

# **Endothelial-specific Gata3 expression is required for hematopoietic stem cell generation**

Nada Zaidan<sup>1,2,†</sup>, Leslie Nitsche<sup>1</sup>, Evangelia Diamanti<sup>2</sup>, Rebecca Hannah<sup>2</sup>, Antonella Fidanza<sup>1</sup>, Nicola K. Wilson<sup>2</sup>, Lesley M. Forrester<sup>1</sup>, Berthold Göttgens<sup>2</sup> and Katrin Ottersbach<sup>1\*</sup>

<sup>1</sup>Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, EH16 4UU, UK

<sup>2</sup>Department of Haematology, Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 0AW, UK

<sup>†</sup>Current address: KACST-BWH Centre of Excellence for Biomedicine, Joint Centers of Excellence Program, King Abdulaziz City for Science and Technology, Riyadh 12354, Saudi Arabia

\*Correspondence: [katrin.ottersbach@ed.ac.uk](mailto:katrin.ottersbach@ed.ac.uk)

## Summary

To generate sufficient numbers of transplantable hematopoietic stem cells (HSCs) *in vitro*, a detailed understanding of how this process takes place *in vivo* is essential. The endothelial-to-hematopoietic transition (EHT), which culminates in the production of the first HSCs, is a highly complex process during which key regulators are switched on and off at precise moments and which is embedded into a myriad of microenvironmental signals from surrounding cells and tissues. We have previously demonstrated an HSC-supportive function for GATA3 within the sympathetic nervous system and the sub-aortic mesenchyme, but show here that it also plays a cell-intrinsic role during the EHT. It is expressed in hemogenic endothelial cells and early HSC precursors, where its expression correlates with a more quiescent state. Importantly, endothelial-specific deletion of *Gata3* shows that it is functionally required for these cells to mature into HSCs, placing GATA3 at the core of the EHT regulatory network.

## Introduction

The generation of the first transplantable HSCs initiates in the dorsal aorta of the aorta-gonad-mesonephros (AGM) region at E10.5 in mouse embryos (de Bruijn et al., 2000; Medvinsky and Dzierzak, 1996). It involves activating a hematopoietic transcriptional program in a subset of endothelial cