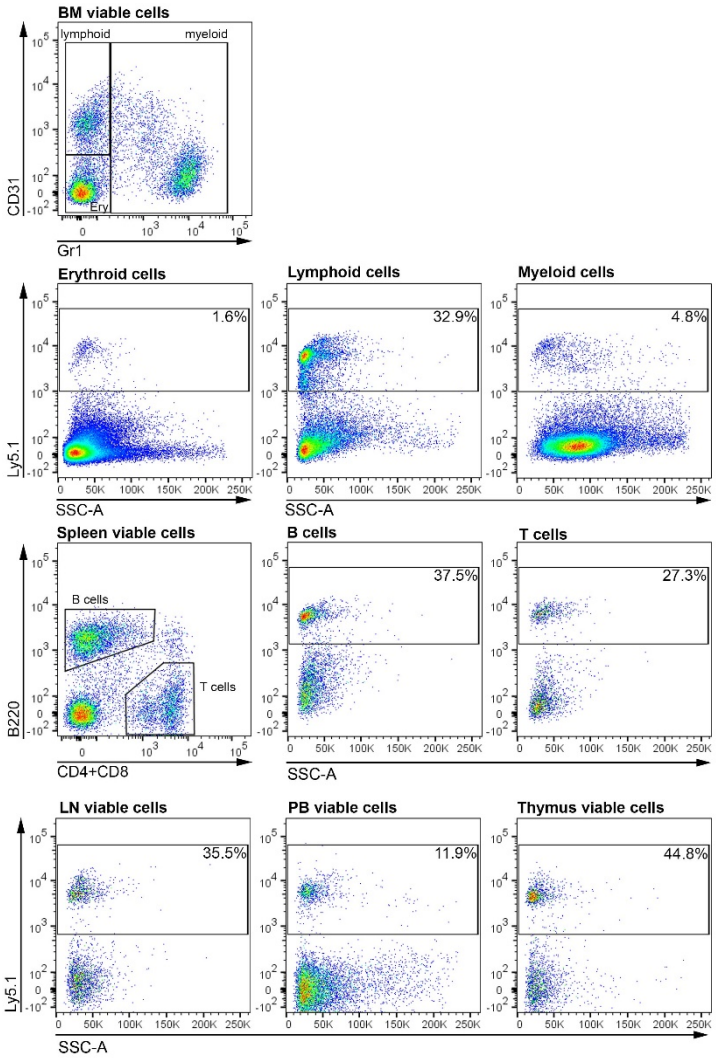
For qRT-PCR, SuperScriptII or III ReverseTranscriptase (LifeTechnologies) was used for first-strand cDNA synthesis. Reaction mix=0.2µM forward (F) and reverse (R) primers, 2µl cDNA, 0.2mM nucleotides and 1U Platinum®*Taq* Polymerase (LifeTechnologies) with 30-40 cycles (Bio-Rad rtPCRThermocycler): with primers as described in supplementary table 1. *β-actin* was used for normalization.

Supplemental Table 1. Primer sequences

Cloning	Fwd	Rev
left	5'AATCGATGCCGAGGGAGTTCAGTGCTAG 3'	5'AGATATCACAGTAATGGCGGCACAAGGC3'
right	5'ATACGTACAGGAAGGAAACATTCTCTGG3	5'ACTCGAGGTCTCAGGCAAGACTATGG3'
middle	5'CATATGGCCTGGAATCTGCGCAGGAC3'	5'GCGGCCGCAATATTTCTAACTGGGCTGC3'
Genotyping		
right arm junction	5'CAGCTTCAGCCTGCTTACTCA3'	5'TATAGACAAACGCACACCGGC3'
Gata2$^{-/-}$	GGAACGCCAACGGGGAC	GCTGGACATCTTCCGATTCCGGGT
qRT-PCR		GATCTCCTGTCATCTCACCTTGCT
Venus	F5'ATCTTCTTCAAGGACGACGG3'	R5'GGCTGTTGTAGTTGTACTCC3'
Gata2	F5' AAGCTGCACAATGTTAACAGG3'	R5'CCTTTCTTGCTCTTCTTGGAC3'
Gata1	F5'TGCCTGTGGCTTGTATCA3'	R5'TGTTGTAG GGTCGTTTGAC3'
Gata3	F5'CCTTATCAAGCCCAAGCGA3'	R5'GATGCCTTCTTTCTTCATAGTCAG3'
Gata4	F5'GAGCTGGTGCTACCAAGAGG3'	F5'GAGCTGGTGCTACCAAGAGG3'
Gata5	F5'CGACGTAGCCCCTTCGTGG3'	R5'GCCACAGTGGTGTAGACAG3'
Gata6	F5'GCCTGCGGCCTCTACATGAA3'	R5'CAGGACCTGCTGGCGTCTTA3'

Supplemental Figure 1

Representative multilineage FACS analysis of GataVenus$^+$ repopulated recipients. Recipients repopulated with Venus$^+$ cells from *G2V* E11 AGM were examined at 4mo post-transplantation for donor-derived cells. Bone marrow (BM), spleen, lymph node (LN), peripheral blood (PB) and thymus cells were isolated and examined for the Ly5.1 donor marker in multiple lineages. (Top panel) BM viable cells were gated into the following fractions - CD31$^+$Gr$^-$ =lymphoid; CD31$^-$Gr1$^-$ =erythroid; CD31$^+$Gr1$^+$ and CD31$^-$Gr1$^+$ =myeloid. (Second row panels) percentages of Ly5.1 (Y-axis) cells in gated erythroid, lymphoid and myeloid lineage fractions are indicated. (Third row, Left panel) Spleen viable cells were gated on B cells (B220$^+$) and T cells (CD4$^+$/CD8$^+$) and (Right panels) show percentages of Ly5.1 (Y-axis) donor-derived B and T cells. (Bottom row panels) Percentages of Ly5.1 donor-derived cells are shown for viable cells from LN, PB and Thymus. X-axis is side scatter (SSC-A) in all Ly5.1 plots.



In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate

Christina Eich,[1] Jochen Arlt,[3] Chris S. Vink,[4] Parham Solaimani Kartalaei,[1] Polynikis Kaimakis,[1] Samanta A. Mariani,[4] Reinier van der Linden,[1] Wiggert A. van Cappellen,[2] and Elaine Dzierzak[1,4]

[1]Department of Cell Biology, Erasmus Stem Cell Institute and [2]Department of Pathology, Erasmus Optical Imaging Centre, Erasmus Medical Center, Rotterdam, Netherlands
[3]School of Physics and Astronomy and [4]Centre for Inflammation Research, Queen's Medical Research Institute, The University of Edinburgh, Edinburgh, Scotland, UK

Cell fate is established through coordinated gene expression programs in individual cells. Regulatory networks that include the Gata2 transcription factor play central roles in hematopoietic fate establishment. Although Gata2 is essential to the embryonic development and function of hematopoietic stem cells that form the adult hierarchy, little is known about the in vivo expression dynamics of Gata2 in single cells. Here, we examine Gata2 expression in single aortic cells as they establish hematopoietic fate in *Gata2Venus* mouse embryos. Time-lapse imaging reveals rapid pulsatile level changes in Gata2 reporter expression in cells undergoing endothelial-to-hematopoietic transition. Moreover, Gata2 reporter pulsatile expression is dramatically altered in *Gata2*$^{+/-}$ aortic cells, which undergo fewer transitions and are reduced in hematopoietic potential. Our novel finding of dynamic pulsatile expression of *Gata2* suggests a highly unstable genetic state in single cells concomitant with their transition to hematopoietic fate. This reinforces the notion that threshold levels of Gata2 influence fate establishment and has implications for transcription factor–related hematologic dysfunctions.

INTRODUCTION

During a short window of developmental time, hematopoietic stem cells (HSCs) arise from the transdifferentiation of specialized endothelial cells (ECs) lining the major embryonic vasculature. In the mouse, this endothelial-to-hematopoietic transition (EHT) occurs at embryonic day (E) 10.5 and is best characterized by the emergence of clusters of hematopoietic stem and progenitor cells (HSPCs) from the aortic endothelium of the aorta-gonad-mesonephros (AGM) region (Dzierzak and Medvinsky, 2008; Dzierzak and Speck, 2008). The transition involves changes in the transcriptional program of a subset of (hemogenic) ECs to a program promoting HSPC identity. RNA-sequencing data from our group and others has shown that expression of a group of "heptad" transcription factors (TFs; Wilson et al., 2010; Lichtinger et al., 2012; Solaimani Kartalaei et al., 2015; Goode et al., 2016) increases during EHT (Solaimani Kartalaei et al., 2015), suggesting that heptad TFs could act as a transcriptional hub for the regulation of EHT.

Gata2, one of the heptad TFs, is crucial for the generation of HSCs. *Gata2* is expressed in the mouse embryo in the primitive streak, some ECs of the paired and midgestation dorsal aorta, and vitelline/umbilical arteries (Minegishi et al., 1999; Robert-Moreno et al., 2005; Kaimakis et al., 2016). At the time of definitive HSPC formation and during EHT, it is expressed in hemogenic ECs (HECs) and intra-aortic hematopoietic cluster cells (IAHCs). *Gata2*$^{-/-}$ embryos suffer from fetal liver anemia and die in midgestation at the time of HSC generation (Ng et al., 1994; Tsai et al., 1994; Orlic et al., 1995; Tsai and Orkin, 1997; Minegishi et al., 1999; Nardelli et al., 1999; Ling et al., 2004; Robert-Moreno et al., 2005; Khandekar et al., 2007; de Pater et al., 2013). *Gata2* heterozygous mutant (*Gata2*$^{+/-}$) embryos are severely affected in the production of early progenitors (Tsai et al., 1994) and have greatly reduced numbers of HECs, IAHCs, HPCs, and HSCs (Ling et al., 2004; Khandekar et al., 2007; de Pater et al., 2013; Gao et al., 2013). *Gata2*$^{+/-}$ HSCs are qualitatively defective (Ling et al., 2004; Rodrigues et al., 2005). Thus, Gata2 has distinct roles during the different stages of hematopoietic development and is a pivotal regulator of EHT cell transition, HSC generation, and function (de Pater et al., 2013). How Gata2 controls these different processes and how levels of Gata2 expression influence cell fate decisions remain elusive.

Recent studies have identified a growing list of TFs that show pulsatile dynamic behavior (Lahav et al., 2004; Nelson et al., 2004; Cai et al., 2008; Cohen-Saidon et al., 2009; Locke et al., 2011; Levine et al., 2013; Purvis and Lahav, 2013; Ryu et al., 2016; Zambrano et al., 2016). A pulse is detected when a critical threshold of TF molecules accumulate and ends when they are degraded/deactivated. The presence of pulsatile expression for various regulators in bacteria (Locke et al., 2011; Young et al., 2013), yeast (Garmendia-Torres et

Correspondence to Elaine Dzierzak: elaine.dzierzak@ed.ac.uk

C. Eich's present address is Translational Nanobiomaterials and Imaging, Dept. of Radiology, Leiden University Medical Center, Leiden, Netherlands.

The Rockefeller University Press
J. Exp. Med. 2018 Vol. 215 No. 1 233–248
https://doi.org/10.1084/jem.20170807

al., 2007; Dalal et al., 2014), and the mammalian stress response and signaling pathways (Lahav et al., 2004; Nelson et al., 2004; Kageyama et al., 2008; Cohen-Saidon et al., 2009; Kholodenko et al., 2010; Tay et al., 2010; Batchelor et al., 2011; Albeck et al., 2013; Yissachar et al., 2013) suggests that it is a common process. Pulsing may provide a time-based mode of regulation, where an input typically modulates the pulse frequency, amplitude, and/or duration of individual TFs to control downstream target gene expression. This dynamic behavior and pulsatile expression of TFs in single cells is implicated in cell transitions and fate decisions (Nelson et al., 2004; Shimojo et al., 2008; Kobayashi et al., 2009; Tay et al., 2010; Pourquié, 2011; Imayoshi et al., 2013; Kueh et al., 2013, 2016; Neuert et al., 2013; Stern and Piatkowska, 2015) and includes, for example the NF-κb and Notch signaling pathways (Kim et al., 2013; Levine et al., 2013; Purvis and Lahav, 2013; Isomura and Kageyama, 2014).

Although much information is emerging on transcriptomic signatures and molecules affecting the development of the hematopoietic system (Lichtinger et al., 2012; Swiers et al., 2013; Solaimani Kartalaei et al., 2015; Goode et al., 2016; Zhou et al., 2016), dynamic expression is still a largely unexplored area. We set out to examine the dynamics of *Gata2* expression during the establishment of hematopoietic fate in the aortic endothelium, because *Gata2* is a downstream target of the Notch pathway (Robert-Moreno et al., 2005; Gama-Norton et al., 2015) and is known to affect EHT (Kumano et al., 2003; Ling et al., 2004; de Pater et al., 2013), and the dosage of *Gata2* is of major importance for normal hematopoietic development (Ling et al., 2004; Khandekar et al., 2007; Tipping et al., 2009; de Pater et al., 2013; Gao et al., 2013). Here, we demonstrate for the first time the pulsatile expression of a *Gata2* reporter during the process whereby hematopoietic cells are generated from HECs. By vital imaging of single cells in the mouse embryonic aorta (WT and *Gata2* heterozygous mutant), we show that cell states during EHT correlate with *Gata2* reporter expression duration, levels (amplitude changes), and pulse periodicity, thus supporting the notion that Gata2 levels and dynamic behavior are linked to hematopoietic fate.

RESULTS

Gata2 reporter expression in single cells is dynamic

Cell populations undergoing EHT in the mouse embryonic aorta are characterized by their localization, morphology, and expression of pivotal markers and TFs, including Gata2 (Kaimakis et al., 2016). To specifically examine Gata2 expression in single cells undergoing EHT, confocal imaging was performed on immunostained E10.5 AGM from *Gata2Venus* (*G2V*) reporter embryos (Kaimakis et al., 2016) in fixed whole-mount (Fig. 1 A) and vital transverse-section preparations (Fig. 1 B). Venus fluorescence reports *Gata2* expression (*IRES-Venus* insertion in *Gata2 3′ UTR*) without disrupting normal *Gata2* transcription, protein stability or function, or hematopoietic development (Kaimakis et al., 2016). Importantly, Gata2 is a short-lived protein (30–60 min; Minegishi et al., 2005; Lurie et al., 2008). Rather than using the GFP reporter (half-life = 9 h), the Venus reporter was chosen for this study because its half-life (60–120 min) closely reflects real-time Gata2 expression dynamics (Li et al., 1998). Vital imaging was performed on 150-µm transversal sections of E10.5 *G2V* embryos preinjected with antibodies against CD31 (endothelial and IAHC marker) and ckit (IAHC marker; Boisset et al., 2010; Yokomizo et al., 2012; Solaimani Kartalaei et al., 2015).

Venus expression was found in single cells of the aortic endothelium, cells bulging from the endothelial wall, and IAHCs (Fig. 1, A and B), all of which are $CD31^+$. ECs ($CD31^+ckit^-$) are flattened $ckit^-$ cells in the vascular wall, and for this study, G2V-expressing (V^+) ECs are referred to as HECs. G2V-expressing $ckit^+$ cells undergoing a change in morphology as they emerge from the wall are referred to as bulging cells (BCs; $V^+CD31^+ckit^+$). IAHCs ($V^+CD31^+ckit^+$) are the rounded cells found in clusters adjacent to the endothelial layer (Fig. 1 C). Flow cytometric analysis (FACS) showed that varying levels (medium and high) of Venus expression could be detected in the $CD31^+ckit^+$ cells (Fig. 1 D). Upon sorting $CD31^+ckit^+$ $Venus^{med}$ and $Venus^{high}$ cells, RNA-sequencing analysis (Fig. 1 E) showed medium and high levels of *Gata2* transcripts, respectively. Furthermore, $Venus^{med}$ and $Venus^{high}$ expression levels correctly reflect Gata2 protein levels, as confirmed by Western blotting of sorted cell fractions from adult *G2V* bone marrow (Fig. S1 A). Equivalent ratios of quantified Gata2 to Venus protein signal were found for all the sorted cell populations. Hence, the G2V reporter allows the accurate tracking of *Gata2* expression in single live cells during EHT.

Vital imaging of *G2V* embryo transversal sections through the AGM was performed at 15-min intervals for 10–15 h (Videos 1, 2, and 3). Imaging data were examined for EHT events in which Venus-expressing (V^+) hematopoietic cells emerge from the aortic wall. In 15 independent time-lapse imaging experiments with a total of 49 sections, we observed 13 EHT events of $V^+CD31^+(ckit^+)$ hematopoietic cells emerging from the $V^+CD31^+(ckit^-)$ endothelium of E10 embryos (32–37 somite pairs [SPs]; Fig. 2 and Videos 1, 2, and 3). Taking into account the thickness of the embryo section and length of the aorta (forelimb to hindlimb where IAHC are found), we calculated that there are ∼20 EHT events per embryo. This is in contrast to the 1.7 EHT events per embryo previously observed in the *Ly6aGFP* reporter aorta imaging studies (Boisset et al., 2010). In the cases in which we imaged an EHT event, visual analysis of time-lapse images revealed changes in the mean fluorescent intensity (MFI) suggestive of flexible and pulsatile Gata2 expression in BCs (Fig. 2, A–C) and, to a lesser extent, IAHCs (Fig. 2 D and Fig. S3 F).

V^+ cells were counted in the first frame image of time-lapse experiments (n = 15; 1,126 cells). When the numbers were calculated per total aorta, 660 ± 87 V^+ cells were found in the endothelial layer (HECs), followed by 305

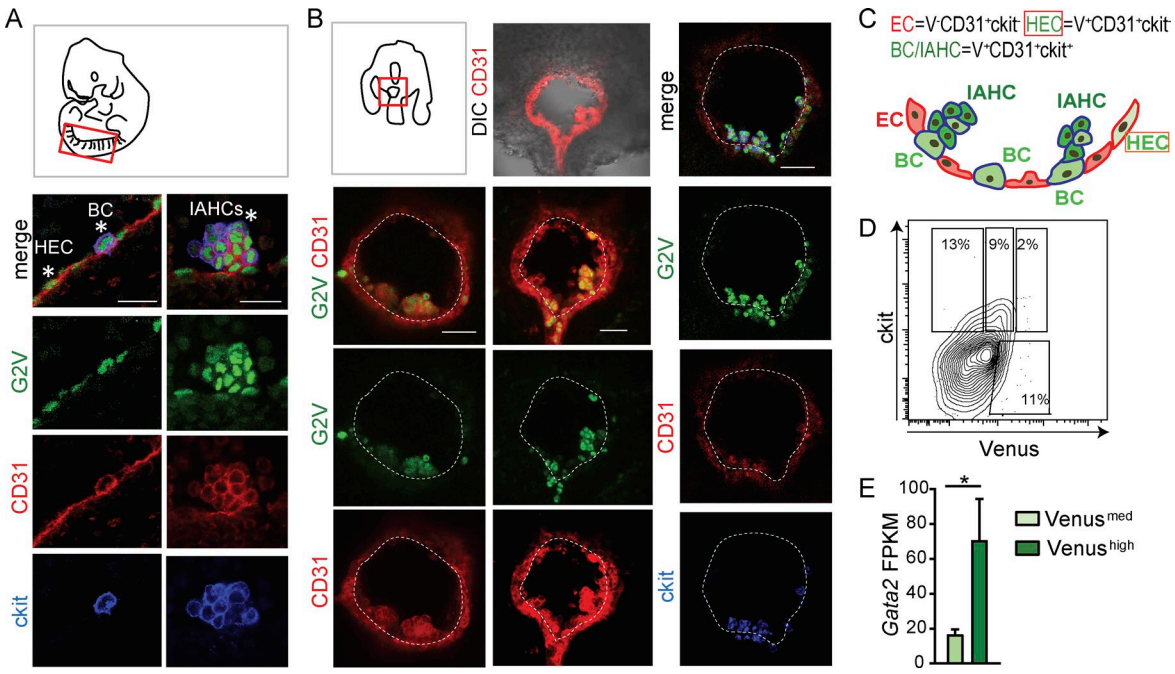


Figure 1. ***G2V* reporter reveals expression of Venus in different EHT cell subsets. (A)** Diagram of a whole mount of a 35-SP embryo with the dorsal aorta outlined in red and confocal images of the ventral aspect of the *G2V* dorsal aorta. HECs and BCs (left) and IAHCs (right) are indicated by asterisks (CD31, red; ckit, blue; and Venus, green). Bars, 20 µm. **(B)** Diagram of an embryo (33–34 SPs) transversal slice prepared for vital confocal time-lapse imaging. Confocal images of representative immunostained *G2V* embryo traversal sections (CD31, red; ckit, blue; and Venus, green). Bars, 40 µm. Ventral side downward. DIC, differential interference contrast. **(C)** Schematic representation of Gata2 (green) expression in the different EHT cell subsets in the E10.5 mouse dorsal aorta (ventral aspect; ECs, red, red outline; HECs, light green, red outline; BCs, medium green, blue outline; and IAHCs, dark green, blue outline). Cells with blue outline express ckit in addition to CD31. **(D)** Flow cytometric contour plot of CD31 gated cells. Percentages of $CD31^+ckit^+Venus^{med}$ and $CD31^+ckit^+Venus^{high}$ expressing cells from E10.5 *G2V* AGMs. **(E)** Bar graph of *Gata2* transcription in E10.5 AGM *G2V* sorted $CD31^+ckit^+Venus^{med}$ (light green) and $CD31^+ckit^+Venus^{high}$ (dark green) cells. y axis shows FPKM values obtained from RNA-sequencing analysis. The data represent the mean ± SEM of three independent experiments and were compared using a Student's *t* test (*, P = 0.0431).

±131 in BCs (7% of which undergo EHT) and 266 ±132 in IAHCs. Highly sensitive FACS of E10.5 *G2V* AGMs confirmed the microscopy results, showing the highest numbers of V^+ cells in the aortic endothelium ($CD31^+ckit^-$; 1,076 HECs) and fewer V^+ cells in the $CD31^+ckit^+$ hematopoietic population (680 BCs and IAHCs; Table 1). These numbers are higher than what has been published previously for the *Ly6aGFP* reporter in the E10.5 AGM (831 HECs, and 261 BCs and IAHCs at 34 SPs; Solaimani Kartalaei et al., 2015), indicating that G2V expression is encompassing more EHT cells than the *Ly6aGFP* reporter (Fig. S1 B). Additionally, the majority of V^+ HECs, BCs, and IAHCs are found on the ventral side of the aorta, with only 23% of V^+ cells on the dorsal side (Fig. 3 A).

Expression dynamics differ among HECs, BCs, and IAHCs

Time-lapse imaging allowed us to follow Venus expression up to 15 h without confounding bleaching effects. To quantitate Gata2 expression levels and dynamics during the transition of HECs to BCs and BCs to IAHCs, confocal images of single aortic cells were deconvoluted to improve the signal to noise ratio and analyzed using commercial and custom-made tools for tracking cells in four dimensions (Fig. S2). Only cells tracked for at least 10 consecutive frames were included, and tracking in three dimensions guaranteed that the observed changes in Venus intensity were not due to cells moving in or out of the imaging plane (Fig. 3 B). Venus MFI values plotted over time (Fig. S2; details in Materials and methods) showed expression in individual cells to be dynamic (Videos 1, 2, and 3), and as shown in Fig. 2, EHT was accompanied by increased, decreased, and/or alternating levels of Venus expression. Changes in Venus MFI were observed in both the raw and deconvoluted data, ruling out the possibility that deconvolution introduced artifacts (Fig. S3, A and B).

To examine whether expression levels in EHT subsets differed, MFI values of individual $Venus^+$ cells in each subset were averaged for 10 consecutive time points (49 sections analyzed, 718 cells tracked). MFI values were significantly higher in IAHCs (50.3 ± 1.2) than in HECs (40.2 ±0.9; ***, P < 0.0001) and BCs (45.1 ± 1.7, **, P < 0.001; Fig. 3 C). MFI plots of EHT subsets according to their ventral or dorsal location revealed significantly higher levels of Venus expression in ventral HECs (*, P = 0.023), ventral BCs (**, P = 0.0036), and ventral IAHCs

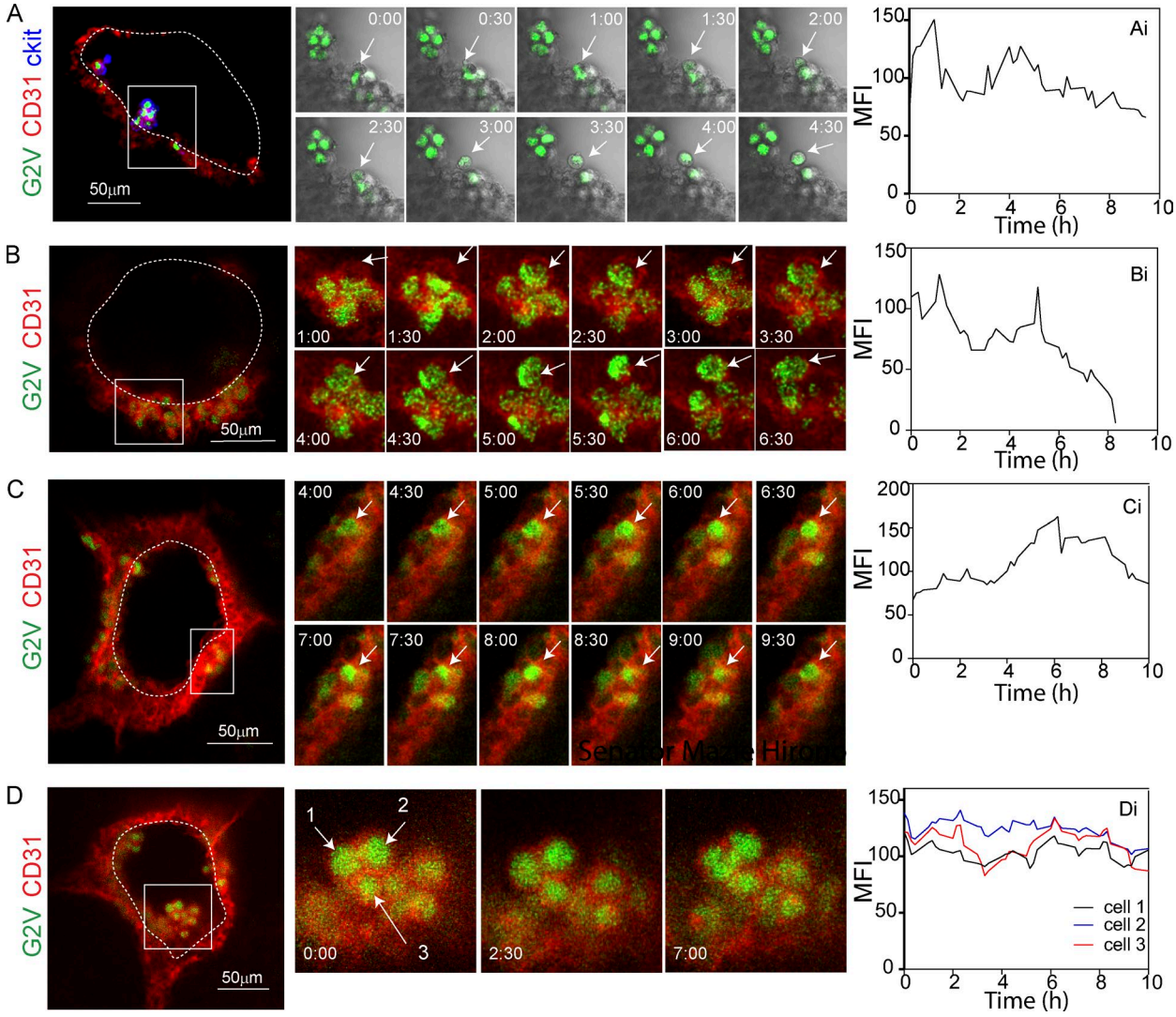


Figure 2. **Dynamic expression of the Gata2 reporter during EHT. (A–C)** Confocal time-lapse imaging of E10 (33–34 SPs) *G2V* embryos (Venus, green) immunostained with anti-CD31 (red) and anti-ckit (blue) antibodies (A) or anti-CD31 (red) antibody only (B and C). Arrows indicate cells undergoing EHT events. **(D)** Confocal time-lapse imaging of IAHC at E10 (33–34 SPs; Venus, green), immunostained with anti-CD31 (red) antibody. Arrows indicate three IAHC cells. **(Ai–Di)** Quantification of Venus MFI over time (hours) corresponding to the highlighted cells during the process of EHT (Ai–Ci) and in IAHC (Di). Transverse aortic sections were imaged for 10 h at 15-min intervals. A–C (middle) show 30-min snapshots, and D shows snapshots at 0, 2.5, and 7 h. Bars, 50 µm.

(***, P = 0.0001) than in dorsal ECs, dorsal BCs, and dorsal IAHCs, respectively (Fig. 3 D). Moreover, Venus MFI was significantly higher in ventral BCs (*, P = 0.0288) and ventral IAHCs (***, P < 0.0001) than in ventral HECs, suggesting that EHT cell identity is related to the level of *Gata2* expression.

To ensure that the differences in Venus MFI were not caused by noise inherent to the microscopy procedure, we

Table 1. **Absolute cell numbers in each EHT phenotypic subset in the AGM as determined by flow cytometry**

Gata2 genotype	Stage (sp)	$CD31^+$	$CD31^+$ $ckit^+V^-$	$CD31^+$ $ckit^-V^+$	$CD31^+$ $ckit^+V^{med}$	$CD31^+$ $ckit^+V^{hi}$	Number of experiments (embryos)
+/+	E10 (28–36)	10,860 ± 1,704	1,257 ± 264	1,076 ± 153	616 ± 223	64 ± 16	*n* = 5 (25)
+/–	E10 (30–36)	11,664 ± 1,992	1,049 ± 258	780 ± 179	186 ± 60	32 ± 6	*n* = 3 (19)

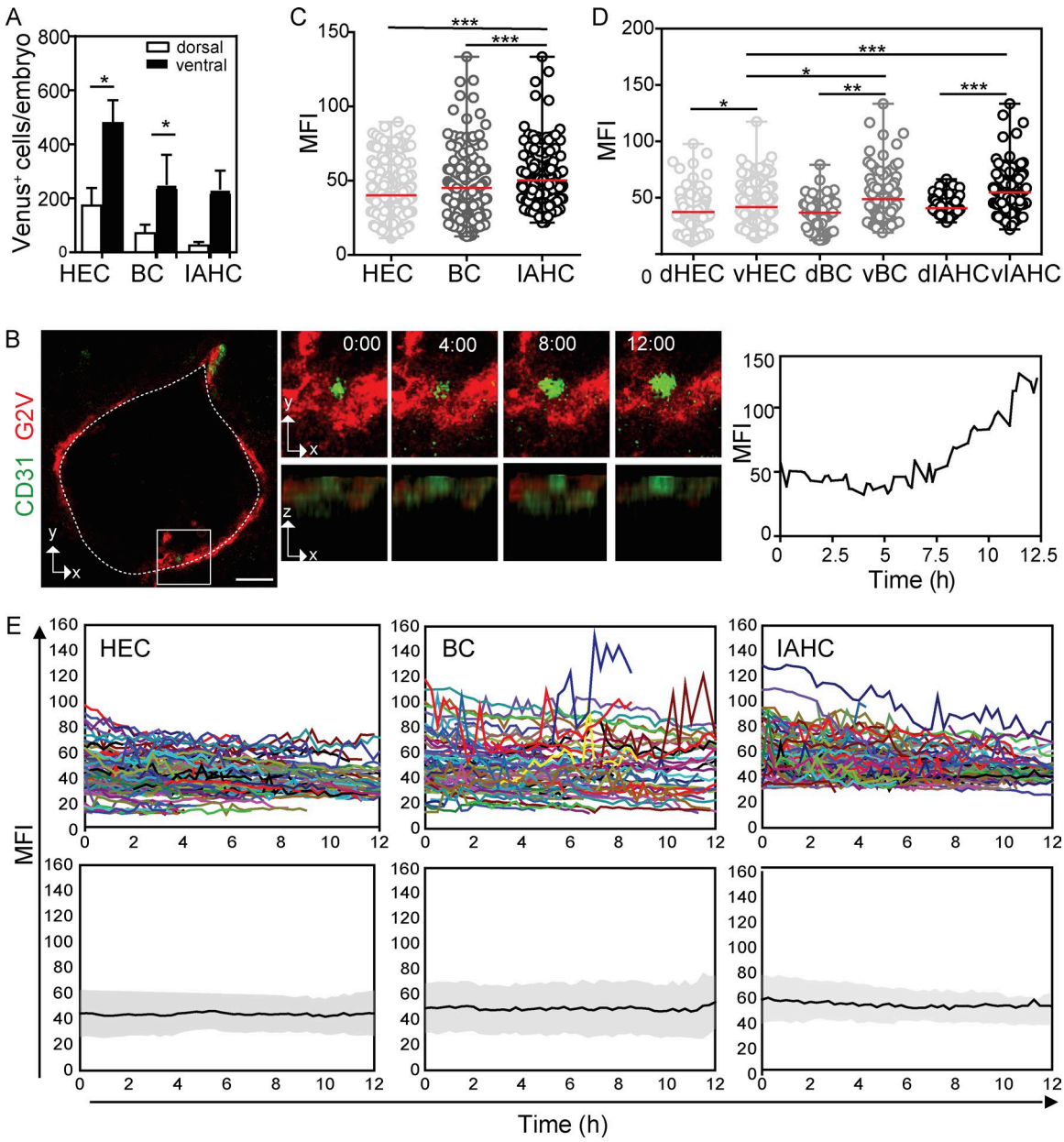


Figure 3. **Time-lapse imaging reveals differences in Gata2 dynamics between ECs, BCs, and IAHCs. (A)** Mean number of Venus$^+$ EHT subset cells per *G2V* embryo according to their ventral and dorsal aortic location, as determined by microscopy. Venus$^+$ HECs, BCs, or IAHCs were counted in the first frame of time-lapse imaging experiments (n = 15, 42 sections) of *G2V* embryo slices of 150 µm thickness and calculated per embryo (E10, 32–37 SPs). The data represent the mean ± SEM of 15 independent experiments, and dorsal and ventral location were compared using two-way ANOVA with Bonferroni post test (*, $P < 0.05$). **(B)** Left: Example of a BC showing decreasing and increasing levels of Venus expression during a 12-h imaging period, imaged at a time interval of 15 min. Middle: Three-dimensional projections (x-y axes and x-z axes) of the same BC as shown in the top panel with time (hours) indicated. Right: Corresponding Venus (green) MFI profile in time. Bar, 25 µm. Sections were stained with anti-CD31 (red) antibody. **(C)** Venus MFI (averaged over frames 3–12) in single EHT subset cells (HECs, BCs, and IAHCs; n = 15, 718 cells). The data represent the mean ± range. Statistical significance was calculated on the pooled data of 15 independent experiments using Mann–Whitney U test (***, $P < 0.0001$). **(D)** Venus MFI in single EHT subset cells plotted according to their dorsal (d) and ventral (v) location (n = 15, 718 cells). The data represent the mean ± range. Statistical significance was calculated on the pooled data of 15 independent experiments using Mann-Whitney U test (*, $P < 0.0288$; **, $P = 0.0020$; ***, $P < 0.0001$). **(E)** Top: Temporal variation of Venus MFI for individual Venus$^+$ cells (colored lines) plotted according to their affiliation to one of the EHT subsets (EC, BC, or IAHC). Bottom: Gray bands represent the standard deviation of the mean MFI of all tracks (black line) over time.

performed time-lapse imaging of aorta sections that were labeled with DRAQ5, a fluorescent DNA-binding dye that should not undergo intensity changes during imaging. DRAQ5 was titrated to reach similar MFIs as Venus (Fig. S3 C). No dramatic intensity changes in DRAQ5 MFI were observed (Fig. S3 D). Visual inspection of the Venus intensity profiles of single HECs, BCs, and IAHCs over time (Fig. 3 E, top; and Fig. S3 F) revealed heterogeneity and differences in expression dynamics that are not observed in population analyses of these EHT subsets (Fig. 3 E, bottom).

The single cell analyses of Venus MFI showed higher amplitude and pulsatile changes in BCs and IAHCs than in HECs, indicating activated but unstable Gata2 expression as cells transit to hematopoietic fate. Because it was previously reported that Gata2 levels decrease during mitosis (Koga et al., 2007), we monitored our time-lapse videos for proliferation events. Although cell division was observed in 14% of IAHCs (Fig. S4), very few BCs (0.3%) and 0% of HECs divided (n = 15). As expected, Venus expression in IAHCs decreased during cell division. Because 3.5-fold more IAHCs (50%) showed fluctuating Venus expression than underwent cell division, and because no or very few proliferation events were detected in HECs and BCs during the imaging period, it is unlikely that cell division is responsible for the pulsatile behavior of Gata2 expression that we observe during the HEC to BC transition. To control for this, our results on Gata2 dynamic expression exclude cells undergoing mitosis.

Gata2 reporter pulse amplitude and periodicity distinguishes EHT subsets

Pulsatile behavior of regulatory molecules that relay information relevant to biological systems are characterized by their amplitude and periodicity of expression and/or activation state (Pourquié, 2011; Purvis and Lahav, 2013; Isomura and Kageyama, 2014). Expression amplitude is the maximal (peak) value a regulatory molecule attains during the observation period, whereas oscillation periodicity indicates the time between two adjacent peaks (Fig. 4 A). To obtain the quantitative changes (trough-to-peak amplitude) in Venus expression, the fold change between the MFI at the peak and at the preceding trough was calculated. We developed an automated data-processing methodology to quantify amplitude and pulse periodicity in individually tracked cells (for details, see Materials and methods).

To discriminate noise from real Venus peaks, we set a threshold in which the peak Venus MFI differs from its neighboring minima by ≥15% of the mean intensity of the track. We could detect zero, one, two, and three peaks in Venus MFI profiles for individual EHT cells tracked over at least 10 consecutive time frames (Fig. 4 B). The majority of cells showed constant Venus MFIs (0 peaks, 57%) through the imaging period, and two peaks were found for 21% of cells (Fig. S3 E). To account for the difference in track lengths of the individual cells imaged, we normalized the pulse data and calculated the occurrence of peaks per 2.5 h (10 frames). Within the EHT subsets (Fig. 4 C), a significantly higher percentage of HECs showed no peaks (68%) compared with BCs (56%: *, $P < 0.05$) and IAHCs (49%: ***, $P < 0.001$). Thus, Gata2 reporter expression showed a greater pulsatile behavior in BCs and IAHCs than in HECs (Fig. 4 C).

The pulse periodicity was calculated for all Venus$^+$ cells (Fig. 4 D) and individual EHT subset cells (Fig. 4 E) showing at least two peaks. Approximately 55% of Venus$^+$ cells showed pulse periodicities of 2–3 h (Fig. 4 D). EHT subset cells showed differing Venus periodicities (Fig. 4 E). As compared with HECs, more BCs showed periodicities of 1 h (20.4% vs. 7.1%: ***, $P < 0.001$) and 2 h (42.9% vs. 30.3%: ***, $P < 0.001$). Fewer BCs and IAHCs showed 3-h pulse periodicities (18.4%: ***, $P < 0.001$; and 25.5%: **, $P < 0.01$, respectively) as compared with HECs (32.1%). The majority of HECs (62%) and IAHCs (56%) showed peak Venus expression periodicities of 2 to 3 h, whereas the majority of BCs (63%) showed shorter pulse periodicities of 1 to 2 h, thus indicating a high rate of dynamic expression specifically in BC after emergence from the endothelium.

Because *Gata2* dosage is known to be important for the normal production of IAHCs and functional HSPCs, the maximal fluorescent protein abundance reached within an expression pulse was also calculated. In individual cells, the mean Venus peak MFI was higher in BCs (53 ± 2.3) and IAHCs (51 ± 2.3) than in HECs (46 ± 1.5). Ventrally localized cells showed higher Venus peak MFIs than dorsal cells (Fig. 4 F). Interestingly, and in contrast to the peak MFIs, trough-to-peak amplitude measurements (Fig. 4 G) showed significantly higher fold changes in expression in BCs than in HECs (***, $P = 0.0008$) and IAHCs (***, $P = 0.0008$). 12.8% of ventral BCs and 1.5% of dorsal BCs showed a trough-to-peak amplitude higher than twofold, in contrast to 0.5% of ventral HECs and 0% dorsal HECs. In IAHCs (ventral and dorsal), only 6% of cells showed trough-to-peak amplitudes higher than twofold. Thus, the degree of trough-to-peak amplitude changes in Venus expression in BCs suggest that the upstream and downstream signals will be variable in this EHT subset and could provide an explanation for the known phenotypic/functional heterogeneity of hematopoietic cells.

Gata2 reporter levels, pulsatile behavior, and hematopoietic functions are interrelated

Recently, we showed that most AGM HPCs and all HSCs are Gata2 expressing (Kaimakis et al., 2016). By FACS, CD31 and ckit expression discriminates ECs (CD31$^+$ckit$^-$V$^-$), HECs (CD31$^+$ckit$^-$V$^+$), and BCs/IAHCs (CD31$^+$ckit$^+$V$^+$; Fig. 1 D). In line with our imaging data, CD31$^+$ckit$^+$V$^+$ cells can be further divided into CD31$^+$ckit$^+$V^{med} and CD31$^+$ckit$^+$V^{high} fractions and assessed for hematopoietic function (Fig. 5, A and B). HPCs in E10 *G2V* AGMs were found to be highly enriched (71%) in the CD31$^+$ckit$^+$V^{med} fraction, with 14% in the CD31$^+$ckit$^+$V^{high} fraction (Fig. 5 A). Importantly, multipotent progenitors were highly enriched in the CD31$^+$ckit$^+$V^{med} fraction, and in vivo transplantations revealed

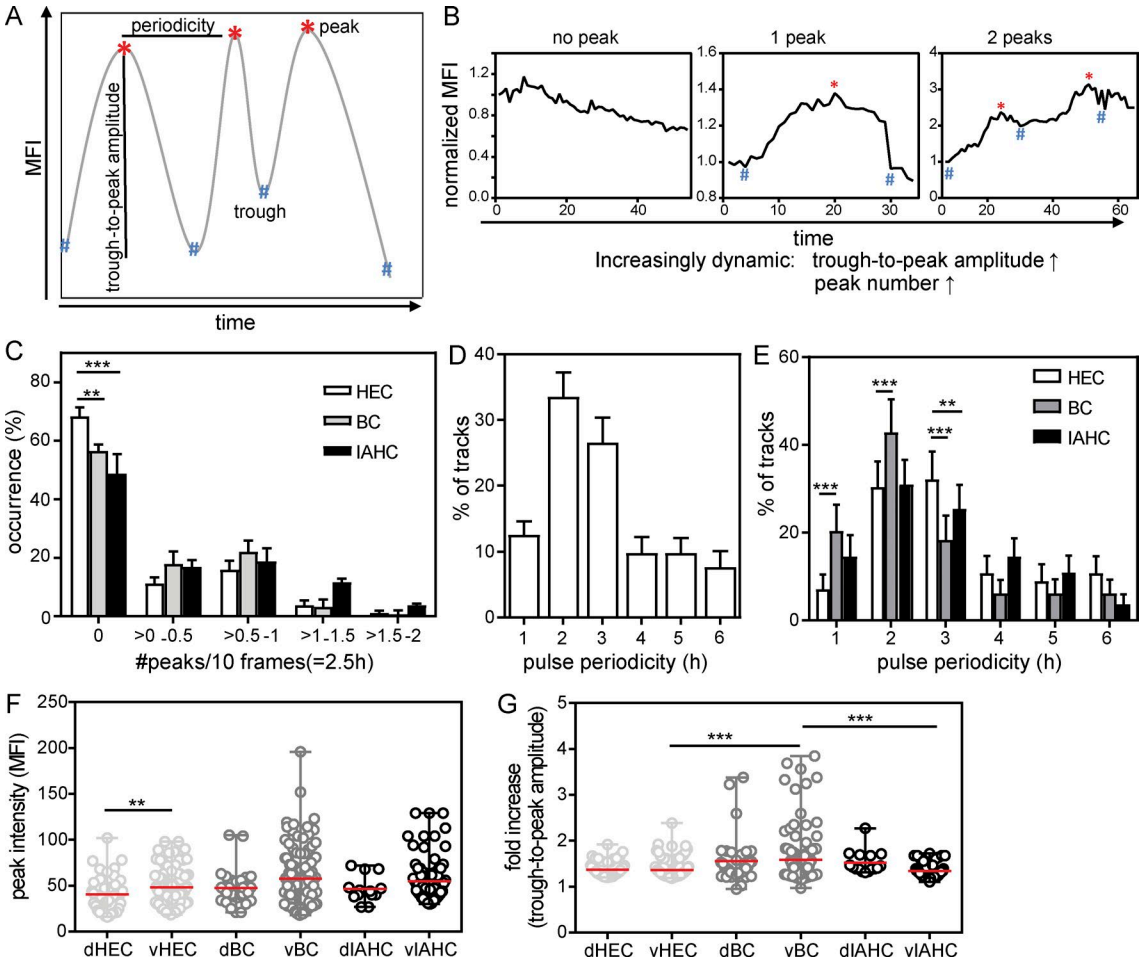


Figure 4. **Pulse frequency and amplitude of Venus expression distinguishes EHT subsets. (A)** Schematic representation of the automatic peak detection code. A local MFI maximum is considered a peak if it has at least a 15% higher intensity than its neighboring minima (see Materials and methods). The pulse period is the time between two adjacent peaks and the trough-to-peak amplitude the change between peak (highest value) and the preceding trough (lowest value). **(B)** Examples of normalized MFI profiles with no peak, one peak, and two pulse peaks showing increasing trough-to-peak amplitudes. **(C)** Distribution of the occurrence (percentage) of normalized pulse peak numbers in ECs, BCs, and IAHCs tracked over at least 10 consecutive frames (718 cells). To normalize for differences in track length, the data are presented as peaks per 10 frames (2.5 h) and represent the mean ± SEM ($n = 15$). Statistical significance was calculated using two-way ANOVA with Bonferroni post test (**, $P < 0.01$; ***, $P < 0.001$). **(D)** Distribution of the pulse periodicities of Venus$^+$ cells showing at least two pulse peaks ($n = 15$, 221 cells). **(E)** Distribution of the pulse periodicities in EHT subset cells showing at least two pulse peaks ($n = 15$, 86 HECs, 80 BCs, and 55 IAHCs). The data represent the mean ± SD. Statistical significance was calculated on the pooled data ($n = 15$) using two-way ANOVA with Bonferroni post test (**, $P \leq 0.01$; ***, $P < 0.001$). **(F and G)** Peak intensity (F) and trough-to-peak amplitude (G) in the EHT cell subsets, plotted according to their ventral (v) or dorsal (d) location in the aorta ($n = 13$, cells showing at least one peak: 170 HECs, 151 BCs, and 65 IAHCs). The data represent the mean ± range. Statistical significance was calculated on the pooled data ($n = 15$) using Mann–Whitney *U* test (**, $P = 0.0054$; ***, $P < 0.0008$).

that only this fraction contained HSCs (Fig. 5 B). Taken together, it is likely that the V^{med} cell fraction contains both BCs and IAHCs with multipotent hematopoietic activity.

Molecular characterization of these cell fractions (RNA sequencing) revealed differential expression of 1,321 genes, of which 1,089 genes showed down-regulated expression and 232 genes up-regulated expression in the CD31$^+$ckit$^+$V^{high} fraction. The Ingenuity Pathway Analysis tool revealed that the 232 up-regulated differentially expressed genes in the CD31$^+$ckit$^+$V^{high} fraction were significantly overrepresented in canonical pathways expressed by mature myeloid cell types (innate immune; Fig. 5 C). In contrast, the CD31$^+$ckit$^+$V^{med} cells showed enrichment for genes involved in leukocyte extravasation and epithelial adherence junction pathways. These data support the functional data to indicate that CD31$^+$ckit$^+$V^{high} cells are more differentiated hematopoietic cells and the CD31$^+$ckit$^+$V^{med} cells are immature progenitors and stem cells. Considering that the Notch signaling pathway is involved in hematopoietic cell development and IAHC formation (Kumano et al., 2003; Guiu et al., 2013) and that *Gata2* is a direct Notch target (Robert-Moreno et al., 2005), pathway-component analysis was performed. Sig-

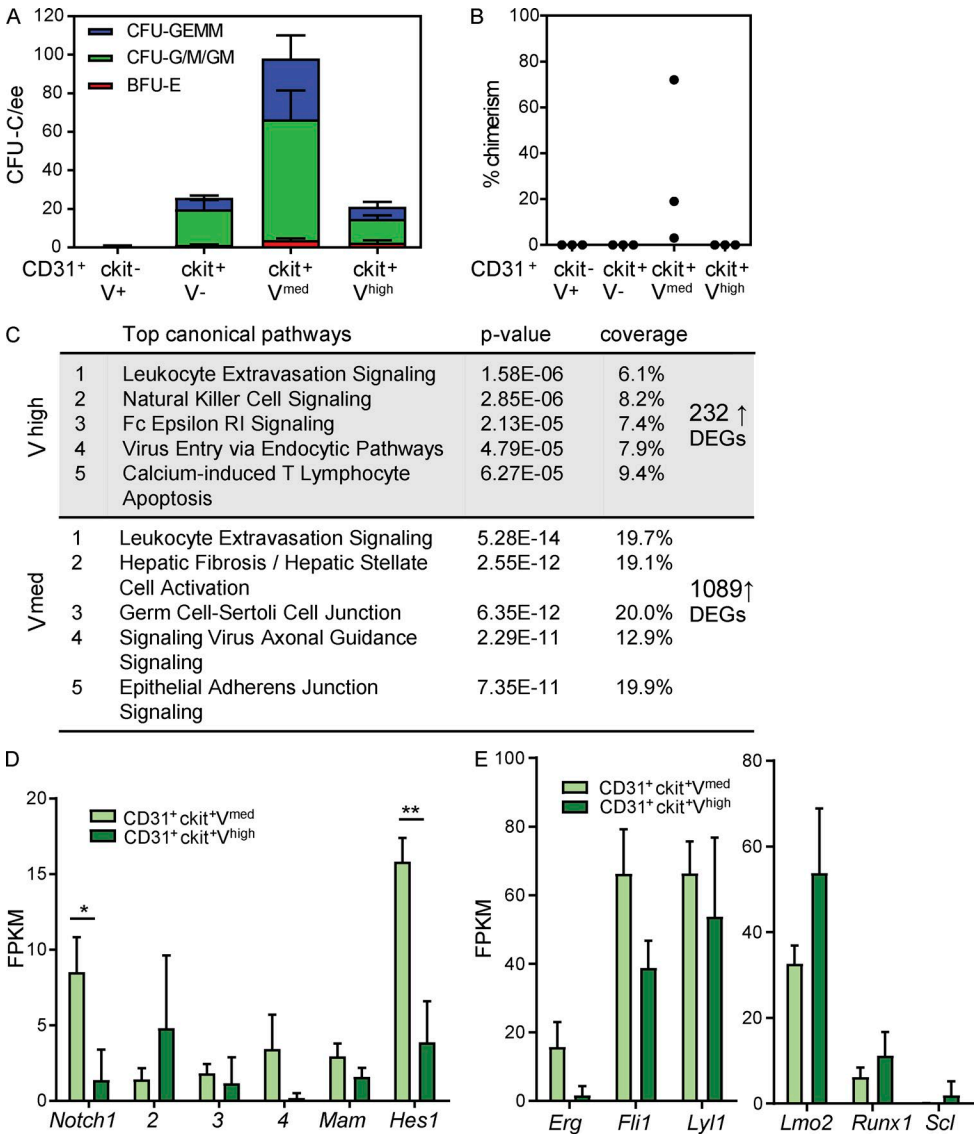


Figure 5. **Hematopoietic potential correlates with Venus expression levels. (A)** Hematopoietic progenitor numbers in E10.5 (30–36 SPs) CD31$^+$ ckit$^+$V^{med} and CD31$^+$ckit$^+$V^{high} AGM sorted cells. Colony-forming unit-culture per embryo equivalent (CFU-C/ee) is shown, with colony types designated by colored bars. BFU-E, burst-forming unit erythroid; CFU-G/M/GM, CFU granulocyte, CFU macrophage, and CFU granulocyte-macrophage; CFU-GEMM, CFU granulocyte, erythroid, macrophage, megakaryocyte. The data represent the mean ± SEM of four independent experiments. **(B)** Percentage donor cell chimerism in recipient mice injected with CD31$^+$ckit$^-$V$^+$, CD31$^+$ckit$^+$V$^-$, CD31$^+$ckit$^+$V^{med}, or CD31$^+$ckitV^{high} sorted E11 (41–49 SPs) AGM cells. Engraftment at 4 mo after transplantation was determined by flow cytometric analysis of Ly5.1/Ly5.2 marker expression of peripheral blood cells. Each dot represents one recipient receiving 1.3 to 4.1 embryo equivalent (ee) of sorted AGM cells. The data represent the mean ± SD. *, P ≤ 0.024; **, P = 0.0085; ***, P = 0.0003). **(C)** Overrepresentation of up-regulated differentially expressed genes (DEGs) in E10.5 CD31$^+$ckit$^+$V^{med} and CD31$^+$ckitV^{high} sorted cells in IPA canonical pathways. **(D and E)** Mean FPKM values for genes in the Notch pathway (D) and heptad factor genes in E10.5 CD31$^+$ckit$^+$V^{med} and CD31$^+$ckitV^{high} sorted E10.5 AGM cells (E). The data were compared using Student's *t* test (*, P = 0.0404; **, P = 0.0096). The data represent the mean ± SEM of three independent experiments.

nificantly higher expression of *Notch1* and its target gene, *Hes1*, was found in the CD31$^+$ckit$^+$V^{med} fraction as compared with the CD31$^+$ckit$^+$V^{high} fraction (Fig. 5 D), supporting a role for Notch in V^{med} BCs and/or IAHCs. As expected, the heptad hematopoietic TF genes were expressed in both fractions (Fig. 5 E); expression of *Erg*, *Fli1*, and *Lyl1* was lower in CD31$^+$ckit$^+$V^{high} cells, and expression of *Lmo2*, *Runx1*, and *Scl* was higher in CD31$^+$ckit$^+$V^{high} cells. These data indicate a degree of molecular heterogeneity within V$^+$ emerging hematopoietic cells.

To further examine whether Gata2 pulse periodicity, trough-to-peak amplitude, and hematopoietic functions in the aorta are related, we crossed *G2V* (*G2$^{V/V}$*) and *Gata2$^{+/-}$* mice (Tsai et al., 1994; C-terminal zinc-finger deletion) to

obtain embryos (*Gata2*$^{V/-}$) with one mutated and one functional allele of *Gata2* (Fig. S5 A). It is known that *Gata2* heterozygous mutant embryos have a greatly reduced number of IAHCs, HPCs, and HSCs (Tsai and Orkin, 1997; Ling et al., 2004; Khandekar et al., 2007; de Pater et al., 2013; Gao et al., 2013), and HSCs are qualitatively defective (Ling et al., 2004; Rodrigues et al., 2005). As found by vital imaging (n = 6, 18 sections), the number of V$^+$ IAHCs and BCs was lower in *Gata2*$^{V/-}$ aortas than in *Gata2*$^{V/+}$ aortas, whereas the number of V$^+$ HECs was similar between *Gata2*$^{V/-}$ and *Gata2*$^{V/+}$ aortas (Fig. 6 A; compare Fig. S5 B with Fig. 3 A). Venus$^+$ BC and IAHC were almost exclusively on the ventral side. In line with our microscopy data, FACS analysis of *Gata2*$^{V/-}$ aortas showed reduced numbers of CD31$^+$ckit$^+$V$^+$ cells (Table 1). To examine Gata2 protein levels, we sorted E10.5 *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ AGM cells into V$^+$ and V$^-$ fractions and performed Western blotting (Fig. S5, C and D). Equal levels of Gata2 protein were found in *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ AGMs. Upon examination of the MFI values of all imaged V$^+$ cells in *Gata2*$^{V/-}$ aortas (Fig. 6 B), no differences in MFI were detected between *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ HECs. However, in contrast to *Gata2*$^{V/+}$ aortas, where Venus expression was highest in IAHCs, Venus expression in *Gata2*$^{V/-}$ aortas was highest in BCs. In the few remaining *Gata2*$^{V/-}$ IAHCs, Venus expression was lower than in *Gata2*$^{V/+}$ IAHCs.

Further analysis of Gata2 reporter pulsatile expression parameters showed no difference in peak number distribution between *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ HECs; in both cases, 30% of HECs showed pulsatile expression (Fig. S5 E). A trend toward reduced numbers of BCs with pulsatile expression was found in *Gata2*$^{V/-}$ embryos (31%) as compared with *Gata2*$^{V/+}$ embryos (46%; Fig. S5 F). Despite no peak number differences in HECs, 51% of *Gata2*$^{V/-}$ HECs showed pulse periodicities of ≤2 h as compared with 25% in *Gata2*$^{V/+}$ HECs (Fig. 6 C), indicating that bursts of Gata2 expression are more frequent in *Gata2*$^{V/-}$ HECs. A similar trend toward reduced pulse periodicities was also found in *Gata2*$^{V/-}$ BCs as compared with *Gata2*$^{V/+}$ BCs (Fig. 6 D). Because IAHC were highly reduced in *Gata2*$^{V/-}$ aortas, we could not image sufficient numbers of IAHCs with pulsatile characteristics to reliably calculate periodicities. Within a pulse, *Gata2*$^{V/-}$ HECs reached higher peak intensities (52.6 ± 2.7) than *Gata2*$^{V/+}$ HECs (45.8 ± 1.5, **, P = 0.0056; Fig. 6 E). BCs and IAHCs showed no peak MFI differences between *Gata2*$^{V/-}$ and *Gata2*$^{V/+}$ cells. The fold increase in trough-to-peak amplitude (Fig. 6 F) in *Gata2*$^{V/-}$ HECs did not change compared with *Gata2*$^{V/+}$ HECs. However, 11% of *Gata2*$^{V/+}$ BC showed trough-to-peak amplitudes higher than twofold, whereas only 4% of *Gata2*$^{V/-}$ BCs showed values above twofold. Among the few *Gata2*$^{V/-}$ IAHCs with at least one peak, the trough-to-peak amplitudes were similar to the values observed in *Gata2*$^{V/+}$ IAHCs. Together, our results show that Venus expression levels and pulsatile characteristics are altered during EHT in *Gata2* heterozygous mutant embryos as compared with embryos with normal levels of *Gata2* expression.

DISCUSSION

We have uncovered a new level of dynamic regulation involving the pulsatile expression of the pivotal *Gata2* TF during the establishment of hematopoietic cell fate in the embryo. Although genetic experiments implicate a role for *Gata2* in EHT cell populations, vital imaging of *G2V* EHT cells reveals for the first time pulsatile expression at the single cell level. Pulse parameters, as characterized by amplitude and periodicity of Venus expression in individual cells differs between the EHT subsets (Fig. 6 G). The HEC to BC transition is accompanied by an increase in reporter expression levels and increased pulsatile behavior. Expression further increases and stabilizes during the transition to IAHCs, with the periodicity and amplitude decreasing. Our results suggest that the high degree of pulsatile Gata2 expression in BCs is linked to cell fate transition during EHT and may reflect an active process involving the partial assembly of counterbalancing regulatory states (Kueh et al., 2016). This is supported by pulsatile and level changes in Gata2 expression that accompany a Gata2 heterozygous mutant state in which EHT is disrupted.

Imaging dynamic cell transitions and Gata2 expression during EHT

Importantly, we used a *G2V* mouse model that does not disrupt *Gata2* expression levels or the function of the Gata2 protein (Kaimakis et al., 2016). The recombination of an *IRES-Venus* fragment into the 3′ UTR avoids hypomorphic *Gata2* expression and protein dysfunction that may result from a fusion protein. We showed previously that mice with two *G2V* alleles are normal in terms of HSC numbers and function. Gata2 protein has a relatively short half-life of 30–60 min (Minegishi et al., 2005; Lurie et al., 2008) as compared with other hematopoietic TFs such as Runx1 (3.3 h; Lorsbach et al., 2004) and Gata1 (4 to >6 h; Minegishi et al., 2005; Lurie et al., 2008). Its instability is related to ubiquitination (Minegishi et al., 2005; Lurie et al., 2008). The Venus reporter used in the *G2V* model has a half-life of ∼120 min (Li et al., 1998) and provides an excellent reporter of promoter activity, as it has a very short fluorescent protein formation (folding) time as compared with GFP (Snapp, 2009).

Our imaging (Fig. 2) and FACS (Table 1) experiments showed more cells undergoing EHT in the *G2V* embryonic aorta than previously described for the *Ly6aGFP* model (Boisset et al., 2010; Solaimani Kartalaei et al., 2015). At E10, Gata2 is expressed in ∼1,076 aortic ECs (CD31$^+$ckit$^-$) and ∼680 IAHC (CD31$^+$ckit$^+$), whereas Ly6aGFP is expressed in two to seven times fewer aortic ECs (190–831) and IAHCs (97–261). This is at the time when IAHCs peak and indicates that Gata2 expression marks more ECs with hemogenic potential and most if not all IAHCs. This is supported by functional data in which ∼80% of E11 AGM CFU-C are V$^+$ (Fig. 5 A), whereas only 33% are GFP$^+$ (Solaimani Kartalaei et al., 2015). However, all AGM HSCs are V$^+$ and GFP$^+$ (Solaimani Kartalaei et al., 2015). Thus, Ly6aGFP is a developmentally later marker, and Ly6aGFP-expressing cells

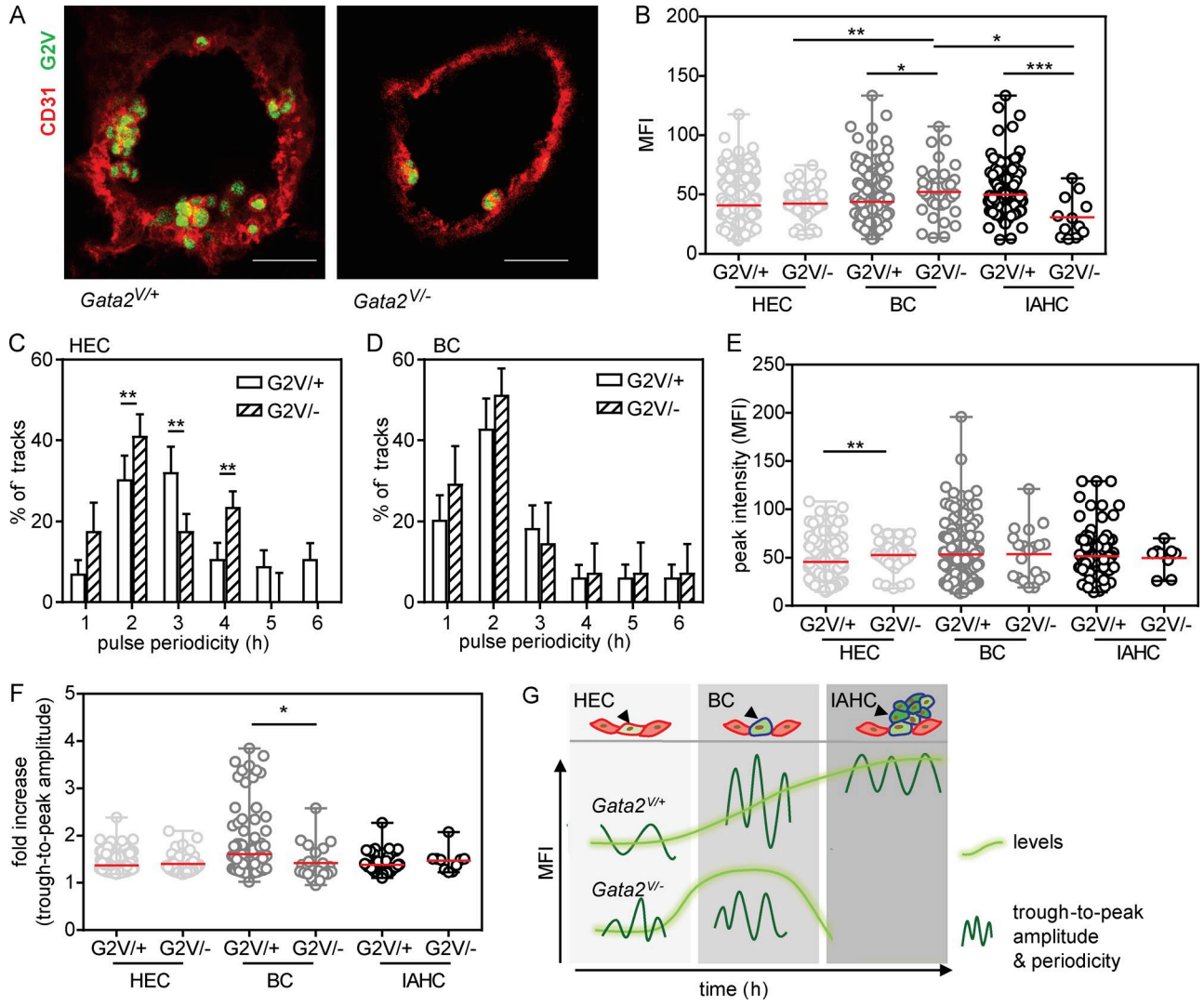


Figure 6. **Gata2 expression parameters, hematopoietic fate, and EHT are interrelated. (A)** Maximum projections of confocal time-lapse images of E10.5 *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ aortas immunostained with anti-CD31 (red; G2V, green). Bars, 40 µm. *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ embryos were harvested from the same litter. Ventral side downward. **(B)** Venus MFI (averaged over frames 3–12) in single *Gata2*$^{V/-}$ EHT subset cells (n = 6; 75 ECs, 37 BCs, and 12 IAHCs). The data were compared with *Gata2*$^{V/+}$ EHT subset cells using one-way ANOVA with Bonferroni post test (mean ± SD; *, P ≤ 0.024; **, P = 0.0085; ***, P = 0.0003). **(C and D)** Distribution of the pulse periodicities in Venus^{+} EHT subset cells: HECs (C) and BCs (D) from E10.5 *Gata2*$^{V/-}$ aortas showing at least two pulsatile peaks (n = 6; 18 HECs and 15 BCs). The data represent the mean ± SD. The data were compared with *Gata2*$^{V/+}$ HECs and BCs using two-way ANOVA with Bonferroni post test (**, P < 0.01). **(E and F)** Peak intensity (E) and trough-to-peak amplitude (F) in the *Gata2*$^{V/-}$ EHT subsets (n = 6, cells showing at least one peak: 36 ECs, 20 BCs, and 6 IAHCs). The data represent the mean ± range. The data were compared with Gata2$^{V/+}$ EHT subset cells using Mann–Whitney *U* test (*, P = 0.0321; **, P = 0.0056). **(G)** Model of Gata2 expression dynamics and pulsatile characteristic during EHT. EHT cell types (top) are shown with accompanying Gata2 dynamic expression changes EHT directly below. G2V MFI (bright green) and pulse parameters (dark green sinusoids) are shown for *Gata2*$^{V/+}$ (middle) and *Gata2*$^{V/-}$ (bottom) EHT subset cells.

are likely to represent a subset of Gata2-expressing cells that will have greater multilineage hematopoietic (including lymphoid) and HSC potential.

Levels of Gata2

The exact relationship between Gata2 levels and cell fate decisions remains unclear. Previous work demonstrated that *Gata2* dosage is important in regulating the quantity and functional quality of HSCs (Ling et al., 2004; Rodrigues et al., 2005; Tipping et al., 2009). It has been shown that *Gata2* expression is down-regulated during lineage commitment (Orlic et al., 1995), suggesting a role for Gata2 in early hematopoietic progenitors and HSCs. More recently it has been shown that Gata2 lies at the core of a network of genes involved in lineage specification and mixed lineage states (Olsson et al., 2016). Transcriptome analysis of CD150^{high} adult

bone marrow HSCs (Guo et al., 2013) showed concurrent high levels of *Gata2* expression and Gata2 occupancy of megakaryocyte-erythroid lineage-related genes. These high expressing HSC showed a bias to the formation of more megakaryocyte-erythroid colonies.

We have shown previously in our *G2V* reporter mouse model that the Venus high-expressing fraction in the embryonic aorta contains differentiated basophilic cell types, myeloid cells and innate immune cells (Kaimakis et al., 2016). Here we confirm that HPCs can be found in the Venus high-expressing fraction and further show that all HSCs and a large number of HPCs are found in the Venus intermediate-expressing cell fraction. Thus, commitment to the myeloid lineage in the early embryo seems to be accompanied by increased levels of *Gata2*. In line with this, low-level overexpression studies in bone marrow HSCs using a tamoxifen-inducible *Gata2 ERT* construct (Tipping et al., 2009), mimicking physiological levels of *Gata2* expression, promoted self-renewal and proliferation of myeloid progenitors. Physiological higher levels of *Gata2* block lymphoid differentiation (Tipping et al., 2009) and negatively correlate with the occupation of Gata2 with lymphoid-related genes (Guo et al., 2013). Although these data suggest that Gata2 levels regulate commitment to specific hematopoietic lineages in the adult bone marrow, during development, when *Gata2* expression is initiated and its expression is increasing in EHT cell populations, its levels are unstable in individual cells. Thus, we propose that *Gata2* pulsatile expression (in combination with the onset/asynchronous expression and stability of other pivotal TFs such as Runx1; Kueh et al., 2016) is likely to play a role in the stochastic commitment to the hematopoietic lineage.

Regulation of pulsatile expression behavior

At the transcriptional level, pathways such as Notch, β-catenin/Wnt, and fibroblast growth factor (Dequéant et al., 2006), form negative feedback loops with appropriate delay time for a pulsatile element to be translated and act at the starting point. The combination of negative and positive feedback loops prevents transcription from reaching a homeostatic steady state and maintains pulsatile expression (Purvis and Lahav, 2013). Transcription of *Gata2* in the AGM is positively regulated by Notch1, which is required for EHT and HSC development (Robert-Moreno et al., 2005; Gama-Norton et al., 2015; Souilhol et al., 2016). Gata2 is autoregulatory and maintains its own transcription (Grass et al., 2003; Burch, 2005; Kobayashi-Osaki et al., 2005; Martowicz et al., 2005). Also, Notch1 activates *Hes1*, and Hes1 represses *Gata2* expression specifically in AGM hematopoietic cells (Guiu et al., 2013). In the absence of Hes1, Gata2 expression is high (Guiu et al., 2013) and the number of cells in intra-aortic hematopoietic clusters is increased. However, persistent high-level Gata2 expression results in nonfunctional HSCs (Tipping et al., 2009).

The positive and negative signals induced by the Notch pathway result in a so-called type I incoherent feed-forward loop (Mangan and Alon, 2003). In the case of *Gata2* in the embryonic aorta, we predict that the Notch–Hes1–Gata2 feed forward loop is responsible for the pulsatile expression of G2V that we observed in the EHT cell subsets: Notch1 would stimulate *Gata2* and *Hes1* transcription, and *Gata2* transcription would be repressed when Hes1 protein reaches a critical threshold (half-life, 24 min; Yoshiura et al., 2007), resulting in a pulse-like dynamics of Gata2 protein levels when Hes levels subsequently drop. That Hes1 is the likely pacemaker of *Gata2* pulsatile expression is supported by our RNA-sequencing data showing that *Hes1* is fourfold and *Notch1* sixfold up-regulated in the $CD31^{+}ckit^{+}Venus^{med}$ fraction, as compared with the $CD31^{+}ckit^{+}Venus^{high}$ fraction. Moreover, the higher expression of *Gata2* compared with *Hes1* in the $CD31^{+}ckit^{+}Venus^{high}$ fraction suggests that critical Notch signaling thresholds will impact *Gata2* expression parameters in BCs versus IAHCs.

The in vivo G2V reporter allows for the first time the unbiased real-time characterization of *Gata2* expression during EHT in the *Gata2*WT and heterozygous mutant state. The mutant Gata2 mouse model has a deletion of the second zinc-finger domain in the *Gata2* gene (Tsai et al., 1994), leaving *Gata2*-binding motifs available on both alleles. Therefore, the altered *Gata2* pulsatile expression behavior in heterozygous mutant ECs and BCs cannot be explained by more Notch1 and Hes1 binding to only one allele of *Gata2*. Only half of the dose of DNA-binding Gata2 protein would be available to bind two alleles of *Gata2*, strongly suggesting that a reduced positive autoregulation manifests itself in altered Gata2 dynamic expression. As yet, we do not have a direct correlation between the levels of Venus protein and Gata2 protein in single cells of the AGM. In the future, mass spectrometry CyTOF (Giesen et al., 2014) could be used to more specifically address this issue. Further in vivo vital molecular studies, coupled with computational modeling of the interplay of *Gata2* regulators will be needed for a detailed understanding of the molecular basis of pulsatile expression. Given the fact that in hematopoietic disorders and malignancies *GATA2* mutations occur in the second zinc finger (Bresnick et al., 2012), the dysregulation of Gata2 pulsatile expression through a feed-forward loop might provide a mechanistic basis for human hematologic pathophysiologies.

MATERIALS AND METHODS

Mice and embryo generation

$Gata2^{V/+}$ conceptuses were generated by crossing $G2^{V/V}$ males (Kaimakis et al., 2016) with C57BL/6J females, and $Gata2^{V/-}$ conceptuses were generated by crossing $G2^{V/V}$ males with $Gata2^{+/-}$ females (five or more generations backcrossed onto C57BL/6J). *G2V* ($G2^{V/V}$) and *Ly6aGFP* ($Ly6a^{GFP/+}$) mice were mated to obtain E10.5 *G2V:Ly6aGFP* embryos (de Bruijn et al., 2002; Boisset et al., 2011). Quick genotyping (Kapa) of $Gata2^{V/-}$ embryos was performed by PCR using the following primers: *mGata2 39* (5′-GGAACGCCA ACGGGGAC-3′), *mGata2 208* (5′-GCTGGACATCTT

CCGATTCCGGGT-3′), and *Neo 503* (5′-GATCTCCTG TCATCTCACCTTGCT-3′; Tsai et al., 1994). Day of plug discovery was considered as E0 of embryonic development. Embryos were staged by SPs as E10 (28–40 SPs), early E10 (28–34 SPs), E10.5 (35–40 SPs), and E11.5 (40–50 SPs). All experiments were conducted according to Dutch and UK law and approved by the Dutch animal experiment committee (Stichting DEC consult) and the UK Animals Scientific Procedures Act 1986 Project License 70/8076, respectively.

Embryonic tissue isolation and cell preparation and flow cytometry

Single cell suspensions of dissected AGM tissues (including part of the vitelline and umbilical arteries) were prepared by 45-min collagenase treatment (type I; Sigma) at 37°C and subsequently dissociated by filtering. For flow cytometric analysis, cells were incubated (30 min, 4°C) in PBS + 10% FCS + 1% PS with directly conjugated antibodies against CD31, CD45, and ckit. Embryos were, if necessary, genotyped (quick genotyping kit Kapa) and analyzed separately. Stained cells were analyzed or sorted on a SORP-FACSAria II flow cytometer (BD) equipped with a 488 blue laser and BB ("B," blue laser; "B," detector B) 454/35-bandpass filter to optimally detect Venus fluorescence. Dead cells were counter labeled with Hoechst and excluded from analysis and sorting.

Whole-mount imaging

G2V whole-mount conceptuses were immunostained for CD31 and ckit as previously described (Yokomizo et al., 2012). Briefly, conceptuses were collected, fixed in 2% PFA and stored in methanol before labeling. Conceptuses were imaged in 50% benzyl alcohol/benzylbenzoate (1:2)/50% methanol using a laser scanning confocal microscope (SP5; Leica). Signals were collected sequentially to avoid bleed through. Three-dimensional images were reconstructed using Fiji imaging software.

Time-lapse imaging and detection of Gata2 dynamics

Aortic transversal sections of E10 *G2V* embryos were prepared as previously described (Boisset et al., 2010). Briefly, nonfixed E10 (32–37 SP) embryos were freed from placenta, yolk sac, amnion, and head. Antibodies against CD31 and ckit (diluted in PBS/10% FCS/1% PS) were directly injected into the embryonic aorta. Transversal aortic slices of 150 µm width were cut with a tissue chopper (McIlwain). Draq5 (BioLegend) staining was performed on transversal G2V sections (15 min, RT, diluted in PBS/10% FCS/1% PS), after which sections were washed twice. Selected sections (from trunk to hindlimb) were subsequently embedded in 1% agarose in PBS and after polymerization overlaid with myeloid long-term culture medium (MyeloCult; StemCell Technologies) containing hydrocortisone and IL-3. Confocal time-lapse imaging was performed using a Leica SP5 microscope, equipped with 405-nm, argon, 561-nm, and 633-nm laser lines using a 20×, 0.7-NA air objective and typically a pinhole of 1–1.5 AU. Videos were recorded at a time interval of 15 min for a total of 12–15 h. For each experiment, three to five aorta slices, with a z-range of 20–50 steps (step size, 0.7–2.5 µm) were imaged. The sample temperature was maintained by a stage heater (37°C) and the sample was kept under constant CO_2 levels (5%). The G2V signal was collected using an avalanche photo diode (APD) with a BP 535–585 emission filter, whereas the CD31-AF647 signal was collected with a photomultiplier tube (PMT) and a BP 650–720 emission filter. ckit-DyLight405 or ckit-BV421 were typically only imaged at the first frame of the time-lapse imaging series and detected with a BP 420–480 emission filter. To ensure that the Venus MFI was comparable between experiments, the microscope settings (laser power, gain, and settings of the emission filters) were kept similar among experiments.

Image processing

To improve signal-to-noise ratios for more accurate tracking and object recognition, time-lapse imaging series were deconvolved using the Huygens Professional (Scientific Volume Imaging) Deconvolution Wizard. Small drifts in z and xy were corrected by the Huygens Professional Object Stabilizer. Deconvolved and stabilized time series were used for further analysis.

Quantification of Gata2 dynamics

To analyze the dynamics of Venus expression in single cells in the aorta, $Venus^+$ cells had to be tracked and the Venus fluorescence signal corresponding to individual single cells had to be extracted. Because no commercial tool was available that reliably tracked Venus expressing cells in the aorta and at the same time extracted the fluorescent signal, we developed a custom-made code to combine two commercial tools (1) tracking $Venus^+$ cells (Huygens Professional Object Tracker) and (2) extracting voxel information (Huygens Professional Object Analyzer) of the Venus fluorescent signals in three dimensions and in time. Object Tracker and Object Analyzer use different algorithms to track and segment objects; therefore each tool assigned a unique identifier to each object (cell). Because both tools use the center of mass to describe the position of the object, our LabVIEW-based custom-made code assigned to each tracked object the closest segmented object with voxel information (within a maximum range of 5 µm). The resulting $Venus^+$ cells with common identifier could be visualized in each time-lapse series by a custom-written Fiji macro. The tracked cells were visually inspected, and incorrectly tracked cells were excluded from further analysis. Moreover, further analysis was limited to cells that could be tracked over at least 10 consecutive frames and did not show any bleaching or overall intensity changes caused by the microscope setup.

To quantify the dynamics of Venus expression, the LabVIEW code also computed the volumetric MFI values of each tracked cell, which was defined as the sum of all intensities divided by the number of voxels representing the

cell (Fig. S2). For further analysis, MATLAB codes (version 2015b) were developed to plot the Venus MFI as a function of time (Fig. S2). As a control, the MFI values were plotted against the voxel values, confirming that quantitative changes in the fluorescent intensity were not due to tracking errors.

Data analysis

To assess whether the Venus signal in the time series data undergoes quantitative changes, we adapted a public-domain MATLAB code (http://www.billauer.co.il/peakdet.html) to automatically detect significant extrema in our Venus MFI time series data (Todd and Andrews, 1999). To discriminate against "noise" (such as fluctuations introduced by imperfections of the image stack segmentation), the code only considered local maximum as significant if it differed from its neighboring minima by more than a predefined threshold (specified as a percentage of the mean intensity of the track). For the analysis in this article, we used a threshold of 15%. Visual inspection of the minima and maxima confirmed that ∼90% of peaks were correctly detected using this threshold. Tracks with incorrectly recognized peaks were excluded from further analysis. Additional codes calculated the number of peaks, oscillation periodicity, peak minimum and peak maximum, and trough-to-peak amplitude. The data were exported from MATLAB to excel for further analysis and plotted in GraphPad Prism 5.

Hematopoietic assays

The methylcellulose colony-forming assay was performed as previously described (Medvinsky et al., 2008). $CD31^+ckit^+V^{med}$ and $CD31^+ckit^+V^{high}$ sorted E10 AGMs (including part of the vitelline and umbilical arteries) were seeded in triplicate in methylcellulose (1 ml per dish; M3434; Stem Cell Technology) with 1% PS and incubated for 10 to 12 d at 37°C, 5% CO_2. Colonies were counted with a bright-field microscope. Transplantation experiments were performed as previously described (Medvinsky et al., 2008). Sorted $CD31^+ckit^+V^-$, $CD31^+ckit^-V^+$, $CD31^+ckit^+V^{med}$, and $CD31^+ckit^+V^{high}$ (Ly5.2/Ly5.2) cells of five to seven E11 AGMs were transplanted into 9.5-Gy irradiated (Ly5.1/Ly5.1) recipients together with 2×10^5 spleen cells from the recipient strain. Peripheral blood was analyzed by flow cytometry for donor contribution by anti-Ly5.1/anti-Ly5.2 labeling 1 and 4 mo after transplantation. Transplanted recipients were scored as positive if the peripheral blood donor chimerism was ≥10%. Multilineage organ chimerism analysis (lymphoid and myeloid) was performed 4 mo after transplantation.

RNA isolation mRNA-sequencing analysis

$CD31^+ckit^+V^{med}$ and $CD31^+ckit^+V^{high}$ E10.5 AGM cells of $G2^{V/+}$ embryos were sorted into PBS/50% FCS/1% PS. After centrifugation and removal of supernatant, cells were lysed, and RNA was isolated using the mirVana miRNA Isolation kit (Ambion) according to the manufacturer's protocol. RNA quality and quantity were accessed by the 2100 Bioanalyzer (Picochip; Agilent Technologies). RNA samples were prepared by SMARTer protocol. Illumina TrueSeq v2 protocol was used on HiSeq2500 with a single-read 50-bp and 9-bp index. Reads were aligned to the mouse genome (GRCm38/mm10) using Tophat/Bowtie, and the generated count table was analyzed by R/Bioconductor package edgeR according to McCarthy et al. (2012). Counts were normalized for mRNA abundance, and differential expression analysis was performed using edgeR. The B-H method was used for p-value correction with a false discovery rate of 0.05 as statistically significant. Variance stabilized counts were calculated by R/Bioconductor package DESeq for all genes (Anders and Huber, 2010). Cufflinks was used to compute transcript abundance estimates in fragments per kilobase per million (FPKM; Trapnell et al., 2013). For differentially expressed genes, the FPKM for each gene across all samples were normalized by division with maximum FPKM observed for that gene. Differentially expressed genes were analyzed for the top five most enriched Ingenuity Pathway Analysis pathways against a background of all mouse genes by right tailed Fisher exact tests in a core analysis calculating the likelihood that this is due to random chance. The accession number for the RNA-sequencing data is Gene Expression Omnibus: GSE106072.

SDS-PAGE and Western blot

4×10^4 E10 $Venus^{med}$ and $Venus^{high}$ AGM cells were sorted from E10.5 $G2^{V/+}$ and $G2^{V/-}$ embryos (littermates), washed and centrifuged, and directly lysed in Laemmli sample buffer. 4.5 to 6.2×10^4 $G2^{V/V}$ bone marrow mononuclear cells were sorted into $Venus^-$, $Venus^{med}$, and $Venus^{high}$ cell fractions, subsequently lysed in RIPA buffer plus protease and phosphatase inhibitor, and sonicated. Then Laemmli buffer was added. Proteins were separated by SDS-PAGE and transferred to PVDF membranes (Millipore). Subsequently, proteins were detected by anti-Gata2, anti-Venus, anti-GAPDH, anti–β-actin, anti-Hsp90, and anti-Cohesin immunoblotting. After labeling, Western blots were scanned using the Odyssey imager (LI-COR Biosciences).

Antibodies

For flow cytometry, cells were stained with anti–CD31-PE-Cy7 (clone 390; BD), anti–cKit-BV421 (clone 2B8; BD), anti–Ly5.1-APC (clone A20, BD), and anti-Ly5.2-Fitc (clone 104, BD) monoclonal antibodies. For microscopy, CD31-AF647 (clone 390; eBioscience) and ckit-DyLight405 (clone 2B8; eBioscience; conjugated by the authors to DyLight 405; Pierce) were injected. In Fig. S1 A, Gata2, Venus, β-actin, GAPDH, and heat-shock protein 90 were detected by mouse monoclonal anti-Gata2 (clone CG2-96; Santa Cruz), rabbit polyclonal anti-Venus (MBL), mouse monoclonal anti–β-actin, rabbit polyclonal anti-GAPDH (gift of C. Hansen; Santa Cruz), and mouse monoclonal anti-Hsp90 (gift of C. Hansen; BD). In Fig. S5 A, Gata2 and Cohesin (subunit SMC3) were detected by rabbit anti-Gata2 (Santa Cruz) and rabbit anti-Cohesin (Abcam) polyclonal antibodies.

Statistics

The data were compared in GraphPad Prism 5 using Mann–Whitney *U* tests, Student's *t* tests, and one- or two-way ANOVA with Bonferroni post test, as indicated. Errors in the frequency of oscillation periodicity were estimated by bootstrapping (resampling residuals approach). Error bars represent two times the standard deviation originated from fitting procedures.

Online supplemental material

Fig. S1 shows a Western blot of sorted Venus^{med} and Venus^{high} cells, demonstrating that Venus protein levels correctly reflect levels of Gata2, and confocal images show the differential expression pattern of Gata2 and Ly6a in *Ly6aGFP:G2V* thick aortic sections at E10.5. Fig. S2 illustrates the image acquisition and processing pipeline to analyze Venus expression in embryonic sections during confocal time-lapse imaging. Fig. S3 shows the visualization of Venus expression peaks throughout the imaging session. Fig. S4 shows two examples in which Venus expression IAHCs undergo mitosis during the imaging session. Fig. S5 shows Gata2 expression characteristics in *Gata2* heterozygous mutant embryos. Videos 1, 2, and 3 show examples of Venus-expressing cells undergoing EHT during G2V time-lapse imaging.

ACKNOWLEDGMENTS

We thank Wilson Poon, Anna Bigas, Frank Grosveld, Sjaak Philipsen, and members of the laboratory for helpful discussions and advice. We thank Anna Bigas and Frank Grosveld for critical reading of the manuscript. We also thank the Erasmus MC Biomics Centre for Illumina sample preparation and RNA sequencing.

This research was supported by the National Institutes of Health (RO37 DK54077), the Netherlands Institute for Regenerative Medicine (FES award NIRM 101675), ZonMw TOP (91211068), and the European Research Council (advanced grant 341096).

The authors declare no competing financial interests.

Author contributions: E. Dzierzak and C. Eich conceived the project and designed the experiments. C. Eich performed the experiments. P. Kaimakis provided the *G2V* mice and preliminary analyses. C.S. Vink performed hematopoietic functional experiments, preparation of samples and RNA isolation for RNA sequencing, and Draq5 live imaging. C.S. Vink and S.A. Mariani performed western blot analysis of bone marrow. W.A. van Cappellen and R. van der Linden assisted with imaging and flow cytometry and P.S. Kartalaei assisted with RNA-sequencing analysis. C. Eich and J. Arlt designed novel code and performed computational analyses. E. Dzierzak and C. Eich wrote the manuscript.

Submitted: 3 May 2017

Revised: 12 September 2017

Accepted: 31 October 2017

REFERENCES

Albeck, J.G., G.B. Mills, and J.S. Brugge. 2013. Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Mol. Cell.* 49:249–261. https://doi.org/10.1016/j.molcel.2012.11.002

Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. https://doi.org/10.1186/gb-2010-11-10-r106

Batchelor, E., A. Loewer, C. Mock, and G. Lahav. 2011. Stimulus-dependent dynamics of p53 in single cells. *Mol. Syst. Biol.* 7:488. https://doi.org/10.1038/msb.2011.20

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

Boisset, J.C., C. Andrieu-Soler, W.A. van Cappellen, T. Clapes, and C. Robin. 2011. Ex vivo time-lapse confocal imaging of the mouse embryo aorta. *Nat. Protoc.* 6:1792–1805. https://doi.org/10.1038/nprot.2011.401

Bresnick, E.H., K.R. Katsumura, H.Y. Lee, K.D. Johnson, and A.S. Perkins. 2012. Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Res.* 40:5819–5831. https://doi.org/10.1093/nar/gks281

Burch, J.B. 2005. Regulation of GATA gene expression during vertebrate development. *Semin. Cell Dev. Biol.* 16:71–81. https://doi.org/10.1016/j.semcdb.2004.10.002

Cai, L., C.K. Dalal, and M.B. Elowitz. 2008. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature.* 455:485–490. https://doi.org/10.1038/nature07292

Cohen-Saidon, C., A.A. Cohen, A. Sigal, Y. Liron, and U. Alon. 2009. Dynamics and variability of ERK2 response to EGF in individual living cells. *Mol. Cell.* 36:885–893. https://doi.org/10.1016/j.molcel.2009.11.025

Dalal, C.K., L. Cai, Y. Lin, K. Rahbar, and M.B. Elowitz. 2014. Pulsatile dynamics in the yeast proteome. *Curr. Biol.* 24:2189–2194. https://doi.org/10.1016/j.cub.2014.07.076

de Bruijn, M.F., X. Ma, C. Robin, K. Ottersbach, M.J. Sanchez, and E. Dzierzak. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity.* 16:673–683. https://doi.org/10.1016/S1074-7613(02)00313-8

de Pater, E., P. Kaimakis, C.S. Vink, T. Yokomizo, T. Yamada-Inagawa, R. van der Linden, P.S. Kartalaei, S.A. Camper, N. Speck, and E. Dzierzak. 2013. Gata2 is required for HSC generation and survival. *J. Exp. Med.* 210:2843–2850. https://doi.org/10.1084/jem.20130751

Dequéant, M.L., E. Glynn, K. Gaudenz, M. Wahl, J. Chen, A. Mushegian, and O. Pourquié. 2006. A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science.* 314:1595–1598. https://doi.org/10.1126/science.1133141

Dzierzak, E., and A. Medvinsky. 2008. The discovery of a source of adult hematopoietic cells in the embryo. *Development.* 135:2343–2346. https://doi.org/10.1242/dev.021279

Dzierzak, E., and N.A. Speck. 2008. Of lineage and legacy: the development of mammalian hematopoietic stem cells. *Nat. Immunol.* 9:129–136. https://doi.org/10.1038/ni1560

Gama-Norton, L., E. Ferrando, C. Ruiz-Herguido, Z. Liu, J. Guiu, A.B. Islam, S.U. Lee, M. Yan, C.J. Guidos, N. López-Bigas, et al. 2015. Notch signal strength controls cell fate in the haemogenic endothelium. *Nat. Commun.* 6:8510. https://doi.org/10.1038/ncomms9510

Gao, X., K.D. Johnson, Y.I. Chang, M.E. Boyer, C.N. Dewey, J. Zhang, and E.H. Bresnick. 2013. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J. Exp. Med.* 210:2833–2842. https://doi.org/10.1084/jem.20130733

Garmendia-Torres, C., A. Goldbeter, and M. Jacquet. 2007. Nucleocytoplasmic oscillations of the yeast transcription factor Msn2: evidence for periodic PKA activation. *Curr. Biol.* 17:1044–1049. https://doi.org/10.1016/j.cub.2007.05.032

Giesen, C., H.A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P.J. Schüffler, D. Grolimund, J.M. Buhmann, S. Brandt, et al. 2014. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods.* 11:417–422. https://doi.org/10.1038/nmeth.2869

Goode, D.K., N. Obier, M.S. Vijayabaskar, M. Lie-A-Ling, A.J. Lilly, R. Hannah, M. Lichtinger, K. Batta, M. Florkowska, R. Patel, et al. 2016. Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation. *Dev. Cell.* 36:572–587. https://doi.org/10.1016/j.devcel.2016.01.024

Grass, J.A., M.E. Boyer, S. Pal, J. Wu, M.J. Weiss, and E.H. Bresnick. 2003. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc. Natl. Acad. Sci. USA.* 100:8811–8816. https://doi.org/10.1073/pnas.1432147100

Guiu, J., R. Shimizu, T. D'Altri, S.T. Fraser, J. Hatakeyama, E.H. Bresnick, R. Kageyama, E. Dzierzak, M. Yamamoto, L. Espinosa, and A. Bigas. 2013. Hes repressors are essential regulators of hematopoietic stem cell development downstream of Notch signaling. *J. Exp. Med.* 210:71–84. https://doi.org/10.1084/jem.20120993

Guo, G., S. Luc, E. Marco, T.W. Lin, C. Peng, M.A. Kerenyi, S. Beyaz, W. Kim, J. Xu, P.P. Das, et al. 2013. Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. *Cell Stem Cell.* 13:492–505. https://doi.org/10.1016/j.stem.2013.07.017

Imayoshi, I., A. Isomura, Y. Harima, K. Kawaguchi, H. Kori, H. Miyachi, T. Fujiwara, F. Ishidate, and R. Kageyama. 2013. Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science.* 342:1203–1208. https://doi.org/10.1126/science.1242366

Isomura, A., and R. Kageyama. 2014. Ultradian oscillations and pulses: coordinating cellular responses and cell fate decisions. *Development.* 141:3627–3636. https://doi.org/10.1242/dev.104497

Kageyama, R., T. Ohtsuka, H. Shimojo, and I. Imayoshi. 2008. Dynamic Notch signaling in neural progenitor cells and a revised view of lateral inhibition. *Nat. Neurosci.* 11:1247–1251. https://doi.org/10.1038/nn.2208

Kaimakis, P., E. de Pater, C. Eich, P. Solaimani Kartalaei, M.L. Kauts, C.S. Vink, R. van der Linden, M. Jaegle, T. Yokomizo, D. Meijer, and E. Dzierzak. 2016. Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. *Blood.* 127:1426–1437. https://doi.org/10.1182/blood-2015-10-673749

Khandekar, M., W. Brandt, Y. Zhou, S. Dagenais, T.W. Glover, N. Suzuki, R. Shimizu, M. Yamamoto, K.C. Lim, and J.D. Engel. 2007. A Gata2 intronic enhancer confers its pan-endothelia-specific regulation. *Development.* 134:1703–1712. https://doi.org/10.1242/dev.001297

Kholodenko, B.N., J.F. Hancock, and W. Kolch. 2010. Signalling ballet in space and time. *Nat. Rev. Mol. Cell Biol.* 11:414–426. https://doi.org/10.1038/nrm2901

Kim, D., D. Grün, and A. van Oudenaarden. 2013. Dampening of expression oscillations by synchronous regulation of a microRNA and its target. *Nat. Genet.* 45:1337–1344. https://doi.org/10.1038/ng.2763

Kobayashi, T., H. Mizuno, I. Imayoshi, C. Furusawa, K. Shirahige, and R. Kageyama. 2009. The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells. *Genes Dev.* 23:1870–1875. https://doi.org/10.1101/gad.1823109

Kobayashi-Osaki, M., O. Ohneda, N. Suzuki, N. Minegishi, T. Yokomizo, S. Takahashi, K.C. Lim, J.D. Engel, and M. Yamamoto. 2005. GATA motifs regulate early hematopoietic lineage-specific expression of the Gata2 gene. *Mol. Cell. Biol.* 25:7005–7020. https://doi.org/10.1128/MCB.25.16.7005-7020.2005

Koga, S., N. Yamaguchi, T. Abe, M. Minegishi, S. Tsuchiya, M. Yamamoto, and N. Minegishi. 2007. Cell-cycle-dependent oscillation of GATA2 expression in hematopoietic cells. *Blood.* 109:4200–4208. https://doi.org/10.1182/blood-2006-08-044149

Kueh, H.Y., A. Champhekar, S.L. Nutt, M.B. Elowitz, and E.V. Rothenberg. 2013. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science.* 341:670–673. https://doi.org/10.1126/science.1240831

Kueh, H.Y., M.A. Yui, K.K. Ng, S.S. Pease, J.A. Zhang, S.S. Damle, G. Freedman, S. Siu, I.D. Bernstein, M.B. Elowitz, and E.V. Rothenberg. 2016. Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. *Nat. Immunol.* 17:956–965. https://doi.org/10.1038/ni.3514

Kumano, K., S. Chiba, A. Kunisato, M. Sata, T. Saito, E. Nakagami-Yamaguchi, T. Yamaguchi, S. Masuda, K. Shimizu, T. Takahashi, et al. 2003. Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. *Immunity.* 18:699–711. https://doi.org/10.1016/S1074-7613(03)00117-1

Lahav, G., N. Rosenfeld, A. Sigal, N. Geva-Zatorsky, A.J. Levine, M.B. Elowitz, and U. Alon. 2004. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet.* 36:147–150. https://doi.org/10.1038/ng1293

Levine, J.H., Y. Lin, and M.B. Elowitz. 2013. Functional roles of pulsing in genetic circuits. *Science.* 342:1193–1200. https://doi.org/10.1126/science.1239999

Li, X., X. Zhao, Y. Fang, X. Jiang, T. Duong, C. Fan, C.C. Huang, and S.R. Kain. 1998. Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* 273:34970–34975. https://doi.org/10.1074/jbc.273.52.34970

Lichtinger, M., R. Ingram, R. Hannah, D. Müller, D. Clarke, S.A. Assi, M. Lie-A-Ling, L. Noailles, M.S. Vijayabaskar, M. Wu, et al. 2012. RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J.* 31:4318–4333. https://doi.org/10.1038/emboj.2012.275

Ling, K.W., K. Ottersbach, J.P. van Hamburg, A. Oziemlak, F.Y. Tsai, S.H. Orkin, R. Ploemacher, R.W. Hendriks, and E. Dzierzak. 2004. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J. Exp. Med.* 200:871–882. https://doi.org/10.1084/jem.20031556

Locke, J.C., J.W. Young, M. Fontes, M.J. Hernández Jiménez, and M.B. Elowitz. 2011. Stochastic pulse regulation in bacterial stress response. *Science.* 334:366–369. https://doi.org/10.1126/science.1208144

Lorsbach, R.B., J. Moore, S.O. Ang, W. Sun, N. Lenny, and J.R. Downing. 2004. Role of RUNX1 in adult hematopoiesis: analysis of RUNX1-IRES-GFP knock-in mice reveals differential lineage expression. *Blood.* 103:2522–2529. https://doi.org/10.1182/blood-2003-07-2439

Lurie, L.J., M.E. Boyer, J.A. Grass, and E.H. Bresnick. 2008. Differential GATA factor stabilities: implications for chromatin occupancy by structurally similar transcription factors. *Biochemistry.* 47:859–869. https://doi.org/10.1021/bi701692p

Mangan, S., and U. Alon. 2003. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. USA.* 100:11980–11985. https://doi.org/10.1073/pnas.2133841100

Martowicz, M.L., J.A. Grass, M.E. Boyer, H. Guend, and E.H. Bresnick. 2005. Dynamic GATA factor interplay at a multicomponent regulatory region of the GATA-2 locus. *J. Biol. Chem.* 280:1724–1732. https://doi.org/10.1074/jbc.M406038200

McCarthy, D.J., Y. Chen, and G.K. Smyth. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40:4288–4297. https://doi.org/10.1093/nar/gks042

Medvinsky, A., S. Taoudi, S. Mendes, and E. Dzierzak. 2008. Analysis and manipulation of hematopoietic progenitor and stem cells from murine embryonic tissues. *Curr. Protoc. Stem Cell Biol.* Chapter 2:Unit 2A.6. https://doi.org/10.1002/9780470151808.sc02a06s4

Minegishi, N., J. Ohta, H. Yamagiwa, N. Suzuki, S. Kawauchi, Y. Zhou, S. Takahashi, N. Hayashi, J.D. Engel, and M. Yamamoto. 1999. The mouse GATA-2 gene is expressed in the para-aortic splanchnopleura and aorta-gonads and mesonephros region. *Blood.* 93:4196–4207.

Minegishi, N., N. Suzuki, Y. Kawatani, R. Shimizu, and M. Yamamoto. 2005. Rapid turnover of GATA-2 via ubiquitin-proteasome protein

degradation pathway. *Genes Cells.* 10:693–704. https://doi.org/10.1111/j.1365-2443.2005.00864.x

Nardelli, J., D. Thiesson, Y. Fujiwara, F.Y. Tsai, and S.H. Orkin. 1999. Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. *Dev. Biol.* 210:305–321. https://doi.org/10.1006/dbio.1999.9278

Nelson, D.E., A.E. Ihekwaba, M. Elliott, J.R. Johnson, C.A. Gibney, B.E. Foreman, G. Nelson, V. See, C.A. Horton, D.G. Spiller, et al. 2004. Oscillations in NF-kappaB signaling control the dynamics of gene expression. *Science.* 306:704–708. https://doi.org/10.1126/science.1099962

Neuert, G., B. Munsky, R.Z. Tan, L. Teytelman, M. Khammash, and A. van Oudenaarden. 2013. Systematic identification of signal-activated stochastic gene regulation. *Science.* 339:584–587. https://doi.org/10.1126/science.1231456

Ng, Y.K., K.M. George, J.D. Engel, and D.I. Linzer. 1994. GATA factor activity is required for the trophoblast-specific transcriptional regulation of the mouse placental lactogen I gene. *Development.* 120:3257–3266.

Olsson, A., M. Venkatasubramanian, V.K. Chaudhri, B.J. Aronow, N. Salomonis, H. Singh, and H.L. Grimes. 2016. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature.* 537:698–702. https://doi.org/10.1038/nature19348

Orlic, D., S. Anderson, L.G. Biesecker, B.P. Sorrentino, and D.M. Bodine. 1995. Pluripotent hematopoietic stem cells contain high levels of mRNA for c-kit, GATA-2, p45 NF-E2, and c-myb and low levels or no mRNA for c-fms and the receptors for granulocyte colony-stimulating factor and interleukins 5 and 7. *Proc. Natl. Acad. Sci. USA.* 92:4601–4605. https://doi.org/10.1073/pnas.92.10.4601

Pourquié, O. 2011. Vertebrate segmentation: from cyclic gene networks to scoliosis. *Cell.* 145:650–663. https://doi.org/10.1016/j.cell.2011.05.011

Purvis, J.E., and G. Lahav. 2013. Encoding and decoding cellular information through signaling dynamics. *Cell.* 152:945–956. https://doi.org/10.1016/j.cell.2013.02.005

Robert-Moreno, A., L. Espinosa, J.L. de la Pompa, and A. Bigas. 2005. RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development.* 132:1117–1126. https://doi.org/10.1242/dev.01660

Rodrigues, N.P., V. Janzen, R. Forkert, D.M. Dombkowski, A.S. Boyd, S.H. Orkin, T. Enver, P. Vyas, and D.T. Scadden. 2005. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood.* 106:477–484. https://doi.org/10.1182/blood-2004-08-2989

Ryu, H., M. Chung, M. Dobrzyński, D. Fey, Y. Blum, S. Sik Lee, M. Peter, B.N. Kholodenko, N. Li Jeon, and O. Pertz. 2016. Frequency modulation of ERK activation dynamics rewires cell fate. *Mol. Syst. Biol.* 12:866. https://doi.org/10.15252/msb.20166982

Shimojo, H., T. Ohtsuka, and R. Kageyama. 2008. Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron.* 58:52–64. https://doi.org/10.1016/j.neuron.2008.02.014

Snapp, E.L. 2009. Fluorescent proteins: a cell biologist's user guide. *Trends Cell Biol.* 19:649–655. https://doi.org/10.1016/j.tcb.2009.08.002

Solaimani Kartalaei, P., T. Yamada-Inagawa, C.S. Vink, E. de Pater, R. van der Linden, J. Marks-Bluth, A. van der Sloot, M. van den Hout, T. Yokomizo, M.L. van Schaick-Solernó, et al. 2015. Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *J. Exp. Med.* 212:93–106. https://doi.org/10.1084/jem.20140767

Souilhol, C., J.G. Lendinez, S. Rybtsov, F. Murphy, H. Wilson, D. Hills, A. Batsivari, A. Binagui-Casas, A.C. McGarvey, H.R. MacDonald, et al. 2016. Developing HSCs become Notch independent by the end of maturation in the AGM region. *Blood.* 128:1567–1577. https://doi.org/10.1182/blood-2016-03-708164

Stern, C.D., and A.M. Piatkowska. 2015. Multiple roles of timing in somite formation. *Semin. Cell Dev. Biol.* 42:134–139. https://doi.org/10.1016/j.semcdb.2015.06.002

Swiers, G., C. Baumann, J. O'Rourke, E. Giannoulatou, S. Taylor, A. Joshi, V. Moignard, C. Pina, T. Bee, K.D. Kokkaliaris, et al. 2013. Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat. Commun.* 4:2924. https://doi.org/10.1038/ncomms3924

Tay, S., J.J. Hughey, T.K. Lee, T. Lipniacki, S.R. Quake, and M.W. Covert. 2010. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. *Nature.* 466:267–271. https://doi.org/10.1038/nature09145

Tipping, A.J., C. Pina, A. Castor, D. Hong, N.P. Rodrigues, L. Lazzari, G.E. May, S.E. Jacobsen, and T. Enver. 2009. High GATA-2 expression inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. *Blood.* 113:2661–2672. https://doi.org/10.1182/blood-2008-06-161117

Todd, B.S., and D.C. Andrews. 1999. The identification of peaks in physiological signals. *Comput. Biomed. Res.* 32:322–335. https://doi.org/10.1006/cbmr.1999.1518

Trapnell, C., D.G. Hendrickson, M. Sauvageau, L. Goff, J.L. Rinn, and L. Pachter. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31:46–53. https://doi.org/10.1038/nbt.2450

Tsai, F.Y., and S.H. Orkin. 1997. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood.* 89:3636–3643.

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

Wilson, N.K., S.D. Foster, X. Wang, K. Knezevic, J. Schütte, P. Kaimakis, P.M. Chilarska, S. Kinston, W.H. Ouwehand, E. Dzierzak, et al. 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell.* 7:532–544. https://doi.org/10.1016/j.stem.2010.07.016

Yissachar, N., T. Sharar Fischler, A.A. Cohen, S. Reich-Zeliger, D. Russ, E. Shifrut, Z. Porat, and N. Friedman. 2013. Dynamic response diversity of NFAT isoforms in individual living cells. *Mol. Cell.* 49:322–330. https://doi.org/10.1016/j.molcel.2012.11.003

Yokomizo, T., T. Yamada-Inagawa, A.D. Yzaguirre, M.J. Chen, N.A. Speck, and E. Dzierzak. 2012. Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nat. Protoc.* 7:421–431. https://doi.org/10.1038/nprot.2011.441

Yoshiura, S., T. Ohtsuka, Y. Takenaka, H. Nagahara, K. Yoshikawa, and R. Kageyama. 2007. Ultradian oscillations of Stat, Smad, and Hes1 expression in response to serum. *Proc. Natl. Acad. Sci. USA.* 104:11292–11297. https://doi.org/10.1073/pnas.0701837104

Young, J.W., J.C. Locke, and M.B. Elowitz. 2013. Rate of environmental change determines stress response specificity. *Proc. Natl. Acad. Sci. USA.* 110:4140–4145. https://doi.org/10.1073/pnas.1213060110

Zambrano, S., I. De Toma, A. Piffer, M.E. Bianchi, and A. Agresti. 2016. NF-κB oscillations translate into functionally related patterns of gene expression. *eLife.* 5:e09100. https://doi.org/10.7554/eLife.09100

Zhou, F., X. Li, W. Wang, P. Zhu, J. Zhou, W. He, M. Ding, F. Xiong, X. Zheng, Z. Li, et al. 2016. Tracing haematopoietic stem cell formation at single-cell resolution. *Nature.* 533:487–492. https://doi.org/10.1038/nature17997

SUPPLEMENTAL MATERIAL

Eich et al., https://doi.org/10.1084/jem.20170807

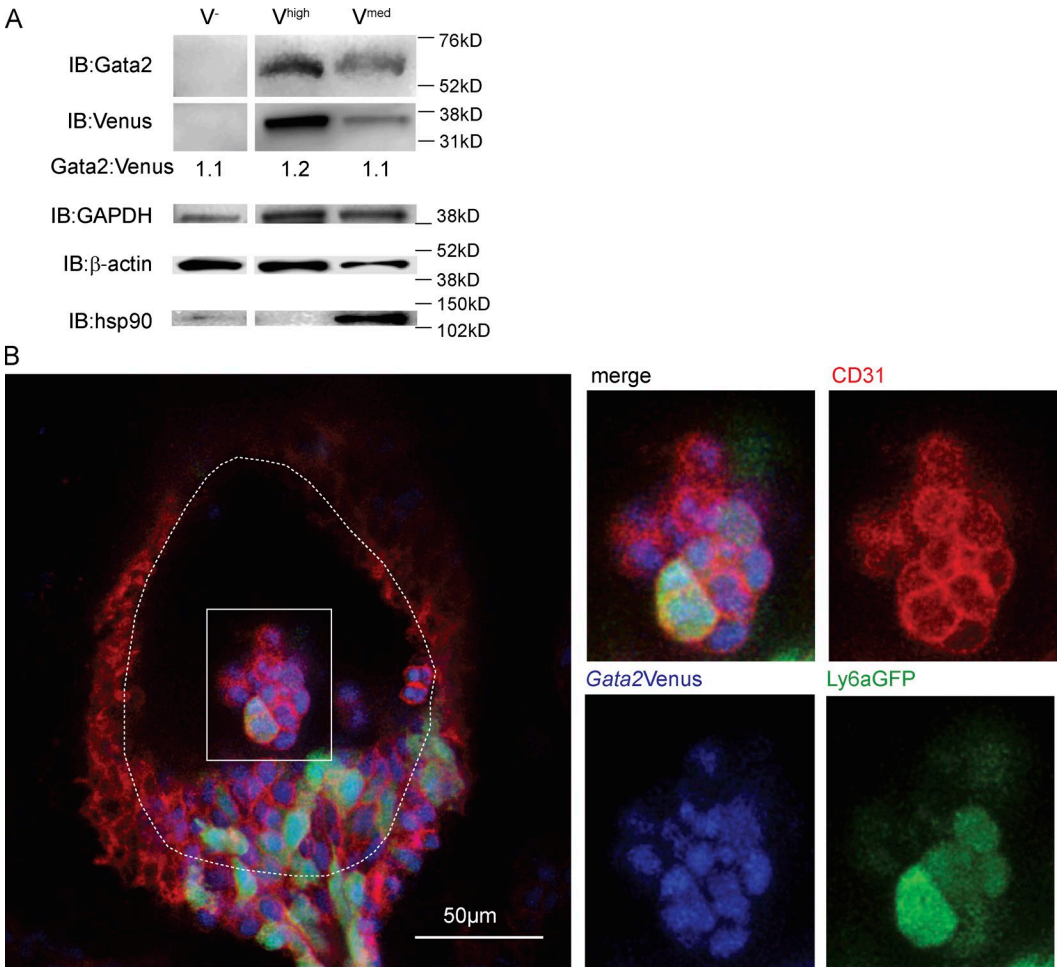


Figure S1. **Expression of Gata2, Venus, and Ly6aGFP. (A)** Representative Gata2 and Venus immunoblot (IB) of sorted Venus^{-}, Venus^{med}, and Venus^{high} cells of Gata2$^{V/V}$ bone marrow (n = 2). Although Venus^{high} and Venus^{med} sorted cells show Gata2 protein, Venus-negative sorted cells do not express Gata2 protein. Venus immunoblotting shows that the ratio of Gata2 protein to Venus protein is close to one in all cases, thus indicating a correlation in the protein levels of Gata2 and Venus. β-Actin, GAPDH, and Hsp90 immunoblottings were performed as protein loading controls but show high variability, suggesting that the sorted fractions contain hematopoietic cell types that differ in metabolism, cytoskeleton, etc. **(B)** Comparison of the *Ly6aGFP* and *G2V* reporter mouse models. Transverse section of *G2V:Ly6aGFP* E10.5 aorta. To compare Gata2 expression to the well-described stem cell marker Ly6aGFP during EHT, we simultaneously imaged both transgene reporters by crossing *G2V* (*G2V/V*) and *Ly6aGFP* (*Ly6aGFP/+*) mice (de Bruijn et al., 2002; Boisset et al., 2011) to obtain E10.5 *G2V:Ly6aGFP* embryos. Anti–CD31 antibody (red) was preinjected into the embryonic aorta (G2V, blue; Ly6aGFP, green). CD31-AF647, Ly6aGFP, and G2V were imaged sequentially. Albeit a similar expression pattern, cross sections of *Ly6aGFP:G2V* embryos at E10.5 showed a broader expression of G2V in EC and in IAHCs, as compared with Ly6aGFP. Although almost all IAHCs cells were Gata2 positive, only a few cells showed Ly6a expression. Bar, 50 µm. Related to Fig. 1 and Table 1.

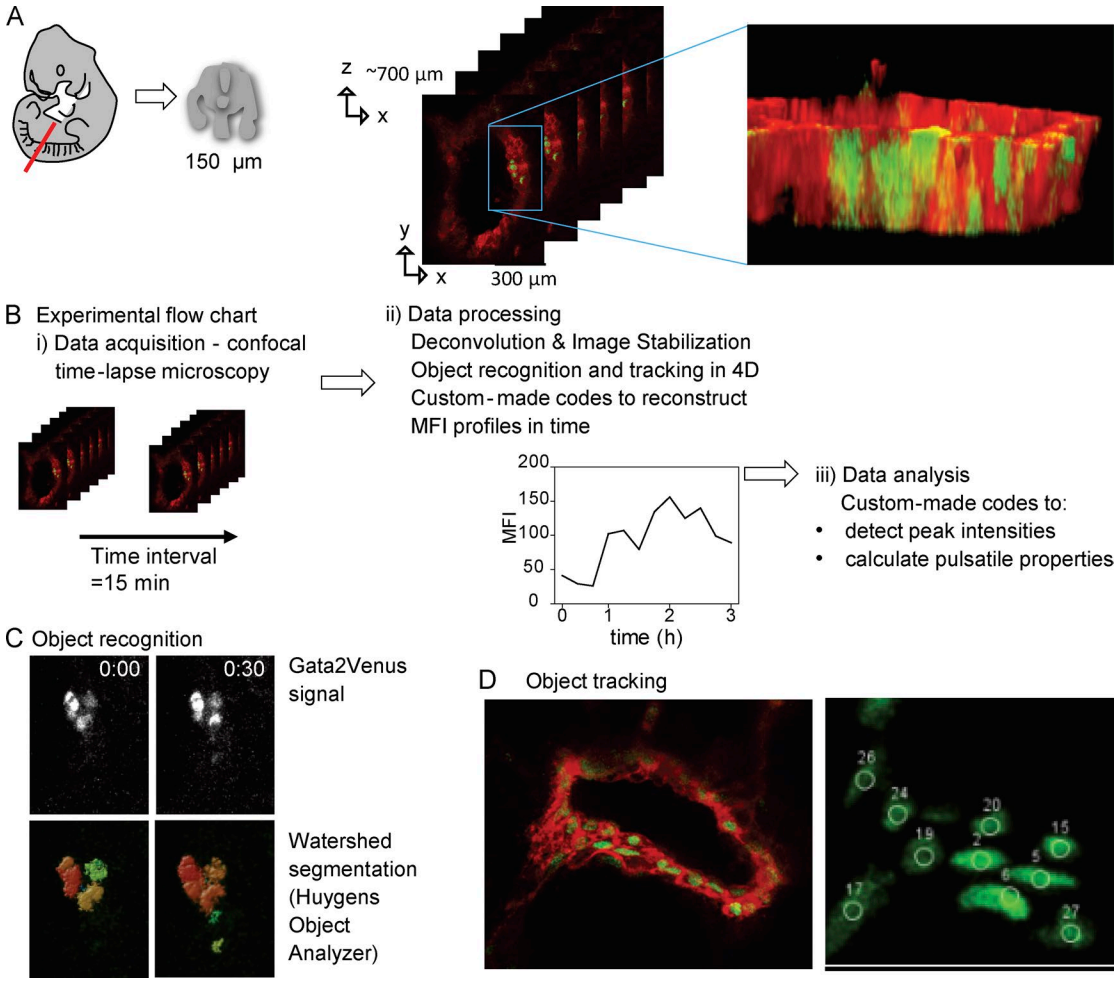


Figure S2. **Image acquisition and processing pipeline to display Venus MFI in four dimensions. (A)** Image acquisition. Images of *G2V* embryo slices were acquired at a spatial resolution of 300 nm in xy and 700–2,500 nm in z, at a time interval of 15 min, for up to 15 h. The mean width that was imaged per section was ∼25–50 µm. **(B)** Image processing and analysis pipeline. Prior to analysis, images were deconvolved using Huygens Professional Deconvolution Wizard to improve the signal-to-noise ratio. To provide a direct view on Gata2 dynamics in single cells during EHT, Venus$^+$ cells were segmented in three dimensions using the Huygens Professional Object Analyzer and tracked in four dimensions using the Huygens Object Tracker. Custom-made codes combined both types of information to display the Venus MFI of single cells in four dimensions, as defined as the sum of all voxels/number of voxels per tracked object (see Materials and methods). MATLAB codes were developed to display the MFI as a function of time. Tracks of Venus$^+$ cells were excluded that showed bleaching or overall intensity changes caused by the microscope setup or cells present in < 10 consecutive frames. Additional MATLAB codes were developed to detect peaks in the Venus MFI data and compute the following parameters: peak number, peak intensity, oscillation periodicity, and trough-to-peak amplitude. **(C)** Example of object recognition using Huygens object analyzer on Venus deconvolved data. Single Venus$^+$ cells could be recognized in three dimensions using a watershed segmentation algorithm. **(D)** Visualization of Venus$^+$ cells recognized in four dimensions by our custom-made code. Our custom-made codes combined information of the object analyzer and object tracker and assigned new identifiers to the recognized cells. The result was visually inspected using a purpose-designed Fiji macro and corrected if necessary.

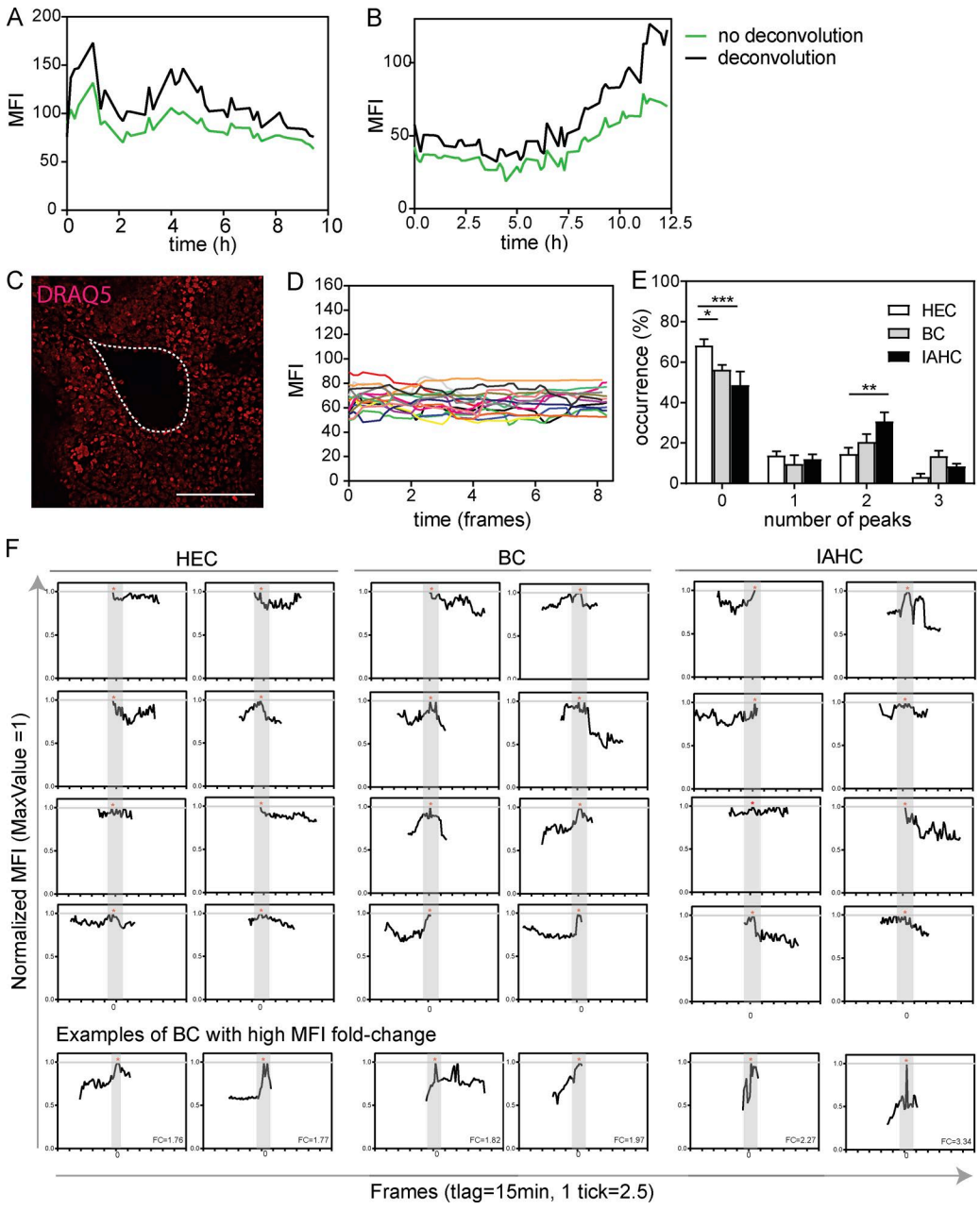


Figure S3. **Visualization of Venus expression peaks in confocal time-lapse imaging of aortic sections. (A and B)** Comparison of temporal variation of Venus MFI of individual Venus$^+$ cells in raw data versus deconvoluted data. A corresponds to Fig. 2 A and B to Fig. 3 B. **(C)** Confocal time-lapse imaging of E10 (33–36 SPs) transverse aortic section stained with DRAQ5 (red). Bar, 100 µm. Dashed line indicates the inner endothelial boundary. Transverse aortic sections were imaged for 10 h at a time interval of 15 min. **(D)** DRAQ5 MFI of individual cells plotted over time. **(E)** Distribution of pulse peak numbers in ECs, BCs, and IAHCs tracked over at least 10 consecutive frames (718 cells). The data represent the mean ± SEM (n = 15). Statistical significance was calculated using two-way ANOVA with Bonferroni post test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). **(F)** Representative normalized Venus MFI tracks of each EHT subset (top) and BCs that undergo high fold change in MFI (bottom). Each track was normalized according to the local MFI maximum (=1, shown on the y axis and highlighted by a red asterisk). The local maxima of all tracks were aligned (shown in the gray bar). Images were acquired at a frame rate of 15 min, and each tick in the x axis represents 10 frames (2.5 h).

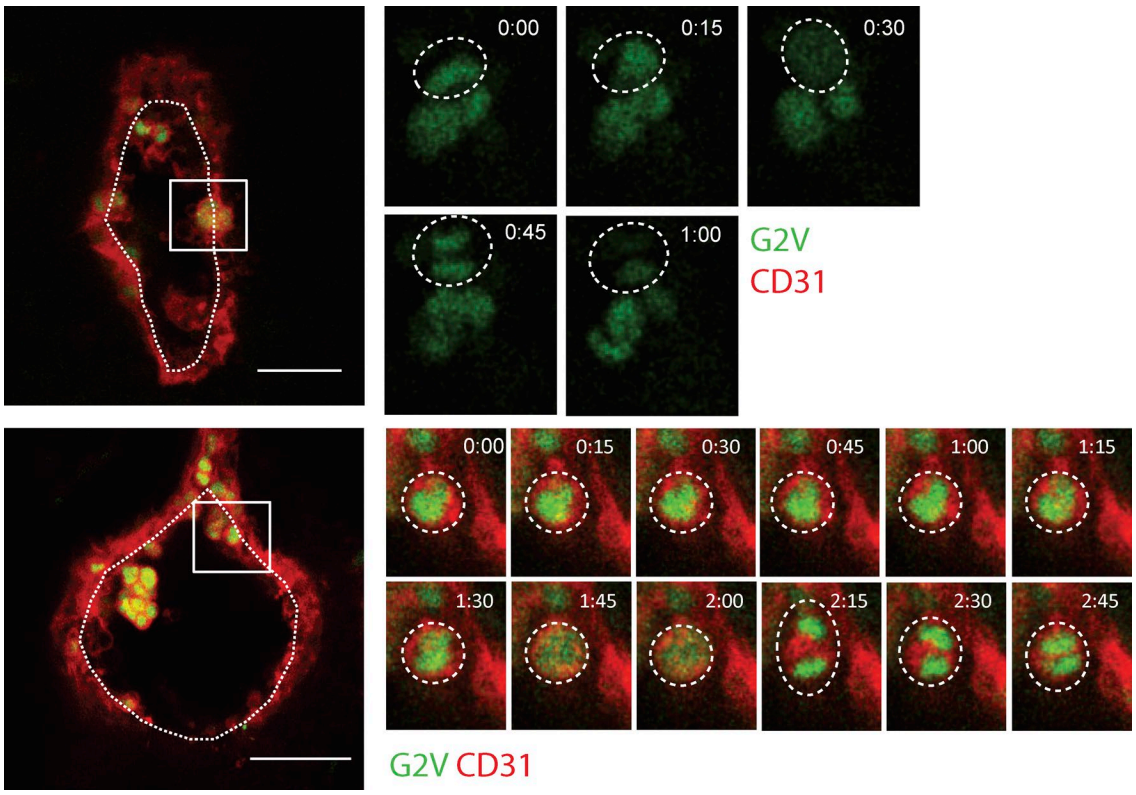


Figure S4. **Proliferation of Venus$^+$ cells observed during time-lapse imaging.** Confocal time-lapse imaging of E10 (33–34 SPs) *G2V* embryos (Venus, green) stained with anti-CD31 (red) antibody. Transverse aortic sections were imaged for 10 h at time intervals of 15 min. Bars, 50 µm. Dashed circles indicate proliferation events observed in IAHCs.

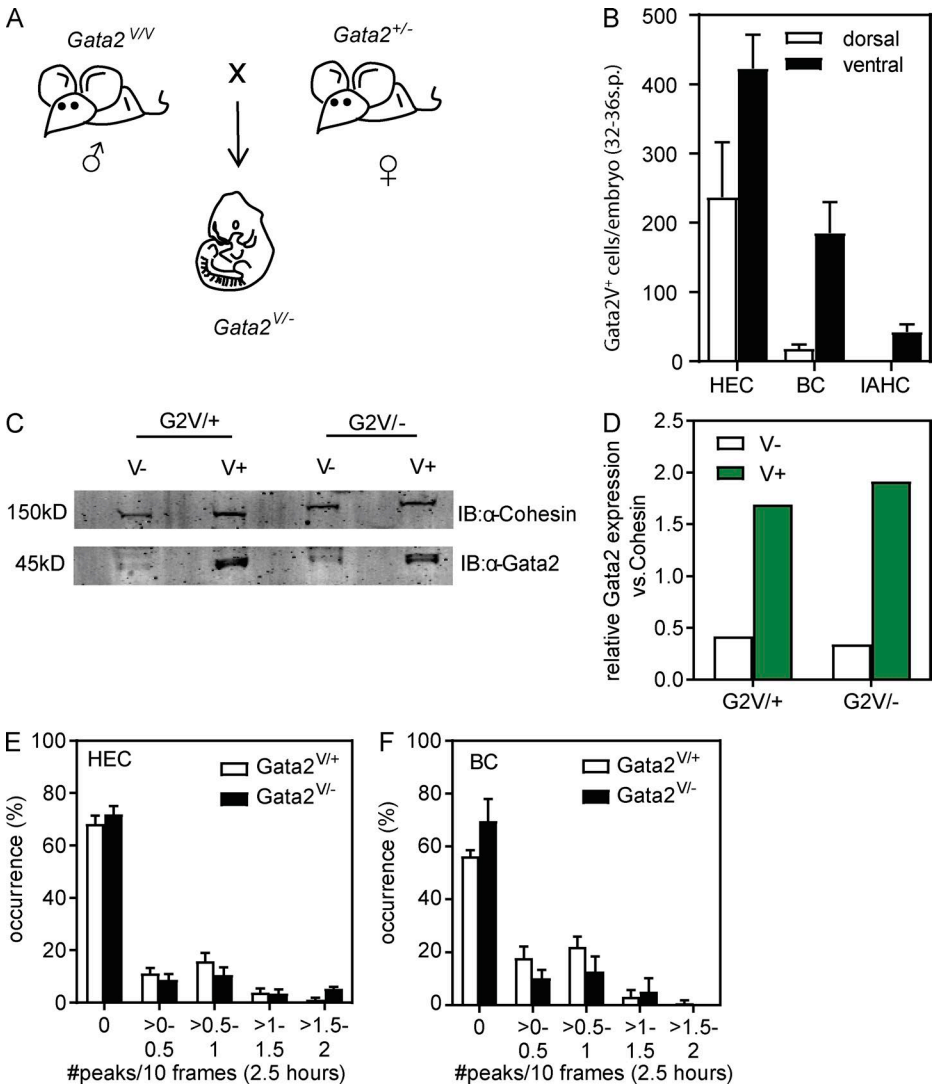
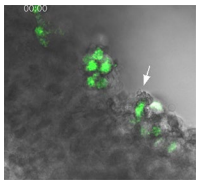
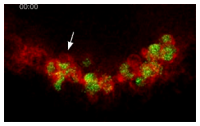


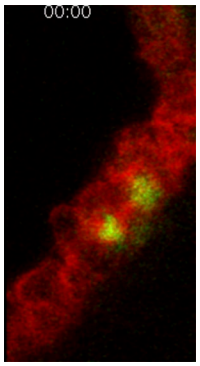
Figure S5. **Gata2 expression characteristics in *Gata2* heterozygous mutant embryos. (A)** Breeding scheme to obtain E10.5 *Gata2*$^{V/-}$ embryos: male G2V (*G2*$^{V/V}$) mice were crossed with *Gata2*$^{+/-}$ females. **(B)** Mean number of Venus^{+} EHT subset cells in *Gata2*$^{V/-}$ embryos per embryo (four to six independent experiments, 194 cells, and 12 slices). **(C)** Gata2 immunoblot of 4×10^4 sorted Venus^{+} cells of *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ E10.5 embryos. Cohesin labeling served as the protein loading control. **(D)** Quantification of Gata2 immunoblot signal relative to Cohesin signal. **(E and F)** Distribution of normalized Gata2 pulse peak numbers in *Gata2*$^{V/+}$ and *Gata2*$^{V/-}$ embryos in HECs (E) and BCs (F). Cells were tracked over at least 10 consecutive frames. To normalize for difference in track length, the data are presented as peaks per 10 frames (2.5 h) and represent the mean ± SEM (n = 15). Related to Fig. 6.



Video 1. **Confocal time-lapse imaging of E10 (33–34 SPs) *G2V* embryo thick section.** Transverse aortic sections were imaged for 10 h at 15-min intervals (Venus, green). Arrow indicates EHT events. Related to Fig. 2 A.



Video 2. **Confocal time-lapse imaging of E10 (33–34 SPs) *G2V* embryo thick section stained with anti–CD31 (red) antibody.** Transverse aortic sections were imaged for 10 h at 15 min intervals (Venus, green). Arrow indicates EHT events. Related to Fig. 2 B.



Video 3. **Confocal time-lapse imaging of E10 (33 SPs) *G2V* thick transverse aortic section stained with anti–CD31 antibody.** Transverse aortic sections were imaged for 10 h at a time interval of 15 min (Venus, green). Arrow indicates EHT events. Related to Fig. 2 C.

REFERENCES

Boisset, J.C., C. Andrieu-Soler, W.A. van Cappellen, T. Clapes, and C. Robin. 2011. Ex vivo time-lapse confocal imaging of the mouse embryo aorta. *Nat. Protoc.* 6:1792–1805. https://doi.org/10.1038/nprot.2011.401

de Bruijn, M.F., X. Ma, C. Robin, K. Ottersbach, M.J. Sanchez, and E. Dzierzak. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity.* 16:673–683. https://doi.org/10.1016/S1074-7613(02)00313-8

Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation

Parham Solaimani Kartalaei,[1]* Tomoko Yamada-Inagawa,[1]* Chris S. Vink,[1] Emma de Pater,[1] Reinier van der Linden,[1] Jonathon Marks-Bluth,[4] Anthon van der Sloot,[2] Mirjam van den Hout,[2] Tomomasa Yokomizo,[5] M. Lucila van Schaick-Solernó,[3] Ruud Delwel,[3] John E. Pimanda,[4] Wilfred F.J. van IJcken,[2] and Elaine Dzierzak[1]

[1]Erasmus MC Stem Cell Institute, Department of Cell Biology, [2]Center for Biomics, and [3]Department of Hematology, Erasmus University Medical Center, 3000 CA Rotterdam, Netherlands
[4]Lowy Cancer Research Centre and Prince of Wales Clinical School, University of New South Wales, Sydney, New South Wales 2052, Australia
[5]Cancer Science Institute of Singapore, National University of Singapore, Singapore 117599

Hematopoietic stem cells (HSCs) are generated via a natural transdifferentiation process known as endothelial to hematopoietic cell transition (EHT). Because of small numbers of embryonal arterial cells undergoing EHT and the paucity of markers to enrich for hemogenic endothelial cells (ECs [HECs]), the genetic program driving HSC emergence is largely unknown. Here, we use a highly sensitive RNAseq method to examine the whole transcriptome of small numbers of enriched aortic HSCs, HECs, and ECs. Gpr56, a G-coupled protein receptor, is one of the most highly up-regulated of the 530 differentially expressed genes. Also, highly up-regulated are hematopoietic transcription factors, including the "heptad" complex of factors. We show that *Gpr56* (mouse and human) is a target of the heptad complex and is required for hematopoietic cluster formation during EHT. Our results identify the processes and regulators involved in EHT and reveal the surprising requirement for Gpr56 in generating the first HSCs.

CORRESPONDENCE
Elaine Dzierzak:
e.dzierzak@erasmusmc.nl
OR
Elaine.Dzierzak@ed.ac.uk

Abbreviations used: AGM, aorta-gonad-mesonephros; BCV, biological coefficient of variation; ChIP, chromatin immunoprecipitation; DEG, differentially expressed gene; EC, endothelial cell; EHT, endothelial to hematopoietic cell transition; FDR, false discovery rate; HC, hematopoietic cell; HEC, hemogenic EC; hpf, hour post fertilization; HSC, hematopoietic stem cell; ISH, in situ hybridization; MO, morpholino oligo; qPCR, quantitative PCR; TF, transcription factor.

Hematopoietic stem cells (HSCs) are responsible for the life-long maintenance and regeneration of the adult vertebrate blood system. HSCs are generated through a natural transdifferentiation process occurring in specialized embryonic vascular cells, known as hemogenic endothelial cells (ECs [HECs]). In mice, the first adult HSCs are generated in the aorta-gonad-mesonephros (AGM) region at embryonic day (E) 10.5 (Müller et al., 1994; Medvinsky and Dzierzak, 1996). The emergence of the definitive hematopoietic system in the mouse embryo correlates with the temporal appearance of clusters of hematopoietic cells (HCs) associated with the aortic endothelium and the major arteries (Garcia-Porrero et al., 1995; North et al., 1999; de Bruijn et al., 2000). Chick embryo dye-marking studies were the first to show that aortic ECs give rise to HCs (Jaffredo et al., 1998). In mammalian embryos, the results of phenotypic and genetic studies, supported by stringent in vivo transplantation studies of enriched cell fractions, demonstrate that HSCs are derived from vascular ECs during a short window of developmental time (de Bruijn et al., 2002; North et al., 2002; Zovein et al., 2008; Chen et al., 2009). This developmental process is known as endothelial to hematopoietic cell transition (EHT).

To facilitate the study of HSC emergence in the mouse embryo, numerous markers have been used individually and/or in combination to identify HSCs and their direct precursors. Immunolocalization of these markers in the

*P. Solaimani Kartalaei and T. Yamada-Inagawa contributed equally to this paper.

E. Dzierzak's present address is University of Edinburgh/ Medical Research Council Centre for Inflammation Research, Queens Medical Research Institute, Edinburgh EH16 4TJ, Scotland, UK.

The Rockefeller University Press $30.00
J. Exp. Med. 2015 Vol. 212 No. 1 93–106
www.jem.org/cgi/doi/10.1084/jem.20140767

AGM highlighted the heterogeneous nature of the cells in the hematopoietic clusters (Ody et al., 1999; Taoudi et al., 2005; Yokomizo and Dzierzak, 2010; Robin et al., 2011). Whereas combinations of these markers allow HSC enrichment, so far no combination of endothelial and/or hematopoietic markers has been able to distinguish hemogenic from nonhemogenic aortic ECs.

The *Ly6aGFP* (Sca1) mouse model, in which all HSCs throughout development are GFP$^+$ (de Bruijn et al., 2002; Ma et al., 2002), has facilitated the study of EHT. Clear proof of EHT was obtained by real-time imaging of the mouse *Ly6aGFP* embryonic aorta (Boisset et al., 2010). In the E10.5 aorta, at the time when the number of hematopoietic clusters peak (Yokomizo and Dzierzak, 2010), flat endothelial GFP$^+$ cells were observed to transition to morphologically round GFP$^+$ cells that begin to express other HSC markers (Boisset et al., 2010). Real-time imaging of transgenic zebrafish embryos similarly revealed the transition of aortic ECs to HCs (Bertrand et al., 2010; Kissa and Herbomel, 2010), indicating that EHT is an evolutionarily conserved process by which the definitive hematopoietic system of vertebrates is generated.

To specifically understand the molecular program involved in EHT, we set out in this study to identify key genes and processes that are functionally relevant in mouse aortic HECs as they transit to HSCs. Based on the vital imaging of EHT, the *Ly6aGFP* reporter is currently the most tractable marker to distinguish and enrich the HECs that are undergoing EHT from other aortic ECs, and also the emerging HSCs from other HCs. Here we present RNA sequencing data obtained from highly enriched small numbers of relevant EHT cells from *Ly6aGFP* embryos, aortic ECs, HECs, and emerging HSCs. Among the few (530) differentially expressed genes (DEGs) during EHT, *Gpr56* is the highest up-regulated gene encoding a cell surface receptor. We show for the first time the functional involvement of Gpr56 in HSC emergence during EHT. In addition, the previously described "heptad" transcription factors (TFs; Wilson et al., 2010) are up-regulated during EHT, bind the Gpr56 enhancer, and regulate its expression. This unique dataset expands our understanding of EHT, identifying the gene networks and processes that are essential for HSC generation in the embryo.

RESULTS

Temporal-spatial and transcriptomic quantitation of aortic hemogenic endothelial and emerging HCs

Ly6aGFP expression marks HCs emerging from hemogenic endothelium at the time of HSC generation in the midgestation mouse aorta. To quantify and localize these cells, we performed confocal imaging of whole and sectioned immunostained E10 *Ly6aGFP* embryos (Fig. 1, A–D). CD31 marks all ECs and HCs, and cKit marks all HCs. However, Ly6aGFP marks only some ECs and some HCs. High-resolution imaging of transverse sections allowed quantitation of four different Ly6aGFP-expressing aortic cell types (Fig. 1 D): flat ECs, bulging cells in the single layer of endothelium, and two differently positioned round cells within the clusters distinguished by the close attachment to (juxtaposed) or a position distal from the endothelium. The total number of GFP$^+$ cells increased from 287 at early E10 (32 sp [somite pairs]) to 1,592 at late E10 (37 sp; Fig. 1 D). From the small fraction of ECs that express GFP (range 13–19%), most aortic GFP$^+$ cells are flat ECs with only 8% of GFP$^+$ cells in hematopoietic clusters. Although by a random distribution more GFP$^+$ cluster cells would be expected in distal positions (as compared with juxtaposed), we observed 70–88% of the GFP$^+$ HCs localized in a juxtaposed position, most likely because GFP$^+$ HCs are emerging from GFP$^+$ ECs and/or are actively maintained at the juxtaposed position.

Because all HSCs are GFP$^+$ and rare HCs have been observed to emerge from GFP$^+$ ECs, Ly6aGFP is the best marker for high enrichment of HECs. Hence, we developed an enrichment method using the Ly6aGFP, cKit, and CD31 markers: ECs (CD31$^+$cKit$^-$GFP$^-$), HECs (CD31$^+$cKit$^-$GFP$^+$), HSCs (CD31$^+$cKit$^+$GFP$^+$), and progenitor/differentiated HCs (CD31$^+$cKit$^+$GFP$^-$; Fig. 1 B). The distinct cell types were sorted (Fig. 1 E) and hematopoietic function was assessed. Hematopoietic progenitors were found in the HC (64%) and HSC (33%) fractions as expected, with a majority of the immature progenitors (CFU-GEMM) in the HSC fraction (Fig. 1 F). In vivo transplantation assays revealed that only the CD31$^+$cKit$^+$GFP$^+$ fraction contained HSCs (Fig. 1 G). These HSCs provided long-term high-level multilineage repopulation of adult irradiated recipients (Fig. S1 A; 6 engrafted of 10 injected with 1–5 ee [embryo equivalents]). Despite injection of high embryo equivalents of cells from the other fractions (4–9 ee), no repopulation was found with the ECs, HECs, or HCs.

RNA sequencing and validation

Sorted ECs, HECs, HSCs, and HCs from three independent biological replicates were used for RNA sequencing. As few as 4–14 E10.5 AGM equivalents (34–41 sp) of sorted cells per replicate were obtained, and cDNA was made from as few as 593 sorted cells (see Table S1 for details). The sequence reads of EHT cell fractions were mapped to the mouse genome (NCBI build 37/mm9), and the generated count table (with 7–57 million unique mapped reads to exons per sample) was normalized and analyzed by *edgeR* (Fig. 2 A; McCarthy et al., 2012).

To confirm that the transcriptome analysis was representative of the sorted EHT fractions, we measured the normalized number of fragments (in FPKMs [fragments per kilobase exon reads per million fragments mapped]) of *CD31*, *cKit*, and *Ly6a* (Fig. 1 H). As expected, *CD31* transcripts were found in all four subsets (ECs, HECs, HSCs, and HCs), *cKit* transcripts were found only in HCs and HSCs, and *Ly6a* transcripts were found in HECs and HSCs.

Gene transcript reads for endothelial genes *Cdh5*, *Tek*, *Esam*, *Kdr*, and *Eng* were highest in HECs as compared with ECs and were higher in HSCs than in ECs or HCs. When the four cell fractions were examined by FACS (Fig. 1 I), cell surface expression correlated significantly with transcript levels ($r^2 = 0.54$, $P = 0.01$). Thus, our datasets reflect a dynamic transcriptional program during EHT.

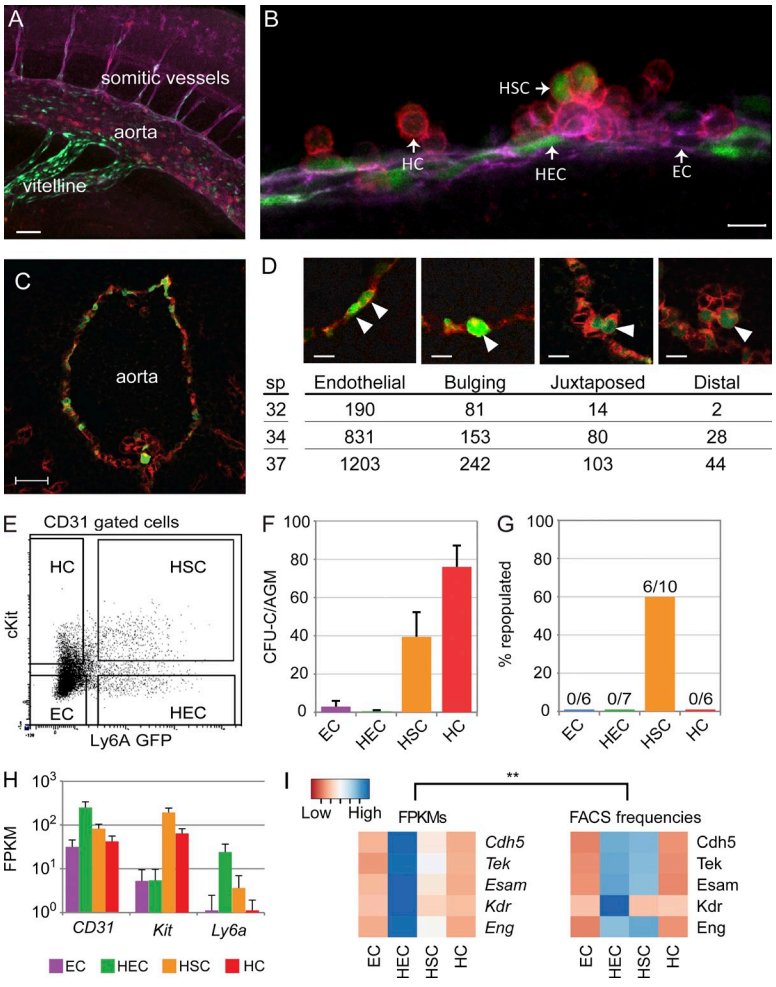


Figure 1. Analysis of EHT cell subsets. (A) Whole-mount image of a 34-sp *Ly6aGFP* embryo showing expression of CD31 (magenta), cKit (red), and GFP (green). The aorta, vitelline artery, and somatic vasculature are indicated. (B) Four types of aortic cells during EHT in a *Ly6aGFP* AGM section (36 sp) stained with anti-CD31 (magenta) and anti-cKit antibodies (red). ECs are $CD31^+cKit^-GFP^-$, HECs are $CD31^+cKit^-GFP^+$, HSCs are $CD31^+cKit^+GFP^+$, and HCs are $CD31^+cKit^+GFP^-$. (C) Transverse section through a 36-sp *Ly6aGFP* embryo showing expression of CD34 (red) and GFP (green). A hematopoietic cluster with some GFP^+ cells is located ventrally. GFP^+ ECs are scattered throughout the aorta. (D) Different GFP^+ cell types (arrowheads) in an E10.5 *Ly6aGFP* aorta. (Endothelial) two flat GFP^+ ECs; (bulging) rounding-up of a GFP^+ EC; (juxtaposed) round HC closely adhering to an EC; (distal) round HC on the distal side of the cluster. The number of cells/aorta is listed below at the 32-, 34-, and 37-sp stages. Bars: (A) 100 µm; (B and D) 10 µm; (C) 50 µm. (E) Scatter plot showing the distribution and sorting gates for EHT subsets EC, HEC, HSC, and HC from E10.5/E11 *Ly6aGFP* AGMs. (F) Hematopoietic progenitor numbers (total CFU-C [CFU-culture]) per AGM. EHT subsets from E10.5 *Ly6aGFP* AGMs (34–39 sp) were plated in methylcellulose, and colonies were counted at day 12 (SD is shown; $n = 4$). (G) HSC long-term repopulating activity in E11 *Ly6aGFP* AGM EHT subsets (40–49 sp). Irradiated adult recipients ($n = 4$) were injected with 5–9 ee of ECs, 4–9 ee of HECs, 1–5 ee of HSCs, and 4–8 ee of HCs together with 2×10^5 spleen cells (recipient type). Percentage of donor cell chimerism at 4 mo after injection is shown. Indicated above each bar is the number of repopulated recipients/number of recipients injected. (H) Normalized number of mapped fragments for genes encoding the markers used for sorting EHT fractions. FPKMs of *CD31*, *cKit*, and *Ly6aGFP* per fraction are shown (error bars are SD). (I) Heat map of FPKMs for genes encoding several relevant cell surface molecules: *Cdh5*, *Tek*, *Esam*, *Kdr*, and *Eng* in each of the sequenced cell fractions and the frequency of cells in each sorted fraction expressing the corresponding protein. Significant positive correlation is observed between FACS and RNAseq data ($r^2 = 0.54$; **, $P = 0.01$).

Global transcriptional differences between the EHT cell subsets

Biological coefficient of variation (BCV) analysis indicates (Fig. 2 B) that EC, HEC, and HSC fractions are closely related but distinct. EC, HEC, and HSC replicate 2 and 3 samples cluster together, whereas replicate 1 EC, HEC and HSC samples show a similar BCV pattern but are further dispersed in the plot. The tighter sample dispersion of replicates 2 and 3 is most likely the result of the higher sequencing depth (Table S1). Hence, distinct transcriptional variation between the EHT fractions is consistent for the three biological replicates.

Dataset comparisons showed a total of 530 DEGs (false discovery rate [FDR] < 0.05; Fig. 2 C and Table S2). The EC to HEC comparison shows 139 DEGs, whereas 340 genes were differentially expressed between HECs and HSCs. Moreover, comparison of ECs with HSCs identified 108 additional genes. MA plots of differential expression analysis show most genes being centered around zero, further confirming the correct normalization of datasets (Fig. S2 A). In the EC to HEC comparison, most DEGs are up-regulated, whereas a majority of the DEGs in HEC to HSC are down-regulated. Hierarchical clustering of DEGs grouped the three biological replicates of each fraction together, suggesting that ECs, HECs, and HSCs have recognizably distinct genetic programs (Fig. 2 D).

Transcriptome analysis reveals processes involved in EHT

DEGs were grouped based on their relative expression levels (H, high; I, intermediate; and L, low expression) into representative patterns for EC genes (HIL, HLL, HHL, and HLI), HEC genes (LHL, IHL, IHI, and LHI), and HSC genes (LHH, LIH, IIH, LLH, and ILH; Fig. 3 and Table S2). Each group was used as input for Gene Ontology (GO), KEGG, and WikiPathways enrichment analysis, and Gene Set Enrichment Analysis (GSEA) was used to detect global shifts of gene sets during each transition.

EC genes show overrepresentation of "focal adhesion," "ECM-receptor interaction," "protein digestion and absorption," "oxidative stress," and "chemokine signaling" terms (Fig. 3 A and Table S3), consistent with EC function (Rajendran et al., 2013). Significant enrichment of "inflammatory response"

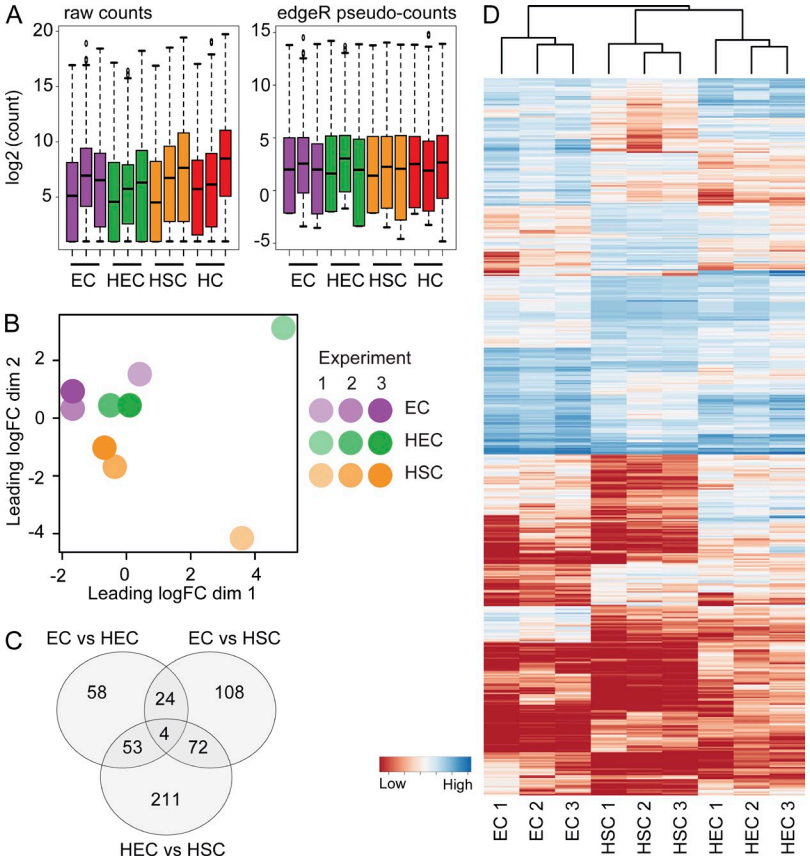


Figure 2. RNAseq data analysis. (A) Distribution of raw counts per sample (left) and *edgeR* internal normalized counts (right). The normalized counts are used in all subsequent analyses. Datasets for three biological replicates are shown for ECs, HECs, HSCs, and HCs. Biological replicate 1 includes two 36-sp and two 37-sp embryos; replicate 2 includes four 34-sp, two 35-sp, three 36-sp, and five 37-sp embryos; replicate 3 includes two 35-sp, one 36-sp, two 38-sp, one 39-sp, one 40-sp, and one 41-sp embryos. (B) BCV in RNAseq samples from three biological replicates of relevant EHT cell fractions: EC, HEC, and HSC. (C) Venn diagram showing numbers of DEGs in comparisons of HECs versus ECs, HSCs versus HECs, and HSCs versus ECs. Total DEGs is 530 (see Table S2 for gene lists). (D) Heat maps showing all 530 DEGs and hierarchical clustering of the genes in each EHT cell fraction from the three biological replicates.

and "TGFbeta signaling" terms suggests ECs to be activated. Whether this is caused by activated endothelium in an actual inflammatory response or by the activation of inflammatory genes that are involved in other signaling pathways in development (Orelio and Dzierzak, 2003; Orelio et al., 2009) requires further study.

HEC genes are most enriched in "cell adhesion" and "migration" gene sets, consistent with changes required for HECs to move out of the endothelial layer and form clusters of HCs (Fig. 3 B and Table S3). Like ECs, 169 out of 191 genes within the "angiogenesis" GO term (GO:0001525) were present in HECs. HEC genes also showed enrichment in "cardiovascular system development," most likely because of the presence of many angiogenic and vascular development genes. Indeed, only 3 (*Jag1*, *Sox17*, and *Fbn1*) out of the 76 published essential cardiovascular genes (Van Handel et al., 2012) are present in HEC genes. "Delta-Notch" and "Notch" pathways known to be important for HSC generation and cluster formation (Kumano et al., 2003) were enriched, and whole transcriptome comparisons using GSEA also show Notch pathway gene sets as up-regulated in HECs (Fig. 4 A and Fig. S2 B). In addition, multiple "VEGF" and "hypoxia-regulated" gene sets were enriched in HECs as compared with HSCs and ECs (Fig. 4 A), including *Hif1α*, which was recently shown to be an important factor for HSC generation (Imanirad et al., 2014). Surprisingly, several previously published HSC gene sets (from MSigDB database v4.0) are significantly enriched in HECs as compared with ECs, whereas no significant enrichment of these genes was found in HSCs as compared with HECs, indicating that the hematopoietic program is already activated in HECs (Fig. 4 A).

HSC genes showed clear overrepresentation of "hematopoietic processes," "cell cycle," and "histone methylation" related genes (Fig. 3 C). Significant enrichment of "cell cycle progression," "DNA replication," and "hematopoietic progenitor" sets was also detected by GSEA (Fig. 4 A and Fig. S2 C). "Hematopoietic progenitor" gene sets are enriched in HSCs as compared with HECs, and detection of "acute myeloid leukemia" from the KEGG database and "pluripotency network" from WikiPathways is in agreement with the acquisition of hematopoietic fate and self-renewal capacity in HSCs (Table S3). This is further supported by significant enrichment of gene sets characteristic of stem cells, such as "telomere lengthening" and "DNA repair," in the HSC fraction by GSEA (Fig. 4 A; Yui et al., 1998; Rossi et al., 2005).

TF expression in cells undergoing EHT

The genetic program directing cell identity is coordinated by TFs, and thus we focused our attention on these genes in our EHT datasets. As compared with ECs, significant up-regulated expression of *Mecom*, *Notch1* and *4*, *Gfi1*, *Sox17*, *Ets2*, and *Elk3* was found in HECs (Table 1), with *Sox7*, *Sox18*, *Runx1*, *Hhex*, and *Lmo2* among the top up-regulated HEC TFs (Table S4).

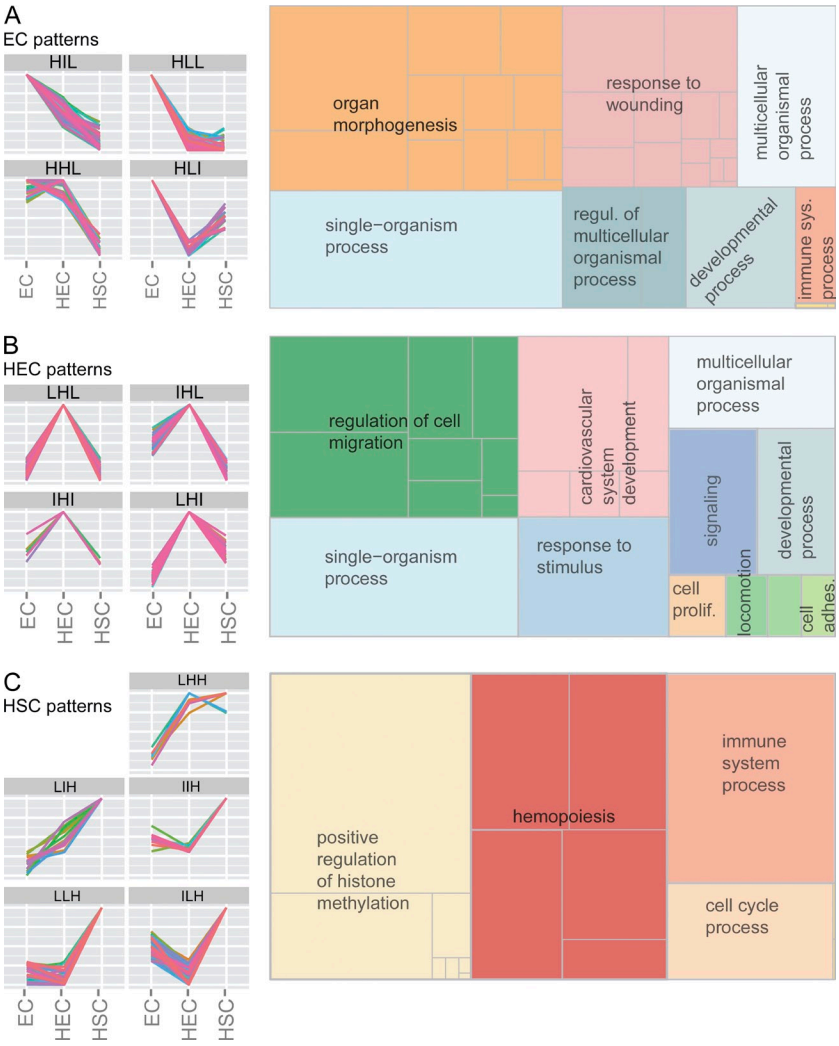


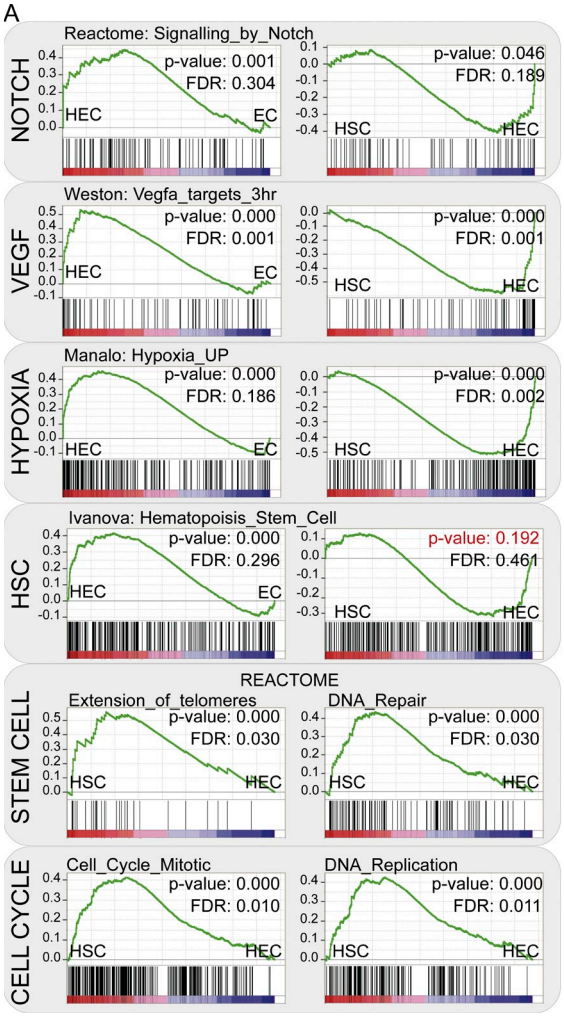
Figure 3. GO terms/processes enriched in EHT subsets. (A) DEG patterns that are EC specific are shown (left): high-intermediate-low (HIL), HHL, HLI, and HLL. GO enrichment analysis was performed using WebGestalt, and enriched GO terms are summarized by REVIGO (right). (B) DEG patterns that are HEC specific are shown (left): LHL, IHL, IHI, and LHI. GO enrichment analysis and GO terms are summarized by REVIGO (right). (C) DEG patterns that are HSC specific are shown (left): LHH, LIH, IIH, LLH, and ILH. GO enrichment and GO terms are summarized by REVIGO. Rectangle size represents the number of DEGs in the accompanying GO term. See Table S3 for enriched ontology terms.

In the HSC fraction, many TFs with known roles in HSC development, including *Etv6*, *Gfi1*, *Gfi1b*, *Myb*, *Myc*, *Hlf*, *Meis1*, *Hhex*, *Runx1*, *Mpl*, and *Ikzf1* and *2* (Table 1 and Table S4), were found to be significantly up-regulated as compared with ECs or HECs. We identified several novel TFs not previously reported to be involved in embryonic HSC generation such as zinc-finger proteins *Zfp106*, *Zfp445*, *Zfp748*, and *Zfp763*, megakaryocyte factor *Nfe2*, transcriptional corepressor *Bcor*, and *Cbfa2t3* (*Eto2*). Also present in the top hits were chromatin-remodeling factors *Suz12*, *Paxip1*, *Kdm5a*, *Smarca4* (*Brg1*), *Ezh1*, *Bptf*, and *Hdac1* and de novo DNA methylation genes *Dnmt3a/Dnmt3b* and *Dnmt1*. The down-regulation of several *Hox*, *Tbx*, and *Fox* genes was observed in the EC to HEC and HEC to HSC transition, whereas only *Hoxa9*, *Hhex*, and *Foxk1* were up-regulated in HSCs as compared with ECs or HECs (Table S4).

It has been shown that a pivotal (heptad) group of TFs work together in transcriptional regulatory complexes to regulate the expression of downstream target genes in hematopoietic progenitor cell lines (Wilson et al., 2010). The heptad TFs could act as one of the transcriptional hubs for the regulation of EHT. Our RNAseq datasets reveal that all heptad TFs increase during EHT (Fig. 5 A). To identify genes encoding novel EHT and embryonic HSC surface markers, the 530 DEGs were compared with the 927 heptad TF targets identified by chromatin immunoprecipitation (ChIP)seq analyses in HPC7 cells. 58 DEGs were found to be targets of heptad TFs, with *CD34* and *Gpr56* as the top hits (Fig. 5, B and C). Interestingly, also in the whole transcriptome analysis of EHT, *Gpr56* was identified as the top differentially expressed receptor gene in the HEC to HSC transition, followed by *cKit* (Fig. 4 B). Because both CD34 and cKit function has been studied in HSCs and these markers are used extensively for HSC isolation (Sánchez et al., 1996), we focused on Gpr56.

The *Gpr56* heptad consensus region in the mouse is located 37 kb upstream of the translational start site. We identified this region as the *Gpr56-37* enhancer. Enrichment of heptad factors at the *Gpr56-37* element was found in mouse HPC7 cell line by quantitative PCR (qPCR; Fig. 5 D). Transactivation assays in hematopoietic progenitor cell lines showed significant activation of *Gpr56-37* enhancer, whereas overexpression of three of the heptad factors (Gata2, Runx1, and Fli1)

A



B

Receptor activity: EC to HEC

Gene	logFC	FDR
Plxnd1	2.70	2.56E-04
Tie1	3.77	4.70E-04
Calcrl	2.58	5.52E-04
*Ramp2**	4.17	1.43E-03
Tek	2.49	1.60E-03
*Flt1**	2.65	2.54E-03
Gpr116	2.91	4.63E-03
S1pr1	3.15	7.73E-03
Eng	3.24	8.16E-03
Eltd1	3.40	9.80E-03
*Kdr**	2.23	2.55E-02
Notch1	2.16	4.17E-02
*Treml1**	-12.78	1.79E-04
F2rl2	-5.54	8.77E-03
Stra6	-6.98	2.84E-02

Receptor activity: HEC to HSC

Gene	logFC	FDR
Gpr56	4.88	1.36E-05
Kit	5.19	4.24E-04
Treml2	7.23	7.56E-03
Nlrp6	10.50	3.47E-02
Grm2	9.61	4.72E-02
*Treml1**	7.38	4.90E-02
Fzd4	-6.53	1.56E-03
*Flt1**	-2.54	3.13E-03
P2ry6	-5.60	8.98E-03
Nrp1	-2.75	1.03E-02
Gpr183	-4.20	1.14E-02
Lifr	-2.60	1.57E-02
Erbb2	-7.29	1.58E-02
Acvrl1	-3.67	1.79E-02
*Ramp2**	-3.18	1.86E-02
Stab2	-3.64	2.39E-02
Olfr31	-9.27	2.41E-02
Qrfpr	-8.34	2.45E-02
Gpr126	-5.34	2.76E-02
Lpar4	-2.26	3.42E-02
Itgb4	-6.30	3.45E-02
Npr1	-3.59	3.47E-02
Mrgprd	-10.01	3.47E-02
Olfr118	-9.16	3.57E-02
*Kdr**	-2.01	3.83E-02
Gpr65	-4.70	4.08E-02
Clec1b	-3.70	4.53E-02

Figure 4. Changing processes and cell surface molecules during EHT. (A) GSEA for VEGF, Notch, Hypoxia up-regulated genes, genes specifically expressed in HSCs, stem cell function–related gene sets like "telomere lengthening" and "DNA repair," cell cycle–related gene sets, and early hematopoietic progenitor–specific genes. (B) Receptor-related genes with significant expression changes in EC to HEC and HEC to HSC transitions. Blue, increased expression; red, decreased expression; gray font, genes with low overall expression levels as defined by edgeR-calculated logCPM of <3 (and higher probability of being false positive); asterisks, genes differentially expressed during both transitions; logFC, log fold change.

showed synergistic activation of the *Gpr56-37* enhancer (Fig. 5, E and F). Moreover, we identified a homologous element 48 kb upstream of the human *GPR56* gene. In human CD34$^+$ HSC-enriched cells, we found binding of all seven heptad TFs to the human *GPR56-48* element (Fig. 5 G; Chacon et al., 2014). These data suggest that the heptad TFs and their downstream target *Gpr56* are important in HSC generation during EHT, as well as in healthy and leukemic human HCs. Because nothing is known concerning Gpr56 in embryonic hematopoietic development, we examined its regulation and role during EHT.

Gpr56 is required during EHT for HSC generation

To confirm localized expression of *Gpr56* in cells undergoing EHT, we performed in situ hybridization (ISH) analysis of the E10.5 AGM. High-level *Gpr56* expression was observed in some aortic HCs (Fig. 6 A), and no/low expression was observed in aortic ECs. An overlap of *Gpr56* expression with some GFP$^+$ HCs was found by ISH of *Ly6aGFP* E10.5 AGM sections (Fig. 6 B). Thus, the localized expression of *Gpr56* is consistent with FPKM values derived from RNAseq datasets and strongly suggests a role in HSC generation.

Gpr56 is highly conserved across different vertebrate species (mean multiple sequencing alignment score = 85%; Fig. 6 C). To validate the involvement of *Gpr56* in HSC generation, we used a zebrafish morpholino oligo (MO) knockdown approach. At 30 and 48 h post fertilization (hpf), morphants were assayed by ISH (30 hpf) for *cmyb*, a marker for emerging HSCs (Jing and Zon, 2011). WT embryos show *cmyb*-expressing cells along the aorta (Fig. 6 D). In contrast, *cmyb*-expressing cells are severely reduced in *gpr56* MO–injected embryos. To validate that this was a defect in HSC generation, we injected the *gpr56* MO into CD41-GFP transgenic embryos (CD41 marks HCs; Lin et al., 2005; Jing and Zon, 2011; Robin et al., 2011). The number of CD41-GFP$^+$ cells in the caudal hematopoietic tissue at 48 hpf is significantly decreased from 36.7 ± 4.0 cells in WT to 12.5 ± 1.8 in *gpr56* morphants (Fig. 6 E), suggesting that Gpr56 is important for the emergence of HSCs. No abnormalities in embryo growth or the structure of the vasculature/aorta were found by ISH for arterial endothelial marker *gridlock* (*grl*; Fig. 6 F; Zhong et al., 2000). To test whether the *gpr56* MO does not show an off-target effect, we performed rescue experiments by injecting *gpr56* mRNA. *Gpr56* morphants could be rescued with zebrafish

Table 1. Differentially expressed TFs

HECs versus ECs			HSCs versus HECs			HSCs versus ECs		
Gene	logFC	FDR	Gene	logFC	FDR	Gene	logFC	FDR
(A) Top 25 up-regulated genes								
Elk3	2.42	1.1E-02	*Myb*	4.23	1.7E-06	*Myb*	4.55	7.1E-08
Mecom	3.21	1.4E-02	*Gfi1b*	5.48	7.7E-05	*Ikzf2*	3.13	2.5E-04
Notch4	5.41	2.9E-02	*Hlf*	7.03	8.9E-05	*Hlf*	5.99	6.7E-04
Notch1	2.16	4.2E-02	*Meis1*	2.20	8.1E-04	*Runx1*	2.68	1.8E-03
Rab11a	2.03	5.2E-02	*Zfp106*	2.43	3.5E-03	*Myc*	2.27	9.5E-03
Gfi1	7.69	6.8E-02	*Ncoa4*	2.05	8.8E-03	*Dnmt3a*	1.73	2.2E-02
Wwtr1	3.20	6.8E-02	*Nop2*	2.35	1.9E-02	*Chd4*	1.62	2.9E-02
Junb	4.90	6.8E-02	*Elf1*	2.51	2.0E-02	*Dnmt1*	1.58	3.2E-02
Ets2	2.50	6.9E-02	*Zfp445*	2.05	2.3E-02	*Gfi1*	8.33	3.3E-02
Mapk3	1.62	7.3E-02	*Nfe2*	5.05	4.1E-02	*Zfp445*	1.91	4.6E-02
Nkx2-3	5.95	7.4E-02	*Zfp763*	4.29	4.2E-02	*Setbp1*	2.78	5.9E-02
Ldb2	2.30	7.5E-02	*Ikzf2*	1.94	5.2E-02	*Ikzf1*	2.60	6.8E-02
Hdac7	2.20	7.6E-02	*Huwe1*	1.55	5.7E-02	*Bcor*	2.21	8.4E-02
Sox17	5.14	8.4E-02	*Orc2*	2.22	6.5E-02	*Trp53bp1*	1.51	8.5E-02
Ctnnb1	1.28	9.5E-02	*Mpl*	2.84	6.6E-02	*Mycn*	2.57	9.9E-02
Hey1	4.24	1.1E-01	*Etv6*	1.99	6.9E-02	*Zfp106*	1.67	1.0E-01
Epas1	3.10	1.2E-01	*Lmo1*	8.87	7.5E-02	*Suz12*	1.57	1.1E-01
Hey2	4.35	1.2E-01	*Paxip1*	1.93	7.6E-02	*Cbfa2t3*	2.63	1.2E-01
Sox7	4.27	1.5E-01	*Cpsf6*	1.28	8.1E-02	*Paxip1*	1.83	1.4E-01
Sox18	4.05	1.5E-01	*Zfp748*	2.68	8.3E-02	*Notch1*	1.72	1.4E-01
Tsc22d1	1.46	1.6E-01	*Polr1a*	2.41	8.3E-02	*Etv6*	1.82	1.5E-01
Nrarp	3.76	1.7E-01	*Trp53bp1*	1.42	8.9E-02	*Kdm5a*	1.38	1.5E-01
Nfic	2.07	1.8E-01	*Dnmt1*	1.29	9.4E-02	*Meis1*	1.22	1.8E-01
Pdlim1	1.96	2.1E-01	*Med23*	1.84	9.8E-02	*Rreb1*	1.49	1.8E-01
Hmg20b	2.46	2.1E-01	*Krr1*	1.45	1.0E-01	*Bptf*	1.17	1.8E-01
(B) Top 25 down-regulated genes								
Pou5f1	−7.70	2.6E-04	*Snai2*	−7.39	7.7E-05	*Snai2*	−8.64	3.0E-06
Utf1	−8.22	1.0E-02	*Id3*	−3.87	5.6E-04	*Rhox6*	−10.09	2.9E-05
Gfi1b	−3.77	1.3E-02	*Hey2*	−7.55	2.3E-03	*Rhox9*	−10.23	4.6E-05
Foxd1	−10.54	6.6E-02	*Rhox6*	−6.49	6.8E-03	*Tgfb1i1*	−5.76	4.6E-05
Lmo1	−10.20	6.7E-02	*Msx2*	−8.54	7.4E-03	*Pou5f1*	−7.16	2.2E-04
Hand1	−6.55	7.4E-02	*Rhox9*	−6.56	9.0E-03	*Etv5*	−3.26	6.0E-04
Zfp612	−7.77	1.0E-01	*Etv5*	−2.60	1.1E-02	*Utf1*	−8.47	3.6E-03
Prrx2	−7.19	1.0E-01	*Isl1*	−4.08	3.6E-02	*Msx2*	−8.98	5.1E-03
Asb12	−9.04	1.1E-01	*Ebf2*	−7.43	4.0E-02	*Id3*	−3.21	5.8E-03
Cdc6	−1.75	1.3E-01	*Ebf1*	−5.39	4.3E-02	*Prrx2*	−10.99	6.9E-03
Ncoa4	−1.48	1.4E-01	*Hey1*	−4.69	4.4E-02	*Prss35*	−5.99	1.8E-02
Krr1	−1.47	1.4E-01	*Epas1*	−3.42	4.6E-02	*Grhl3*	−6.02	2.6E-02
Alx4	−5.93	1.5E-01	*Rarb*	−3.81	5.2E-02	*Foxd1*	−10.54	2.8E-02
Wt1	−4.76	2.0E-01	*Hoxd9*	−7.63	5.5E-02	*Ripk4*	−9.34	3.0E-02
Rhox2e	−3.29	2.0E-01	*Sox17*	−5.00	5.7E-02	*Ebf2*	−8.08	3.0E-02
Pax8	−8.85	2.1E-01	*Creb3l1*	−4.59	5.8E-02	*Foxp2*	−3.93	3.4E-02
Rhox6	−3.59	2.2E-01	*Ugp2*	−1.79	6.5E-02	*Zim1*	−7.30	3.4E-02
Neurod6	−6.02	2.2E-01	*Nr3c1*	−2.93	6.9E-02	*Creb3l1*	−5.25	3.5E-02
Hoxc4	−8.73	2.3E-01	*Notch4*	−4.21	7.1E-02	*Hoxd9*	−8.66	3.6E-02
Rhox9	−3.68	2.3E-01	*Myt1*	−6.02	7.6E-02	*Sall4*	−2.66	3.8E-02
Six1	−7.30	2.3E-01	*Elf3*	−6.86	7.6E-02	*Id2*	−2.69	3.9E-02

(A and B) Top 25 up-regulated TFs (A) and top 25 down-regulated TFs (B) in HEC versus EC, HSC versus HEC, and HSC versus EC comparisons. FDR, FDR corrected p-value; logFC, log fold change. All genes with FDR < 0.05 except genes with underlining.

Table 1. *(Continued)*

HECs versus ECs			HSCs versus HECs			HSCs versus ECs		
Gene	logFC	FDR	Gene	logFC	FDR	Gene	logFC	FDR
Tcf21	−5.63	2.3E-01	*Ankrd1*	−4.84	7.7E-02	*Hoxc10*	−5.83	5.0E-02
Runx1t1	−1.79	2.4E-01	*Onecut3*	−8.32	8.3E-02	*Myt1*	−6.64	6.7E-02
Klf1	−6.56	2.6E-01	*Hivep3*	−5.19	8.4E-02	*Elf3*	−7.56	6.9E-02
Aff3	−3.18	3.1E-01	*Zfp36l1*	−1.90	8.6E-02	*Isl1*	−3.79	7.0E-02

(A and B) Top 25 up-regulated TFs (A) and top 25 down-regulated TFs (B) in HEC versus EC, HSC versus HEC, and HSC versus EC comparisons. FDR, FDR corrected p-value; logFC, log fold change. All genes with FDR < 0.05 except genes with underlining.

gpr56 mRNA, as well as with a mouse *Gpr56* mRNA to yield full restoration of aortic *cmyb* expression (Fig. 6 G). Interestingly, some ectopic expression of *cmyb* is observed in the region ventral to the aorta. These data indicate that *gpr56* is an essential player in the HSC generation program and that its functional domains are maintained between mouse and zebrafish. Ectopic generation of phenotypic HSCs in zebrafish by Gpr56 overexpression further highlights the unexpected function of this molecule in induction of HSC generation.

To further study the function of Gpr56 in HSCs, we used the 32D-CSF3R cell line, a unipotent mouse stem cell differentiation model in which colony-stimulating factor-3 stimulates their differentiation to neutrophils. (Fig. 6 H). When stimulated with CSF3, cells transduced with an empty vector or WT human *GPR56* vector lost their blast characteristics and differentiated. However, cells transduced with constitutively active (MUT) human *GPR56* resulted in an increase in blast-like cells and in significantly fewer differentiated

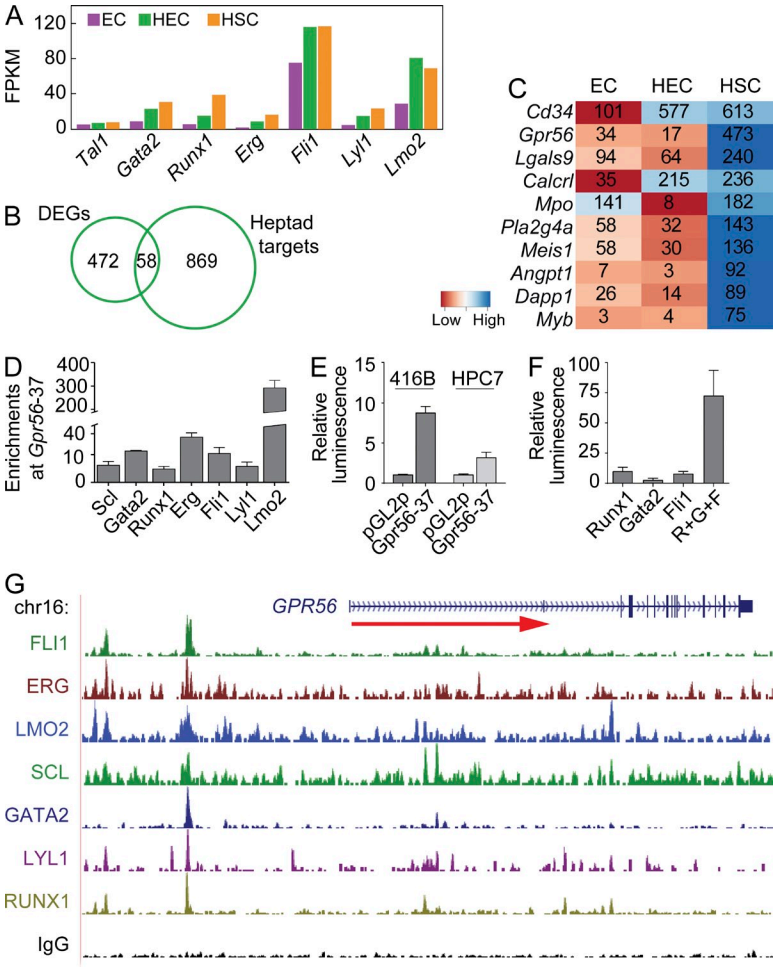


Figure 5. ***Gpr56* is a heptad target in mouse and human blood progenitors.** (A) Mean FPKM values of heptad factors in EC, HEC, and HSC fractions. (B) A Venn diagram showing the overlap between sites with combinatorial binding of Scl, Gata2, Runx1, Erg, Fli1, Lyl1, and Lmo2 in HPC7 cells (Heptad targets) and 530 DEGs during EHT. (C) Heat map of top 10 heptad target DEGs based on highest expression in HSCs and with respective mean FPKM values inside heat map. (D) qPCR for TF enrichment at *Gpr56-37* as compared with IgG and control in HPC7 mouse myeloid progenitor cells ($n = 4$). (E) Transfection assays in 416B and HPC7 mouse progenitors show enhancer activity of *Gpr56-37* ($n = 3$). (F) Transactivation assays in Cos7 cells showing synergistic responsiveness of the *Gpr56-37* element to Runx1, Gata2, and Fli1 ($n = 4$). (D–F) Error bars show SD. (G) TF binding at Hs*GPR56-48* (corresponding region to Mm*Gpr56-37*) in primary human CD34 HSCs.

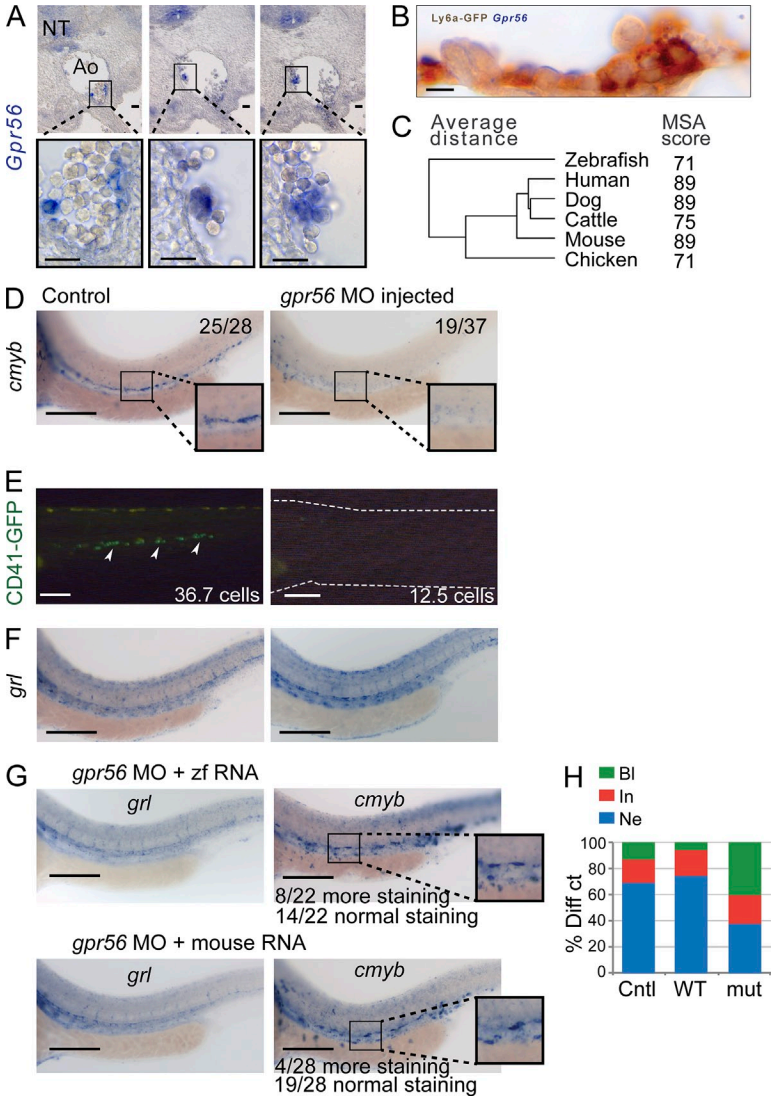


Figure 6. In silico and in vivo analysis of *Gpr56*. (A) ISH of WT mouse E10.5 AGM sections shows specific expression of *Gpr56* in some HCs, a few cells lining the aorta (Ao), and the notochord (NT). The top images show low magnification of AGM cross-section, and the bottom images show high magnification of the boxed areas. (B) ISH of E10.5 *Ly6aGFP* AGM shows coexpression of GFP and *Gpr56* in some HCs. (C) Homology relationships of the *Gpr56* coding sequence of different vertebrate species. (D–F) Analysis of WT and *gpr56* MO zebrafish for the presence of HSCs. (D) ISH with the HSC marker *cmyb* at 30 hpf. (E) Fluorescent analysis of WT and MO-injected CD41:GFP transgenic embryos at 48 hpf. Numbers in the panels indicate the number of embryos with the depicted phenotype. Arrowheads (left) indicate CD41-expressing HCs in the aorta. The dashed lines (right) indicate the outline of the morphant zebrafish embryo for orientation purposes. (F) ISH with arterial cell marker *grl*. No vascular or developmental abnormalities can be observed in *gpr56* morphant embryos. (G) HSC rescue of *gpr56* morphant zebrafish with *gpr56* RNA (zebrafish and mouse) as shown by ISH for *cmyb*. Ectopic *cmyb* expression in the posterior cardinal vein is clearly visible. No vascular abnormalities can be observed by *grl* ISH. (D and G) Insets show boxed areas at higher magnification. Bars: (A) 30 µm; (B) 10 µm; (D–G) 100 µm. (H) Effect of human GPR56 activity in neutrophil differentiation of the 32D-CSF3R unipotent stem cell line. 32D-CSF3R cells cultured in medium containing CSF3 efficiently differentiated into neutrophils. Only constitutive active mutant GPR56 (MUT) could block differentiation. Diff ct, differential count; Cntl, empty vector control; WT, WT human GPR56 vector; mut, constitutively active human GPR56 mutant vector; Bl, blast morphology; In, intermediate morphology; Ne, neutrophil morphology.

cells (eight- to fourfold fewer), suggesting that GPR56 is also essential for the maintenance of an undifferentiated cell state.

DISCUSSION

RNA sequencing analyses of EHT and developing HSCs provide a new perspective on the molecules and processes. Whereas previous methods of transcriptome analysis have identified many of the obvious regulators of hematopoietic development, this method provides an accurate accounting of all expressed genes and also small gene expression level changes between the rare, relevant cell types. We have shown here that HECs (the precursors of the earliest emerging HSCs in the midgestation mouse aorta), as distinguished from other ECs by Ly6aGFP expression, have closely related but distinct transcriptional programs. Comparisons between ECs and HECs reveal a developing program indicative of cell migration and changing cell morphology in HECs, while they retain an angiogenic program. The up-regulation of the HSC hematopoiesis program begins in HECs, further distinguishing them from ECs. Of the 530 DEGs, important receptors and TF genes were identified, including *Gpr56*, which is required for HSC formation. Moreover, the heptad (hematopoietic) TFs were found to be up-regulated during EHT. These factors bind the *Gpr56* enhancer and regulate its expression, thus providing a proof of principle for in silico bioinformatical predictions of the combinatory role of the heptad TFs in the emergence of HSCs during EHT. Thus, our datasets are predictive of functionally relevant EHT genes and processes.

RNAseq analysis of small numbers of physiologically relevant cells

The *Ly6aGFP* transgenic marker in combination with CD31 and c-Kit cell surface markers allowed the high enrichment of HSCs, HECs, and ECs isolated from the aorta at the developmental time when HSCs begin to emerge. Imaging experiments verified the correlation between these markers, the expected cell type, and localization within the AGM region.

Moreover, we confirmed that functional adult-repopulating HSCs are exclusively contained within the HSC fraction (0.002% of AGM cells) and that both endothelial fractions (ECs and HECs) do not contain hematopoietic progenitors or HSCs. Thus, the Ly6aGFP marker currently allows the highest level of enrichment for HECs (Fig. S1 B) that will undergo transition to HSCs, as compared with previously used markers.

Previous comparative HSC gene expression profiling (microarray) studies identified several new regulators of AGM HSCs, but the genetic program of HECs was not examined (Mascarenhas et al., 2009; McKinney-Freeman et al., 2012). During preparation of this manuscript, a new microarray study of EHT-relevant populations was performed based on cells expressing the *Runx1+23*-enhancer marker (Swiers et al., 2013). Runx1+23GFP marks 68% of VE-cadherin$^+$ (endothelial and hematopoietic) cells at E8.5, marking many HECs that are not as yet exhibiting EHT. However, Ly6aGFP expression marks only a small fraction (13–19%) of CD31$^+$ aortic cells and is probably more specifically marking the active HECs at E10.5.

Given the limited number of cells in our enriched aortic EC, HEC, and emerging HSC fractions, RNA sequencing provides the most efficient and sensitive method for analysis of EHT-relevant cells. Only 4–14 embryos (aortas) were used per sequencing experiment to isolate sufficient quantities of total RNA from sorted cell populations. With as few as 593 sorted cells, we successfully applied RNAseq technology with the SMARTER protocol, recently shown to be the best RNAseq method for low numbers of cells (Bhargava et al., 2014). Additionally, RNAseq analysis has the great advantage over microarrays in not only providing the whole transcriptome, but also revealing isoform-specific transcripts in the sequenced samples. For example, *Gpr56* expresses two transcript variants. We found variant 1 (GenBank accession no. NM_001198894) of *Gpr56* to be expressed exclusively in HSCs, whereas variant 2 (GenBank accession no. NM_018882) was expressed in ECs, HECs, and HSCs. Variant 2 was more highly represented in HSCs (FPKM = 323) as compared with variant 1 (FPKM = 52).

Identification of processes involved in EHT

For the first time, datasets from aortic ECs, HECs, and HSCs provide an overview of the general processes involved during EHT. Quantitative levels of gene expression between EHT-enriched cell fractions show only a small number of significant DEGs: 139 between HECs versus ECs and 340 between HSCs versus HECs. Not surprisingly, the genes with high expression in midgestation aortic ECs are mainly those involved in "general developmental processes." These and other GO categories related to cell migration and focal adhesion are highly represented in HECs, highlighting the fact that HECs must change their adhesive properties to bulge out of endothelial lining of the aorta, undergo morphological changes as they become HCs, adhere to other HCs within the clusters, and take on hematopoietic identity and function. GO analysis of DEGs with the highest expression in HSCs shows enrichment of "hematopoiesis" and "positive regulation of histone methylation" terms.

HECs are a transcriptionally dynamic cell type at the interface of EHT. Concurrent with the initiation of the hematopoietic program and HC formation, hematopoietic genes become activated in HECs, whereas endothelium-specific cell adhesion molecules and TFs are down-regulated in HSCs. Our RNAseq data are in agreement with the recent single-cell high-throughput qPCR analysis results for 18 known endothelial and hematopoietic genes during EHT (Fig. S2 D; Swiers et al., 2013). We also identified several genes involved in angiogenesis by selection for GO term "receptor activity" in HEC versus EC comparisons (Fig. 4 A). These include *Plxnd1*, *Eltd1*, *Calcrl*, *Ramp2*, and *S1pr1*; *Plxnd1* and *Eltd1* are both induced by VEGF (Kim et al., 2011; Masiero et al., 2013). *Calcrl*, a GPCR, induces angiogenesis upon association with Ramp2 and Kdr/Vegf-r2, both of which are significantly induced in HECs (Guidolin et al., 2008). Collectively, these findings suggest a role for angiogenesis-related receptors in activation of hematopoietic potential and generation of HECs.

GPR56: a novel EHT regulator

Gpr56, one of the top hits in our HSC versus HEC comparison (30-fold increase) and bound by all heptad TFs, is indeed a novel regulator for emerging HSCs in the embryonic vasculature. Contrary to expectations raised by the lack of HSC defects in mouse *Gpr56* KO embryos (generated by deletion of the first two exons [Saito et al. 2013]), our RNAseq data suggested a strong role for Gpr56 in emergence of HSCs. In the E10.5 mouse aorta, we localized *Gpr56* expression to a few HCs/HSCs (Ly6aGFP$^+$). Upon *gpr56* knockdown, zebrafish embryos showed severe reduction in HSCs (*cmyb*) and CD41$^+$ hematopoietic stem/progenitor cells, revealing a requirement for Gpr56 in HSC generation. Our rescue experiments in *gpr56* morphants show that both zebrafish and mouse *Gpr56* RNA can restore aortic hematopoietic stem/progenitor generation. Moreover, *Gpr56* overexpression resulted in ectopic hematopoietic progenitor/stem cell formation in the axial vein, suggesting that the Gpr56 signaling axis may be useful for inducing new HSCs.

We propose that the lack of embryonic lethality in *Gpr56* KO embryos could be the result of redundancy by other GPCRs or residual Gpr56 activity in the mouse transgenic model. Our RNAseq and RT-qPCR validation (Fig. S2 E) data show an increase in the expression of *Gpr114* (77 kb upstream of *Gpr56*) and *Gpr97* (48 kb downstream) during EHT. The ligand binding N-terminal part of Gpr114 has 47% amino acid similarity (and 27% identity) with Gpr56. *Gpr114* is present only in mammals. Also, assays testing a human *Gpr56* variant missing a large part of the second exon and the complete third exon showed that it partially retains the ability to activate *SRE*, *E2F*, *NFAT*, and *iNOS* promoters (Kim et al., 2010). Thus, Gpr56 is an unexpected novel EHT regulator essential for HSC generation and maintenance, and its function is conserved between mouse and zebrafish.

How Gpr56 acts in HECs as they transdifferentiate to HSCs is unknown, but it could affect physical properties such as adherence, cluster formation, signal transduction, migration, and/or self-renewal. Some of these features are consistent with findings in neuronal stem cells, BM HSCs, and leukemic cells, in which it has been proposed that Gpr56 functions in cell adhesion, migration, and/or repression of apoptosis (Iguchi et al., 2008; Saito et al., 2013). We found that Gpr56 functions in the maintenance of the undifferentiated state of a unipotential HSC line. The conservation of Gpr56 across species will allow for future high-throughput study of the mechanism by which Gpr56 affects EHT and generation of HSCs.

Our results on heptad TF binding to the Gpr56 enhancer suggest that other heptad targets in the overlapping list are likely to be relevant in EHT. However, not all genes that we identified as highly up-regulated during EHT are targets of the heptad complex, for example *cKit*. Because EHT regulation is likely to be multilayered, we are using our whole transcriptome dataset as a resource to identify other candidate transcriptional hubs.

In summary, novel and known EC, HEC, and HSC genes were identified in our RNAseq datasets. These comparative quantitative data have high predictive value for identifying functionally important molecules that direct the cellular processes involved in EHT and could instruct methods for de novo HSC generation either by direct somatic cell conversion or pluripotent stem cell differentiation.

MATERIALS AND METHODS

Cell preparation and flow cytometry. *Ly6aGFP* and WT mouse embryos were dissected as described previously (Robin and Dzierzak, 2010), and single cells were prepared by collagenase treatment (0.125%, 45 min, 37°C) and washed with PBS, 10% heat-inactivated FCS, and 1% penicillin/streptomycin (PS). Cells were stained with RαMCD31-AF647 (1:400; BioLegend) and RαMcKit-PE (1:1,200; BD) for 30 min at 4°C, washed with PBS/10% FCS/1% PS, and analyzed/sorted on a FACSAria III or SORP FACSAria II (BD).

Mouse embryo immunostaining and imaging. 10-µm cryosections were prepared as described in Ling et al. (2004; except last dehydration steps were omitted), stained with RαMCD34-biotin (1:100; BD) and Streptavidin-Cy5 (1:500; Jackson ImmunoResearch Laboratories, Inc.), and imaged on an SP5 confocal microscope (Leica).

Whole-mount embryos were prepared as described previously (Yokomizo et al., 2012), stained with RαM-cKit (1:500; eBioscience) and αRat–Alexa Fluor 647 (1:5,000; Invitrogen), RαMCD31-biotin (1:500; BD) and Streptavidin–Alexa Fluor 555 (1:500; Invitrogen), RαGFP (1:2,000; MBL) and GαR–Alexa Fluor 488 (1:1,000; Invitrogen), and imaged on an SP5 microscope. 1.48 µm between stacks; 17 stacks (23.7 µm) merged.

For ISH, embryos (36 sp) were fixed and rotated overnight, 4°C in 4% paraformaldehyde, washed three times in PBS at 4°C, and embedded in paraffin, followed by overnight ethanol dehydration and two xylene washes using a Histokinette (Microm HMP110). 10-µm sections were obtained using a microtome. For *Gpr56* cRNA probes, the mouse coding sequence (942 bp) was primed from BM cDNA with FW 5′-TTGCAGCAGCTTAGCAGGTA-3′ and RV 5′-GATAGCCGGGCACATAGGTA-3′ oligos, and the fragment was ligated to pGEM-T Easy (Promega) and linearized before sense and α-sense probes synthesis: overnight at room temperature with DIG-dUTP mix (Roche) and SP6/T7 polymerases (Roche). Hybridization was performed as described previously (Ciau-Uitz et al., 2000), without incubation in H_2O_2. After developing color for several days, sections were washed five times for 30 min in PBS-0.1% Tween (Tw) and mounted in Kaiser's Glycerol gelatin (Merck). *Ly6aGFP* sections were incubated overnight at room temperature with rabbit polyclonal anti-GFP (1:1,000; Abcam) in PBS^{Block} (1% BSA, 0.05% Tw), washed in PBS^{+} (0.05% Tw), and incubated for 30 min at room temperature with polyclonal biotinylated GantiRIg (1:400; Dako) in PBS^{Block}, washed, incubated for 30 min at room temperature with Streptavidin-HRP (PK-7100; Vector Laboratories), washed, and color developed at room temperature in the dark by 6-min incubation in 4 ml of 5% (wt/vol) diaminobenzidine (Fluka) and 200 ml PBS. 70 µl of 35% H_2O_2 was added to start the reaction. Slides were rinsed with tap water, mounted, and imaged on a BX40F4 microscope and Colorview IIIu camera (Olympus). Sense control probes showed no signal.

Hematopoietic assays. Sorted cells were plated in triplicate in methylcellulose (MethoCult GF; STEMCELL Technologies) with 1% PS and incubated at 37°C, 5% CO_2 for 12 d. Hematopoietic colony types were distinguished by morphology and counted with an inverted microscope.

For all transplantations, 9-Gy irradiated (split dose) C57BL/6 female recipients were used. 2×10^5 C57BL/6 spleen cells were coinjected with the sorted cell samples. Chimerism in hematopoietic tissues was assessed by semi-qPCR for the *GFP* transgene (eGFP FW 5′-AAACGGCCACAAGTTCAGCG-3′ and RV 5′-GGCGGATCTTGAAGTTCACC-3′), normalized to myogenin (Myo FW 5′-TTACGTCCATCGTGGACAGC-3′ and RV 5′-TGGGCTGGGTGTTTAGTCTTA-3′). Control mixes of *Ly6aGFP* and WT DNA were used to make a standard curve, and the trend line formula was used to calculate the percentage of reconstitution of each sample. Peripheral blood cell donor chimerism was assayed at 1 and 4 mo after injection, and mice were sacrificed for analysis of donor chimerism in all hematopoietic tissues. Recipients considered reconstituted are ≥10% donor chimerism positive. All experiments have been conducted according to Dutch law and have been approved by the animal experiments committee (Stichting DEC consult, Dier Experimenten commissie, protocol numbers 138-11-01 and 138-12-13).

RNA isolation. Cells (see Table S1) were directly sorted into PBS/10% FCS/1% PS and centrifuged, and supernatant was removed. Cells were lysed, and RNA was isolated using the *mir*Vana miRNA Isolation kit (Ambion) according to the manufacturer's protocol. RNA quality and quantity were measured by the 2100 Bioanalyzer (Agilent Technologies).

mRNA sequencing analysis. RNA samples (Table S1) were prepared by SMARTer protocol. Illumina TrueSeq v2 protocol was used on HiSeq2000 with single read of 36 bp + 7 bp index. Reads were aligned to the mouse genome (NCBI37/mm9) using Tophat/Bowtie and mapped to the mouse genome (NCBI37/mm9), and the generated count table was analyzed by R/Bioconductor package edgeR according to McCarthy et al. (2012). Counts were normalized for mRNA abundance, and differential expression analysis was performed using edgeR (Fig. 2 A). B-H method was used for p-value correction with an FDR of 0.05 as statistically significant. Variance stabilized counts were calculated by R/Bioconductor package "DESeq" for all the genes (Anders and Huber, 2010). Heat maps were generated from the log-scaled variance stabilized counts of DEGs. GSEA was performed using the preranked option in combination with log fold change values of each comparison calculated by edgeR. Cufflinks was used to compute transcript abundance estimates in FPKMs (Trapnell et al., 2013). For DEGs, the FPKMs for each gene across all samples were normalized by division with maximum FPKM observed for that gene. Patterns were generated based on normalized FPKM, with expression levels lower than 1/3 assigned as low (L), between 1/3 and 2/3 as intermediate (I), and more than 2/3 as high (H). Patterns were then categorized as ECs, HECs, or HSCs. Genes corresponding to EC, HEC or HSC patterns were separately used for GO, KEGG, and Phenotype ontology enrichment analysis using the WebGestalt web application (Wang et al., 2013). GO terms were summarized using the REVIGO tool (Supek et al., 2011).

Cell lines. CHOK3 cells transfected with an expression vector for the mouse SCF gene were grown initially in DMEM (Gibco) until they became confluent. They were then grown in Stem Cell Pro media (Gibco) supplemented with 1% PS, L-glutamine, and 0.5% FBS. HPC-7 mouse hematopoietic progenitor cells were grown in IMDM (Gibco) supplemented with 10% CHOK3 conditioned media, 1% PS, 10% FBS, and 1.5×10^{-4} monothioglycerol. See Knezevic et al. (2011) and Wilson et al. (2010) for details.

ChIP. ChIP assays were performed in HPC-7 cells. 2×10^7 cells per antibody were treated with 0.4% formaldehyde, and cross-linked chromatin was sonicated to fragments of 300–500 bp. Cross-linked, sonicated chromatin was distributed evenly for immunoprecipitation. SYBR Green RT-PCR was performed on a Stratagene Mx3000p and analyzed using the MxPro software. Relative enrichment levels were calculated by normalizing results to the IgG control. For details see Knezevic et al. (2011). CD34 and HPC7 ChIPseq data were downloaded from the BloodChIP database (Beck et al., 2013; Chacon et al., 2014) and from Wilson et al. (2010).

Luciferase and LacZ assays. Transfection was performed by electroporation of $5–10 \times 10^6$ cells with 10 µg vector DNA using a GenePulser Excell (Bio-Rad Laboratories). The luciferase assay was performed using a modified version of the Dual-Luciferase Reporter Assay System (Promega). Transient transfections were cotransfected with the pEFBOS-LacZ vector, and luciferase data were normalized to the lacZ data. For stable transfection assays, cells were cotransfected with pGK Neo and resistant cells were used for luciferase assays as described in Knezevic et al. (2011).

Transactivation. Cos7 cells were cultured in 6-well plates (5×10^5 cells/well) overnight. *Runx1.pcDNA3*, *CBFb.pcDNA3*, *Fli1.pcDNA3*, and *Gata2.pMSCV-PIG* or empty vectors were transfected along with the pEFBOS-*lacZ* control vector and the *Gpr56-37*.pGL2prom enhancer construct in varying combinations (0.5 µg DNA/well) using the ProFection Mammalian Transfection system (Promega). After 48 h, luciferase and *lacZ* assays were performed as detailed in Knezevic et al. (2011).

Zebrafish. Zebrafish (*Danio rerio*) embryos were raised at 28.5°C (Westerfield, 1995). Heterozygous *-6.0itga2b:EGFP* embryos (CD41-GFP; Lin et al., 2005) were maintained by crosses with WT zebrafish. For ISH 0.003% 1-phenyl-2-thiourea (PTU)–treated embryos (30 hpf) were fixed over night with 4% PFA in PBS containing 3% sucrose, transferred to MeOH, hybridized with a *cmyb* probe (gift of R. Patient, University of Oxford, Oxford, England, UK) according to Chocron et al. (2007), and imaged with BX40 (Olympus) and AX10 (Carl Zeiss) fluorescent microscopes with an AxioCam MRm camera (Carl Zeiss).

Human GPR56 expression constructs and differentiation of 32D/G-CSF-R cells. WT *hGPR56* and truncated constitutively active *hGPR56* (as described by Paavola et al., 2011) were amplified, and a Kozak sequence and two N-terminal flag tags were added. The WT and constitutively active *hGPR56* (*wtGPR56/caGPR56*) were subcloned into *pEGFPN1* to generate an eGFP fusion protein. The *wt/caGPR56-eGFP* fusion inserts were PCR amplified and cloned into the *pLNCX2* retroviral vector (Takara Bio Inc.).

The 32D/G-CSF receptor (32D/G-CSF-R) cell line was cultured at a density $<10^6$ cells/ml in RPMI 1640 medium (Life Technologies), supplemented with 1% PS, 10% FCS, and 10 ng/ml mouse IL-3 (CHO conditioned). 10^6 cells were transduced (RetroNectin; Takara Bio Inc.) with *pLNCX2-EGFP*, *pLNCX2-wtGPR56-EGFP*, or *pLNCX2-caGPR56-EGFP* and selected in 0.8 mg/ml G418 (Life Technologies) 2 d after infection (efficiency 100% at day 4). For the differentiation assay, cells were washed twice in HBSS (Life Technologies), cultured in the same medium, except mIL3 was replaced with 10 ng/ml human CSF3 (Amgen), and placed in 6-well plates (2×10^5 cells/ml). Cell density was adjusted to 2×10^5 cells/ml on a daily basis. Morphology was determined by microscopy on May-Grünwald-Giemsa–stained cytospins at day 7 and scored based on phenotype (blast, intermediate, neutrophil, >100 cells).

Primers. For cloning: Gpr56-37F, 5′-GAGGATCCTCCATGAGGGACATCTTCAA-3′; Gpr56-37R, 5′-AGTCGACACGGGCTTATCACGAGAAAT-3′. For ChIP-qPCR: Gpr56-37F, 5′-AATGTTATCAACCGTCTGC-3′; Gpr56-37R, 5′-CCTCACCTAATCAAGATATGTC-3′. For RT-qPCR validations: Gpr56 FW, 5′-GCAGAACACCAAAGTCACCA-3′; Gpr56 RV, 5′-TGTCTCTGCTCACTGTCTCG-3′; Gpr97 FW, 5′-CTGGGATATGGCTAAAGGAGAC-3′; Gpr97 RV, 5′-AAGGCGAAGAAGGTCAAGTG-3′; Gpr114 FW, 5′-TCACTGCTCAATAACTATGTCC-3′; Gpr114 RV, 5′-ACTGTATACCCTTCCAGACTC-3′; Ikzf2 FW, 5′-AGCCCTTCAAATGTCCTTTCTG-3′; Ikzf2 RV, 5′-CAGCGTTCCTTGTGTTCCTC-3′; Meis1 FW, 5′-CATCTTTCCCAAAGTAGCCAC-3′; Meis1 RV, 5′-GTAAGTCCTGTATCTTGTGCC-3′; Mpl FW, 5′-TTGGACTTCAGTGCTTTACCT-3′; Mpl RV, 5′-CTCCTCTTCACATTTCTCCCA-3′; Mycn FW, 5′-GGAGAGGATACCTTGAGCGA-3′; Mycn RV, 5′-GGTTACCGCCTTGTTGTTAGAG-3′; Gata2 FW, 5′-CACCCCTAAGCAGAGAAGCAA-3′; Gata2 RV, 5′-TGGCACCACAGTTGACACACT-3′; Gata3 FW, 5′-TGTGGGCTGTACTACAAGCT-3′; Gata3 RV, 5′-TCGATTTGCTAGACATCTTCCG-3′; Runx1 FW, 5′-CAGGTAGCGAGATTCAACGA-3′; Runx1 RV, 5′-TTTGATGGCTCTATGGTAGGTG-3′; Hlf FW, 5′-CGCAAAGTCTTCATTCCCGA-3′; Hlf RV, 5′-GCTCCTTCCTTAAATCAGCCA-3′; Gapdh FW, 5′-GACTTCAACAGCAACTCCCA-3′; Gapdh RV, 5′-GCCGTATTCATTGTCATACCAG-3′.

Online supplemental material. Fig. S1 provides FACS characterization of $CD31^+$ AGM cells and a representative multilineage repopulation analysis of transplanted mice. Further details of RNAseq analyses and a comparison with the published single-cell qPCR analysis of EHT cells (Swiers et al. 2013) and adult BM HSCs (Riddell et al. 2014) is provided in Fig. S2. Table S1 contains details of the material used for RNAseq analysis. Table S2 contains the 530 DEGs and their expression patterns and groupings. Ontology enrichment analysis results for each group are provided in Table S3. Table S4 contains a list of up- and down-regulated TFs as found by differential expression analysis with edgeR. All sequencing data has been uploaded to GEO repository under accession number GSE63316. Online supplemental material is available at http://www.jem.org/cgi/content/full/jem.20140767/DC1.

We thank all laboratory members and N. Speck for stimulating discussions, P. Kaimakis for technical assistance, R. Patient for (*cmyb* and *grl*) zebrafish ISH probes, and the Experimental Animal Center of Erasmus MC.

This study was supported by a National Institutes of Health grant (R37 DK054077), ZonMw (Netherlands Scientific Research Organization) equipment grant (91109036), ZonMw TOP Award (40-00812-98-11068), NIRM (Netherlands Institute for Regenerative Medicine) FES award, Erasmus MC PhD grant, and Landsteiner Society for Blood Research grant (0407).

The authors declare no competing financial interests.

Submitted: 23 April 2014
Accepted: 5 December 2014

REFERENCES

Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. http://dx.doi.org/10.1186/gb-2010-11-10-r106

Beck, D., J.A. Thoms, D. Perera, J. Schütte, A. Unnikrishnan, K. Knezevic, S.J. Kinston, N.K. Wilson, T.A. O'Brien, B. Göttgens, et al. 2013. Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood.* 122:e12–e22. http://dx.doi.org/10.1182/blood-2013-03-490425

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

Bhargava, V., S.R. Head, P. Ordoukhanian, M. Mercola, and S. Subramaniam. 2014. Technical variations in low-input RNA-seq methodologies. *Sci. Rep.* 4:3678. http://dx.doi.org/10.1038/srep03678

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

Chacon, D., D. Beck, D. Perera, J.W. Wong, and J.E. Pimanda. 2014. BloodChIP: a database of comparative genome-wide transcription factor binding profiles in human blood cells. *Nucleic Acids Res*. 42:D172–D177. http://dx.doi.org/10.1093/nar/gkt1036

Chen, M.J., T. Yokomizo, B.M. Zeigler, E. Dzierzak, and N.A. Speck. 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature*. 457:887–891. http://dx.doi.org/10.1038/nature07619

Chocron, S., M.C. Verhoeven, F. Rentzsch, M. Hammerschmidt, and J. Bakkers. 2007. Zebrafish Bmp4 regulates left-right asymmetry at two distinct developmental time points. *Dev. Biol*. 305:577–588. http://dx.doi.org/10.1016/j.ydbio.2007.03.001

Ciau-Uitz, A., M. Walmsley, and R. Patient. 2000. Distinct origins of adult and embryonic blood in *Xenopus*. *Cell*. 102:787–796. http://dx.doi.org/10.1016/S0092-8674(00)00067-2

de Bruijn, M.F., N.A. Speck, M.C. Peeters, and E. Dzierzak. 2000. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J*. 19:2465–2474. http://dx.doi.org/10.1093/emboj/19.11.2465

de Bruijn, M.F.T.R., X. Ma, C. Robin, K. Ottersbach, M.J. Sanchez, and E. Dzierzak. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity*. 16:673–683. http://dx.doi.org/10.1016/S1074-7613(02)00313-8

Garcia-Porrero, J.A., I.E. Godin, and F. Dieterlen-Lièvre. 1995. Potential intraembryonic hemogenic sites at pre-liver stages in the mouse. *Anat. Embryol. (Berl.)*. 192:425–435. http://dx.doi.org/10.1007/BF00240375

Guidolin, D., G. Albertin, R. Spinazzi, E. Sorato, A. Mascarin, D. Cavallo, M. Antonello, and D. Ribatti. 2008. Adrenomedullin stimulates angiogenic response in cultured human vascular endothelial cells: involvement of the vascular endothelial growth factor receptor 2. *Peptides*. 29:2013–2023. http://dx.doi.org/10.1016/j.peptides.2008.07.009

Iguchi, T., K. Sakata, K. Yoshizaki, K. Tago, N. Mizuno, and H. Itoh. 2008. Orphan G protein-coupled receptor GPR56 regulates neural progenitor cell migration via a Gα12/13 and Rho pathway. *J. Biol. Chem*. 283:14469–14478. http://dx.doi.org/10.1074/jbc.M708919200

Imanirad, P., P. Solaimani Kartalaei, M. Crisan, C. Vink, T. Yamada-Inagawa, E. de Pater, D. Kurek, P. Kaimakis, R. van der Linden, N. Speck, and E. Dzierzak. 2014. HIF1α is a regulator of hematopoietic progenitor and stem cell development in hypoxic sites of the mouse embryo. *Stem Cell Res. (Amst.)*. 12:24–35. http://dx.doi.org/10.1016/j.scr.2013.09.006

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

Jing, L., and L.I. Zon. 2011. Zebrafish as a model for normal and malignant hematopoiesis. *Dis. Model. Mech*. 4:433–438. http://dx.doi.org/10.1242/dmm.006791

Kim, J.E., J.M. Han, C.R. Park, K.J. Shin, C. Ahn, J.Y. Seong, and J.I. Hwang. 2010. Splicing variants of the orphan G-protein-coupled receptor GPR56 regulate the activity of transcription factors associated with tumorigenesis. *J. Cancer Res. Clin. Oncol*. 136:47–53. http://dx.doi.org/10.1007/s00432-009-0635-z

Kim, J., W.J. Oh, N. Gaiano, Y. Yoshida, and C. Gu. 2011. Semaphorin 3E-Plexin-D1 signaling regulates VEGF function in developmental angiogenesis via a feedback mechanism. *Genes Dev.* 25:1399–1411. http://dx.doi.org/10.1101/gad.2042011

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

Knezevic, K., T. Bee, N.K. Wilson, M.E. Janes, S. Kinston, S. Polderdijk, A. Kolb-Kokocinski, K. Ottersbach, N. Pencovich, Y. Groner, et al. 2011. A Runx1-Smad6 rheostat controls Runx1 activity during embryonic hematopoiesis. *Mol. Cell. Biol*. 31:2817–2826. http://dx.doi.org/10.1128/MCB.01305-10

Kumano, K., S. Chiba, A. Kunisato, M. Sata, T. Saito, E. Nakagami-Yamaguchi, T. Yamaguchi, S. Masuda, K. Shimizu, T. Takahashi, et al. 2003. Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. *Immunity*. 18:699–711. http://dx.doi.org/10.1016/S1074-7613(03)00117-1

Lin, H.F., D. Traver, H. Zhu, K. Dooley, B.H. Paw, L.I. Zon, and R.I. Handin. 2005. Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood*. 106:3803–3810. http://dx.doi.org/10.1182/blood-2005-01-0179

Ling, K.W., K. Ottersbach, J.P. van Hamburg, A. Oziemlak, F.Y. Tsai, S.H. Orkin, R. Ploemacher, R.W. Hendriks, and E. Dzierzak. 2004. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J. Exp. Med*. 200:871–882. http://dx.doi.org/10.1084/jem.20031556

Ma, X., C. Robin, K. Ottersbach, and E. Dzierzak. 2002. The Ly-6A (Sca-1) GFP transgene is expressed in all adult mouse hematopoietic stem cells. *Stem Cells*. 20:514–521. http://dx.doi.org/10.1634/stemcells.20-6-514

Mascarenhas, M.I., A. Parker, E. Dzierzak, and K. Ottersbach. 2009. Identification of novel regulators of hematopoietic stem cell development through refinement of stem cell localization and expression profiling. *Blood*. 114:4645–4653. http://dx.doi.org/10.1182/blood-2009-06-230037

Masiero, M., F.C. Simões, H.D. Han, C. Snell, T. Peterkin, E. Bridges, L.S. Mangala, S.Y.Y. Wu, S. Pradeep, D. Li, et al. 2013. A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell*. 24:229–241. http://dx.doi.org/10.1016/j.ccr.2013.06.004

McCarthy, D.J., Y. Chen, and G.K. Smyth. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40:4288–4297. http://dx.doi.org/10.1093/nar/gks042

McKinney-Freeman, S., P. Cahan, H. Li, S.A. Lacadie, H.T. Huang, M. Curran, S. Loewer, O. Naveiras, K.L. Kathrein, M. Konantz, et al. 2012. The transcriptional landscape of hematopoietic stem cell ontogeny. *Cell Stem Cell*. 11:701–714. http://dx.doi.org/10.1016/j.stem.2012.07.018

Medvinsky, A., and E. Dzierzak. 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*. 86:897–906. http://dx.doi.org/10.1016/S0092-8674(00)80165-8

Müller, A.M., A. Medvinsky, J. Strouboulis, F. Grosveld, and E. Dzierzak. 1994. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity*. 1:291–301. http://dx.doi.org/10.1016/1074-7613(94)90081-7

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

North, T.E., M.F. de Bruijn, T. Stacy, L. Talebian, E. Lind, C. Robin, M. Binder, E. Dzierzak, and N.A. Speck. 2002. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity*. 16:661–672. http://dx.doi.org/10.1016/S1074-7613(02)00296-0

Ody, C., P. Vaigot, P. Quéré, B.A. Imhof, and C. Corbel. 1999. Glycoprotein IIb-IIIa is expressed on avian multilineage hematopoietic progenitor cells. *Blood*. 93:2898–2906.

Orelio, C., and E. Dzierzak. 2003. Identification of 2 novel genes developmentally regulated in the mouse aorta-gonad-mesonephros region. *Blood*. 101:2246–2249. http://dx.doi.org/10.1182/blood-2002-07-2260

Orelio, C., M. Peeters, E. Haak, K. van der Horn, and E. Dzierzak. 2009. Interleukin-1 regulates hematopoietic progenitor and stem cells in the midgestation mouse fetal liver. *Haematologica*. 94:462–469. http://dx.doi.org/10.3324/haematol.13728

Paavola, K.J., J.R. Stephenson, S.L. Ritter, S.P. Alter, and R.A. Hall. 2011. The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity. *J. Biol. Chem*. 286:28914–28921. http://dx.doi.org/10.1074/jbc.M111.247973

Rajendran, P., T. Rengarajan, J. Thangavel, Y. Nishigaki, D. Sakthisekaran, G. Sethi, and I. Nishigaki. 2013. The vascular endothelium and human diseases. *Int. J. Biol. Sci*. 9:1057–1069. http://dx.doi.org/10.7150/ijbs.7502

Riddell, J., R. Gazit, B.S. Garrison, G. Guo, A. Saadatpour, P.K. Mandal, W. Ebina, P. Volchkov, G.C. Yuan, S.H. Orkin, and D.J. Rossi. 2014. Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell*. 157:549–564. http://dx.doi.org/10.1016/j.cell.2014.04.006

Robin, C., and E. Dzierzak. 2010. Preparation of hematopoietic stem and progenitor cells from the human placenta. *Curr. Protoc. Stem Cell Biol.* Chapter 2:Unit 2A.9. http://dx.doi.org/10.1002/9780470151808.sc02a09s14

Robin, C., K. Ottersbach, J.C. Boisset, A. Oziemlak, and E. Dzierzak. 2011. CD41 is developmentally regulated and differentially expressed on mouse hematopoietic stem cells. *Blood*. 117:5088–5091. http://dx.doi.org/10.1182/blood-2011-01-329516

Rossi, D.J., D. Bryder, J.M. Zahn, H. Ahlenius, R. Sonu, A.J. Wagers, and I.L. Weissman. 2005. Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc. Natl. Acad. Sci. USA*. 102:9194–9199. http://dx.doi.org/10.1073/pnas.0503280102

Saito, Y., K. Kaneda, A. Suekane, E. Ichihara, S. Nakahata, N. Yamakawa, K. Nagai, N. Mizuno, K. Kogawa, I. Miura, et al. 2013. Maintenance of the hematopoietic stem cell pool in bone marrow niches by EVI1-regulated GPR56. *Leukemia*. 27:1637–1649. http://dx.doi.org/10.1038/leu.2013.75

Sánchez, M.J., A. Holmes, C. Miles, and E. Dzierzak. 1996. Characterization of the first definitive hematopoietic stem cells in the AGM and liver of the mouse embryo. *Immunity*. 5:513–525. http://dx.doi.org/10.1016/S1074-7613(00)80267-8

Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*. 6:e21800. http://dx.doi.org/10.1371/journal.pone.0021800

Swiers, G., C. Baumann, J. O'Rourke, E. Giannoulatou, S. Taylor, A. Joshi, V. Moignard, C. Pina, T. Bee, K.D. Kokkaliaris, et al. 2013. Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat. Commun.* 4:2924. http://dx.doi.org/10.1038/ncomms3924

Taoudi, S., A.M. Morrison, H. Inoue, R. Gribi, J. Ure, and A. Medvinsky. 2005. Progressive divergence of definitive haematopoietic stem cells from the endothelial compartment does not depend on contact with the foetal liver. *Development*. 132:4179–4191. http://dx.doi.org/10.1242/dev.01974

Trapnell, C., D.G. Hendrickson, M. Sauvageau, L. Goff, J.L. Rinn, and L. Pachter. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31:46–53. http://dx.doi.org/10.1038/nbt.2450

Van Handel, B., A. Montel-Hagen, R. Sasidharan, H. Nakano, R. Ferrari, C.J. Boogerd, J. Schredelseker, Y. Wang, S. Hunter, T. Org, et al. 2012. Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell*. 150:590–605. http://dx.doi.org/10.1016/j.cell.2012.06.026

Wang, J., D. Duncan, Z. Shi, and B. Zhang. 2013. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* 41:W77–W83.

Westerfield, M. 1995. The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (*Danio rerio*). Third edition. University of Oregon Press, Eugene, OR. 385 pp.

Wilson, N.K., S.D. Foster, X. Wang, K. Knezevic, J. Schütte, P. Kaimakis, P.M. Chilarska, S. Kinston, W.H. Ouwehand, E. Dzierzak, et al. 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell*. 7:532–544. http://dx.doi.org/10.1016/j.stem.2010.07.016

Yokomizo, T., and E. Dzierzak. 2010. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development*. 137:3651–3661. http://dx.doi.org/10.1242/dev.051094

Yokomizo, T., T. Yamada-Inagawa, A.D. Yzaguirre, M.J. Chen, N.A. Speck, and E. Dzierzak. 2012. Whole-mount three-dimensional imaging of internally localized immunostained cells within mouse embryos. *Nat. Protoc.* 7:421–431. http://dx.doi.org/10.1038/nprot.2011.441

Yui, J., C.P. Chiu, and P.M. Lansdorp. 1998. Telomerase activity in candidate stem cells from fetal liver and adult bone marrow. *Blood*. 91:3255–3262.

Zhong, T.P., M. Rosenberg, M.A.P.K. Mohideen, B. Weinstein, and M.C. Fishman. 2000. *gridlock*, an HLH gene required for assembly of the aorta in zebrafish. *Science*. 287:1820–1824. http://dx.doi.org/10.1126/science.287.5459.1820

Zovein, A.C., J.J. Hofmann, M. Lynch, W.J. French, K.A. Turlo, Y. Yang, M.S. Becker, L. Zanetta, E. Dejana, J.C. Gasson, et al. 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell*. 3:625–636. http://dx.doi.org/10.1016/j.stem.2008.09.018

SUPPLEMENTAL MATERIAL

Solaimani Kartalaei et al., http://www.jem.org/cgi/content/full/jem.20140767/DC1

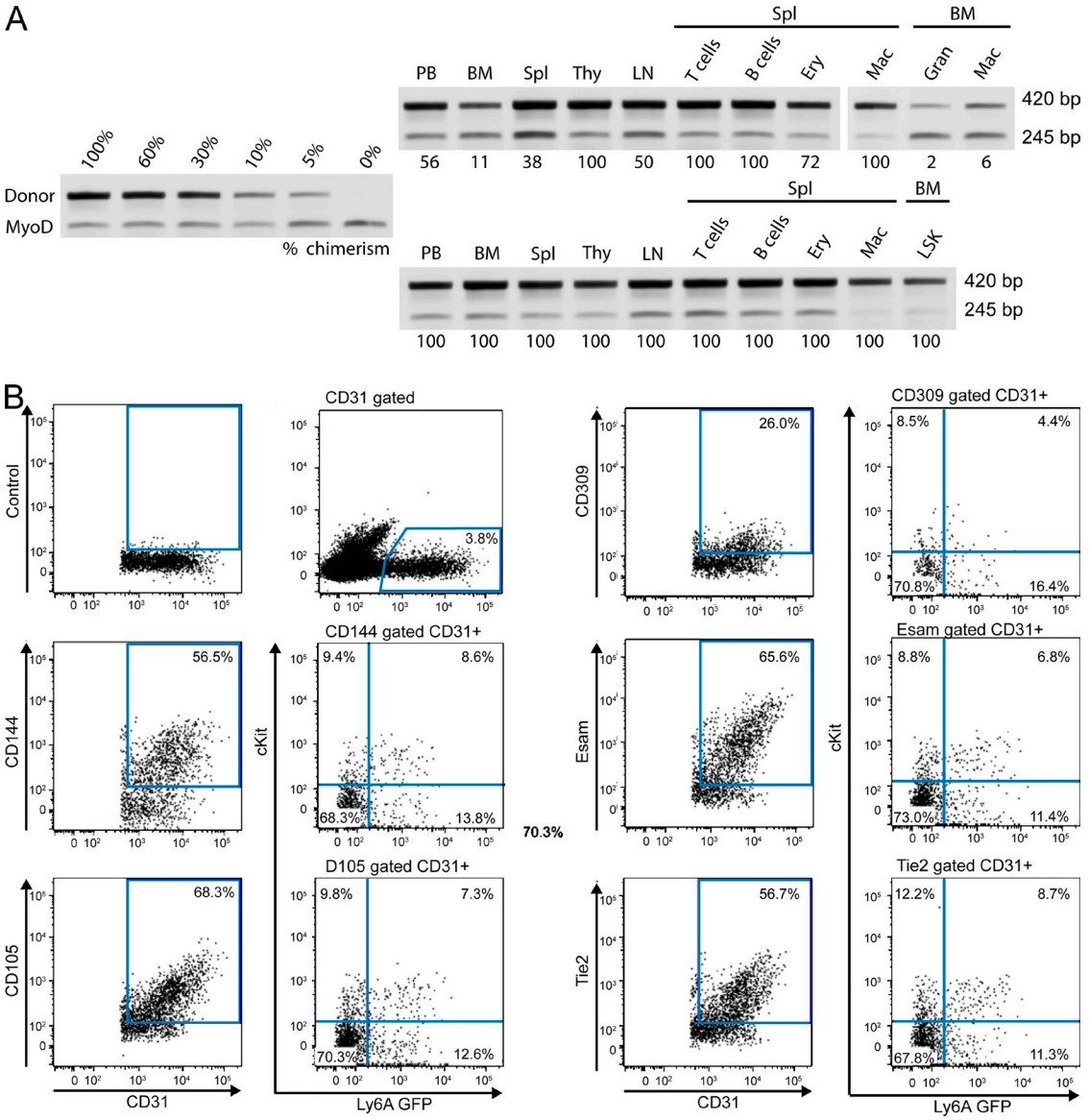


Figure S1. Analysis of Ly6aGFP⁺ AGM cell frequency and hematopoietic potential. (A) Multilineage analysis of two of the adult recipient mice transplanted with the CD31⁺cKit⁺GFP⁺ HSC fraction of E10.5 AGM cells. At >4 mo after injection, hematopoietic tissues were harvested and cells were sorted by lineage. DNA was prepared from PB (peripheral blood), BM, Spl (spleen), Thy (thymus), and LNs and T, B, Ery (erythroid), Mac (macrophage), and Gran (granulocyte) cells. PCR for the donor *Gfp* marker gene and control *MyoD* gene (DNA normalization) was performed. Percentage of engraftment is indicated under each lane, as determined after *MyoD* normalization by fragment intensity comparison with percentage of chimerism standards. Transplantation of BM from primary recipients showed engraftment of 2/3 and 3/3 secondary recipients (not depicted; related to Fig. 1). (B) Frequency of Ly-6aGFP⁺ cells in endothelial marker–expressing AGM cells as defined by FACS. AGM cells were first gated for CD31 and then for CD144, CD105, CD309, Esam, or Tie2, followed by Ly6aGFP and cKit expression.

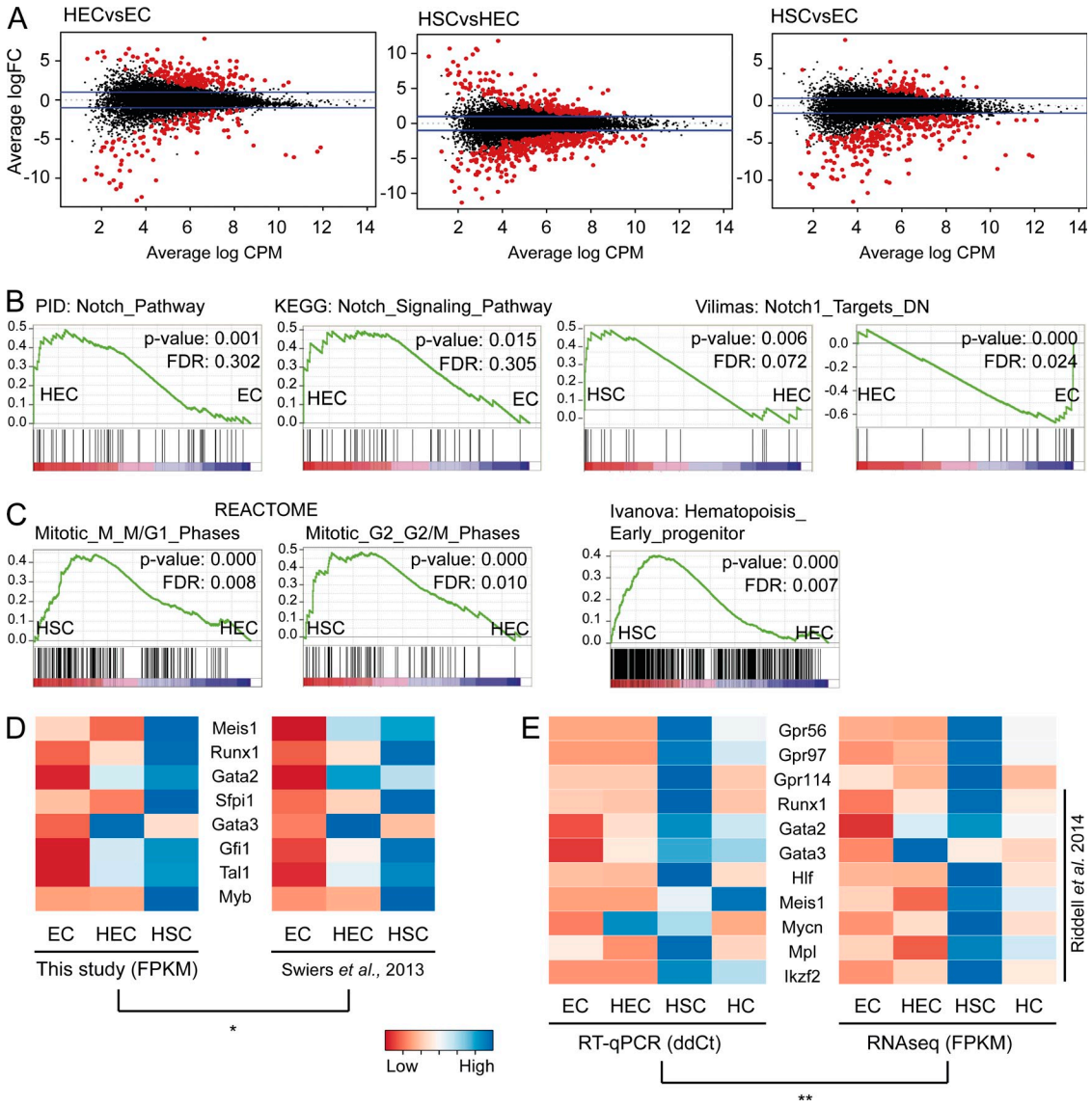


Figure S2. RNAseq analysis and validation. (A) MA plots of all three comparisons after normalization with *edgeR*. Each dot represents a gene, with DEGs in red. (B) GSEA plots for Notch signaling showing higher Notch activity in HECs as compared with ECs or HSCs. (C) GSEA plots for cell cycle genes in HECs as compared with ECs and HSCs and early hematopoietic progenitor genes in HSCs versus HECs. (D) Pearson's correlation of gene-scaled mean FPKMs found by this study and gene-scaled mRNA levels detected by high-throughput qPCR analysis in Swiers et al. (2013) for the three populations EC, HEC, and HP/SC (related to Fig. 2; *, $\rho = 0.83$; p-value = 4.97e-7) . (E) Pearson's correlation of gene-scaled RT-qPCR on four sorted populations ($n = 2$) and mean gene-scaled FPKM values for genes highly expressed in adult BM HSCs as found by Riddell et al. (2014) (**, $\rho = 0.791$; p-value = 1.60e-10).

Table S1. Details of the samples used for RNAseq (related to Fig. 1)

Exp	HCs	HSCs	ECs	HECs
Number of sorted cells per fraction				
1	2,486	774	7,140	593
2	1,150	993	9,591	873
3	3,713	1,776	35,967	2,985
Amount of total RNA (ng)				
1	27.8	18.2	52.0	21.6
2	6.2	36.0	86.8	34.2
3	24.3	8.8	99.4	6.1
Sequencing depth (number of mapped reads $\times 10^6$)				
1	9.2	8.6	8.7	7.8
2	19.1	22.4	21.1	6.5
3	56.7	48.4	23.2	20.4

Table S2, included as a separate Excel file, contains the 530 DEGs and their expression patterns and groupings.

Table S3, included as a separate Excel file, contains ontology enrichment analysis results for each group.

Table S4, included as a separate Excel file, contains a list of up- and down-regulated TFs as found by differential expression analysis with edgeR.

REFERENCES

Riddell, J., R. Gazit, B.S. Garrison, G. Guo, A. Saadatpour, P.K. Mandal, W. Ebina, P. Volchkov, G.C. Yuan, S.H. Orkin, and D.J. Rossi. 2014. Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell*. 157:549–564. http://dx.doi.org/10.1016/j.cell.2014.04.006

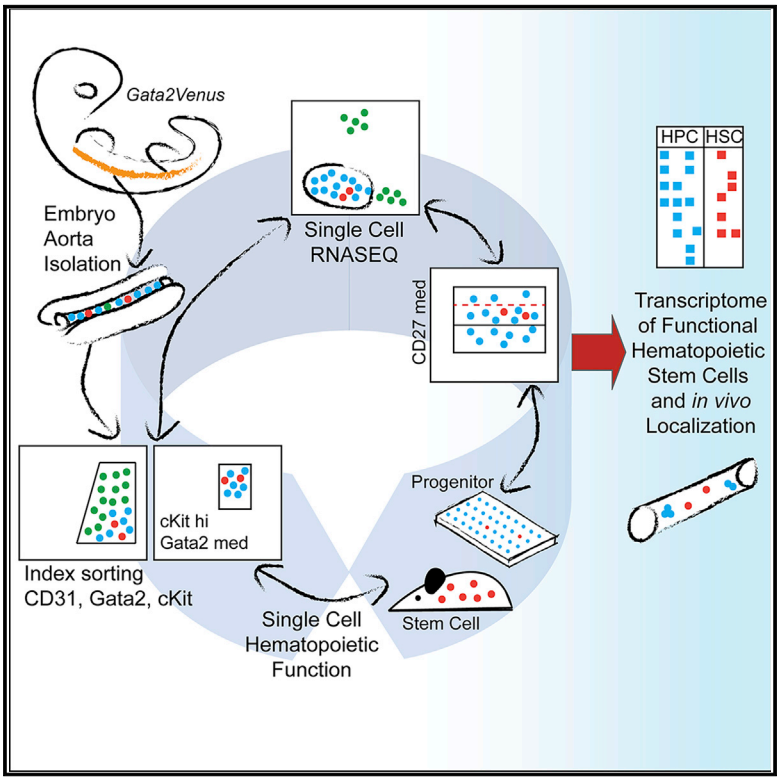
Swiers, G., C. Baumann, J. O'Rourke, E. Giannoulatou, S. Taylor, A. Joshi, V. Moignard, C. Pina, T. Bee, K.D. Kokkaliaris, et al. 2013. Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat. Commun.* 4:2924. http://dx.doi.org/10.1038/ncomms3924

Cell Reports

Article

Iterative Single-Cell Analyses Define the Transcriptome of the First Functional Hematopoietic Stem Cells

Graphical Abstract



Highlights

- Single-cell iterations capture the transcriptome of the first functional HSCs in mouse
- Gata2 connects the "inner" HSC regulatory network with single-cell function
- Specific Gata2, cKit, and CD27 levels define all HSCs in embryonic aortic clusters
- HSCs emerge from endothelium as single cells within aortic clusters of 1–2 cells

Authors

Chris Sebastiaan Vink, Fernando Jose Calero-Nieto, Xiaonan Wang, ..., Wajid Jawaid, Berthold Göttgens, Elaine Dzierzak

Correspondence

elaine.dzierzak@ed.ac.uk

In Brief

Vink and colleagues capture the transcriptome of the first functional HSCs in mouse by single-cell RNA-seq, index-sorting, and *in vivo* and *in vitro* hematopoietic analyses. The HSC transcriptome is unique compared to HPCs, and heterogeneous expression of pivotal genes suggests that establishment of functional hematopoietic fate during cell emergence from embryonic aortic endothelium is stochastic.

Vink et al., 2020, Cell Reports *31*, 107627
May 12, 2020
https://doi.org/10.1016/j.celrep.2020.107627

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Iterative Single-Cell Analyses Define the Transcriptome of the First Functional Hematopoietic Stem Cells

Chris Sebastiaan Vink,[1] Fernando Jose Calero-Nieto,[2] Xiaonan Wang,[2] Antonio Maglitto,[1] Samanta Antonella Mariani,[1] Wajid Jawaid,[2] Berthold Göttgens,[2] and Elaine Dzierzak[1,3,*]

[1]Centre for Inflammation Research, Queens Medical Research Institute, University of Edinburgh, Edinburgh, Midlothian, Scotland EH16 4TJ, UK

[2]Department of Haematology, Wellcome & MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, Cambridgeshire, England CB2 0AW, UK

[3]Lead Contact

*Correspondence: elaine.dzierzak@ed.ac.uk

https://doi.org/10.1016/j.celrep.2020.107627

SUMMARY

Whereas hundreds of cells in the mouse embryonic aorta transdifferentiate to hematopoietic cells, only very few establish hematopoietic stem cell (HSC) identity at a single time point. The Gata2 transcription factor is essential for HSC generation and function. In contrast to surface-marker-based cell isolation, Gata2-based enrichment provides a direct link to the internal HSC regulatory network. Here, we use iterations of index-sorting of Gata2-expressing intra-aortic hematopoietic cluster (IAHC) cells, single-cell transcriptomics, and functional analyses to connect HSC identity to specific gene expression. Gata2-expressing IAHC cells separate into 5 major transcriptomic clusters. Iterative analyses reveal refined CD31, cKit, and CD27 phenotypic parameters that associate specific molecular profiles in one cluster with distinct HSC and multipotent progenitor function. Thus, by iterations of single-cell approaches, we identify the transcriptome of the first functional HSCs as they emerge in the mouse embryo and localize them to aortic clusters containing 1–2 cells.

INTRODUCTION

Hematopoietic stem cells (HSCs) arise through a developmental fate process known as endothelial-to-hematopoietic transition (EHT) (Dzierzak and Bigas, 2018; Dzierzak and Speck, 2008; Jaffredo et al., 2005). EHT is best characterized by the emergence of intra-aortic hematopoietic cluster (IAHC) cells within the aorta-gonad-mesonephros (AGM) region that transdifferentiate from specialized hemogenic endothelial cells (de Bruijn et al., 2002; North et al., 2002; Yokomizo and Dzierzak, 2010; Zovein et al., 2008). Between mouse embryonic day (E)10.5 and E12.5, IAHCs contain HSCs, hematopoietic progenitor cells (HPCs), and other hematopoietic cells (HCs). Time-lapse imaging gave the first visual *in vivo* proof of EHT in the mouse and zebrafish embryonic aorta (Bertrand et al., 2010; Boisset et al., 2010; Kissa and Herbomel, 2010). This transition involves changes from an endothelial transcriptional program to a program promoting HC identity, morphology, and function (Swiers et al., 2013). Although the precise program directing the generation of HSCs during EHT is as yet unknown, important molecular and physiological aspects of HSC generation are conserved between zebrafish, chick, *Xenopus*, mouse, and human (Ciau-Uitz and Patient, 2019; Dzierzak and Bigas, 2018; Ivanovs et al., 2017).

Quantitation of IAHC cells in whole mouse embryos (Yokomizo and Dzierzak, 2010) shows a peak of ~700 phenotypic (cKit-positive) HCs at E10.5. Multicolor lineage tracing shows that clusters are polyclonal, with about 545 Vav^+ nascent hematopoietic progenitor/stem cells (HP/SCs) emerging between E11.5 and 14.5 to establish lifelong hematopoiesis (Ganuza et al., 2017b). However, in the E10.5 aorta at the time of embryo isolation, very few of these cells function as bona fide HSCs (Kumaravelu et al., 2002; Medvinsky and Dzierzak, 1996; Müller et al., 1994), as determined by direct *in vivo* transplantation and long-term hematopoietic reconstitution of adult recipients (a standard test for HSC function and clinical relevance). Recent *in vivo* marking methods question the precise role of phenotypic HSCs defined by *in vivo* repopulating activity and, instead, suggest that multipotent progenitors are responsible for adult steady-state hematopoiesis (Busch et al., 2015; Pei et al., 2017; Rodriguez-Fraticelli et al., 2018; Schoedel et al., 2016; Sun et al., 2014). Out of the hundreds of IAHC cells in E10.5 embryos, colony-forming unit-culture (CFU-C) studies show that about half (350) of the cluster cells are HPCs (de Pater et al., 2013). Thus, the very low frequency of HSCs in the IAHCs highlights the complexity in programming HSC identity as defined by repopulating function and raises the questions: what processes influence acquisition of this rare HSC identity rather than the more abundant HPC or

HC identities, and can we capture the transcriptome of the first HSCs?

Many studies, including our own, have set out to describe the genetic program of HC transdifferentiation from that of embryonic aortic endothelium. Published transcriptome databases are available for phenotypically (surface marker) enriched endothelial cells, hemogenic endothelial cells, transitioning cells, IAHCs, and HP/SCs, at both the population and single-cell levels (Baron et al., 2018; Li et al., 2014, 2017; Lichtinger et al., 2012; McKinney-Freeman et al., 2012; Moignard et al., 2015; Solaimani Kartalaei et al., 2015; Swiers et al., 2013; Zhou et al., 2016). However, no unique transcriptional profile has yet been ascribed to AGM HSCs, since the HSC transcriptome is represented in the single-cell datasets only at a very low frequency compared with the high frequency of HPCs and HCs. Interestingly, all these datasets show the upregulated expression of several hematopoietic (''heptad'') transcription factors (TFs) (Wilson et al., 2010) during EHT, including *Gata2*, which is pivotal to HSC and HPC generation.

Gata2 is required for generation of IAHCs and functional HSCs (de Pater et al., 2013; Ling et al., 2004; Tsai et al., 1994; Tsai and Orkin, 1997). Haploinsufficiency perturbs EHT and the timing and quantitative generation of HSCs (de Pater et al., 2013; Ling et al., 2004), and its overexpression blocks HSC function (Guiu et al., 2013; Tipping et al., 2009). Our recent demonstration of pulsatile Gata2 expression level changes in aortic endothelial cells undergoing EHT (Eich et al., 2018) reveals a previously unexplored dynamic regulatory aspect in hematopoietic fate acquisition. Studies in other systems have shown an unstable program of gene expression in cells as they take on specific lineage fate. These include regulators nuclear factor κB (NF-κB), Hes, and Nanog, which show pulsatile expression behavior in single cells (Abranches et al., 2014; Chang et al., 2008; Cohen-Saidon et al., 2009; Huang, 2009; Kueh et al., 2016; Lahav, 2004; Miyanari and Torres-Padilla, 2012; Purvis and Lahav, 2013; Ryu et al., 2016).

To date, transcriptomic analyses of HSCs have relied on cell-surface-marker-mediated enrichments. As TFs are responsible for promoting the expression of these markers, hematopoietic identities may be better understood through direct enrichment based on TF expression. Our previously described *Gata2Venus* (*G2V*) mouse model reports real-time transcription and protein expression of Gata2 *in vivo* without affecting hematopoietic development or function (Kaimakis et al., 2016). To understand whether the transcriptome of single IAHC cells (Baron et al., 2018; Zhou et al., 2016) correlates with hematopoietic function, and whether HSC identity is based on specific combinations of genes and expression levels, *G2V* embryos were used to examine IAHC cells by iterative single-cell transcriptomics and functional analyses. Here, we show that functional heterogeneity in highly enriched single index-sorted IAHC cells expressing medium levels of Gata2 ($G2^{med}$) corresponds with heterogeneity in TF expression levels. Following this, we reveal that specific expression levels of Gata2, cKit, and CD27 define all *in vivo* repopulating HSCs and that HSCs localize to IAHCs containing 1–2 cells. This iterative methodology captures the transcriptome of the first functional HSCs and closely related HPCs and suggests a stochastic process of quantitative molecular states leading to the establishment of HSC identity.

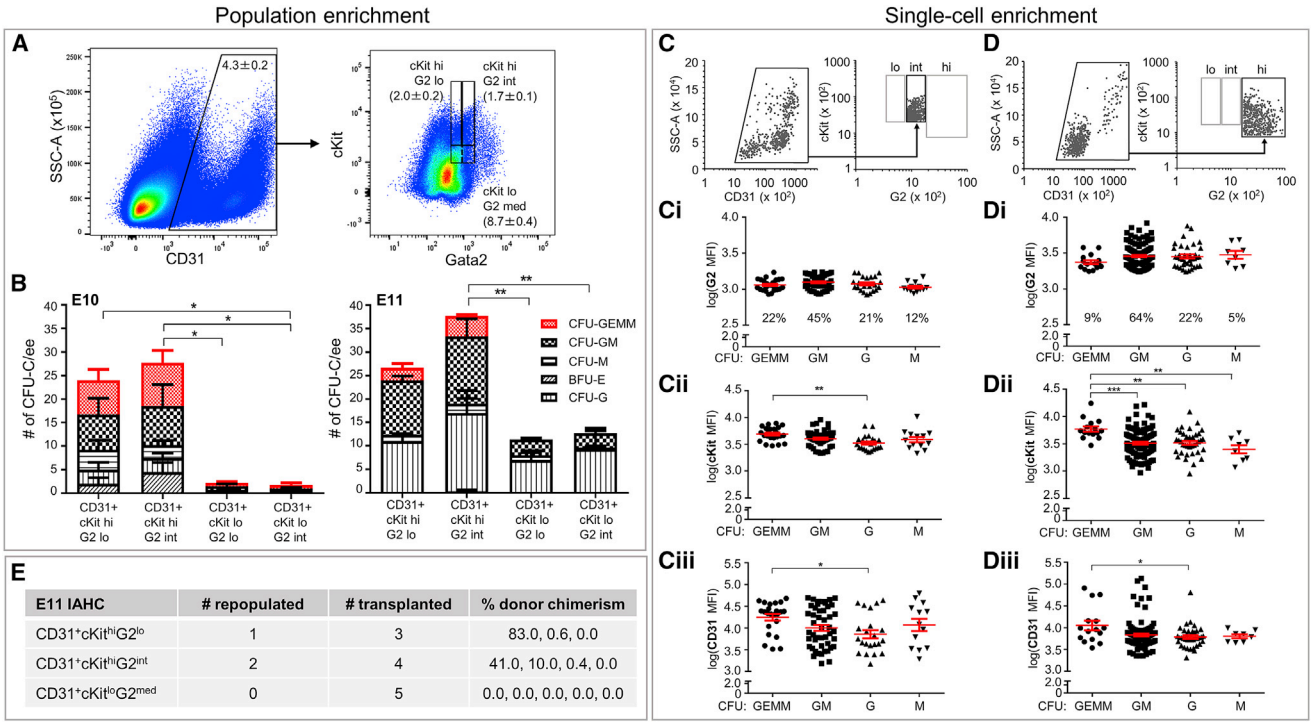
RESULTS

Specific Gata2 and cKit Levels Define All HSCs and the Majority of Multipotent HPCs

Previously, we showed that Venus fluorescence faithfully reports Gata2 expression in *Gata2IRESVenus* reporter (*G2V*) mice (Eich et al., 2018; Kaimakis et al., 2016). In *G2V* embryos, the spatiotemporal expression of Gata2 and the development of EHT and HSC are normal. At E10.5, all HSCs and 70% of HPCs express $G2^{med}$. Some HPCs are Gata2 negative ($G2^{-}$) or high ($G2^{hi}$) (Eich et al., 2018; Kaimakis et al., 2016), and mature mast cells express the highest levels (Kauts et al., 2018; Tsai and Orkin, 1997). To focus on the most potent Gata2-expressing progenitors and HSCs, we followed a multi-pronged approach combining the refinement of sorting gates and single-cell index-sorting with functional *in vivo* and *in vitro* assays. $G2^{med}CD31^{+}ckit^{+}$ IAHC cells (all IAHC cells co-express CD31 and cKit) were separated into low ($G2^{lo}$) and intermediate ($G2^{int}$) Gata2 fractions and further partitioned into $cKit^{lo}$ and $cKit^{hi}$ fractions (Figure 1A). *In vitro* progenitor assays (Figure 1B) of the four fractions revealed that most HPCs were distributed equally between the $CD31^{+}cKit^{hi}G2^{lo}$ and -$G2^{int}$ fractions at both E10 and E11. Whereas, at E10, few or no CFU-Cs were found in the $cKit^{lo}$ fractions, these fractions at E11 contained some CFU-Cs. Multipotent CFU-GEMMs (colony forming unit-granulocyte-erythroid-macrophage-megakaryocytes) were highly enriched or exclusively found in the $CD31^{+}cKit^{hi}G2^{lo}$ and -$G2^{int}$ fractions (the latter containing slightly more GEMMs) at both embryonic stages.

Index-sorted $CD31^{+}cKit^{hi}G2^{int}$ ($G2^{int}$) cells (Figures 1B and 1C) and $CD31^{+}cKit^{+}G2^{hi}$ ($G2^{hi}$) cells (as control; Figure 1D) were examined in single-cell CFU assays to test whether HPC lineage potential was related to Gata2 level. After 9–11 days, both $G2^{int}$ and $G2^{hi}$ cells produced CFU-GEMMs, -GMs (granulocyte-macrophages), -Gs (granulocytes), and -Ms (macrophages) (Figures 1Ci and 1Di; plating efficiency, 1:5–6 and 1:2–4, respectively). However, the proportion of GEMMs was higher in the $G2^{int}$ fraction (22% versus 9%). The CFU-C type arising from each single index-sorted cell was plotted against log-normalized mean fluorescent intensities (MFIs) for G2 (Figures 1Ci and 1Di), cKit (Figures 1Cii and 1Dii), and CD31 (Figures 1Ciii and 1Diii). Cells that produced CFU-GMs and -Gs were widely varying in levels of G2 expression (MFI = 842–8,315) compared to those that produced CFU-GEMMs and -Ms (MFI = 850–4,845). The average MFI for cKit expression on $G2^{int}$ cells yielding CFU-GEMMs was significantly higher than for cells yielding CFU-Gs (Figure 1Cii), and the average MFI for cKit expression on $G2^{hi}$ cells yielding CFU-GEMMs was significantly higher than for the other three CFU types (Figure 1Dii). Both $G2^{int}$ and $G2^{hi}$ fractions contained cells yielding CFU-GEMMs with a significantly higher average CD31 MFI than that for cells yielding CFU-Gs (Figures 1Ciii and 1Diii). Thus, the most immature HPCs are $cKit^{hi}$ and $CD31^{hi}$.

Examination of *in vivo* repopulating activity within the sorted E11 IAHC fractions showed that, despite $G2^{int}$ containing slightly



E

E11 IAHC	# repopulated	# transplanted	% donor chimerism
CD31^{+}cKit^{hi}G2^{lo}	1	3	83.0, 0.6, 0.0
CD31^{+}cKit^{hi}G2^{int}	2	4	41.0, 10.0, 0.4, 0.0
CD31^{+}cKit^{lo}G2^{med}	0	5	0.0, 0.0, 0.0, 0.0, 0.0

Figure 1. The Most Immature AGM HPCs Are cKit^{hi}CD31^{hi}

(A) Flow-cytometric dot plots of representative E11 *G2V* AGM cells analyzed for side scatter (SSC-A) and CD31 expression (left) and cKit and Gata2 (G2) expression of CD31^{+} cells (right). Cell percentage (±SEM) within gated regions. Hi, high; lo, low; int, intermediate; med, medium (lo+int).

(B) Hematopoietic progenitor numbers per lineage per embryo equivalent (ee) of E10 (32–37 somite pairs, sp; n = 4; embryos/experiment: 4, 4, 4, and 4) and E11 (43–48 sp; n = 3; embryos/experiment: 3, 6, 5, and 7) sorted AGM fractions in the colony-forming unit-culture (CFU-C) assay. GEMM, granulocyte-erythroid-macrophage-megakaryocyte; GM, granulocyte-macrophage; M, macrophage; G, granulocyte; BFU-E, burst-forming unit-erythroid. Error bars represent mean ± SEM. For E10 GEMMs: *cKit^{hi}G2^{int} versus cKit^{lo}G2$^{lo/int}$. For E11 GEMMs: *cKit^{hi}G2^{lo} versus cKit^{lo}G2$^{lo/int}$, and **cKit^{hi}G2^{int} versus cKit^{lo}G2$^{lo/int}$. For E11 GMs: *cKit^{hi}G2^{int} versus cKit^{lo}G2$^{lo/int}$.

(C and D) E11 *G2V* AGM single cells were sorted for (C) CD31^{+}cKit^{hi}G2^{int} (11 embryos, 41–47 sp) and (D) CD31^{+}cKit^{+}G2^{hi} cells (20 embryos; 40–49 sp) and deposited in methylcellulose. CFU-Cs were counted at days 9–11 (total CD31^{+}cKit^{hi}G2^{int} = 169; total CD31^{+}cKit^{+}G2^{hi} = 105). Each CFU (individual dot) was examined for colony type and plotted against $\log_{10}$-normalized mean fluorescence intensities (MFIs) for expression of (Ci and Di) G2, (Cii and Dii) cKit, and (Ciii and Diii) CD31. Red lines indicate mean ± SEM.

(E) HSC activity in sorted AGM fractions as measured by *in vivo* transplantation (peripheral blood donor chimerism at 4 months post-injection). Cell numbers injected: for cKit^{hi}G2^{lo}, 950–1963, 2.9–5.4 ee; for cKit^{hi}G2^{int}, 912–2,048, 2.5–6.2 ee; for cKit^{lo}G2^{med}, 4,530–12,523, 3.3–5.9 ee (42–50 sp; n = 4; embryos/experiment: 7, 9, 12, and 14).

more (multipotent) HPCs, both CD31^{+}cKit^{hi}G2^{lo} and -G2^{int} cells provided HSC engraftment at 4 months post-transplantation (Figure 1E). No HSCs were detected in the CD31^{+}cKit^{lo}G2^{med} fraction (Figure 1E) or, as we previously reported, in the G2^{hi} fraction (Eich et al., 2018).

Transcriptome Heterogeneity and Clustering of Single CD31^{+}cKit^{hi}G2^{med} Cells

Considering that, at E11, all GEMMs were found in the G2^{lo}, G^{int}, and G2^{hi} fractions and that HSCs are exclusively in the G2^{lo} and G2^{int} fractions, we index-sorted CD31^{+}cKit^{hi}G2^{med} (G2^{lo}+G2^{int} = G2^{med}) IAHC cells and analyzed them by single-cell RNA sequencing (RNA-seq) using SmartSeq2 (Picelli et al., 2014). Out of 1,015 CD31^{+}cKit^{hi}G2^{med} cells, 968 passed quality control (95%). Of note, the mapping of reads was performed using a custom reference that included the mouse genome, the sequence of *Venus* and External RNA Controls Consortium (ERCC) sequences, allowing confirmation of the co-regulation of *Gata2* and *Venus* expression levels (Figure S1A; correlation coefficient, 0.7). We obtained the SPRING visualization (Weinreb et al., 2018) of the 968 cells and defined 11 clusters (L1.1–L1.11) using the Louvain method (Figure 2A; Figure S1Bi). Datasets obtained from several biological samples, or analyzed on different days, integrated well and showed no batch effect (Figure S1Bii). Hence, cell acquisition parameters and results were consistent between experiments.

To investigate cluster composition, cells were colored according to the expression of well-known hematopoietic (Figure 2B) and endothelial (Figure 2C) genes as well as several pivotal hematopoietic TFs (heptad TFs; Figure 2D). In this way, we could merge clusters to define 5 main groups of cells within our highly enriched HSC population (Figure 2A). The endothelial-like cluster (EC; L1.1 and L1.2) contained 335 cells expressing endothelial genes such as *Flk1* (*VEGFR-2*), *Cdh5* (*VE-cadherin*), and *ApInR*. *Runx1*, which would be expected to be expressed in all cells (Bee et al., 2009), was downregulated in the EC when compared

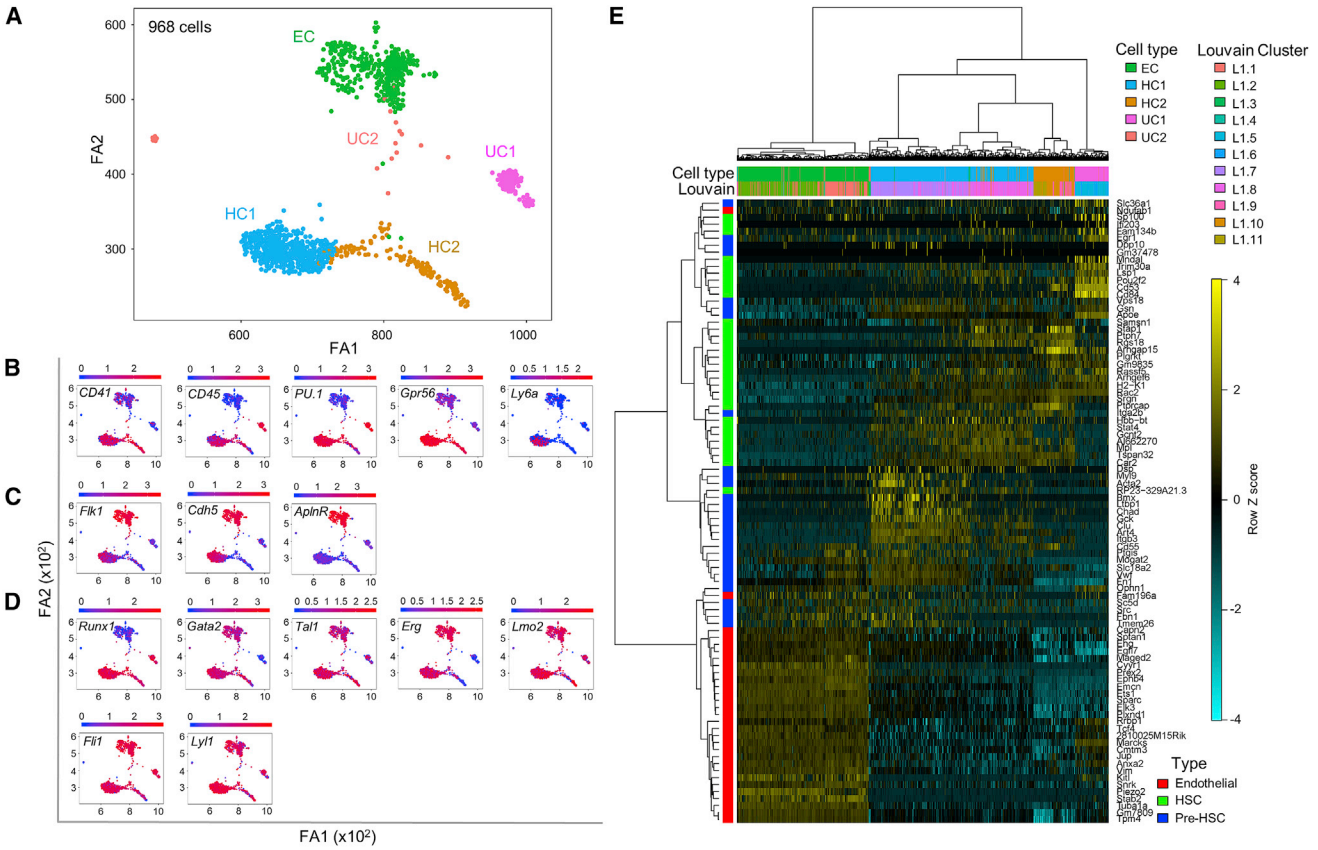


Figure 2. Five Distinct Transcriptomic Clusters Are Defined by Single-Cell RNA Sequencing of Highly Enriched AGM HSCs

(A) SPRING visualization of merged Louvain clusters corresponding to 968 single CD31^{+}cKit^{hi}G2^{med} IAHC cells (n = 6 at E11, 27.6 *G2V* embryos, 41–49 somite pairs; n = 1 at E10, 2.7 *G2V* embryos, 33–34 sp). See Figure S1B for Louvain clustering. Five clusters emerge based on the cell-lineage expression characteristics and hierarchical clustering distances. EC, endothelial-like cluster; HC1, hematopoietic cluster 1; HC2, hematopoietic cluster 2; UC1, unknown cluster 1; UC2, unknown cluster 2.

(B–D) Cells were colored according to the expression of (B) hematopoietic markers, (C) endothelial markers, and (D) hematopoietic "heptad" transcription factors.

(E) Heatmap and hierarchical clustering of the dataset of 968 CD31^{+}cKit^{hi}G2^{med} index-sorted cells using endothelial-, pre-HSC-, and HSC-specific genes (indicated on the left in red, blue, and green, respectively) obtained from a published dataset of phenotypically enriched endothelial cells, pre-HSCs (T1+T2), and E12/E14 HSCs (Zhou et al., 2016). Louvain identity (L1.1 to L1.11) and merged clusters (EC, HC1, HC2, UC1, and UC2) are also indicated (top).

to the other clusters. The hematopoietic cluster (HC) contained 529 cells (L1.7–L1.11) expressing well-known hematopoietic marker genes such as *CD41*, *CD45*, *PU.1*, and *Gpr56*. HCs could be subdivided into HC1 (405 cells; L1.7 and L1.8) and HC2 (124 cells; L1.9–L1.11) cells, according to the expression of genes such as *VE-cadherin* and *Ly6a*. Finally, two smaller unknown clusters were defined: UC1 and UC2. UC1 contained 85 cells (L1.5–L1.6) expressing some hematopoietic markers such as *CD45* and *PU.1* and lacking expression of *CD41*, *Gpr56*, and endothelial markers *Flk1*, *Cdh5*, and *AplnR*. Cells within this cluster also expressed high levels of *Runx1*. UC2 contained 19 cells (L1.3–L1.4) that expressed *Gata2* consistently, although there were too few cells for confident analysis. In a more detailed analysis of the expression of the heptad TFs, *Gata2* showed some heterogeneity in expression levels in all clusters except UC1, in which it was downregulated (see Discussion). *Tal1* also showed heterogeneous expression in all clusters. *Erg* and *Lmo2* were downregulated in HC2. These data demonstrate heterogeneity in heptad TF expression among the single cells of highly enriched HPCs and HSCs. While some cell clusters show consistent levels of one TF, other TFs vary in expression within a cluster, suggesting an unstable regulation and/or dynamic expression between highly similar single cells.

We also calculated the similarity of each cell to the gene expression profile of validated adult bone marrow HSCs using hscScore (Hamey and Göttgens, 2019) (Figure S1C). The cells within HC1 obtained the highest score, suggesting that this cluster most likely contains the HSC fraction. Our datasets were further examined in the context of a published single-cell dataset (Zhou et al., 2016) of phenotypically enriched AGM endothelial cells, pre-HSCs (type 1 and type 2), and E12 and E14 fetal liver HSCs. First, 3 groups of genes specifically expressed in endothelial cells, pre-HSCs, and HSCs were defined by comparing the different cell types represented in the published dataset (Figure S1D). These genes were used to cluster our dataset (Figure 2E). Our EC correlated well with the endothelial profile. HC1 contained cells that expressed both pre-HSC-specific genes (*Procr*) and HSC-specific genes

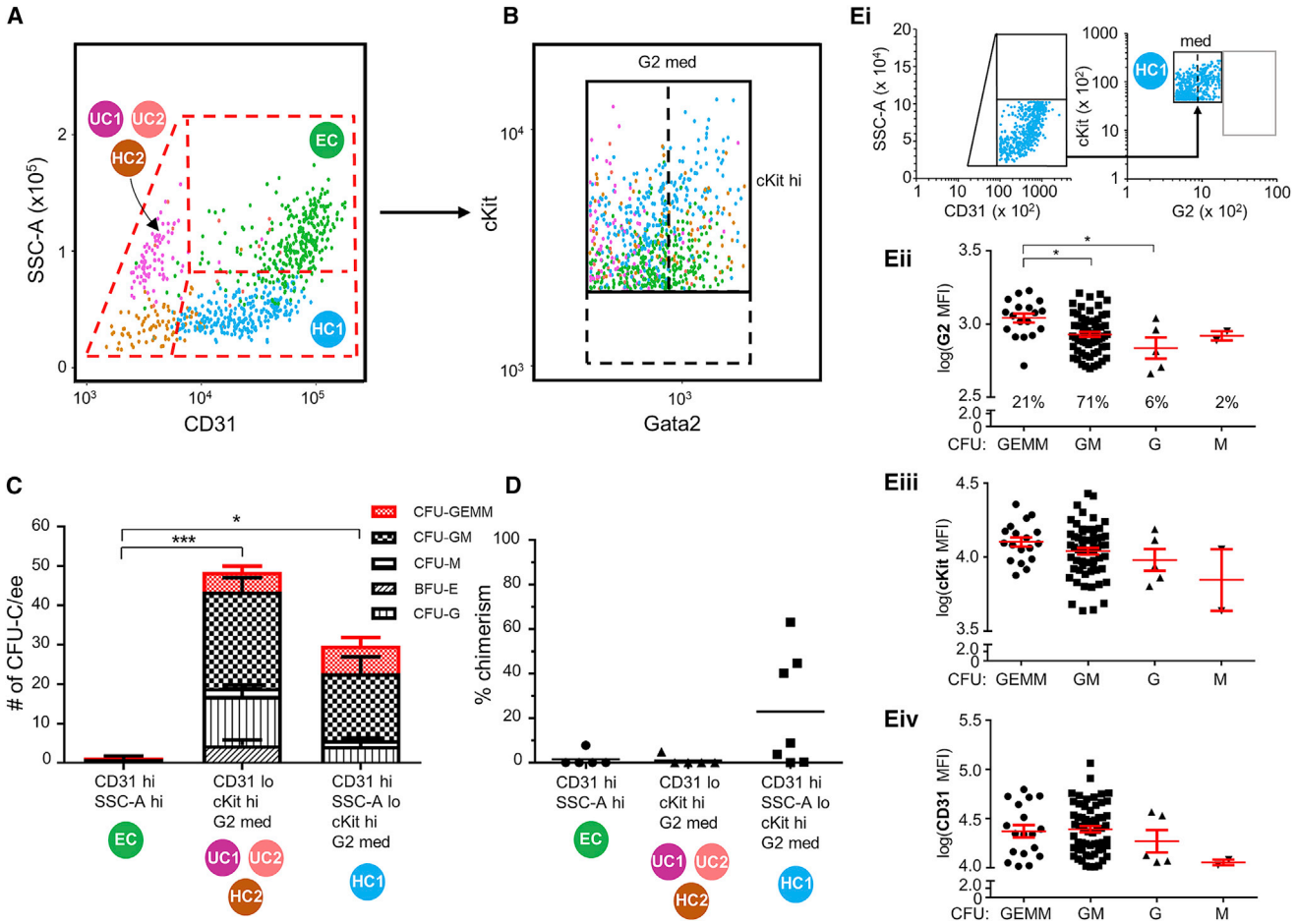


Figure 3. Index-Sorting Parameters Separate Transcriptomic Clusters and Guide Further Enrichment of Functional Hematopoietic Stem and Multipotent Progenitor Cells

(A and B) Flow-cytometric plots showing (A) side-scatter (SSC-A) and CD31 expression and (B) cKit and Gata2 (G2) expression of the 968 index-sorted single cells. The color of each cell indicates presence in specific clusters: EC (endothelial-like cluster, green), HC1 (hematopoietic cluster 1, blue), HC2 (hematopoietic cluster 2, ochre), UC1 (unkown cluster 1, magenta), and UC2 (unkown cluster 2, rose).

(C) Hematopoietic progenitor number and type per embryo equivalent (ee) of bulk-sorted E11 IAHC cells (n = 4; 41–50 somite pairs, sp; embryos/experiment: 8, 8, 3, 8) corresponding to transcriptomic clusters, as determined by colony-forming unit-culture (CFU-C) assay. GEMM, granulocyte-erythroid-macrophage-megakaryocyte; GM, granulocyte-macrophage; M, macrophage; G, granulocyte; BFU-E, burst-forming unit-erythroid. For GEMMs: *EC versus HC1; for GMs: *EC versus HC1, and **UC1+2+HC2 versus EC. Mean ± SEM.

(D) HSC activity in sorted E11 IAHC cells (n = 5; 40–49 sp; embryos/experiment: 15, 14, 11, 8, and 19) corresponding to transcriptomic clusters by *in vivo* transplantation. Each point indicates percentage of donor-derived cells in peripheral blood of recipients at 4 months post-injection. Cell numbers injected: for HC1, 721–2,811, 4.8–8 ee; for EC, 5,654–14,208, 4.7–8.1 ee; for UC1+2+HC2, 1,312–2,828, 4.8–8.4 ee. Black line indicates mean.

(E) In (Ei), cytometric gates were used for sorting HC1 cells from E11 *G2V* IAHCs (16.7 embryos, 41–49 sp). Single cells deposited for CFU-Cs were counted at 12 days (total CFU-Cs = 86). Each CFU (individual dot) was examined for colony type and plotted against $\log_{10}$-normalized mean fluorescence intensities (MFIs) for (Eii) G2, (Eiii) cKit, and (Eiv) CD31 expression. GEMM, granulocyte-erythroid-macrophage-megakaryocyte; GM, granulocyte-macrophage; G, granulocyte; M, macrophage. Red lines indicate mean ± SEM.

(*CD41, Mpl*, and *Ly6a*) (mainly cells within L1.7) and cells that expressed predominantly only HSC-specific genes (mainly cells within L1.8). HC2 represented a more mature hematopoietic profile where cells expressed all HSC-specific and no pre-HSC-specific genes. UC1 had some endothelial and hematopoietic features and expressed mature hematopoietic genes, including *Cd53*, *Cd84*/SLAM, *Mndal*/myeloid, and *Sp100*/*PML*; almost no pre-HSC-specific genes; and a subset of endothelial-specific genes (*Pmp22*, *Anxa2*, *Anxa3*, *Ophn1*, *Marcks*, and *Tcf4*).

Phenotypic Characteristics Derived from Index-Sorting Guide Further Enrichment of Multipotent HCs

To examine whether the 5 transcriptomic clusters of CD31^{+} cKit^{hi}G2^{med} cells correspond to specific hematopoietic activities, index-sorting data were analyzed for phenotypic characteristics that would provide for cluster-distinct cell isolation. EC cells were CD31^{hi} and side-scatter high (SSC^{hi}), UC1+UC2+HC2 cells were CD31^{lo}, and HC1 cells were CD31^{hi} and SSC^{lo} (Figure 3A). No clear further separation was observed within the cKit^{hi} fraction (Figure 3B), although HC1 cells expressed higher

levels of cKit than EC and UC1+UC2+HC2 cells, which showed variable levels. Based on these new sorting parameters ($CD31^{hi}SSC^{hi}$, $CD31^{lo}$, and $CD31^{hi}SSC^{lo}$), three fractions were sorted and tested *in vitro* for CFU-C formation and *in vivo* for long-term HSC repopulation.

No CFU-Cs were generated from EC cells ($CD31^{hi}SSC^{hi}$; Figure 3C) as expected. $CD31^{lo}cKit^{hi}G2^{med}$ cells (UC1+UC2+HC2) contained 50 CFU-Cs per embryo equivalent (ee) of various CFU-C types. $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}$ cells (HC1) contained 30 CFU-Cs per ee. Interestingly, a trend of more multipotent HPCs (7 GEMMs per ee), fewer bi-potent HPCs (17 GMs per ee), and few or no unipotent HPCs (Gs and burst-forming unit-erythroid [BFU-Es]) was detected in HC1 cells as compared to the UC1+UC2+HC2 cells (5 GEMMs per ee and 24.5 GMs per ee). The results of long-term *in vivo* transplantations showed that the HC1 fraction contains all HSCs (Figure 3D) and is 10.5-fold enriched compared to the $CD31^{+}cKit^{+}G2^{med}$ fraction.

To provide scope for further enrichment of HSCs and multipotent HPCs, HC1 cells (Figure 3Ei) were index-sorted and analyzed for single-cell CFU-C lineage output and G2, cKit, and CD31 MFIs. A total of 86 CFU-Cs was obtained (plating efficiency, 1:9–11). Ninety-two percent (21% GEMMs + 71% GMs) of the CFU-Cs were multipotent/bipotent (Figure 3Eii). Cells yielding GEMMs showed significantly higher average MFIs for G2 expression as compared to GM and G cells (Figure 3Eii). MFIs for cKit (Figure 3Eiii) and CD31 (Figure 3Eiv) were not significantly different when GEMM and GM cells were compared. Together, these data indicate that all HSCs and the most multipotent HPCs are enriched within the $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}$ (HC1) fraction and that cells that produce GEMMs had higher G2 MFIs as compared to the more widely varying G2 MFIs for GM-producing cells.

Iterative Analyses Link Expression of Specific Genes with HSC and Multipotent HPC Function

Gene lists were examined for new candidate markers of HC1 cells as compared to the other clusters (Table S1). Violin plots showed significantly higher *Tnfrsf7* (CD27) expression in HC1 (Figure 4A; Figure S2A), as well as *Selp* and *Mpl* (Figure S2B). Index-sorting parameters revealed that 70% of E11 $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}$ (HC1) cells were $CD27^{+}$ (Figure 4B). Using this wide (+W) gate (MFI = 150–7,000), $CD27^{+W}$ cells were isolated and transplanted into irradiated recipients. High-level, long-term multilineage *in vivo* repopulation by the $CD27^{+W}$ HC1 donor cells was detected in 2 of 2 recipients (Figure 4C, open squares). Secondary transplantations confirmed that these were self-renewing HSCs (Figure S3A). For further enrichment, a higher and narrower (+N) CD27 gate (MFI = 500–7,000) was set (Figure 4B), and $CD27^{+N}$ (19%) and $CD27^{-/lo}$ (81%) HC1 cells were transplanted. The $CD27^{+N}$ HC1 cells yielded high-level, long-term multilineage repopulation in both primary (Figure 4C, black squares) and secondary (Figure S3A) transplantations. No HSCs were found in the $CD27^{-/lo}$ HC1 fraction, even with injection of high numbers of cells. As compared to the $CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}$ sorted fraction, the $CD27^{+N}$ gate enriches HSCs by a further 5.2-fold.

Single-cell RNA-seq of 119 $CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N}$ cells (from 6 AGMs) was performed, and data were combined with the 968 cells to produce a new SPRING plot of 1,087 cells. Louvain clustering showed a cluster distribution (L2.1–L2.12) of the 1,087 cells (Figures S3B and S3C) almost identical to that obtained from the 968 cells (Figure S1Bi). The majority (71%) of the 119 $CD27^{+N}$ cells (Figure 4D, red) localized to HC1, and 58% localized to L2.8 (Figure 4D, circled).

Interestingly, when index-sorted $CD27^{+N}$ cells were analyzed in single-cell CFU-C assays (plating efficiency, 1:3–4 cells), GEMM-forming cells showed a significantly lower average CD27 MFI (767.1 ± 41.8; Figure 4Ei) and a significantly higher G2 MFI (1,022.0 ± 60.1; Figure 4Eii) than those of GM-forming cells (945.2 ± 44.6 and 858.3 ± 31.8, respectively). Average cKit and CD31 MFIs were not significantly different (Figures 4Eiii and 4Eiv). Notably, 95% of immature/HSC-like progenitors forming GEMMs fall within the smaller CD27 MFI range of 513–1,084 (Figure 4F, red dotted area), which is now referred to as $CD27^{med}$. The $CD27^{med}$ fraction is 15.5% ± 1.3% of HC1 and 86% of $CD27^{+N}$ cells. There are 20 ± 2 $CD27^{med}$ HC1 cells per AGM, of which 1–2 were HSCs at the time of analysis. In addition to our HC1 enrichment (10.5-fold), a further 6.5-fold enrichment of HSCs was achieved in the $CD27^{med}$ fraction to yield a combined enrichment of 68.3-fold (summarized in Figure 4G).

Defining the AGM HSC Transcriptome

The increased enrichment of multipotent cells within the $CD27^{med}$ MFI range provided a means for dataset refinement in which the molecular profiles of different functional subsets could be examined. Out of the 119 $CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N}$ cells previously isolated for single-cell RNA-seq (Figure 4D), 85 belonged to HC1, of which 14 showed $CD27^{hi}$ and 71 $CD27^{med}$ expression. We analyzed the 85 cells using Iterative Clustering and Guide-gene Selection (ICGS) (Olsson et al., 2016), which can define groups of genes that behave similarly and identify groups of cells with similar gene-expression patterns. Three groups of genes and 3 groups of cells that we named ICGS1, ICGS2, and ICGS3 contained 54, 29, and 2 cells, respectively (Figures S4 and S5). When the distribution of the ICGS groups within the previously defined L2.7 and L2.8 clusters was examined, 81% of the 85 $CD27^{+N}$ cells fell within L2.8, and 19% fell within L2.7. ICGS1 cells were mainly in L2.8, whereas ICGS2 cells were equally distributed in L2.7 and L2.8, suggesting that these cells are acquiring HSC identity. The hscScore for each cell was calculated and plotted against its CD27 MFI value (Figure 5A). ICGS2 cells were almost exclusively in the $CD27^{med}$ and high hscScore range, whereas ICGS1 cells were within the full CD27 MFI range and showed a low hscScore. Together, these data confirm that ICGS2 cells are the most multipotent or HSC-like.

To more precisely identify the transcriptome of the few HSCs within the 85 cell profiles, other known HSC characteristics were examined. *Ly6a* is known to be expressed by all *in vivo* adult repopulating AGM HSCs (de Bruijn et al., 2002). *Ly6a* can be detected in 27 of the 85 cells within HC1 (Figure 5Bi), 23 of 71 $CD27^{med}$ cells (Figure 5Bii), and 4 of 14 $CD27^{hi}$ cells (Figure 5Biii), consistent with previous findings that *Ly6a* expression is not exclusive to HSCs. We then focused our attention on the 23 $Ly6a^{+}$ cells within the "multipotent/HSC" $CD27^{med}$ MFI range (Figure 5C). These cells separated into two ICGSs: 15 belonged

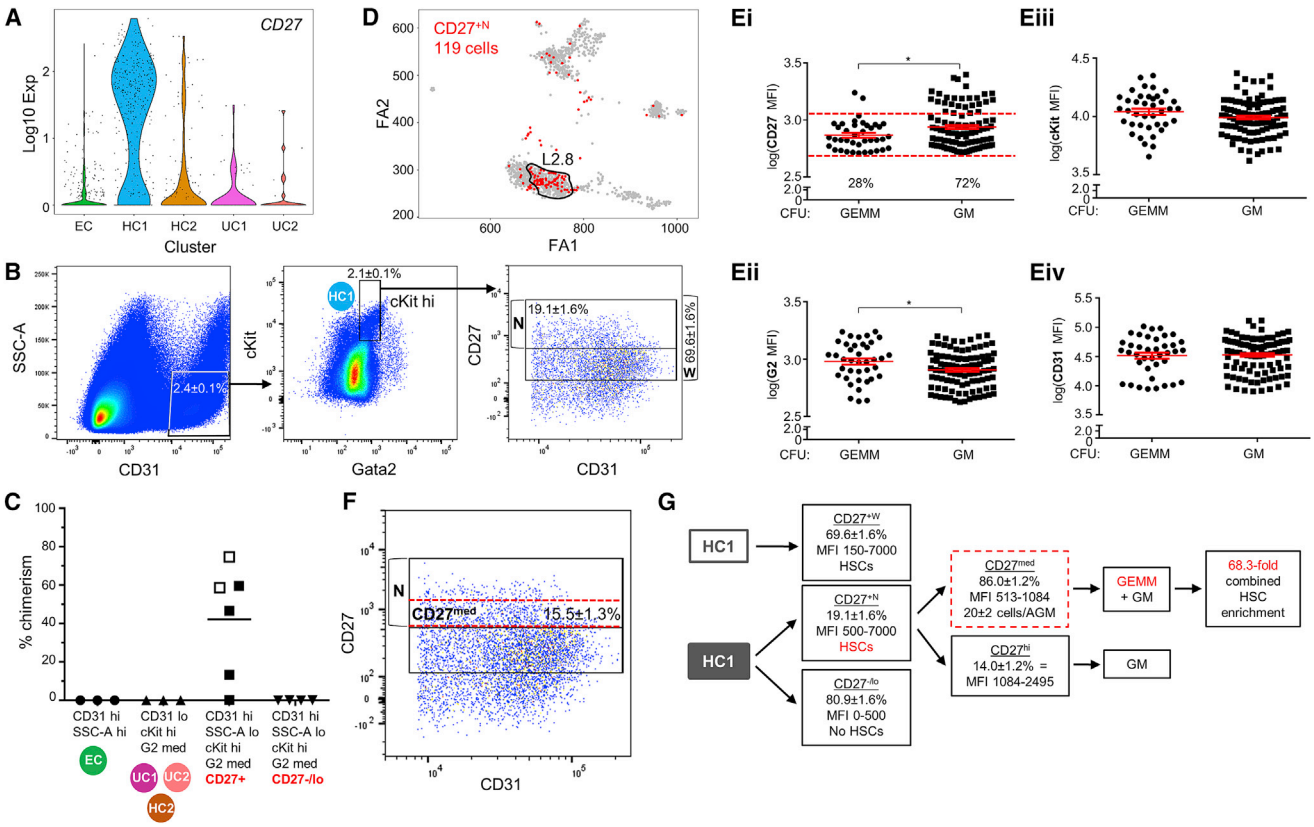


Figure 4. All Functional AGM HSCs Are within the HC1 CD27^{med} Fraction

(A) Violin plots of *Tnfrsf7* (CD27) expression in the 5 major transcriptomic clusters. EC, endothelial-like cluster; HC1, hematopoietic cluster 1; HC2, hematopoietic cluster 2; UC1, unknown cluster 1; UC2, unknown cluster 2.

(B) Enrichment sorting parameters for CD27^{+} HC1 cell isolation. Side scatter (SSC-A), CD31, cKit, Gata2 (G2), and CD27 sorting gates and percentage (±SEM) of cells. CD27^{+} cells (%) are indicated within a wide (+W) and a narrow (+N) mean fluorescent intensity (MFI) gate.

(C) Analysis of HSC function within CD27 fractions (CD27^{+W}, CD27^{+N}, and CD27$^{-/lo}$) of HC1 cells by *in vivo* transplantation. Open squares denote CD27^{+W} HC1 cell recipients (n = 2) of 442 cells (2.5ee) and 479 cells (3.3 ee) from 15 and 9 (42–50 somite pairs, sp) embryos, respectively. Closed squares (n = 4; 42–49 sp; embryos/experiment: 25, 11, 22, 28) denote recipients injected with CD27^{+N} HC1 cells (238–957; 3.7–9.8 ee), EC cells (899–2,333; 3.9–8.6 ee), CD27$^{-/lo}$ HC1 cells (1,042–4,325; 4.1-9.8 ee), and UC1+2+HC2 cells (1,361–3,741; 4.2–9.9 ee). Percentage of donor cell chimerism in peripheral blood of recipients at 4 months post-injection. Black line indicates mean.

(D) Single-cell RNA-seq of 119 CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N} sorted cells (all passed quality control) from 6 embryos (41–47 sp). These data combined with the previous dataset (gray) resulted in an extended dataset SPRING plot of 1,087 cells in total. Overlay of CD27^{+N} cells (red) showing localization to one HC1 subcluster (circled).

(E) Single CD27^{+N} HC1 cells from E11 *G2V* IAHCs (50 embryos; 42–49 sp) deposited into methylcellulose and colony-forming unit-culture (CFU-C) counts at 8–10 days (138 total CFU-Cs). Colony type (individual dot) plotted against $\log_{10}$-normalized mean fluorescence intensities (MFIs) for (Ei) CD27, (Eii) G2, (Eiii) cKit, and (Eiv) CD31 expression. GEMM, granulocyte-erythroid-macrophage-megakaryocyte; GM, granulocyte-macrophage. Red lines indicate mean ± SEM.

(F) Representative FACS plot showing newly defined CD27^{med} MFI range (red dotted line) within the CD27^{+N} fraction. Percentage ±SEM.

(G) Summary of enrichment strategy. The HC1 fraction containing all HSCs is 10.5-fold enriched compared to previous sorting strategies (Eich et al., 2018). Further enrichments are based on gating of specific CD27^{+N} and CD27^{med} fractions to give an additional 6.5-fold enrichment. After iterative analyses and refinement of sorting gates, a total combined HSC enrichment of 68.3-fold is achieved.

to ICGS1, and 8 belonged to ICGS2. The 8 cells contained in ICGS2 had the highest hscScores: 4 belonged to L2.7, and 4 belonged to L2.8. Thus, the stratified results of single-cell index-sorting, transcriptome profiling, and functional analyses strongly suggest a profile for HSCs that is unique but closely related to that of multipotent HPCs.

Considering the expression of several genes relevant to HSC generation, the profile for the 8 *Ly6a*$^{+}$CD27^{med} ICGS2 single cells stands out (Figure 5C). Notch pathway genes (Bigas and Porcheri, 2018) are highly represented in these cells and include *Hes1* and *Hey1* (negative regulators), *Dll4* (Notch ligand), and *Notch4*. *Sox17*, which is known to be expressed in fetal HSCs and maintain stem cell potency (He et al., 2011), is highly expressed, as is *Il6st*, which is part of the receptor complex for hematopoietic and stem cell factors interleukin (IL)-6, leukemia inhibitory factor (LIF), and Oncostatin M (Mihara et al., 2012). Also, *Nos3* (endothelial nitric oxide synthetase), which functions in AGM HSC and HP/SC generation (North et al., 2009), is highly expressed. *VEGF-R*, *Flt1*, and *Kdr* are expressed along with other HSC genes

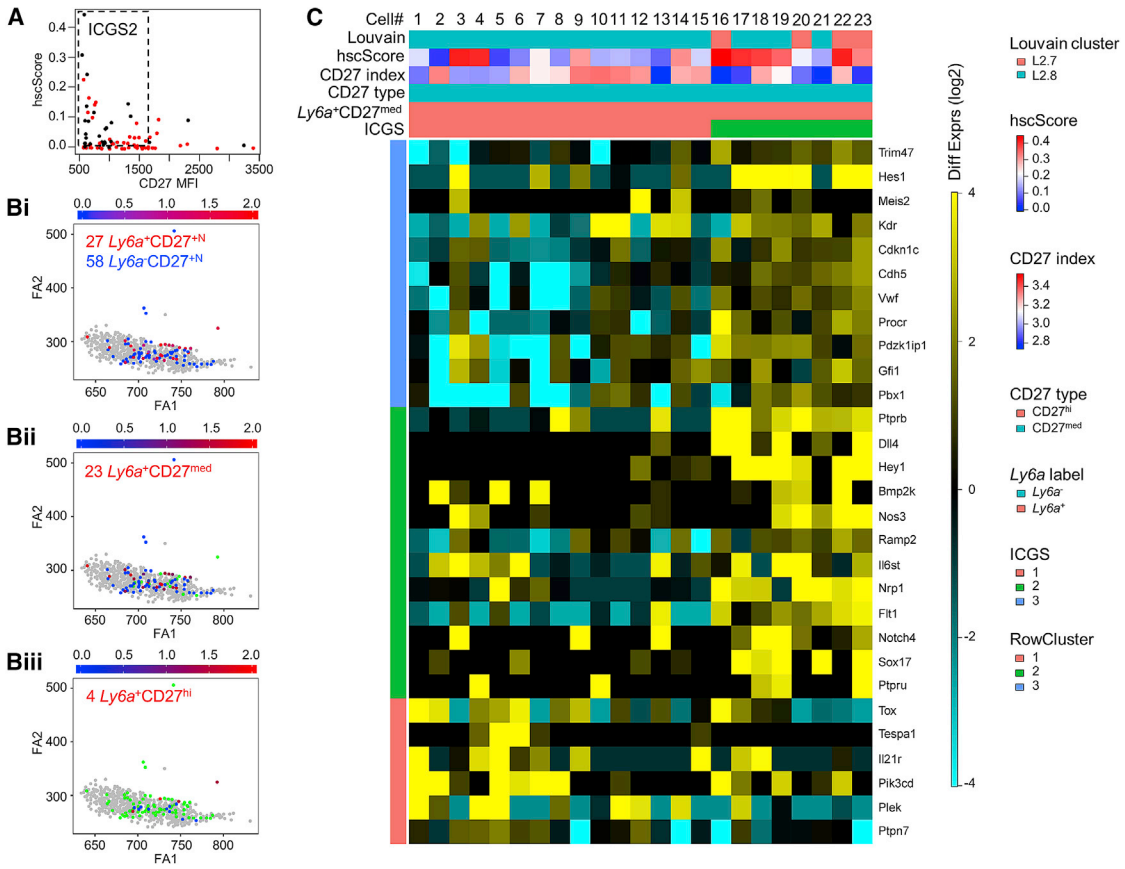


Figure 5. Defining the Transcriptome of the First Functional HSCs

(A) Plot of hscScore against CD27 MFI for 85 CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N} cells. Majority of ICGS2 cells (black) found in the CD27^{med} and higher hscScore range and ICGS1 cells (red) within the full CD27 MFI and mostly low hscScore range. ICGS3 cells are indicated in gray.

(B) *Ly6a* expression was plotted on the SPRING visualization of (Bi) 85 CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N} cells, (Bii) 71 CD27^{med} cells (CD27^{hi} in green), and (Biii) 14 CD27^{hi} cells (CD27^{med} in green) within HC1 (gray).

(C) Replotting of 85-cell heatmap in Figure S4 (maintaining the same order and position) with subset of ICGS output (Figure S5) for the 23 *Ly6a*$^{+}$CD27^{med} cells belonging to HC1 (L2.7+L2.8). 3 gene clusters (RowCluster) and 2 cell clusters (ICGS1 and ICGS2) can be defined. HscScore, $\log_{10}$ of the CD27 index-sorting (protein) levels and Louvain identity (L2.7, L2.8) are also depicted.

such as *Cdkn1c*, *Pdzk1ip1*, *Procr*, *Ramp2*, *Trim47*, and *Vwf*. Thus, the near-equivalence of the transcriptomic profiles of the 8 *Ly6a*$^{+}$CD27^{med} ICGS2 cells provides a likely genetic identity for the first cells achieving HSC function in the mouse embryo.

Functional HSCs Are Localized in 1- to 2-Cell IAHCs

Having defined the transcriptome of functional HSCs, we examined the localization of CD27-, Ly6a-, and Gata2-expressing cells within IAHCs. E11 *G2V* cryosections immunostained for CD27 and CD34 showed rare G2^{+}CD34^{+}CD27^{+} cells almost exclusively within the smallest ventrally located IAHCs. Figure 6A shows two CD34^{+}G2^{+}CD27^{+} cells (arrowhead) attached to one CD34^{+}G2^{+}CD27^{-} cell within the endothelium, and a single rounded/bulging CD34^{+}G2^{+}CD27^{+} cell (arrow). Another section shows a ventral cluster (Figure 6Bi) with two CD34^{+}G2^{med}CD27^{+} cells (one appearing to be bulging out of the endothelium) surrounded by 3–4 CD34^{+}G2^{+}CD27^{-} cells. Also found ventrally (Figure 6Bii) was a 2-cell cluster with a single CD34^{+}G2^{med}CD27^{+} cell closely associated with a CD34^{+}G2^{lo}CD27^{-} cell (Video S1). Figure 6Biii shows a single rounded bulging CD34^{+}G2^{+}CD27^{+} cell on the dorsal aspect of the aorta (likely an HPC; Taoudi and Medvinsky, 2007). Examination of CD27^{+} cell-containing IAHCs revealed that 80% (56 out of 70; 42 in 1-cell clusters and 14 in 2-cell clusters; n = 4 embryos) were 1- to 2-cell clusters (Figure S6). Since all HSCs and most multipotent progenitors are CD27^{+}, we conclude that these functional cells are localized closely to the ventral aortic endothelium within small or single-cell IAHCs.

E11 *Ly6aGFP* transgenic embryo sections were also immunostained to examine whether this HSC marker could subset and/or provide further validation into the localization of CD27^{+} HSCs. A single GFP^{+}CD27^{+} cell was observed at the apex of a bulging GFP^{+} cell within the ventral endothelium (Figure 6Ci, arrowhead); another GFP^{+} cell (arrow) was observed in the ventral endothelium, but it was CD27^{-}. Figure 6Cii indicates two round adjacent CD27^{+} cells, of which

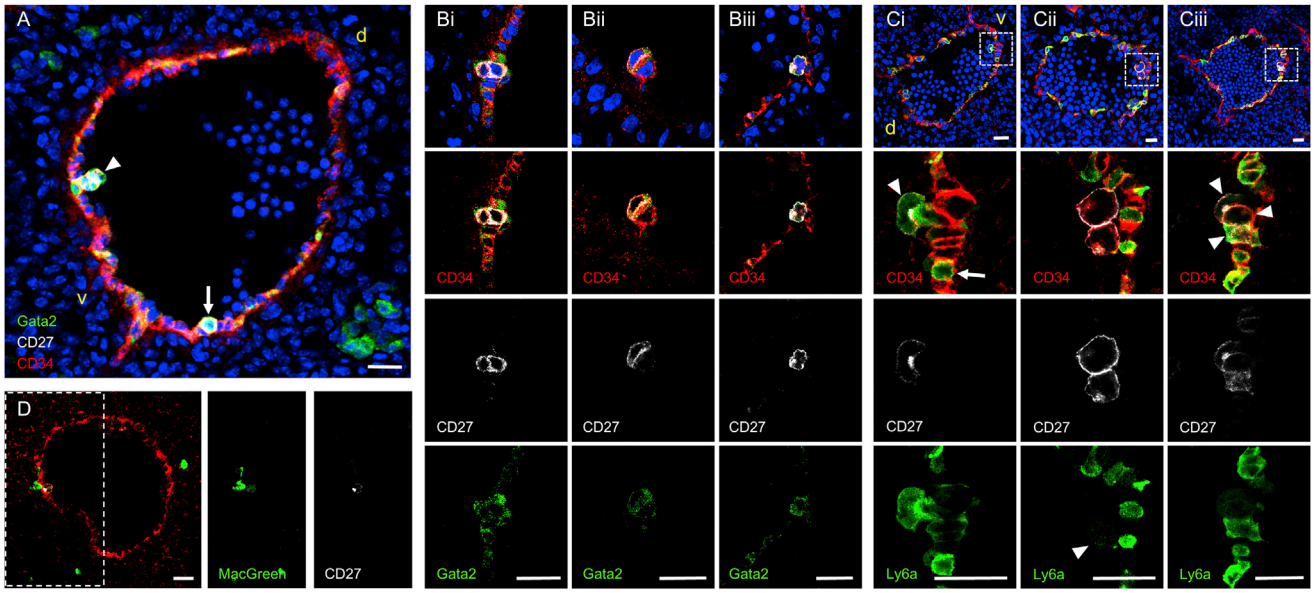


Figure 6. Functional HSCs Are Localized in 1- to 2-Cell IAHC Cells

(A–D) Sections through dorsal aorta of E11 (43 somite pairs, sp) *Gata2Venus* (green in A and B) and *Ly6aGFP* (green in C) and of E10.5 (36sp) *Csf1rGFP* (MacGreen; green in D) embryos. Nuclei are indicated in blue (DAPI), endothelial and IAHC cells are indicated in red (CD34) and CD27$^+$ cells are indicated in white. v, ventral, d, dorsal.

(A) Arrowhead indicates small cluster containing two G2$^+$CD27$^+$ cells attached to one G2$^+$CD27$^-$ cell. Arrow indicates a single rounded/bulging G2$^+$CD27$^+$ cell.

(B) In (Bi), ventral cluster with two G2^{med}CD27$^+$ cells (one bulging out of endothelium). (Bii) Ventral two-cell cluster (G2^{med}CD27$^+$ and G2^{lo}CD27$^-$ cells). (Biii) Single rounded bulging G2$^+$CD27$^+$ cell on dorsal side of aorta.

(C) In (Ci), single Ly6aGFP$^+$CD27$^+$ cell (arrowhead) at apex of bulging Ly6aGFP$^+$ cell within ventral endothelium. A neighboring GFP$^+$ cell in the ventral endothelium is CD27$^-$ (arrow). (Cii) Two round CD27$^+$ cells (one is Ly6aGFP^{lo}, indicated by arrowhead) adjacent to each other and attached to interspersed Ly6aGFP$^+$ endothelium. (Ciii) Three CD27$^+$ cells (indicated by arrowheads; one is Ly6aGFP$^+$) in very close association with ventral endothelium. White dashed boxes indicate magnified areas.

(D) One CD27$^+$CD34$^+$ cell on dorsal endothelial wall closely associated with a macrophage. White dashed box indicates area for which MacGreen and CD27 expression is shown.

Scale bars, 20 μm.

one is GFP^{lo} (arrowhead) and attached to interspersed GFP$^+$ endothelial cells. Three CD27$^+$ cells, one of which is GFP$^+$ (Figure 6Ciii, arrowheads), were found in very close association with the ventral endothelium. The juxtaposition of single CD27$^+$Ly6aGFP$^+$ cells to the ventral aortic endothelium indicates that HSCs emerge directly from embryonic endothelium as single cells within intra-aortic clusters of 1–2 cells.

DISCUSSION

Here, we identified the transcriptome of the first functional adult repopulating HSCs. Starting from Gata2-based single-cell index-sorting and iterative single-cell analyses, the "inner" HSC transcriptional regulatory network is connected directly with multipotent hematopoietic function *in vivo* and *in vitro*. Specific levels of Gata2, cKit, and CD27 expression define all functional *in vivo* repopulating HSCs, and, together with Ly6aGFP expression, CD27 localizes the first HSC to the smallest intra-aortic clusters of 1–2 cells. The surprising variability in gene expression levels within highly enriched single cells suggests that establishment of hematopoietic function is stochastic, occurring through dynamic regulation and/or unstable expression of groups of pivotal genes.

Emerging HCs Are Heterogeneous in Phenotype, Transcriptome, and Function

EHT occurs during a short 2- to 3-day period during mouse midgestation (Yokomizo and Dzierzak, 2010; Zovein et al., 2008) and extends over the length of the aorta to produce IAHCs that are predominantly ventral (Peeters et al., 2009; Souilhol et al., 2016; Taoudi and Medvinsky, 2007) and more abundant near the juncture of the vitelline artery (Mascarenhas et al., 2009). Individual intra-aortic clusters range from 1 to 19 cells at E10.5 and from 1 to 11 cells at E11.5 (Yokomizo and Dzierzak, 2010). Single-cell quantitation of functional HPCs and HSCs within the clusters shows that the number of lineage-restricted HPCs is the highest, followed by that of multipotent HPCs and rare HSCs. This low frequency of HSCs implies that stem cell identity/function is a difficult hematopoietic fate to achieve.

Whereas all functional HSCs are in ventral clusters and near the vitelline junction, it cannot be predicted which emerging cell in the aorta is an HSC. The variability in phenotypic marker expression, and the degree of heterogeneity found in the transcriptomes of individual cells during the embryonic development of HPCs, pre-HSCs, and HSCs (Zhou et al., 2016), has precluded the identification of the transcriptome of rare HSCs. Rather than dependence on surface molecules not required for HSC

function, Gata2-based index-sorting more directly connected with the "inner" hematopoietic regulatory network. G2^{med}CD31^{hi} cKit^{hi} cells highly enriched in multipotent HPCs and containing all functional repopulating HSCs show unexpected heterogeneity between single-cell transcriptomic profiles. Louvain analysis of >1,000 sequenced cells revealed 5 major cell clusters. Unexpectedly, within this highly enriched set of IAHC cells, one major cluster still expressed endothelial markers and was deemed endothelial-like (an EC). Surprisingly, it did not express *Runx1*. The second major cluster exhibited a hematopoietic gene-expression profile, and other minor clusters showed mixed or mature HC profiles.

Considering that the G2^{med}CD31^{hi}cKit^{hi} fraction contains a high frequency of multipotent HPC and HSC activity, it begs the question whether the EC and/or the other minor cell clusters possess some degree of hematopoietic identity but either are not yet functional or are the predecessors of HPCs and HSCs. Data from several laboratories have described embryonic HCs lacking HSC activity and have shown that these cells can exhibit HSC function after *in vitro* culturing (Ganuza et al., 2017a; Rybtsov et al., 2016) or neonatal transplantation (Yoder et al., 1997), perhaps by stabilization of a genetic program. While some of these cells could represent pre-HSCs, others could be products of EHT that never successfully achieve HSC function. They may be in a "neutral" or "naive" state awaiting establishment of another functional hematopoietic identity, perhaps to a multipotent progenitor that others have shown to maintain steady-state hematopoiesis (Busch et al., 2015; Pei et al., 2017; Rodriguez-Fraticelli et al., 2018; Schoedel et al., 2016; Sun et al., 2014).

Heptad Factors and Gata2 Heterogeneous Expression

We discovered variability in the expression levels of the hematopoietic heptad TFs in the clusters. The requirements for *Gata2*, *Runx1*, and other heptad TFs *Lmo2*, *Fli1/Erg*, and *Tal/Lyl1* in EHT and HSC generation (Chen et al., 2009; de Pater et al., 2013) have been extensively demonstrated in the mouse embryo (Dzierzak and Speck, 2008). The TFs work in a combinatorial manner to activate transcription of downstream hematopoietic effector genes (Pimanda and Göttgens, 2010; Wilson et al., 2010). Whereas the single haploinsufficient state of *Gata2* and *Runx1* affects HC development, the double haploinsufficiency of these TFs has more profound effects (Wilson et al., 2010), indicating the exquisite sensitivity of hematopoietic development to combinatorial TF levels.

The fact that, within the highly enriched IAHC fraction, the EC was negative for *Runx1* expression and UC1 was negative for *Gata2* expression indicates that TF level changes may result in non-functional HCs. However, it is important to note that *Runx1*- and *Gata2*-deleted embryos still produce some functional HCs (Kaimakis et al., 2016; Tsai et al., 1994; Wang et al., 1996), suggesting that there is more than one way to make a HC. The absence of complex definitive hematopoietic activity in such embryos confirms the pivotal role of these TFs in the generation of functional HSCs.

When linked to the recently demonstrated pulsatile expression of Gata2 in cells undergoing EHT (Eich et al., 2018), a dynamic combinatorial TF code may affect hematopoietic fate and function acquisition in individual cells. For example asynchronous fluctuations of levels of several TFs (due to transcriptional pulsing, protein processing, and/or half-life) could result in a variety of different molecular states, as shown in T cells (Kueh et al., 2016). Only when a sweet spot of combinatorial factor levels is reached within a single cell will HSC or HPC identity and function be achieved. This would be consistent with the low frequency of HSCs (1–2 HSCs/~700 cluster cells) and the higher frequency of HPCs (350/~700) found in the *in vivo* E10.5 aorta at the time of embryo isolation. It also could explain why the simple combined transduction and overexpression of TFs in reprogramming experiments does not robustly induce the transition to HSC fate (Dzierzak and Bigas, 2018).

Novel Combinations of Iterative Single-Cell Methods Identify the HSC Transcriptome

Iterative single-cell RNA-seq, index-sorting, functional analyses, and bioinformatics led us to progressively higher enrichments of HSCs. Several putative HSC markers were identified within the HC1 single-cell transcriptomes and included *CD27*, *Selp/P-selectin*, and *Mpl*. Expression level distribution of single cells in Louvain clusters indicated that CD27 and Selp were likely to be most useful. Previously, CD27 was reported to be expressed on bone marrow (Vazquez et al., 2015) and AGM HSCs (Li et al., 2017). Although CD27 is not exclusive to HSCs (and is expressed by some cells in the EC Louvain cluster), we discovered that cells expressing specific medium levels of CD27 within the G2^{med}SSC^{lo}CD31^{hi} cKit^{hi} fraction yielded a further 6.5-fold enrichment on top of the HC1-only enrichment. While multipotent/bipotent HPCs were also enriched in this CD27^{med} fraction, all unipotent HPCs were absent. As *Selp* is most prominently expressed by HC1 cells, it also could offer an opportunity for further enrichment of aortic HSCs (Sullivan et al., 2011). Finally, Mpl, the receptor for thrombopoietin is known to regulate the generation and self-renewal of AGM HSCs (Petit-Cocault et al., 2007). Together, these markers offer additional opportunities for the isolation/identification of single functional HSCs as they are first generated in the embryo.

Stratifications of single-cell transcriptomic data revealed the expression of a combination of other genes expressed by functional HSCs. These include *Ly6a*, *DLL4*, *Sox17*, and *eNOS*, as well as others. Ly6aGFP has been shown to be expressed on all AGM HSCs (de Bruijn et al., 2002; Solaimani Kartalaei et al., 2015); DLL4 is involved in arterial endothelial identity and, most recently, has been shown to govern the dynamics of cluster formation and HSC production (Bonkhofer et al., 2019; Gama-Norton et al., 2015; Porcheri et al., 2020); Sox17 maintains stem cell potency (He et al., 2011); and the deficiency of eNOS affects the development of IAHCs and functional HSCs (North et al., 2009). The fact that the combined expression of these genes occurs in the 8 cells within the 119 sequenced G2^{med}SSC^{lo}CD31^{hi} cKit^{hi}CD27^{+N} single cells (isolated from 6 embryos) has important implications for understanding the regulation of HSC fate/function establishment and provides a useful reference for the *ex vivo* generation and/or growth of HSCs.

Emerging Functional HSCs Are Localized in 1- to 2-Cell Ventral Aortic Clusters

Since HSC fate acquisition is rare, there has been much speculation on the specific localization of HSCs within the IAHCs. Do they take on HSC fate immediately upon emergence from the hemogenic endothelium, or do they emerge as pre-HSCs? Do they emerge as a single cell and remain as a 1-cell cluster, or are they within larger clusters of proliferative cells?

Time-lapse imaging of the zebrafish and mouse embryonic aorta proved that the AGM region is a HC-generative niche (Bertrand et al., 2010; Boisset et al., 2010; Kissa and Herbomel, 2010). Single HCs emerge directly from ventral aortic endothelial cells, and results of genetic studies support that at least some of the cells taking on hematopoietic identity are functional HSCs. Here, by co-expression of CD27 and Ly6aGFP, we show the localization of HSCs to the small 1- to 2-cell aortic clusters. Additionally, our RNA-seq data show DLL4 expression in the single HSC profiles, whereas the other single G2^{med} SSC^{lo}CD31^{hi}cKit^{hi}CD27^{+N} aortic cells are DLL4 negative/low. These results are in agreement with a recent study showing that, in the absence of DLL4, larger and more abundant aortic clusters are established (Porcheri et al., 2020), and, together, they suggest that DLL4 may be maintaining HSC quiescence and/or function.

Finally, we observed single MacGreen^{+} macrophages interacting with single aortic CD27^{+} (Figure 6D) and Ly6a^{+} emerging cells (Mariani et al., 2019). The presence of these macrophages in the AGM enhances HSC and HPC output (Mariani et al., 2019) and further reinforces the notion that intersecting multiple signals are needed for establishing HSC fate identity and function in single emerging cells. In conclusion, the iterative approaches used in our study have enabled the correspondence of single-cell transcriptomic profiles with the first functional HSCs generated in mouse.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mouse models
- METHOD DETAILS
  - Mouse embryo generation
  - Preparation of AGM single cell suspensions for flow cytometry
  - Hematopoietic stem cell transplantation and (single-cell) CFU-C assay
  - Immunohistochemistry on cryosections
  - Single-cell RNA sequencing
  - Bioinformatic Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2020.107627.

ACKNOWLEDGMENTS

The authors thank all lab members for helpful comments, particularly Carmen Rodriguez-Seoane and Sarah J. Kinston for technical assistance and the QMRI Flow Cytometry and Cell Sorting Facility (S. Johnston, W. Ramsay, and M. Pattinson) for cell sorting. These studies were supported by ERC Advanced Grant 341096 (to E.D.), BBSRC research grant BB/S01845X/1 (to E.D.), and Bloodwise research grant 18010 (to E.D.). Work in the B.G. group was supported by Wellcome (206328/Z/17/Z), MRC (MR/M008975/1), CRUK (C1163/A21762), Bloodwise (18002), and NIH NIDDK (R24 DK106766-01A1), as well as by core infrastructure support to the Cambridge Stem Cell Institute by Wellcome and MRC (097922/Z/11/Z).

AUTHOR CONTRIBUTIONS

C.S.V. performed the embryo dissections, cell preparations, flow cytometry, *in vitro/in vivo* hematopoietic assays, and immunofluorescence staining and microscopy. A.M. helped with the collection of cells for RNA-seq and mouse irradiations. F.J.C.-N., W.J., and X.W. performed RNA-seq and bioinformatics analyses. S.A.M. performed dissection, cell preparation, and isolation for CD27 RNA-seq. C.S.V., B.G., and E.D. designed the study and interpreted data, and C.S.V. and E.D. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 27, 2020
Revised: March 18, 2020
Accepted: April 18, 2020
Published: May 12, 2020

REFERENCES

Abranches, E., Guedes, A.M., Moravec, M., Maamar, H., Svoboda, P., Raj, A., and Henrique, D. (2014). Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. Development *141*, 2770–2779.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

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.

Becker, James, T., Huff, Jacob, Nebes, Robert, D., Holland, Audrey, Boller, Francois, et al. (1988). Neuropsychological Function in Alzheimer's Disease: Pattern of Impairment and Rates of Progression. Arch Neurol. *45* (*3*), 263–268.

Bee, T., Liddiard, K., Swiers, G., Bickley, S.R., Vink, C.S., Jarratt, A., Hughes, J.R., Medvinsky, A., and de Bruijn, M.F. (2009). Alternative Runx1 promoter usage in mouse developmental hematopoiesis. Blood Cells Mol. Dis. *43*, 35–42.

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

Bigas, A., and Porcheri, C. (2018). Notch and Stem Cells. Adv. Exp. Med. Biol. *1066*, 235–263.

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.

Bonkhofer, F., Rispoli, R., Pinheiro, P., Krecsmarik, M., Schneider-Swales, J., Tsang, I.H.C., de Bruijn, M., Monteiro, R., Peterkin, T., and Patient, R. (2019).

Blood stem cell-forming haemogenic endothelium in zebrafish derives from arterial endothelium. Nat. Commun. *10*, 3577.

Brennecke, P., Anders, S., Kim, J.K., Kołodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., and Heisler, M.G. (2013). Accounting for technical noise in single-cell RNA-seq experiments. Nat. Methods *10*, 1093–1095.

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.

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature *453*, 544–547.

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.

Ciau-Uitz, A., and Patient, R. (2019). Gene Regulatory Networks Governing the Generation and Regeneration of Blood. J. Comput. Biol. *26*, 719–725.

Cohen-Saidon, C., Cohen, A.A., Sigal, A., Liron, Y., and Alon, U. (2009). Dynamics and variability of ERK2 response to EGF in individual living cells. Mol. Cell *36*, 885–893.

de Bruijn, M.F., 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.

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

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

Eich, C., Arlt, J., Vink, C.S., Solaimani Kartalaei, P., Kaimakis, P., Mariani, S.A., van der Linden, R., van Cappellen, W.A., and Dzierzak, E. (2018). In vivo single cell analysis reveals Gata2 dynamics in cells transitioning to hematopoietic fate. J. Exp. Med. *215*, 233–248.

Gama-Norton, L., Ferrando, E., Ruiz-Herguido, C., Liu, Z., Guiu, J., Islam, A.B., 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.

Ganuza, M., Hadland, B., Chabot, A., Li, C., Kang, G., Bernstein, I., and McKinney-Freeman, S. (2017a). Murine hemogenic endothelial precursors display heterogeneous hematopoietic potential ex vivo. Exp. Hematol. *51*, 25–35.e6.

Ganuza, M., Hall, T., Finkelstein, D., Chabot, A., Kang, G., and McKinney-Freeman, S. (2017b). Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny. Nat. Cell Biol. *19*, 1153–1163.

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

Hamey, F.K., and Göttgens, B. (2019). Machine learning predicts putative hematopoietic stem cells within large single-cell transcriptomics data sets. Exp. Hematol. *78*, 11–20.

He, S., Kim, I., Lim, M.S., and Morrison, S.J. (2011). Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. Genes Dev. *25*, 1613–1627.

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. Development *136*, 3853–3862.

Ivanovs, A., Rybtsov, S., Ng, E.S., Stanley, E.G., Elefanty, A.G., and Medvinsky, A. (2017). Human haematopoietic stem cell development: from the embryo to the dish. Development *144*, 2323–2337.

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

Kaimakis, P., de Pater, E., Eich, C., Solaimani Kartalaei, P., Kauts, M.L., Vink, C.S., van der Linden, R., Jaegle, M., Yokomizo, T., Meijer, D., and Dzierzak, E. (2016). Functional and molecular characterization of mouse Gata2-independent hematopoietic progenitors. Blood *127*, 1426–1437.

Kauts, M.L., De Leo, B., Rodríguez-Seoane, C., Ronn, R., Glykofrydis, F., Maglitto, A., Kaimakis, P., Basi, M., Taylor, H., Forrester, L., et al. (2018). Rapid Mast Cell Generation from Gata2 Reporter Pluripotent Stem Cells. Stem Cell Reports *11*, 1009–1020.

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

Kueh, H.Y., Yui, M.A., Ng, K.K., Pease, S.S., Zhang, J.A., Damle, S.S., Freedman, G., Siu, S., Bernstein, I.D., Elowitz, M.B., and Rothenberg, E.V. (2016). Asynchronous combinatorial action of four regulatory factors activates Bcl11b for T cell commitment. Nat. Immunol. *17*, 956–965.

Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., Zuyev, S., Ansell, J., and Medvinsky, A. (2002). Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. Development *129*, 4891–4899.

Lahav, G. (2004). The strength of indecisiveness: oscillatory behavior for better cell fate determination. Sci. STKE *2004*, pe55.

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.

Li, Y., Gao, L., Hadland, B., Tan, K., and Speck, N.A. (2017). CD27 marks murine embryonic hematopoietic stem cells and type II prehematopoietic stem cells. Blood *130*, 372–376.

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

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.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Mariani, S.A., Li, Z., Rice, S., Krieg, C., Fragkogianni, S., Robinson, M., Vink, C.S., Pollard, J.W., and Dzierzak, E. (2019). Pro-inflammatory Aorta-Associated Macrophages Are Involved in Embryonic Development of Hematopoietic Stem Cells. Immunity *50*, 1439–1452.e5.

Mascarenhas, M.I., Parker, A., Dzierzak, E., and Ottersbach, K. (2009). Identification of novel regulators of hematopoietic stem cell development through refinement of stem cell localization and expression profiling. Blood *114*, 4645–4653.

McKinney-Freeman, S., Cahan, P., Li, H., Lacadie, S.A., Huang, H.T., Curran, M., Loewer, S., Naveiras, O., Kathrein, K.L., Konantz, M., et al. (2012). The transcriptional landscape of hematopoietic stem cell ontogeny. Cell Stem Cell *11*, 701–714.

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

Medvinsky, A., Taoudi, S., Mendes, S., and Dzierzak, E. (2008). Analysis and manipulation of hematopoietic progenitor and stem cells from murine embryonic tissues. Curr. Protoc. Stem Cell Biol. Chapter 2, Unit 2A.6.

Mihara, M., Hashizume, M., Yoshida, H., Suzuki, M., and Shiina, M. (2012). IL-6/IL-6 receptor system and its role in physiological and pathological conditions. Clin. Sci. (Lond.) *122*, 143–159.

Miyanari, Y., and Torres-Padilla, M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470–473.

Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A.J., Tanaka, Y., Wilkinson, A.C., Buettner, F., Macaulay, I.C., Jawaid, W., Diamanti, E., et al. (2015). Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nat. Biotechnol. *33*, 269–276.

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

North, T.E., de Bruijn, M.F., 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.

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.

Olsson, A., Venkatasubramanian, M., Chaudhri, V.K., Aronow, B.J., Salomonis, N., Singh, H., and Grimes, H.L. (2016). Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. Nature *537*, 698–702.

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

Pei, W., Feyerabend, T.B., Rössler, J., Wang, X., Postrach, D., Busch, K., Rode, I., Klapproth, K., Dietlein, N., Quedenau, C., et al. (2017). Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. Nature *548*, 456–460.

Petit-Cocault, L., Volle-Challier, C., Fleury, M., Péault, B., and Souyri, M. (2007). Dual role of Mpl receptor during the establishment of definitive hematopoiesis. Development *134*, 3031–3040.

Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.

Pimanda, J.E., and Göttgens, B. (2010). Gene regulatory networks governing haematopoietic stem cell development and identity. Int. J. Dev. Biol. *54*, 1201–1211.

Porcheri, C., Golan, O., Calero-Nieto, F.J., Thambyrajah, R., Ruiz-Herguido, C., Wang, X., Catto, F., Guillén, Y., Sinha, R., González, J., et al. (2020). Notch ligand Dll4 impairs cell recruitment to aortic clusters and limits blood stem cell generation. EMBO J. *39*, e104270.

Purvis, J.E., and Lahav, G. (2013). Encoding and decoding cellular information through signaling dynamics. Cell *152*, 945–956.

Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M., Sun, J., Calogero, R.A., Klein, A.M., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. Nature *553*, 212–216.

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

Ryu, H., Chung, M., Dobrzyński, M., Fey, D., Blum, Y., Sik Lee, S., Peter, M., Kholodenko, B.N., Li Jeon, N., and Pertz, O. (2016). Frequency modulation of ERK activation dynamics rewires cell fate. Mol. Syst. Biol. *12*, 866.

Sasmono, R.T., Oceandy, D., Pollard, J.W., Tong, W., Pavli, P., Wainwright, B.J., Ostrowski, M.C., Himes, S.R., and Hume, D.A. (2003). A macrophage colony-stimulating factor receptor-green fluorescent protein transgene is expressed throughout the mononuclear phagocyte system of the mouse. Blood *101*, 1155–1163.

Schoedel, K.B., Morcos, M.N.F., Zerjatke, T., Roeder, I., Grinenko, T., Voehringer, D., Gothert, J.R., Waskow, C., Roers, A., and Gerbaulet, A. (2016). The bulk of the hematopoietic stem cell population is dispensable for murine steady-state and stress hematopoiesis. Blood *128*, 2285–2296.

Solaimani Kartalaei, P., Yamada-Inagawa, T., Vink, C.S., de Pater, E., van der Linden, R., Marks-Bluth, J., van der Sloot, A., van den Hout, M., Yokomizo, T., van Schaick-Solernó, M.L., et al. (2015). Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. J. Exp. Med. *212*, 93–106.

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

Sullivan, C., Chen, Y., Shan, Y., Hu, Y., Peng, C., Zhang, H., Kong, L., and Li, S. (2011). Functional ramifications for the loss of P-selectin expression on hematopoietic and leukemic stem cells. PLoS ONE *6*, e26246.

Sun, J., Ramos, A., Chapman, B., Johnnidis, J.B., Le, L., Ho, Y.J., Klein, A., Hofmann, O., and Camargo, F.D. (2014). Clonal dynamics of native haematopoiesis. Nature *514*, 322–327.

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.

Taoudi, S., and Medvinsky, A. (2007). Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. Proc. Natl. Acad. Sci. USA *104*, 9399–9403.

Tipping, A.J., Pina, C., Castor, A., Hong, D., Rodrigues, N.P., Lazzari, L., May, G.E., Jacobsen, S.E., and Enver, T. (2009). High GATA-2 expression inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. Blood *113*, 2661–2672.

Tsai, F.Y., and Orkin, S.H. (1997). Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. Blood *89*, 3636–3643.

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.

Vazquez, S.E., Inlay, M.A., and Serwold, T. (2015). CD201 and CD27 identify hematopoietic stem and progenitor cells across multiple murine strains independently of Kit and Sca-1. Exp. Hematol. *43*, 578–585.

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. Proc. Natl. Acad. Sci. USA *93*, 3444–3449.

Weinreb, C., Wolock, S., and Klein, A.M. (2018). SPRING: a kinetic interface for visualizing high dimensional single-cell expression data. Bioinformatics *34*, 1246–1248.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell *7*, 532–544.

Wu, T.D., Reeder, J., Lawrence, M., Becker, G., and Brauer, M.J. (2016). GMAP and GSNAP for Genomic Sequence Alignment: Enhancements to Speed, Accuracy, and Functionality. Methods Mol. Biol. *1418*, 283–334.

Yoder, M.C., Hiatt, K., and Mukherjee, P. (1997). In vivo repopulating hematopoietic stem cells are present in the murine yolk sac at day 9.0 postcoitus. Proc. Natl. Acad. Sci. USA *94*, 6776–6780.

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

Zhou, F., Li, X., Wang, W., Zhu, P., Zhou, J., He, W., Ding, M., Xiong, F., Zheng, X., Li, Z., et al. (2016). Tracing haematopoietic stem cell formation at single-cell resolution. Nature *533*, 487–492.

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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat anti-Mouse CD31-BV605 (clone 390)	BioLegend	Cat# 102427; RRID:AB_2563982
Rat anti-Mouse CD117-BV421 (c-Kit) (clone 2B8)	BD Biosciences	Cat#562609; RRID:AB_11154585
Armenian hamster anti-Mouse CD27-APC (clone LG.7F9)	eBioscience/Thermo Fisher Scientific	Cat# 17-0271-82; RRID: AB_469370
Rat anti-Mouse CD34-biotin (clone RAM34)	eBioscience/Thermo Fisher Scientific	Cat# 13-0341-82; RRID: AB_466425
Rabbit anti-GFP	MBL International	Cat# 598; RRID: AB_591819
Armenian hamster anti-Mouse CD27 (clone LG.7F9)	eBioscience/Thermo Fisher Scientific	Cat# 14-0271-82; RRID: AB_467183
Streptavidin-Alexa Fluor 555	Invitrogen/Thermo Fisher Scientific	Cat# s21381; RRID: AB_2307336
Goat anti-Rabbit IgG (H+L)-Alexa Fluor 488	Invitrogen/Thermo Fisher Scientific	Cat# A-11008; RRID: AB_143165
Goat anti-Armenian hamster IgG-Alexa Fluor 647 (clone Poly4055)	BioLegend	Cat# 405510, RRID:AB_2566695
APC anti-Mouse CD45.1 (clone A20)	BD Biosciences	Cat#558701; RRID:AB_1645214
PE anti-Mouse CD45.2 (clone 104)	BD Biosciences	Cat#560695; RRID:AB_1727493
Chemicals, Peptides, and Recombinant Proteins		
Collagenase Type I from *Clostridium histolyticum*	Sigma-Aldrich	Cat#C0130; Lot#02M1425V
Foetal Bovine Serum Standard Quality	Brunschwig Chemie BV	Cat#BS01500; Lot#212-223650
10% Triton X-100	Sigma-Aldrich	Cat#93443-100 ml
SUPERase-In RNase Inhibitor	Ambion	Cat#AM2694
dNTP mix	ThermoFisher	Cat#10319879
ERCC RNA Spike-In Mix	ThermoFisher	Cat#4456740
Superscript II Reverse Transcriptase	ThermoFisher	Cat#18064-014
Betaine	Sigma-Alrich	Cat#B0300
Magnesium chloride	ThermoFisher	Cat#AM9530G
KAPA HiFi Hotstart ReadyMix	KAPA Biosystems	Cat#KK2601
Agencourt AMPure XP beads	Beckman Coulter	Cat#A63881
EB solution	QIAGEN	Cat#19086
Critical Commercial Assays		
High-sensitivity DNA chip	Agilent	Cat#5067-4626
Quant-iT Picogreen double stranded DNA assay kit	ThermoFisher	Cat#P7589
Nextera XT DNA sample preparation kit 96 samples	Illumina	Cat#FC-131-1096
Nextera XT 96-Index kit, 384 samples	Illumina	Cat#FC-131-1002
KAPA qPCR quantification kit	KAPA Biosystems	Cat# KK4824
Deposited Data		
Single-cell RNaseq	This paper	Accession number: GSE143637
Experimental Models: Organisms/Strains		
Mouse: Gata2Venus	Laboratory of Elaine Dzierzak; Kaimakis et al., 2016, Blood	N/A
Mouse: B6.Cg-Tg(Ly6a-EGFP)G5Dzk/J	The Jackson Laboratory	https://www.jax.org/strain/012643
Mouse: B6.Cg-Tg(Csf1r-EGFP)1Hume/J (also known as MacGreen)	The Jackson Laboratory	https://www.jax.org/strain/018549

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyCrl (also known as Ly5.1)	Charles River	Strain code: 494 https://www.criver.com/products-services/find-model/ly51-mouse?region=3671
Mouse: C57BL/6JOlaHsd	Envigo	Order code: 057 https://www.envigo.com/products-services/research-models-services/models/research-models/mice/inbred/c57bl-6-inbred-mice/c57bl-6jolahsd/u.k.aspx
Oligonucleotides		
TSO: AAGCAGTGGTATCAACG CAGAGTACATrGrG+G-3′	Picelli et al., 2014	N/A
Oligo-dT30VN: AAGCAGTGGTATCA ACGCAGAGTAC(T30)VN-3′	Picelli et al., 2014	N/A
ISPCR: AAGCAGTGGTATCAACG CAGAGT	Picelli et al., 2014	N/A
Software and Algorithms		
Prism Version 7	GraphPad	http://www.graphpad.com/scientific-software/prism
BD Facs Diva version 8	BD Biosciences	http://www.bdbiosciences.com/en-us
FlowJo Version 10.1r5	Tree Star, Inc	https://www.flowjo.com/solutions/flowjo/downloads/previous-versions
GSNAP	Wu et al., 2016	http://research-pub.gene.com/gmap/
HTSeq	Anders et al., 2015	https://htseq.readthedocs.io/en/release_0.11.1/install.html
PCA	Becker et al., 1988	prcomp in R stats package
Selection of Highly Variable Genes	Brennecke et al., 2013	https://static-content.springer.com/esm/art%3A10.1038%2Fnmeth.2645/MediaObjects/41592_2013_BFnmeth2645_MOESM313_ESM.pdf
ICGS	Olsson et al., 2016	https://altanalyze.readthedocs.io/en/latest/ICGS/
DESeq2	Love et al., 2014	http://bioconductor.org/packages/release/bioc/html/DESeq2.html
SPRING	Weinreb et al., 2018	https://github.com/AllonKleinLab/SPRING
hscScore	Hamey and Göttgens, 2019	https://github.com/fionahamey/hscScore
Other		
Interactive website	This paper	https://gottgens-lab.stemcells.cam.ac.uk/DZIERZAK/
Single-cell RNaseq	Zhou et al., 2016	Accession number: GSE67120

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Elaine Dzierzak (Elaine.Dzierzak@ed.ac.uk).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The single-cell RNA sequencing data generated in this study have been deposited in GEO (https://www.ncbi.nlm.nih.gov/geo/) with accession number: GSE143637. The code used is available on request. An interactive website containing the single-cell RNASeq results is available at https://gottgens-lab.stemcells.cam.ac.uk/DZIERZAK/.

Our datasets were examined in the context of a published single cell dataset (Zhou et al., 2016) which can be found in GEO under accession number: GSE67120.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse models

All mice were housed and bred in animal facilities at the University of Edinburgh, in compliance with UK Home Office regulations. All animal experiments were conducted according to a UK Home Office Project License (PPL 70/8076 and continuation PC6D479E0) and approved by the University of Edinburgh Ethical Review Committee.

Adult C57BL/6JOlaHsd (Ly5.2) wild-type females were obtained from Envigo (order code: 057) and used at 2-4 months old. *Gata2*$^{V/V}$ (Kaimakis et al., 2016), *Ly6aGFP*$^{+/-}$ (de Bruijn et al., 2002) and *Csf1rGFP*$^{+/-}$ (MacGreen, (Sasmono et al., 2003) males (> 10 backcrosses onto C57BL/6JOlaHsd) were used at 2-12 months old.

All transplantations were with 2-4 months old B6.SJL-*Ptprc*^{a}*Pepc*^{b}/BoyCrl (Ly5.1/5.1; Charles River; strain code: 494) recipients (males and females mixed).

METHOD DETAILS

Mouse embryo generation

Gata2$^{V/+}$, *Ly6aGFP*$^{+/-}$ and *Csf1rGFP*$^{+/-}$ embryos were generated by timed matings of C57BL/6JOlaHsd wild-type females with *Gata2*$^{V/V}$, *Ly6aGFP*$^{+/-}$ and *Csf1rGFP*$^{+/-}$ males respectively. Genotyping was performed by Venus or GFP fluorescence screening of the placenta or embryonic tail. Day of plug discovery was embryonic day (E)0 and staging was by somite pair (sp) counts (E10 = 30-39sp; E11 = 40-49sp).

Preparation of AGM single cell suspensions for flow cytometry

AGM (aorta-gonad-mesonephros) regions (including trunk portion of the vitelline and umbilical arteries) of E10/11 embryos were dissected as previously described (Medvinsky et al., 2008). Single cell suspensions were prepared by enzymatically digesting tissues at 37°C in 0.125% type 1 collagenase (C0130, Sigma-Aldrich) for 45 min, after which tissues were further mechanically disrupted with a P1000 pipette. Cells were subsequently passed over a 0.35 μm strainer, collagenase washed away with PBS+10%FBS+1%P/S (PFP), spun down (234*g*) and resuspended in PFP for flow cytometric antibody staining. Cells were incubated at 4°C for 30 min in PFP with directly conjugated antibodies against CD31 (CD31-BV605, 1:100, cat# 102427, Biolegend), cKit (cKit-BV421, 1:200, cat.# 562609, BD) and CD27 (CD27-APC, 1:75, cat.# 17-0271-82. eBioscience). Excess of antibodies was washed away with PFP before fluorescence activated cell sorting (FACS) on an AriaII (BD). Dead cell exclusion was by Hoechst 33342 and gates were set based on unstained WT and fluorescence-minus-one (FMO) controls. Sorted single cells were collected in 50% FBS/PBS for functional analyses (transplantation + bulk CFU-C assay). Both collagenase and FCS were batch tested for optimal AGM conditions.

Hematopoietic stem cell transplantation and (single-cell) CFU-C assay

For transplantation, sorted cell populations (Ly5.2/CD45.2 background) were resuspended in PBS together with 2x10^{5} Ly5.1 spleen carrier cells and intravenously injected into irradiated Ly5.1 (CD45.1) recipients that had received two split-doses of 4.5 Gy (Cs-137 ɣ-irradiator). Peripheral blood was sampled for CD45.1- and CD45.2 flow cytometric analysis at 4 and 16 weeks post-transplantation. Mice were considered repopulated with long-term reconstituting cells at ≥ 5% donor chimerism. For population/bulk CFU-C assays, sorted cells were resuspended and cultured in Methylcellullose medium (M3434; Stem Cell Technologies) in a humidified chamber at 37°C and 5% CO_2 for 9-11 days after which colonies were counted. For single-cell CFU-C assays, cells were sorted directly into 96-well plates with one (P1000) drop of Methylcellulose medium deposited per well and cultured and counted under the same conditions as for population CFU-C assays.

Immunohistochemistry on cryosections

Gata2$^{V/+}$ and *Ly6aGFP*$^{+/-}$ and MacGreen$^{+/-}$ embryos were staged and fixed in 2% PFA/PBS for 1 hr and quickly washed in PBS. Embryos were then equilibrated overnight (O/N) in 20% sucrose/PBS at 4°C, after which they were washed in Tissue-Tek O.C.T. embedding matrix (Sakura). Tissue paper was used to remove remaining sucrose and each embryo was placed in an individual mold filled with embedding matrix. Forelimbs were orientated exactly above hind limbs before being frozen on dry ice in absolute ethanol. 10 μm transversal cryosections were cut through the AGM region and collected on Superfrost Plus slides (Thermo Fisher Scientific). For immunohistochemistry, sections were fixed for 10 mins with acetone, washed three times with 0.05% Tween20/PBS (PBST) and then blocked with a 1:1 mixture of Streptavidin solution (Vector Laboratories Inc) and PBSblock (0.05% Tween20, 1% bovine serum albumin/PBS) for 15 mins. After three PBST washes, slides were incubated with Biotin solution (Vector Laboratories Inc) for 15 mins and washed another three times in PBST. Sections were incubated with rat anti-mouse-CD34-biotin (1:100, 13-0341-82, eBioscience), rabbit anti-GFP (1:1000, 598, MBL) and Armenian hamster anti-mouse-CD27 (1:100, 14-0271-82, eBioscience) diluted in PBSblock for 1hr at RT. Unbound primary antibodies were washed away with PBST after which sections were incubated with Streptavidin-AlexaFluor555 (1:500, S21381, Invitrogen), anti-rabbit AlexaFluor488 (1:500, A11008, Invitrogen) and anti-Armenian hamster AlexaFluor647 (1:500, 405510, Biolegend) diluted in PBSblock for 30 mins at RT. Slides were washed three times in PBST and incubated in DAPI (1:1000) diluted in PBSblock for 10 mins at RT. Afterward, slides were washed once in PBST and

then mounted with SlowFade Diamond Anti-Fade Mountant (Invitrogen, Thermo Fisher Scientific). Sections were imaged with a Leica SP8 confocal microscope and images and videos generated with the LAS X software (Leica).

Single-cell RNA sequencing

Cells were sorted in 2.3 μl of lysis buffer containing 0.2% Triton X-100 (Sigma-Aldrich) and 1U of Superase-In RNase Inhibitor (Ambion) and processed following the Smart-Seq2 protocol (Picelli et al., 2014) Following preamplification, all samples were purified using Ampure XP beads (Beckman Coulter) at a ratio of 1:0.6. The cDNA was then quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher) and size distributions were checked on high-sensitivity DNA chips (Agilent Bioanalyzer). Samples were used to construct Nextera XT libraries (Illumina) from 125 pg of preamplified cDNA. Libraries were purified and size selected (0.5X-0.7X) using Ampure XP beads. Then, libraries were quantified using KAPA qPCR quantification kit (KAPA Biosystems), pooled and sequenced in an Illumina HiSeq 4000 instrument.

Bioinformatic Analysis

Reads were mapped to a custom reference index that included the *Mus musculus* genome (EMSEMBL GRCm38.p4 Release 81) and the sequences of ERCC spike-ins (Life Technologies), IRES and *Venus* using GSNAP (version 2014-10-07) (Wu et al., 2016) with parameters: -A sam –B 5 –t 24 –n 1 –Q –N 1. HTseq-count was used to count reads mapped to each gene with parameters: –s no (Anders et al., 2015). For further analyses, we only retained samples that had: (1) more than 100,000 reads mapped to nuclear mRNAs; (2) a proportion of reads mapped to genes (nuclear + mitochondrial) relative to total reads above 30%; and (3) less than 20% of mapped reads allocated to mitochondrial genes. Overall, 47 cells (4.1% of total) failed the QC under those criteria and 1087 cells were kept for further analysis distributed as follows: 87 E10 $CD31^{+}cKit^{hi}G2^{int}$ cells, 409 E11 $CD31^{+}cKit^{hi}G2^{int}$ cells, 472 E11 $CD31^{+}cKit^{hi}G2^{lo}$ cells and 119 E11 $CD31^{hi}SSC^{lo}cKit^{hi}G2^{med}CD27^{+N}$ cells. The mean of nuclear reads mapped per cell was 1,213,237. Data were normalized for sequencing depth and RNA quantity using size factors calculated on endogenous genes (Brennecke et al., 2013). Highly variable genes were identified as previously described (Brennecke et al., 2013), using a false discovery rate threshold equal to 0.1. Only highly variable genes were considered to perform PCA analysis, using the *prcomp* function in R. SPRING visualization (Weinreb et al., 2018) and was obtained using the top 90 principal components and the 4 nearest neighbors. Clusters were defined using the *cluster_louvain* function in *R igraph* package. Violinplots were obtained using the *ggplot2* package in R. hscScores were calculated as previously described (Hamey and Göttgens, 2019). Differential expression analysis was performed using DESeq2 (Love et al., 2014).

$CD27^{+}$ cells (Protein level) was extracted from HSC clusters (clusters L2.7 and L2.8). In total, there are 85 cells, with 4445 highly variable genes selected. Highly variable genes were used for ICGS clustering (AltAnalyzer (Olsson et al., 2016)) with 'stringent' with cell cycle and all other parameters as default.

For integration with previously published dataset (Zhou et al., 2016), the data was downloaded from GEO repository (GSE67120) and processed as indicated above. Differential expression analysis was performed between either endothelial cells or pre-HSCs (including T1 and T2) or HSCs (including E12 and E14) and the remaining cells to obtain cell type specific genes using DESeq2. Only genes with log2 fold change > 2, adjusted p value (BH method) < 0.1 and base mean > 50 were used for further analysis. Only the top 30 upregulated genes were selected to use in the heatmap.

QUANTIFICATION AND STATISTICAL ANALYSIS

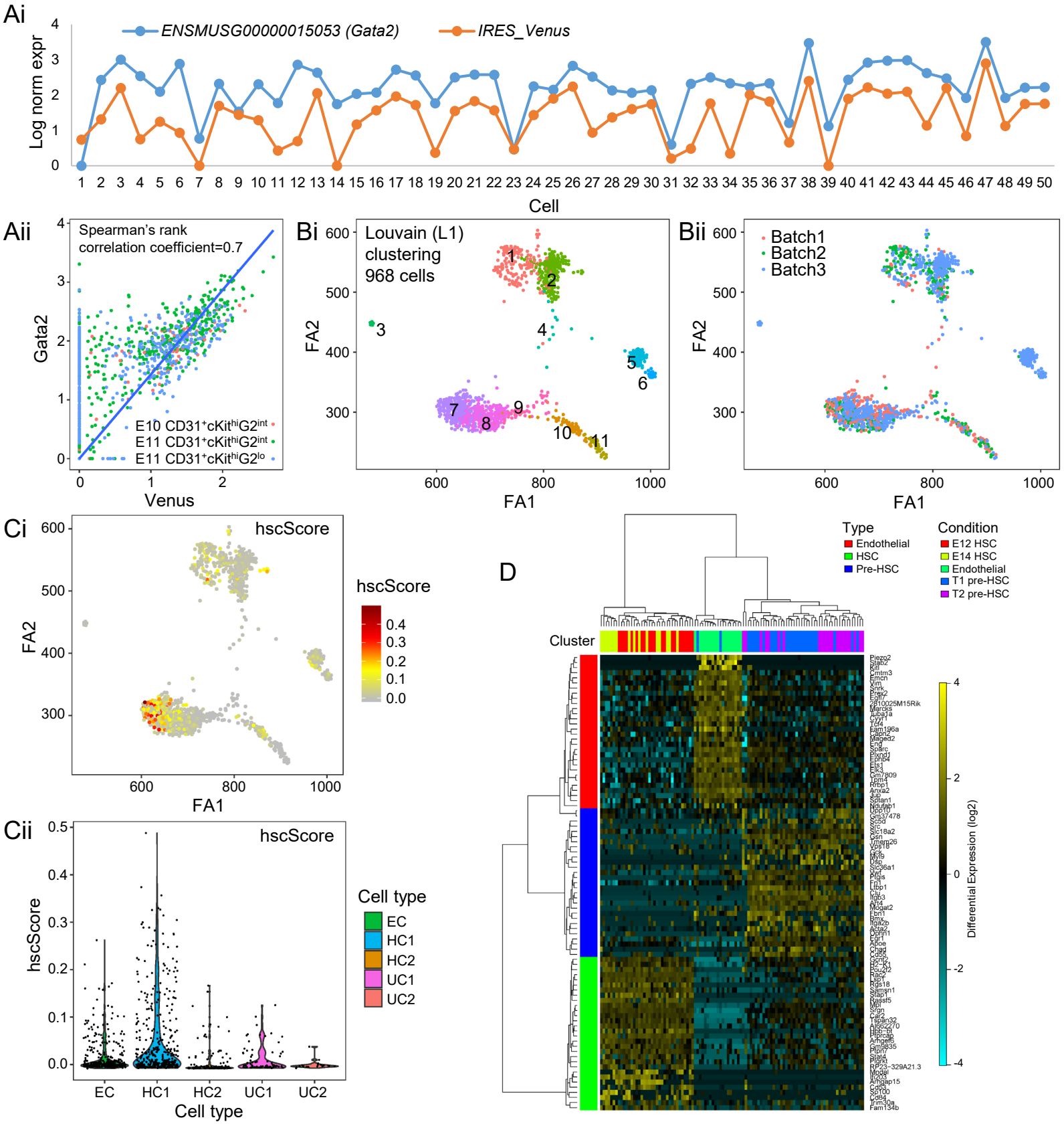
All graphs were generated using GraphPad Prism 7. We performed Bartlett's test for homogeneity of variances to ensure variances for all samples were equal. With the assumption of normality, the data were compared for the same means using one-way ANOVA, followed by a pairwise comparison using a Student's t test, and Bonferroni correction. Significance was determined as $p < 0.05$. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. FACS percentages, averages (mean; red horizontal lines in Figures 1C, 1D, 3E, and 4E) and error bars are ± standard error of the mean (SEM). Further statistical details of experiments can be found in the figure legends. The number of biological replicates is indicated with 'n'.

Cell Reports, Volume *31*

Supplemental Information

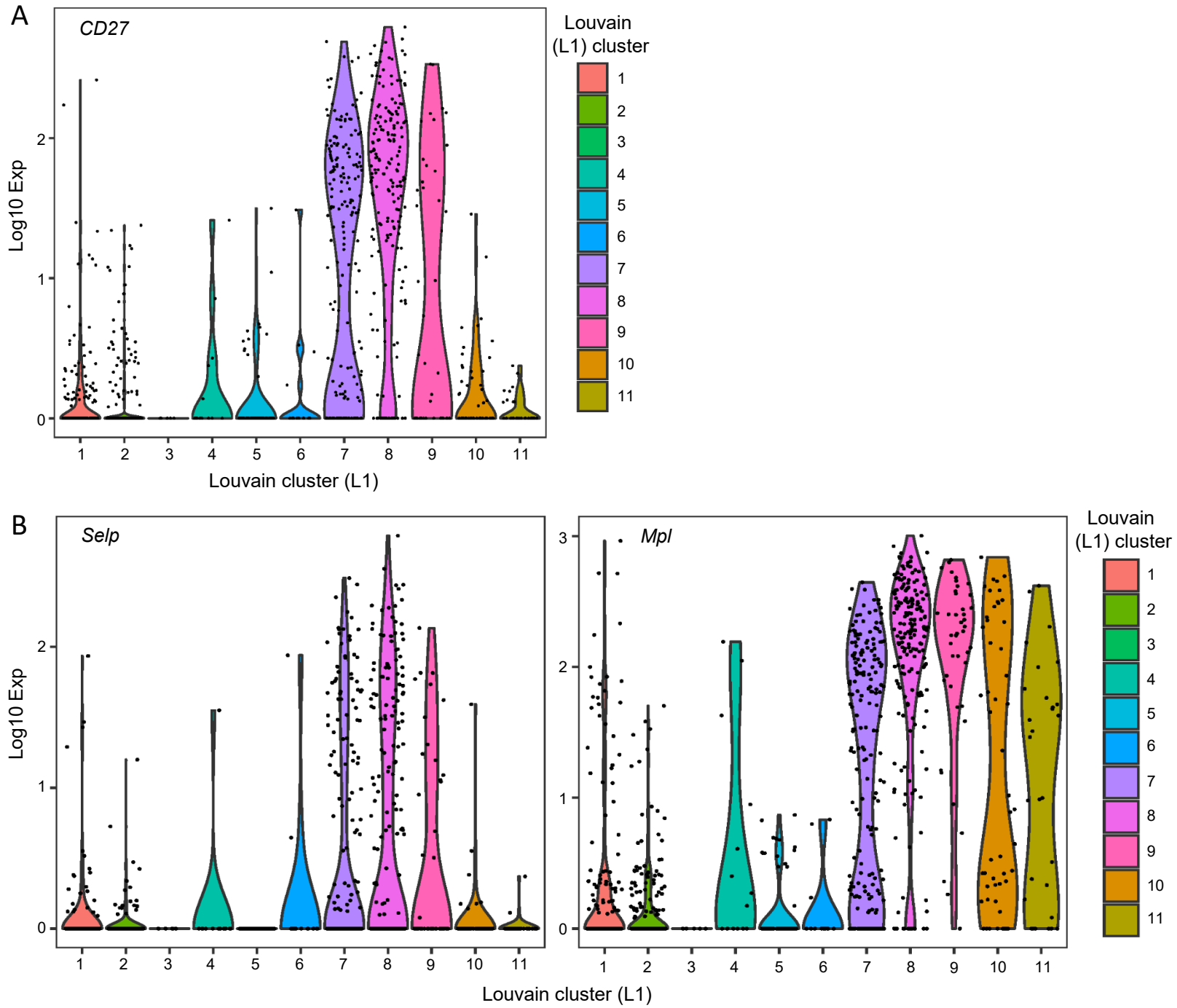
Iterative Single-Cell Analyses Define the Transcriptome of the First Functional Hematopoietic Stem Cells

Chris Sebastiaan Vink, Fernando Jose Calero-Nieto, Xiaonan Wang, Antonio Maglitto, Samanta Antonella Mariani, Wajid Jawaid, Berthold Göttgens, and Elaine Dzierzak

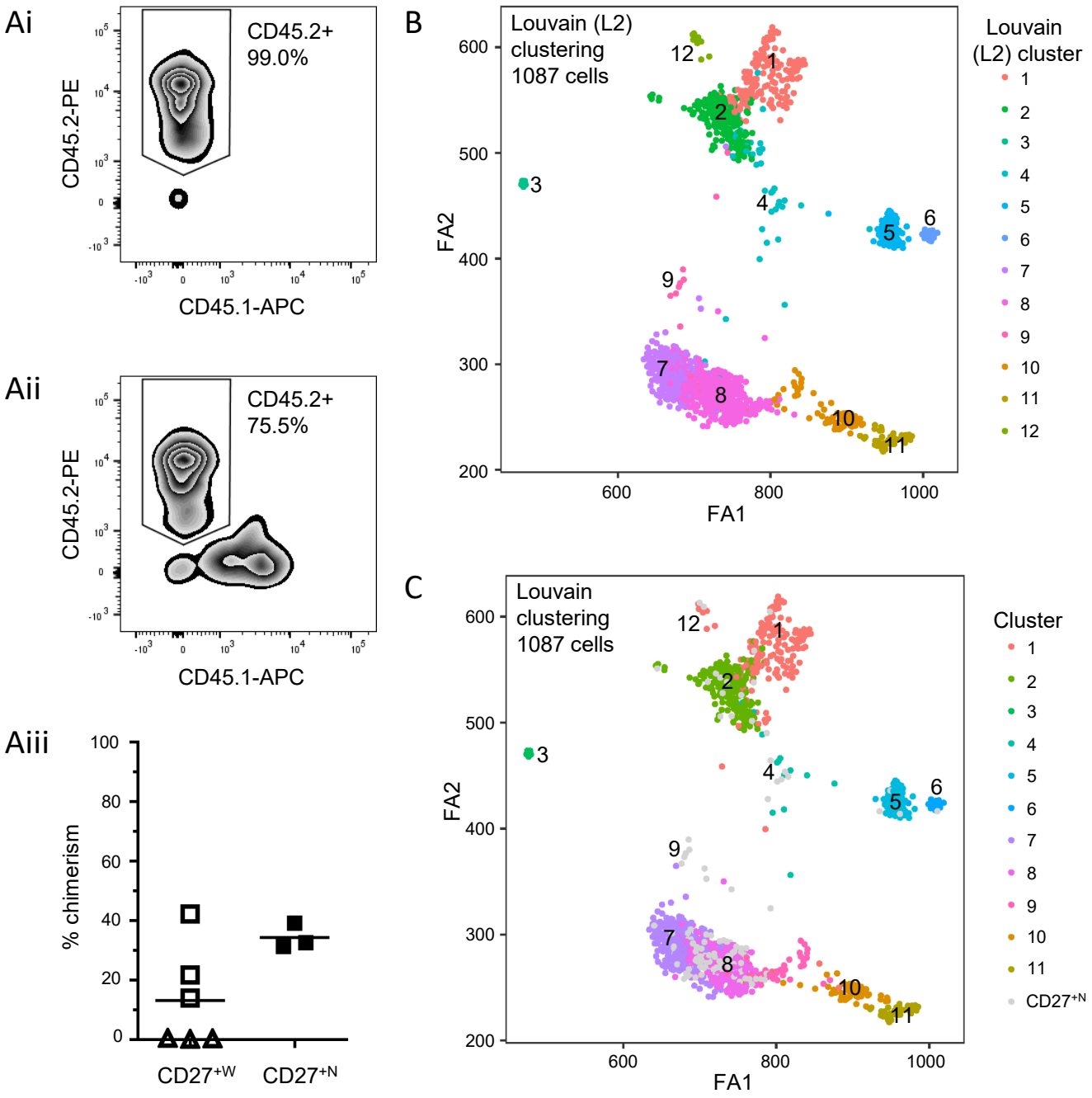


Supplementary Figure 1 – Related to Figure 2. Single cell RNA sequencing analysis of highly enriched (index-sorted) multipotent progenitors and HSCs from E10/E11 IAHCs.

Ai) Graph of log-normalised expression for *Gata2* (ENSMUSG00000015053) and *IRES_Venus* in 50 randomly picked single-cell sorted CD31^{+}cKit^{hi}G2^{med} cells. Vertically aligned dots represent the same cell for *Gata2* (blue) and *Venus* (orange) expression. **Aii**) Correlation between *Gata2* and *IRES_Venus* for all 968 single-cell sorted CD31^{+}cKit^{hi}G2^{med} cells that passed QC. *Gata2* and *IRES_Venus* follow a similar trend and have a correlation coefficient of +0.70. **B**) Force-directed graph (SPRING) visualisation using highly variable genes, the top 90 principal components and the 4 nearest neighbours for all 968 CD31^{+}cKit^{hi}G2^{med} cells, coloured by **Bi**) batch, **Bii**) Louvain cluster (L1.1-11) and **Ci**) hscScore (Hamey and Gottgens, 2019) **Cii**) violin plot showing hscScore obtained for merged Louvain clusters. EC=endothelial cluster; HC1=hematopoietic cluster 1; HC2=hematopoietic cluster 2; UC1=unknown cluster 1; UC2=unknown cluster 2. **D**) Specific genes were obtained for endothelial, pre-HSCs (including both T1 and T2) and fetal liver HSCs (including E12 and E14) of a previously published dataset (Zhou et al., 2016) using DESeq2. These specific genes (depicted on the left in red, blue and green for endothelial, pre-HSC and HSC, respectively) were then used to perform hierarchical clustering of the cells in the Zhou et al. published dataset.

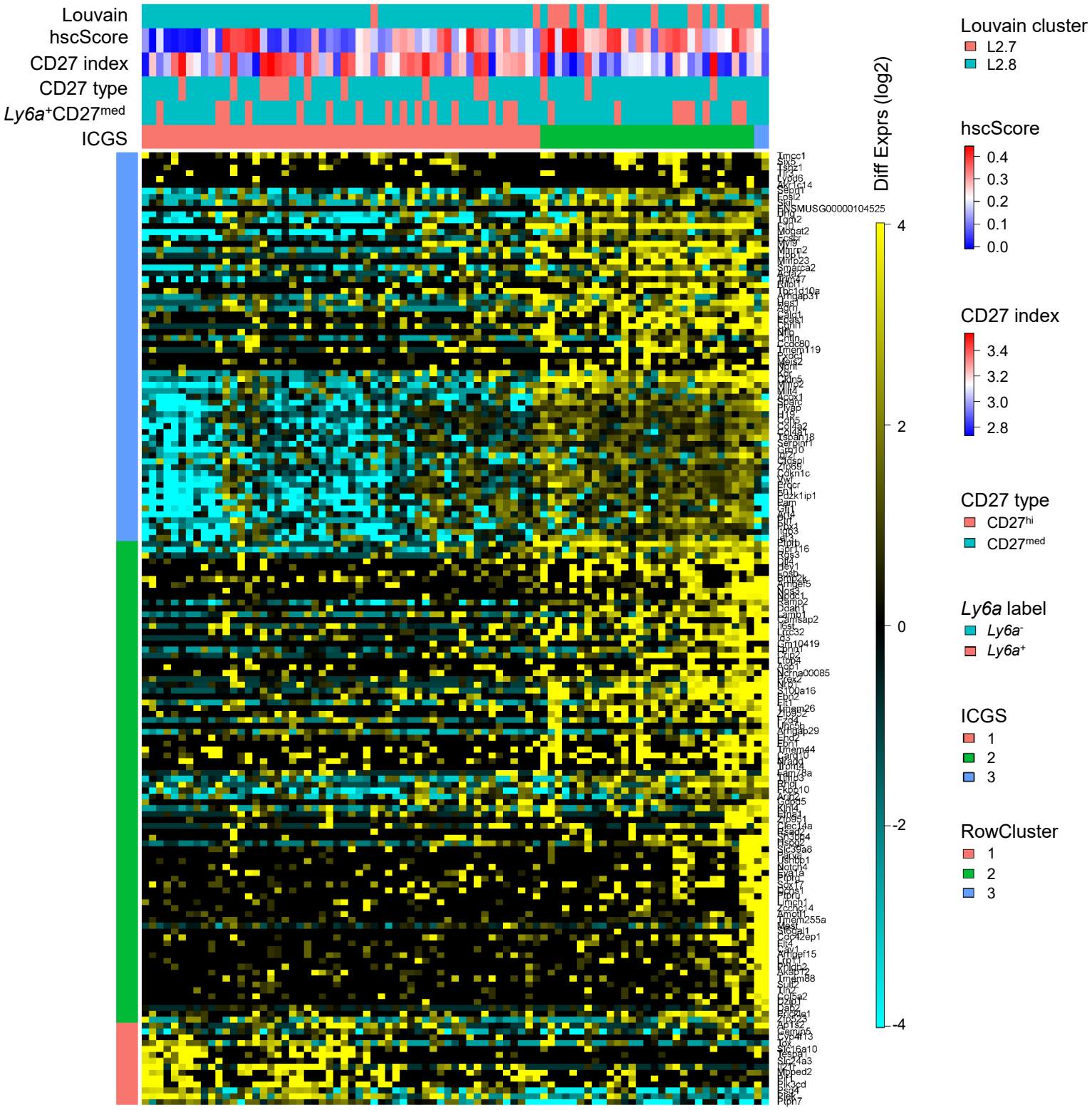


Supplementary Figure 2 – Related to Figure 4. Identification of additional markers for further enrichment of functional multipotent progenitors and HSCs.
Violin plots with expression of **A**) *Tnfrsf7* (CD27), **B**) *Selp* (CD62P) and *Mpl* in L1 Louvain clusters.

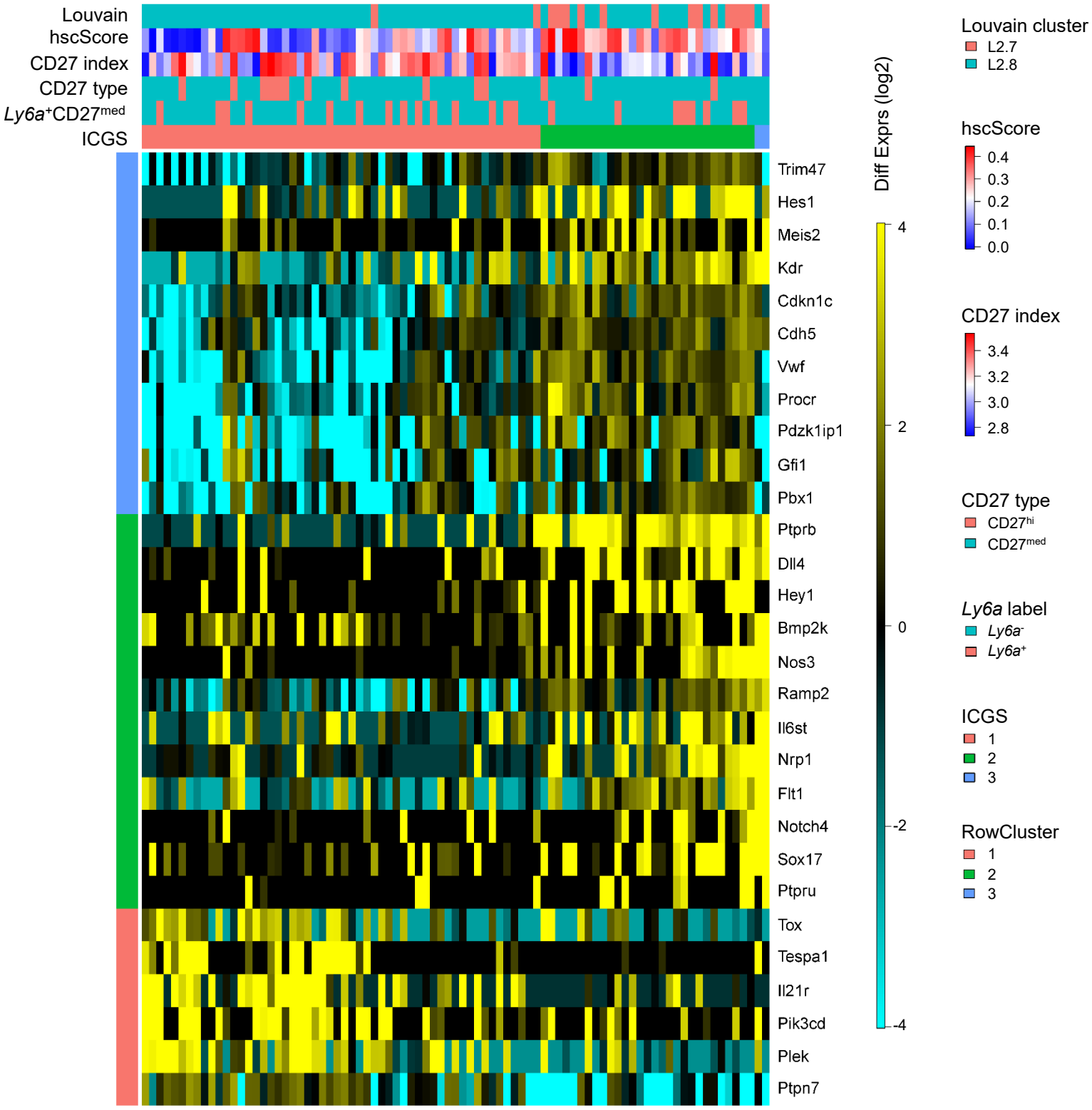


Supplementary Figure 3 – Related to Figure 4. Secondary Transplantations and Louvain 2 clustering of 1087 cells.

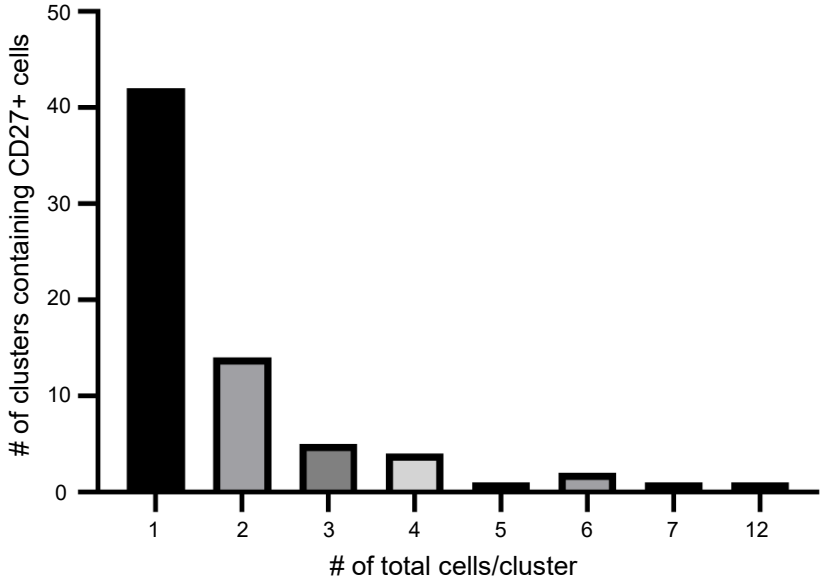
A) Donor chimerism was determined by CD45.1 and CD45.2 flow cytometric analysis of peripheral blood (PB). Flow cytometric zebra plots of representative **Ai**) WT (CD45.2/CD45.2) PB control and **Aii**) 4 months PB $CD27^{+N}$ (CD45.2+) transplant sample into CD45.1 recipients are shown. **Aiii)** Secondary transplants were performed by transplanting $5x10^6$ whole bone marrow (BM) cells harvested at 4 months post primary transplant from recipients that had received: $CD27^{+W}$ (442 cells; 2.5 embryo equivalent, ee; 74.7% PB and 48.0% BM chimerism; open squares), $CD27^{+W}$ (479 cells; 3.3ee; 58.7% PB and 10.0% BM chimerism; open triangles) or $CD27^{+N}$ (385 cells; 8.6ee; 75.5% PB and 58.3% BM chimerism; closed squares) cells. Secondary chimerism was as follows: $CD27^{+W}$: 14.0, 21.8, 42.3, 0.2, 0.4, 0.5%; $CD27^{+N}$: 31.4, 32.6, 39.2%. **B)** Both datasets of 119 $CD31^{hi}SSC^{lo}ckit^{hi}Gata2^{med}CD27^{+N}$ and 968 $CD31^{+}cKit^{hi}G2^{med}$ cells were combined and Louvain analysis repeated (L2.1-L2.12) to produce a new SPRING visualisation plot containing 1087 cells. **C)** Correspondence between L1 and L2 clustering. Cells in the SPRING visualisation are coloured according to L1 clustering, showing $CD27^{+N}$ sorted cells in grey.



Supplementary Figure 4 – Related to Figure 5. Heatmap showing all differential gene expression revealed by unsupervised clustering, reveals subgroups within the CD27^{+N} enriched single cells. ICGS output of the 85 CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N} cells belonging to HC1 (L2.7 and L2.8). 3 gene clusters (RowCluster) and 3 cell clusters (named ICGS1, ICGS2 and ICGS3) were defined. 23 *Ly6a*$^{+}$CD27^{med} cells are indicated in rose. Based on index data, 14 cells were identified as CD27Type CD27^{hi} and 71 as CD27^{med}. HscScore, log10 of the CD27 index-sorting (protein) levels and Louvain identity (L2.7 and L2.8) are also depicted.



Supplementary Figure 5 – Related to Figure 5. Heatmap showing all differential gene expression revealed by unsupervised clustering, reveals subgroups within the CD27^{+N} enriched single cells. Subset of ICGS output (selected genes) for the 85 CD31^{hi}SSC^{lo}cKit^{hi}Gata2^{med}CD27^{+N} cells belonging to HC1 (L2.7 and L2.8). 3 gene clusters (RowCluster) and 3 cell clusters (named ICGS1, ICGS2 and ICGS3) were defined. 23 *Ly6a*$^{+}$CD27^{med} cells are indicated in rose. Based on index data, 14 cells were identified as CD27Type CD27^{hi} and 71 as CD27^{med}. HscScore, log10 of the CD27 index-sorting (protein) levels and Louvain identity (L2.7 and L2.8) are also depicted.



Supplementary Figure 6 – Related to Figure 6. Correlation of CD27 expression with IAHC size.
A total of 62 transversal AGM sections (10µm) from genetically different embryos (*Gata2Venus*, *Ly6aGFP* and *Csf1rGFP*/MacGreen) were analysed for CD27 expression (n=4; 35, 36, 42, 43 somite pairs). $CD27^+$ cells were quantified and correlated to the size of the cluster in which they were localised. 80% (56 out of 70) of clusters containing $CD27^+$ cells had a total size of 1-2 cells.

nature COMMUNICATIONS

ARTICLE

Received 21 Jan 2015 | Accepted 10 Jul 2015 | Published 18 Aug 2015 | Updated 29 Oct 2015

DOI: 10.1038/ncomms9040 OPEN

BMP signalling differentially regulates distinct haematopoietic stem cell types

Mihaela Crisan[1,2], Parham Solaimani Kartalaei[1,2], Chris S. Vink[1,2], Tomoko Yamada-Inagawa[1], Karine Bollerot[1], Wilfred van IJcken[3], Reinier van der Linden[1], Susana M. Chuva de Sousa Lopes[4], Rui Monteiro[4], Christine Mummery[4] & Elaine Dzierzak[1,2]

Adult haematopoiesis is the outcome of distinct haematopoietic stem cell (HSC) subtypes with self-renewable repopulating ability, but with different haematopoietic cell lineage outputs. The molecular basis for this heterogeneity is largely unknown. BMP signalling regulates HSCs as they are first generated in the aorta-gonad-mesonephros region, but at later developmental stages, its role in HSCs is controversial. Here we show that HSCs in murine fetal liver and the bone marrow are of two types that can be prospectively isolated—BMP activated and non-BMP activated. Clonal transplantation demonstrates that they have distinct haematopoietic lineage outputs. Moreover, the two HSC types differ in intrinsic genetic programs, thus supporting a role for the BMP signalling axis in the regulation of HSC heterogeneity and lineage output. Our findings provide insight into the molecular control mechanisms that define HSC types and have important implications for reprogramming cells to HSC fate and treatments targeting distinct HSC types.

[1] Department of Cell Biology, Erasmus MC Stem Cell Institute, Erasmus Medical Center, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands. [2] University of Edinburgh, Centre for Inflammation Research, Queens Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ, UK. [3] Center for Biomics, Erasmus Medical Center, Wytemaweg 80, 3015 CN Rotterdam, Netherlands. [4] Department of Anatomy and Embryology, Leiden University Medical Center, Building 2, Einthovenweg 20, 2333 ZC Leiden, The Netherlands. Correspondence and requests for materials should be addressed to E.D. (email: Elaine.Dzierzak@ed.ac.uk).

Understanding the signalling pathways and mechanisms by which haematopoietic stem cells (HSCs) sustain their robust homoeostatic and regenerative characteristics is important for disease treatments that may differentially affect HSC subtypes. Long-term repopulating HSC subtypes have been defined by their haematopoietic lineage output—myeloid–lymphoid balanced (Bala) HSCs, myeloid-biased (My) HSCs and lymphoid-biased (Ly) HSCs. The clonal composition of the HSC compartment is age-dependent[1–4]. Bala-HSCs are found throughout ontogeny, Ly-HSCs predominate in young individuals and My-HSCs predominate in older individuals[2–4]. The molecular basis for HSC subtypes is hindered by the lack of prospective isolation. Since bone morphogenetic protein (BMP) affects the development of HSCs in the embryo, HSC subtypes may be differentially controlled by the BMP signalling pathway. It is known that high concentrations of BMP4 maintain proliferation and repopulating activity of human cord blood HSCs[5]. However, it has also been shown that conditional inactivation of *BMPR1a* increases the number of bone marrow (BM) HSCs[6], that the canonical BMP pathway is dispensable in embryonic day 14 (E14) fetal liver (FL) and adult BM[7] and that inhibition of Smad-dependent BMP signalling enhances homing and engraftment of BM HSCs[8]. Since the role of BMP in HSC regulation is as yet uncertain, we examined whether HSCs in the embryo, fetus and adult are directly responsive to BMP signalling and the relationship of BMP signalling to HSC heterogeneity.

Here we show in BMP responsive element (*BRE*) *GFP* transgenic mouse embryos[9] that all HSCs emerging *in vivo* in the aorta-gonad-mesonephros (AGM) region are BMP activated. In contrast, HSCs in murine FL and BM are of two types—BMP activated and non-BMP activated. The initially high proportion of BMP-activated HSCs decreases through ontogeny, and is surpassed by HSCs that are non-responsive to BMP. Clonal transplantation of the two BM HSC types demonstrates that HSC lineage output correlates with the state of BMP activation. Moreover, the two HSC types differ in intrinsic genetic programs thus supporting a role for the BMP signalling axis in the regulation of HSC heterogeneity and lineage output.

Results

All AGM HSCs *in vivo* are BMP-activated. The localized production of BMP4 in the AGM[10–12] and *BMPR2* expression by enriched AGM haematopoietic progenitors and stem cells (HPSCs)[10] suggests that BMP may act directly on HSCs. We examined the BMP activation status of HSCs during ontogeny in *BRE–GFP* transgenic mice. GFP is expressed in *BRE–GFP* mice when BMP and the BMP receptor signal through Smad1/5 to activate transcription from the *BRE* sequence (Fig. 1a). GFP expression denotes BMP activation at the moment of cell observation/isolation and does not represent previous BMP activation history. Importantly, all BRE–GFP-expressing cells also express pSmad1/5/8 (ref. 13). Three-dimensional imaging of whole-mount immunostained transgenic E10.5 embryos shows that BMP-activated (GFP^+) cells are predominantly distributed in the ventral aspect of the aorta (lumenal hematopoietic, endothelial and underlying mesenchymal cells; Fig. 1b). HPSCs are known to reside in the haematopoietic clusters closely associated with the aorta[14–16]. Some cluster cells are GFP^+ (Fig. 1c), suggesting heterogeneity in BMP activation within the HPSC compartment at the time of HSC generation. To test the BMP activation status of HSCs, E11 AGM GFP^+ and GFP^- cells were sorted ($6.2 \pm 3.1\%$(mean ± s.d.) of E11 AGM cells are GFP^+; Fig. 1d, Supplementary Fig. 1A) and transplanted into adult irradiated recipient mice. All AGM HSCs were found in the BMP-activated fraction (Fig. 1e): 3 out of 7 recipients receiving GFP^+ AGM cells were long term, high level repopulated (>10% chimerism) by donor cells, whereas no recipients (0 out of 7) receiving GFP^- AGM cells were donor engrafted ($P = 0.05$), even when high embryo equivalents of cells were injected. These data show that BMP signalling is activated in HSCs when they are first detected in the AGM.

Two HSC types can be prospectively isolated from FL and BM . Examination of the BMP activation status of the cells in the *BRE–GFP* E14 FL and adult BM revealed that $3.7 \pm 0.5\%$ (mean ± s.d.) and $5.5\% \pm 1.8$ (mean ± s.d.) of cells, respectively, were GFP^+ (Fig. 2a,c; Supplementary Fig. 1A). When sorted *BRE–GFP* E12–E14 FL or adult BM cells were transplanted, HSCs were found in both fractions, GFP^+ and GFP^-. Recipient mice were found to be long-term, high level repopulated (Fig. 2b,d) and secondary transplantation of BM from primary recipients showed that the HSCs self-renew (Fig. 2b,d). Thus, the FL and BM contain two types of HSCs—BMP activated and non-BMP activated. The percentage of phenotypic HSCs (LSK-SLAM; $Lin^- Sca1^+ cKit^+ CD150^+ CD48^-$) in E14 FL is significantly higher in the BMP-activated fraction (73%) than in the non-BMP-activated fraction (27%; $n = 5$, T-test $P < 0.001$; Fig. 2e,f). In contrast, in the BM, the majority (92%) of the phenotypic LSK-SLAM HSCs is non-BMP activated (Fig. 2g,h). BMP activation as detected by BRE–GFP expression allows prospective isolation of the two HSC types to enable study of their molecular and functional characteristics.

The expression profiles of the BMP-activated and non-activated HSCs as examined by RNA sequencing show distinct genetic programs (Fig. 3). FL and BM LSK-SLAM GFP^+ HSCs express *Bmpr* genes, whereas GFP^- LSK-SLAM HSCs do not (Fig. 3a). *Smad* genes are expressed in both fractions. Genes upregulated in the GFP^+ FL and BM HSC fractions are significantly enriched in BMPR2 downstream targets, confirming BMP signalling activation (Fig. 3b). Moreover, other categories/gene sets were found to be differentially expressed between the two HSC types. For example, sets common to FL and BM upregulated in the BMP-activated HSCs are significantly enriched for genes involved in homoeostasis and metabolism, and MYC and STAT5B target genes (Fig. 3c). Upregulated sets in the non-BMP-activated HSCs are significantly enriched for genes involved in haematopoietic system development, NFKB1, SP1 and NFE2 target genes (Fig. 3d; see figure legend for false discovery rate corrected Fisher exact test P values). As some of these transcription factors affect normal and malignant haematopoietic cells and specific haematopoietic lineages, the distinct programs may influence the functional characteristics of the two HSC types.

BM HSC output correlates with BMP activation status. Limiting dilution and clonal transplantations were performed to examine the frequency and functional properties of BMP-activated and non-activated HSCs. We found that a high frequency of FL GFP^+ cells are HSCs (1 out of 180), whereas only 1 out of 20,545 FL GFP^- cells is a HSC (Fig. 4a; Supplementary Table 1). Given that the FL contains a mean of 12.7×10^6 cells, of which 3.7% are GFP^+, our data demonstrate that 81% of FL HSCs are BMP activated, with 19% being non-BMP activated (Fig. 4b). This corresponds well with the percentages of GFP^+ and GFP^- cells found in the FL LSK-SLAM phenotypic HSC compartment. In contrast, the non-BMP-activated HSC number is significantly higher in the adult BM (Fig. 4c,d). One in 10,053 BM GFP^+ cells and 1 in 17,760 BM GFP^- cells is an HSC (Supplementary Table 2). Taking into account an average of 2×10^7 BM cells per mouse (four long bones), of which 5.5% are GFP^+, we find that 9% of

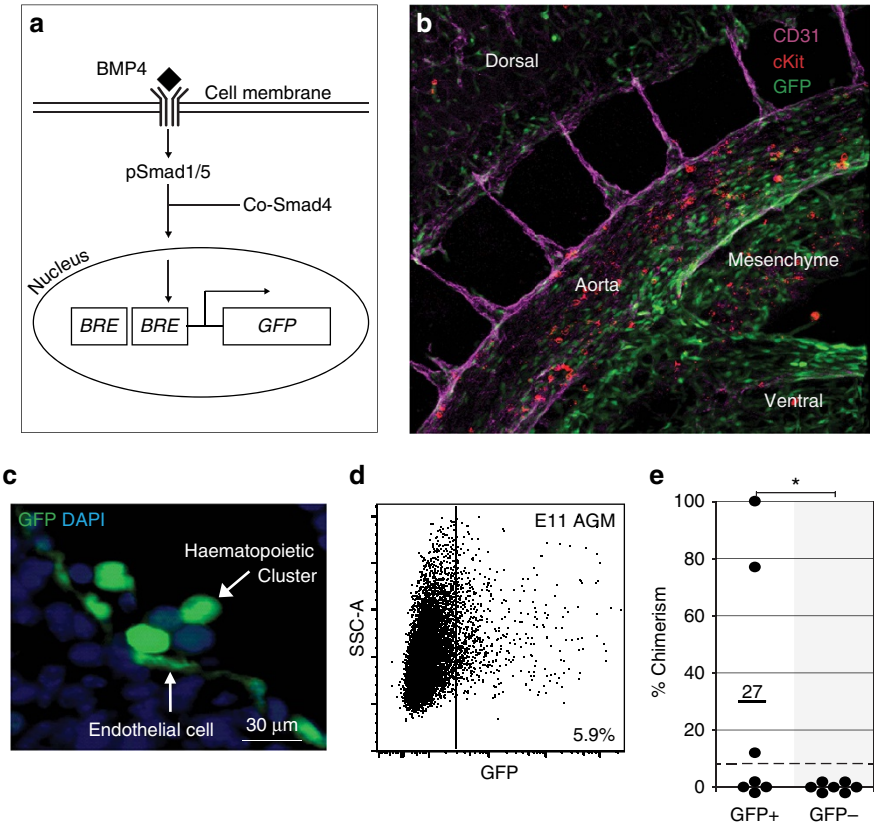
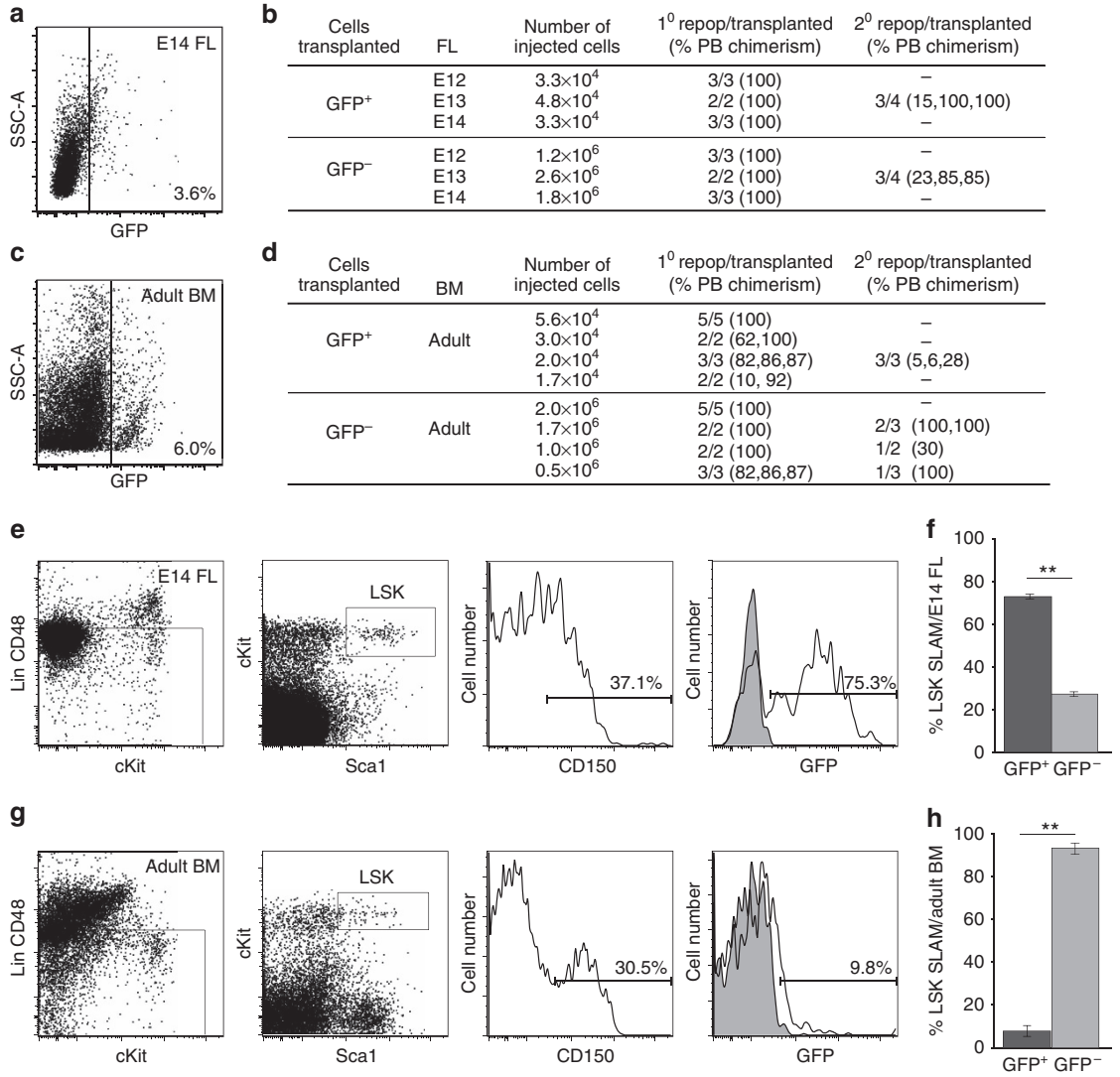


Figure 1 | Aorta-gonad-mesonephros HSCs are BMP activated. (**a**) Scheme showing activation of the canonical BMP signalling pathway through phospo-Smad1/5 binding of the double BMP responsive element (*BRE*) and the resulting transcription of *GFP*. (**b**) Three-dimensional whole-mount image of an E10.5 immunostained *BRE-GFP* mouse embryo (37 somite pairs). Anti-CD31 (magenta), cKit (red) and GFP (green) antibody staining shows the predominantly ventral distribution of GFP$^+$ cells in various cell types within the aorta and underlying mesenchymal cells. (**c**) High-magnification image of a transverse section of an E11 *BRE-GFP* dorsal aorta stained with anti-GFP-antibody and 4,6-diamidino-2-phenylindole (DAPI). Intra-aortic hematopoietic clusters contain GFP$^+$ and GFP$^-$ cells. Some endothelial cells are GFP$^+$. (**d**) Representative FACS plot with side scatter on the *y* axis showing percentage of E11 AGM GFP$^+$ cells (see Supplementary Fig. 1A for control). (**e**) Percentage donor cell chimerism in the peripheral blood of adult irradiated transplant recipients at 4 months after injection of E11 AGM GFP$^+$ and GFP$^-$ cells (2–4.5 embryo equivalents (ee) of AGM cells injected per mouse; $n=3$ independent transplantation experiments). *$P=0.05$ by z-Test for proportions.

BM HSCs are BMP activated and 91% are non-BMP activated. This is in correspondence with the percentages of GFP$^+$ and GFP$^-$ cells found in the BM LSK-SLAM phenotypic HSC compartment. The inversed percentages of BMP-activated HSCs in the BM as compared with the FL are likely due to time- and/or niche-related ontogenic changes.

Considering this shift in the number of BMP-activated HSCs with developmental time and the data of others showing that there are changes in the clonal composition of the HSC compartment associated with lineage output[2], we tested the lineage output of BMP-activated and non-activated HSCs. Clonal transplantations of GFP$^+$ and GFP$^-$ HSCs from FL and adult BM (injected cell number was calculated by limiting dilution experiments) were performed and the peripheral blood lineage output data at 4 months post-transplantation was analysed according to Cho *et al.*[2] (Supplementary Fig. 1B). Only primary recipients that were reconstituted in both lymphoid and myeloid lineages were considered. Our data in the FL show slight differences in the lineage output composition of HSCs, with more Bala-HSCs in the BMP-activated fraction. However, this did not reach significance (Supplementary Fig. 1C,D; Supplementary Table 3). In contrast, BMP-activated and non-BMP-activated HSCs in the BM were significantly different in terms of HSC lineage output composition ($P=0.007$). Bala-HSCs were significantly enriched in the BMP-activated HSC fraction: 7 out of 17 in the GFP$^+$ as compared with 0 out of 15 in the GFP$^-$ fraction ($P=0.005$) (Fig. 4e; Supplementary Table 3). My-HSCs were significantly enriched in the non-BMP-activated fraction: 7 out of 15 in GFP$^-$ compared with 2 out of 17 in the GFP$^+$ ($P=0.031$; Fig. 4f; Supplementary Table 3). Ly-HSCs were equally distributed (8 out of 17 and 8 out of 15; $P=0.727$) (Fig. 4e,f; Supplementary Table 3). Thus, prospective isolation of predominantly balanced and myeloid-biased HSCs is achieved in the BM by cell sorting based on BMP activation.

The two HSC types differ in their intrinsic molecular programs. Previously, it has been suggested that common lymphoid progenitor cells (CLPs) derived from clonal transplantation of lymphoid–myeloid HSCs are poised for expression of lymphoid regulator genes[17]. We examined whether the expression of such lymphoid genes are already present in LSK-SLAM HSCs and are associated with BMP activation status. Whereas 88% of GFP$^+$ BM HSCs have enhanced lymphoid differentiation potential (Bala + Ly), this potential is present in only 53% of HSCs in the GFP$^-$ fraction (Supplementary Table 3). Moreover, RNA sequencing data showed that the GFP$^+$ BM HSCs have higher



Cells transplanted	FL	Number of injected cells	1^0 repop/transplanted (% PB chimerism)	2^0 repop/transplanted (% PB chimerism)
GFP^+	E12	3.3×10^4	3/3 (100)	–
	E13	4.8×10^4	2/2 (100)	3/4 (15,100,100)
	E14	3.3×10^4	3/3 (100)	–
GFP^-	E12	1.2×10^6	3/3 (100)	–
	E13	2.6×10^6	2/2 (100)	3/4 (23,85,85)
	E14	1.8×10^6	3/3 (100)	–

Cells transplanted	BM	Number of injected cells	1^0 repop/transplanted (% PB chimerism)	2^0 repop/transplanted (% PB chimerism)
GFP^+	Adult	5.6×10^4	5/5 (100)	–
		3.0×10^4	2/2 (62,100)	–
		2.0×10^4	3/3 (82,86,87)	3/3 (5,6,28)
		1.7×10^4	2/2 (10, 92)	–
GFP^-	Adult	2.0×10^6	5/5 (100)	–
		1.7×10^6	2/2 (100)	2/3 (100,100)
		1.0×10^6	2/2 (100)	1/2 (30)
		0.5×10^6	3/3 (82,86,87)	1/3 (100)

Figure 2 | The fetal liver (FL) and bone marrow contain two distinct HSC types. Representative FACS plots showing percentage of GFP^+ cells with side scatter shown on the *y* axis in the (**a**) E14 FL and (**c**) adult bone marrow (BM) of *BRE-GFP* mice (see Supplementary Figure S1A for controls). Transplantation results of adult irradiated recipients injected with GFP^+ and GFP^- cells sorted from (**b**) E12, E13 or E14 FL ($n=3$) or (**d**) adult BM ($n=4$). Secondary irradiated adult recipients received BM cells, unsorted (3×10^6) or GFP^+ (1.5×10^5) and GFP^- (0.8, 1.8 and 11.8×10^6) sorted cells from repopulated primary recipients. Data from all secondary recipients were pooled. Peripheral blood (PB) donor chimerism is shown for all recipients. Representative FACS plots showing the sorting strategy for LSK-SLAM GFP^+ and GFP^- cells from (**e**) E14 FL and (**g**) adult BM. Gates are shown for the Lin- and LSK fractions. Histogram analysis shows percentages of $CD150^+$ cells in the LSK fractions and the percentages of GFP^+ cells in the LSK-SLAM fractions of FL and BM. Grey curves show control non-transgenic LSK SLAM cells. Percentages of LSK SLAM cells that are GFP^+ (black) or GFP^- (grey) as found in (**f**) E14 FL ($n=3$) and (**h**) adult BM ($n=3$). $^{**}P<0.001$ *T*-test.

Fragments per kilobase of transcript per million mapped reads (FPKM) values for Ikaros (*Ikzf1*), E2A (*Tcf3*) and *Flt3* genes than GFP^- HSCs, supporting a pro-lymphoid transcriptional programme for BM BMP-activated HSCs. This is not observed in the FL where similar proportions of HSCs with enhanced lymphoid differentiation potential exist in both GFP^+ (96%) and GFP^- (97%) fractions (Supplementary Table 3). FPKM values of the lymphoid genes are comparable in these fractions (Supplementary Fig. 1E). These data suggest that the BM microenvironment promotes a pro-lymphoid gene programme in subset of HSCs activated by the BMP signalling pathway. These results on HSC lineage output, together with the clear differences in the genetic programs between BMP-activated and non-BMP-activated HSCs, suggest the BMP signalling axis as an underlying basis for HSC heterogeneity.

Discussion

The BMP activation marker that we used in this study represents a robust method by which functionally distinct types of HSCs can be prospectively enriched. Other methods previously used such as label retention[18], ckit expression levels[19], Hoechst dye efflux levels[20] and CD229 provide limited separation[17]. It was shown that TGFβ1 differentially controls Ly- and My-HSCs[20]. However, the variation between the individual HSC clones found in this study suggests that HSC subtype segregation is not absolute. Indeed, as we show here only a subset of Ly-HSCs are controlled by the BMP pathway. Interestingly, TGFβ1 appears to be a general stimulatory factor for My-HSC proliferation[20]. This is in line with our study showing that in BM, in absolute number, there are significantly more My-HSCs in the non-BMP-activated fraction and these are likely to be controlled by the TGFβ pathway.

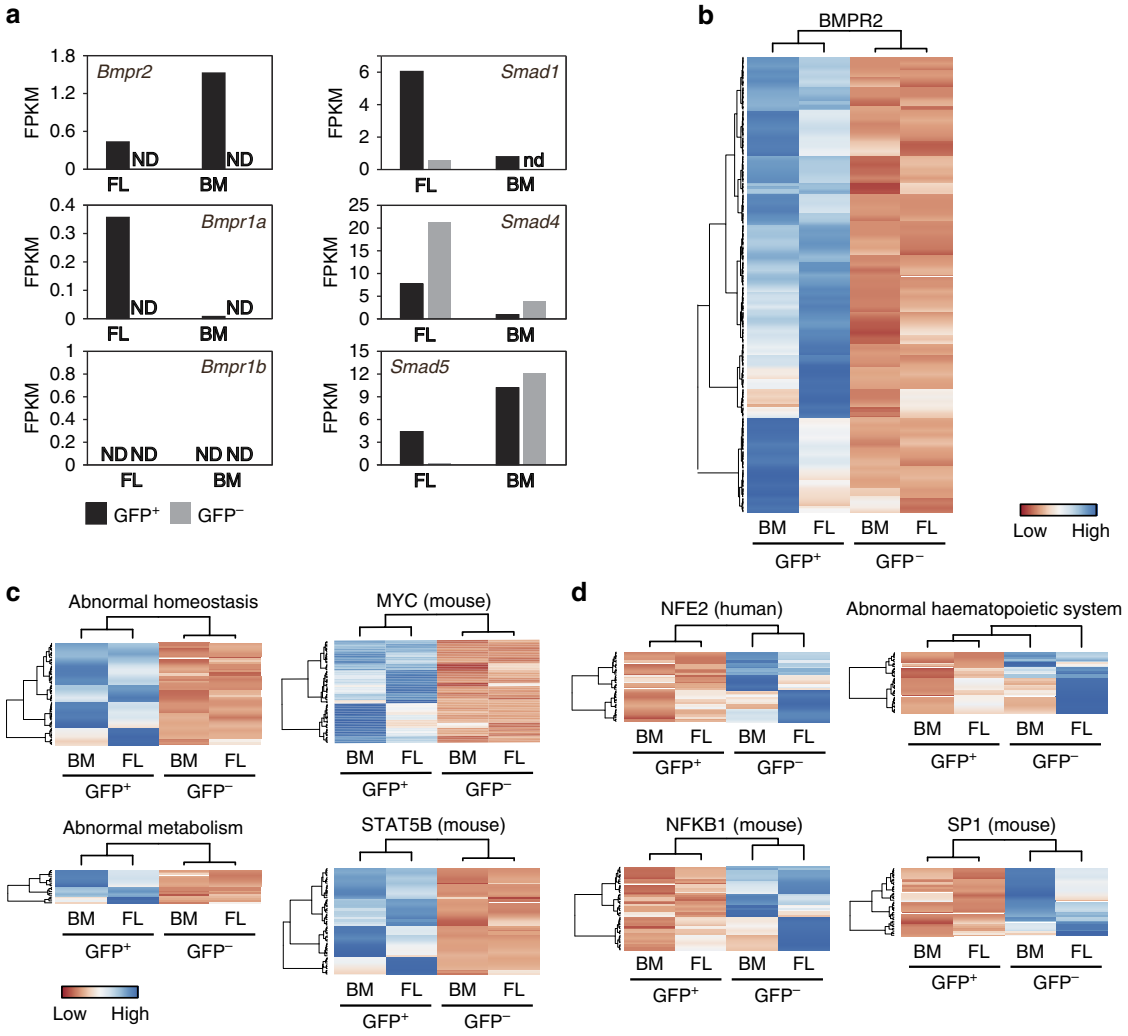


Figure 3 | Transcriptome differences between BMP-activated and non-BMP-activated FL and BM HSCs. (**a**) FPKM values for *BMP* receptor and *Smad* genes expressed by GFP$^+$ and GFP$^-$ LSK SLAM sorted HSCs from E14 FL and adult BM. ND, not detected. (**b**–**d**) Gene list enrichment analysis on genes with more than two fold difference in expression level between GFP$^+$ and GFP$^-$ LSK SLAM sorted cells from FL and BM using *Enrichr* web-application (threshold false discovery rate (FDR) = 0.01). GFP$^+$ high genes are enriched in (**b**) BMPR2 gene targets (FDR < 1.0e^{-13}; Gene Expression Omnibus (GEO) Kinase perturbation gene sets). Gene lists available on request. (**c**) Also, overrepresented in GFP$^+$ high genes are genes involved in homoeostasis (FDR = 9.8e^{-4}; Mouse Genome Informatics (MGI) mammalian phenotype term MP0001764), metabolism (FDR = 5.4e^{-4}; MGI mammalian phenotype term MP0005266), MYC (FDR = 2.7e^{-4}; TRANSFAC and JASPAR Position Weight Matrix (PWM) genesets), and STAT5B (FDR = 2.7e^{-4}; TRANSFAC and JASPAR PWM gene sets); (**d**) GFP$^-$ high genes are enriched for NFE2 (FDR = 9.6e^{-3}; TRANSFAC and JASPAR PWM gene sets), NFKB1 (FDR = 2.7e^{-3}; TRANSFAC and JASPAR PWM gene sets), SP1 (FDR = 2.7e^{-3}; TRANSFAC and JASPAR PWM gene sets), and haematopoietic system (FDR = 5.2e^{-7}; MGI mammalian phenotype term MP0002396)-related genes.

The ability to prospectively identify the distinct HSC types based on BMP activation and directly measure the clonal composition of the HSC pool highlights the importance of the BMP signalling axis in directing the intrinsic HSC programme during ontogeny. MYC and STAT5b target genes that are enriched in the BMP-activated HSCs are, respectively, known to be important in the balance between differentiation and self-renewal, and for conferring efficient lympho–myeloid repopulation and quiescence properties[21,22]. Of the (NFE2, SP1 and NFKB1) targets upregulated in the non-BMP-activated HSCs, NFE2 is often overexpressed in the myeloproliferative disorders, which spontaneously transform to acute myeloid leukaemia and polycythemia[23]. SP1 and NFKB1 are known to regulate HSC specification and are required for normal HSC function and differentiation[24,25]. Interestingly, aged HSCs express high levels of NFKB[3,26]. This is in line with our data showing that the non-BMP-activated BM HSCs are high in NFKB and are predominantly myeloid biased. These transcriptomic profiles together with further studies should provide insight into how HSC clonal composition may be regulated.

Our finding of the existence of non-BMP-activated HSCs now provides an explanation for the absence of a haematopoietic phenotype when the canonical BMP signalling pathway is impaired[6–8]. In the absence of BMP-activated HSCs, the non-activated HSCs are not only sufficient to assume the normal haematopoietic function, but may also expand. The earliest AGM HSCs are BMP activated and as development progresses and HSCs reside in other tissues, there is a shift in the ratio of HSC types from being predominantly BMP activated in the FL to being predominantly non-BMP activated in the adult BM. This implies quantitative changes in BMP-producing niches during the migration process from the AGM (BMP^{high}) to FL and eventually multiple BM niches (osteoblastic, perivascular and neural) with variable BMP levels (BMP^{high} versus BMP^{low})

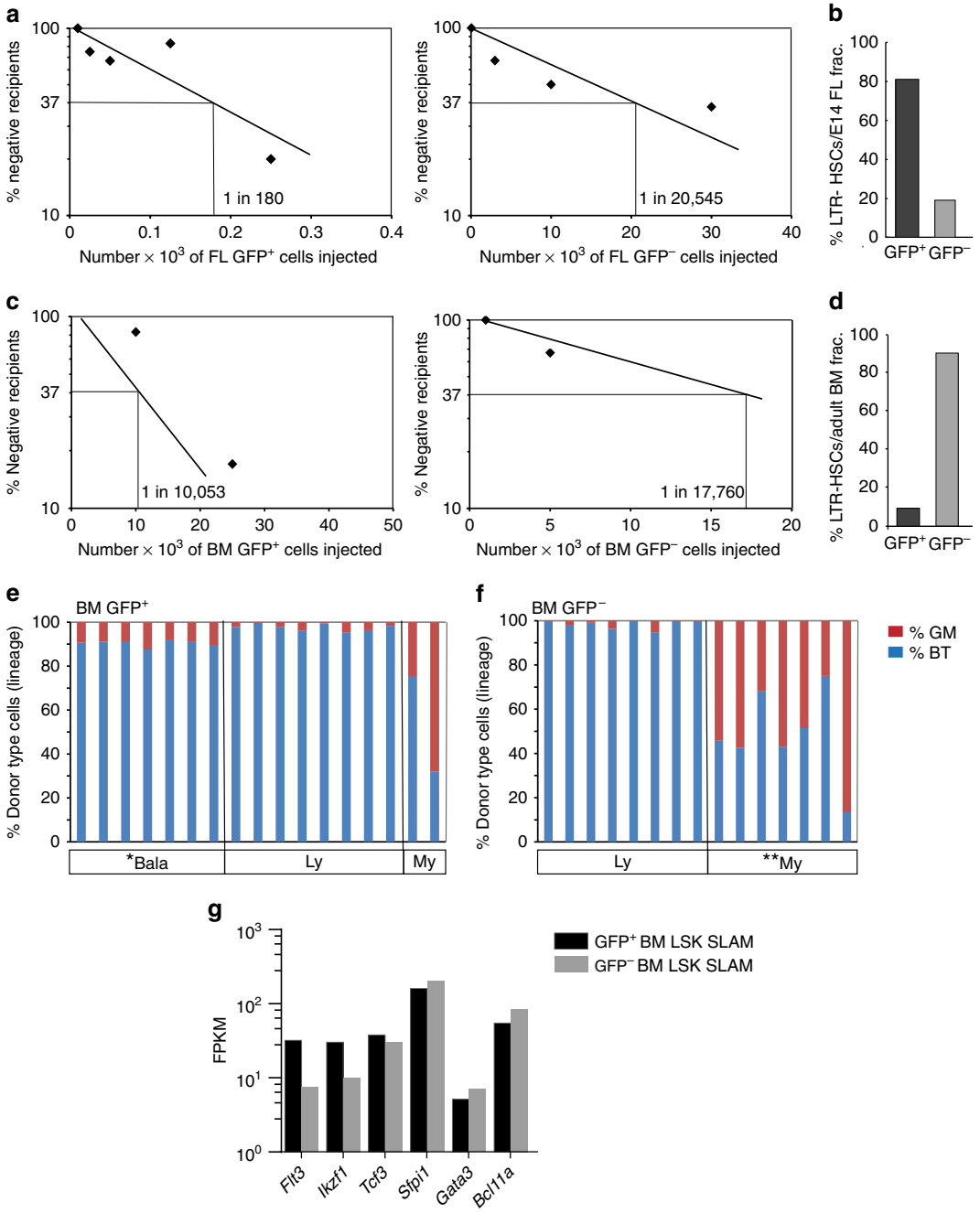


Figure 4 | Relative percentages and lineage output of the BMP-activated and non-BMP-activated HSCs. Limiting dilution transplantation experiments show the frequencies of HSCs in GFP$^+$ and GFP$^-$ cell fractions from (**a**) E14 FL and (**c**) adult BM. See Supplementary Tables 1 and 2 for detailed results. Bar graphs showing percentages of long-term repopulating (LTR) HSCs that are GFP$^+$ (black) or GFP$^-$ (grey) in (**b**) E14 FL ($n=5$) and (**d**) adult BM ($n=5$) as calculated from limiting dilution transplantations. Summary of lineage output analysis of repopulated recipient mice injected with BM (**e**) GFP$^+$ and (**f**) GFP$^-$ cells at the clonal HSC level ($n=3$). Sorted GFP$^+$ and GFP$^-$ adult BM cells from 14–32-week-old *BRE-GFP* (Ly5.2) mice were injected intravenously into Ly5.1 irradiated (9 Gy) recipients and donor chimerism measured at 4 months post transplantation. Peripheral blood was analysed for donor Ly5.2 cell marker in granulocyte–macrophage (GM) lineages (Gr-1 and CD11b) and B- and T-cell lineages (B220 and CD3) by flow cytometry. Per cent donor contribution and donor percentage of GM and B + T were calculated according to Cho *et al.*[2], to reveal HSC type in each recipient. *P* values were calculated by Fisher exact test between the GFP$^+$ and the GFP$^-$ fractions ($P=0.007$) and by two sample test for proportion (χ^2) between balanced (*$P=0.005$), My (**$P=0.031$) and Ly-HSCs ($P=0.727$) of the two groups, GFP$^+$ and GFP. (**g**) Fragments per kilobase of transcript per million mapped reads (FPKM) values for selected lymphoid genes expressed by GFP$^+$ and GFP$^-$ LSK SLAM sorted HSCs from adult BM.

or intrinsic changes in HSC BMP receptor expression with developmental time. Future studies should examine the specific niches supporting the two HSC types in the FL and BM for BMP production, and examine whether other signalling molecules regulate the non-BMP-activated HSCs. Deciphering the behaviour and lineage output of HSCs during ontogeny, understanding whether the two HSC types are interchangeable or have different origins and investigating the underlying intrinsic and extrinsic mechanisms that dictate these changes are of importance for ensuring that HSCs created through

reprogramming strategies will be capable of generating all the desired cell types, and for understanding why specific haematological malignancies are prevalent in paediatric versus elderly patients[27]. Our findings may be of particular interest to drug resistance in acute myeloid leukaemia patients[28], raising the question of whether a combination of drugs affecting the distinct HSC types would be more efficient to control blood-related disorders.

Methods

Mice and embryo generation. Mice were bred and housed at Erasmus MC. *BRE GFP* transgenic mice[9] were maintained in C57BL/6 background. Two independently derived transgenic lines were used in this study. Data from line 1 are shown. Matings were set-up between heterozygous *BRE GFP* transgenic male and wild-type C57BL/6 female mice. The day of vaginal plug detection was designated as E0. All animal procedures were approved by the local ethics committee and performed in compliance with Standards for Care and Use of Laboratory Animals.

Whole-mount immunostaining and immunohistochemistry. Whole-mount embryo immunostaining and three-dimensional imaging were performed as described previously[16,29]. Embryos were fixed for 20 min with 2% PFA/PBS at 4 °C; dehydrated in graded concentrations of methanol; stained with primary antibodies-unconjugated rabbit anti-GFP (MBL, 1:2000), biotinylated rat anti-CD31 (BD, 1:500) and subsequently with rat anti-cKit (BD, 1:500) antibodies in blocking buffer ON at 4 °C; washed; incubated with secondary antibodies goat anti-rabbit IgG-Alexa Fluor488 (Invitrogen, 1:1,000), goat anti-rat IgG-Alexa Flour647 (Invitrogen, 1:5,000) and donkey Streptavidin-Cy3, (Jackson ImmunoResearch, 1:1,000); made transparent in BABB; and analysed with a confocal microscope (Zeiss LSM 510META JNI, Plan-Neofluar × 10/0.3, Eppiplan-Neofluar × 20/0.50). Three-dimensional reconstructions were generated from z-stack with LSM Image Browser (Zeiss) and Amira (VISAGE IMAGING).

Flow cytometry. AGMs were dissected with needles under the microscope as previously described[30] and tissues were dissociated in PBS/FCS/PS with collagenase type I (Sigma, 0,125%) in water bath at 37 °C for 45 min. Cells were further mechanically separated by flushing with a pipette, washed with PBS/FCS/PS and centrifuged at 1,000 r.p.m. for 10 min at 4 °C. Pellets were resuspended in PBS/FCS/PS. E14 FL were mechanically disrupted until a homogeneous cell suspension was obtained. Adult BM was harvested from tibias and femurs by flushing with PBS/FCS/PS through a 1-ml syringe with a 25-G needle. Ficoll fractionation of BM cells was performed at 2,000 r.p.m., 20 min at room temperature. Mononuclear cells were harvested and washed with PBS/FCS/PS followed by centrifugation at 1,000 r.p.m., 10 min at 4 °C. Viable cells were counted by Trypan Blue exclusion (Sigma, 0.4%).

FL and adult BM LSK/SLAM cell stainings included LinPE cocktail (anti-CD3/1:400, anti-B220/1:2000, anti-Ter119/1:400, anti-Ly6c/1:4000 antibodies), anti-Sca1PE-Cy7(1:800), anti-cKitAPC-Alexa Fluor780(1:600) or BV421(1:500), anti-CD48PE(1:500) and anti-CD150APC(1:500) antibodies. Anti-CD11b-PE antibody (1:2,000) was added in the Lin staining for adult BM. Cells were analysed on a FACSAria SORP or FACSAria III (BD) with FloJo software. Dead cell exclusion was with Hoechst 33258 (1:10,000, Molecular Probes). Lineage output was tested on primary recipient peripheral blood at 4 months post transplantation using myeloid lineage (GM) markers, anti-CD11b PE and anti-Ly6c PE antibodies, lymphoid lineage (B + T) markers, anti-CD3APC(1:400) and anti-B220APC(1:2,000) antibodies and the donor marker with anti-CD45.2 FITC(1:500) antibody. All antibodies used were from BD Pharmingen except the anti-CD150 antibody (eBiolegend). Haematopoietic lineage output in peripheral blood of clonal transplant recipients (4 months post transplantation) was considered balanced when the $B+T^{value}/GM^{value}/$ was 4–16, myeloid-biased when the ratio was 0–4 and lymphoid-biased when the ratio was >16.

Transplantation assay. Single cell suspensions were injected intravenously into female ((129SV × C57BL/6)F1 or Ly5.1) mice (3–6 months old) irradiated with 9 Gy (split dose, ^{137}Cs source). A total of 2×10^5 recipient background spleen cells were co-injected. BM cells from primary recipients were injected into irradiated (9 Gy, split dose) secondary recipients. For (129SV × C57BL/6) F1 recipients, peripheral blood peripheral blood chimerism was quantified by DNA PCR for *gfp* and calculated by DNA normalization (*myoD*) and comparison with *gfp* contribution controls with Image Quant software. Mice showing >10% donor chimerism were considered repopulated. For Ly5.1 recipients, peripheral blood chimerism was quantified by CD45.1 and CD45.2 flow cytometric analysis. Mice with >1% $CD45.2^+CD45.1^-$ donor chimerism were considered repopulated. Secondary recipients received either 3×10^6 unsorted BM cells or GFP^+ (1.5×10^5 cells per mouse) and GFP^- ($0.8 - 11.8 \times 10^6$ cells per mouse) from the reconstituted primary recipients. For lineage output analysis, clonal HSC were transplanted based on frequencies found in the limiting dilution experiments. Briefly, GFP^+ and GFP^- fractions from the BM and FL were first injected at different cell numbers in irradiated recipients and the repopulation activity measured at 4 months post injection. Using the following parameters, cell number transplanted and number of mice repopulated per number of mice transplanted, frequencies of HSCs were calculated with the L-Calc Software (StemCellTechnology). Subsequent transplants were performed with the cell numbers containing 1 HSC.

RNA preparation and RNA sequencing. RNA was isolated from E14 and adult BM with the mirVana miRNA Kit and prepared according to SMARTER protocol[31] for the Illumina HiSeq2000 sequencer. The sequencing depth (unique mapped reads) was for BM (GFP^+ $1.21e^{07}$; GFP^- $1.11e^{07}$) and for the FL (GFP^+ $1.22e^{07}$; GFP^- $9.63e^{06}$). Sequences were mapped to the mouse (NCBI37/mm9) genome and FPKMs were calculated using Bowtie (v2.2.3), TopHat (v2.0.12) and Cufflinks (v2.2.1). Differential expression was analysed using Cuffquant with fragment-bias and multi-read corrections and normalized across all samples using Cuffnorm with geometric library-size normalization[32]. Genes are 'high in the BMP-activated HSC fraction (GFP^+)' when FPKM in $GFP^+ > 2\times$ FPKM GFP^- in both BM and FL LSK-SLAM cells, and 'high in the non-BMP-activated fraction (GFP^-)' when FPKM in $GFP^- > 2\times$ FPKM GFP^+ in both BM and FL LSK-SLAM cells. For gene list enrichment analysis, genes were applied to *Enrichr* web application[33]. Gene sets with enrichment false discovery rate of 0.01 were selected and heatmaps (row scaled) were generated with 'GFP^+ high' or 'GFP^- high' genes present in given enriched genelists.

References

1. Geiger, H., de Haan, G. & Florian, M. C. The ageing haematopoietic stem cell compartment. *Nat. Rev. Immunol.* **13,** 376–389 (2013).
2. Cho, R. H., Sieburg, H. B. & Muller-Sieburg, C. E. A new mechanism for the aging of hematopoietic stem cells: aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111,** 5553–5561 (2008).
3. Benz, C. *et al.* Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell* **10,** 273–283 (2012).
4. Verovskaya, E. *et al.* Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding. *Blood* **122,** 523–532 (2013).
5. Bhatia, M. *et al.* Bone morphogenetic proteins regulate the developmental program of human hematopoietic stem cells. *J. Exp. Med.* **189,** 1139–1148 (1999).
6. Zhang, J. *et al.* Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425,** 836–841 (2003).
7. Singbrant, S. *et al.* Canonical BMP signaling is dispensable for hematopoietic stem cell function in both adult and fetal liver hematopoiesis, but essential to preserve colon architecture. *Blood* **115,** 4689–4698 (2010).
8. Khurana, S. *et al.* SMAD signaling regulates CXCL12 expression in the bone marrow niche, affecting homing and mobilization of hematopoietic progenitors. *Stem Cells* **32,** 3012–3022 (2014).
9. Monteiro, R. M. *et al.* Real time monitoring of BMP Smads transcriptional activity during mouse development. *Genesis.* **46,** 335–346 (2008).
10. Durand, C. *et al.* Embryonic stromal clones reveal developmental regulators of definitive hematopoietic stem cells. *Proc. Natl Acad. Sci. USA* **104,** 20838–20843 (2007).
11. Kaimakis, P., Crisan, M. & Dzierzak, E. The biochemistry of hematopoietic stem cell development. *Biochim. Biophys. Acta* **1830,** 2395–2403 (2013).
12. Marshall, C. J., Kinnon, C. & Thrasher, A. J. Polarized expression of bone morphogenetic protein-4 in the human aorta-gonad-mesonephros region. *Blood* **96,** 1591–1593 (2000).
13. Moya, I. M. *et al.* Stalk cell phenotype depends on integration of Notch and Smad1/5 signaling cascades. *Dev. Cell* **22,** 501–514 (2012).
14. de Bruijn, M. F. *et al.* Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* **16,** 673–683 (2002).
15. Taoudi, S. & Medvinsky, A. Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. *Proc. Natl Acad. Sci. USA* **104,** 9399–9403 (2007).
16. Yokomizo, T. & Dzierzak, E. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development* **137,** 3651–3661 (2010).
17. 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 (2013).
18. Kiel, M. J. *et al.* Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU. *Nature* **449,** 238–242 (2007).
19. Grinenko, T. *et al.* Clonal expansion capacity defines two consecutive developmental stages of long-term hematopoietic stem cells. *J. Exp. Med.* **211,** 209–215 (2014).
20. Challen, G. A., Boles, N. C., Chambers, S. M. & Goodell, M. A. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* **6,** 265–278 (2010).

21. Laurenti, E. *et al.* Hematopoietic stem cell function and survival depend on c-Myc and N-Myc activity. *Cell Stem Cell* **3,** 611–624 (2008).
22. Wang, Z. & Bunting, K. D. STAT5 in hematopoietic stem cell biology and transplantation. *JAKSTAT* **2,** e27159 (2013).
23. Kaufmann, K. B. *et al.* A novel murine model of myeloproliferative disorders generated by overexpression of the transcription factor NF-E2. *J. Exp. Med.* **209,** 35–50 (2012).
24. Gilmour, J. *et al.* A crucial role for the ubiquitously expressed transcription factor Sp1 at early stages of hematopoietic specification. *Development* **141,** 2391–2401 (2014).
25. Stein, S. J. & Baldwin, A. S. Deletion of the NF-kappaB subunit p65/RelA in the hematopoietic compartment leads to defects in hematopoietic stem cell function. *Blood* **121,** 5015–5024 (2013).
26. Chambers, S. M. *et al.* Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. *PLoS Biol.* **5,** e201 (2007).
27. Rodriguez-Abreu, D., Bordoni, A. & Zucca, E. Epidemiology of hematological malignancies. *Ann. Oncol.* **18**(Suppl 1): i3–i8 (2007).
28. Zahreddine, H. A. *et al.* The sonic hedgehog factor GLI1 imparts drug resistance through inducible glucuronidation. *Nature* **511,** 90–93 (2014).
29. Yokomizo, T., Ng, C. E., Osato, M. & Dzierzak, E. Three-dimensional imaging of whole midgestation murine embryos shows an intravascular localization for all hematopoietic clusters. *Blood* **117,** 6132–6134 (2011).
30. Robin, C. & Dzierzak, E. Hematopoietic stem cell enrichment from the AGM region of the mouse embryo. *Methods Mol. Med.* **105,** 257–272 (2005).
31. Ramskold, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30,** 777–782 (2012).
32. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7,** 562–578 (2012).
33. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14,** 128 (2013).

Acknowledgements

We thank our lab colleagues and Derk ten Berge for critical discussions of our data and Danny Huylebroeck for comments on the manuscript. This work was supported by EMBO LongTerm Fellowship (ALTF 260-2009), ZonMW Dutch Medical Research Council (VENI 916-12-088, VICI 911-09-036), NIRM (Dutch Innovation Grant FES0908), NIH (RO37 DK54077), Erasmus MC Fellowship (103.494) and Landsteiner Society for Blood Research (1109).

Author contributions

The study was designed and developed by E.D. and M.C. M.C., C.S.V., T.Y.I., and K.B. performed research. W.v.I. performed RNA sequencing and P.S.K. analysed and interpreted the RNAseq data and part of the flow cytometry data. R.v.d.L. collected data. R.M., S.d.S. and C.M. provided reagents. M.C. and E.D. designed the experiments, analysed and interpreted the data and wrote the manuscript.

Additional information

Accession codes: The RNA-seq data have been deposited in the Gene Expression Omnibus (NCBI) database with accession code GSE70132.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Crisan, M. *et al.* BMP signalling differentially regulates distinct hematopoietic stem cell types. *Nat. Commun.* 6:8040 doi: 10.1038/ncomms9040 (2015).

nature COMMUNICATIONS

DOI: 10.1038/ncomms9793 OPEN

Corrigendum: BMP signalling differentially regulates distinct haematopoietic stem cell types

Mihaela Crisan, Parham Solaimani Kartalaei, Chris S. Vink, Tomoko Yamada-Inagawa, Karine Bollerot, Wilfred van IJcken, Reinier van der Linden, Susana M. Chuva de Sousa Lopes, Rui Monteiro, Christine Mummery & Elaine Dzierzak

Nature Communications 6:8040 doi: 10.1038/ncomms9040 (2015); Published 18 Aug 2015; Updated 29 Oct 2015

In the original version of this Article, the middle initial of the author Chris S. Vink was omitted from the author information. This has now been corrected in both the PDF and HTML versions of the Article.

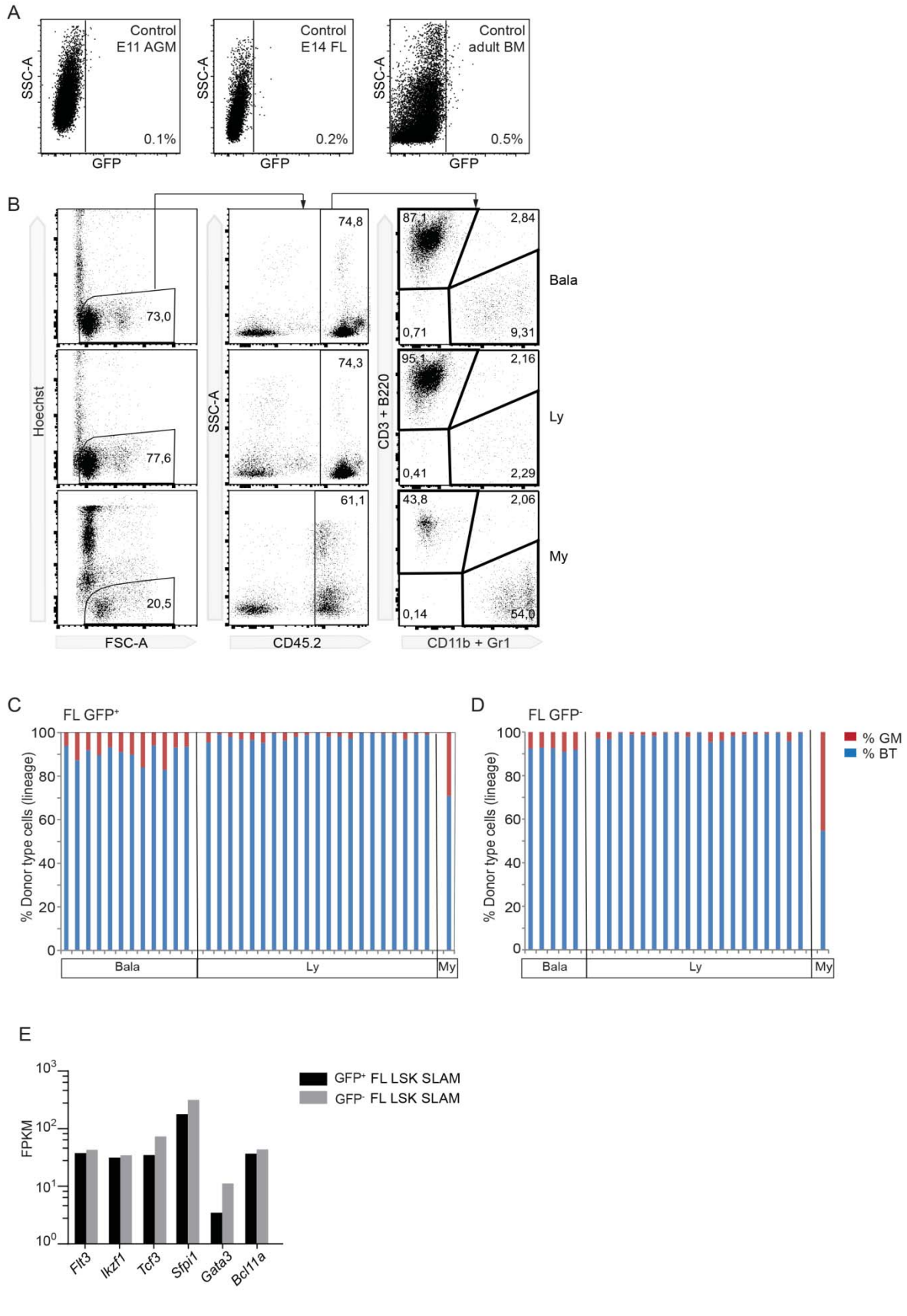


Figure S1

Supplementary Fig. 1

Supplementary Fig. **1**. **Analysis of lineage output composition and gene expression of the two hematopoietic stem cell types. (A)** Flow cytometric negative control from non-transgenic E11 AGM, E14 FL and adult BM. **(B)** Representative example of flow cytometric analysis of a myeloid-lymphoid balanced HSC (Bala, upper row), lymphoid-biased HSC (Ly, middle row) and myeloid-biased HSC (My, bottom row) . Sorted GFP^+ and GFP^- adult BM (from 14-32 week old) or E14 FL cells from *BRE-GFP* (Ly5.2) mice were injected intravenously into Ly5.1 irradiated (9Gy) recipients and donor chimerism measured at 4 months post-transplantation. Peripheral blood was analyzed for granulocyte-macrophage (GM) lineages (Gr-1 and CD11b) and B and T cell lineages (B220 and CD3) in the donor Ly5.2 cell marker expressing cells. Ratio between the Lymphoid (B+T) and Myeloid (GM) donor-derived GM and B+T (donor-derived) was calculated according to Cho et al[1] to reveal the clonal composition of the HSCs donor-derived in each recipient: My (0-4) Bala (4-16) and Ly (>16). Summary of lineage output analysis of repopulated recipient Ly5.1 mice (9Gy) injected with E14FL BRE-GFP Ly5.2 (C) GFP+ and (D) GFP- cells at the clonal HSC level (n=4). Donor chimerism and lineage output in peripheral blood was measured at 4 months post-transplantation. (E) FPKM values for selected lymphoid genes expressed by GFP^+ and GFP^- LSK SLAM sorted HSCs from E14 FL. SSC-A=side scatter analysis

SupplementaryTables

Supplementary **Table 1.** *In vivo* limiting dilution transplantation results of primary recipients receiving GFP^+ and GFP^- cells sorted from E14 fetal liver.

Cell Type	Cells ($x10^3$) injected	Repopulated/ transplanted mice	% Donor Chimerism	HSC Frequency	Average number HSCs/E14 FL
GFP^+	3	5/5	100	1/180*	2611**
	1	3/3	43-100		
	0.5	3/3	85-100		
	0.25	8/10	13-100		
	0.125	1/6	12		
	0.050	2/6	68-100		
	0.025	1/4	24		
	0.010	0/2	0		
GFP^-	388	3/3	100	1/20545*	595
	100	3/3	100		
	30	5/8	17-100		
	10	3/6	19-100		
	6	2/6	22-100		
	3	1/3	69		
	1	0/3	0		
	0.5	0/3	0		
	0.05	0/3	0		

* 95% Confidence Interval (lower/upper) frequency: for GFP^+HSC (1 in 309/1 in 104) and for GFP^- HSC (1 in 38314/1 in 11017), n=4.
** p = 0.0001
FL=fetal liver

Supplementary **Table 2.** *In vivo* limiting dilution transplantation results of primary recipients receiving GFP^+ and GFP^- cells sorted from adult bone marrow.

Cell Type	Cells ($x10^4$) injected	Repopulated/ transplanted mice	% Donor Chimerism	HSC Frequency	Average number HSCs/BM
GFP^+	1	6/6	10-75	1/10053*	110
	0.85	1/6	34		
	0.65	2/6	11,40		
	0.50	2/6	21,65		
	0.10	0/3	0		
GFP^-	10	6/6	22-89	1/17760*	1064
	5	5/6	22-79		
	2.50	5/6	5-27		
	1	1/6	11		

* 95% Confidence Interval (lower/upper) frequency: for GFP^+HSC (1 in 16995/1 in 5947) and for GFP^- HSC (1 in 32925/1 in 9580), n=3. BM=bone marrow

Supplementary **Table 3.** Lineage output HSC distribution (percentages and absolute number) in the E14 FL and BM (BRE-GFP$^+$, BRE-GFP$^-$ and Total).

E14 FL	BRE GFP+			BRE GFP-			Total	
HSC subtype	subtype/total reconstituted	% subtype	Absolute cell no.	subtype/total reconstituted	% subtype	Absolute cell no.	Absolute cell no.	% subtype
Balanced	12/34	35	921	5/25	20	119	1042	33
Lymphoid	21/34	62	1612	19/25	76	452	2064	64
Myeloid	1/34	3	78	1/25	4	23	101	3
Total	34	100	2611	25	100	596	3207	100

BM	BRE GFP+			BRE GFP-			Total	
HSC subtype	subtype/total reconstituted	% subtype	Absolute cell no.	subtype/total reconstituted	% subtype	Absolute cell no.	Absolute cell no.	% subtype
Balanced	7/17	41	45	0/15	0	0	45	4
Lymphoid	8/17	47	52	8/15	53	567	619	53
Myeloid	2/17	12	13	7/15	47	497	510	43
Total	17	100	110	15	100	1064	1174	100

Supplementary Reference

1 Cho, R. H., Sieburg, H. B. & Muller-Sieburg, C. E. A new mechanism for the aging of hematopoietic stem cells: aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111**, 5553-5561, (2008).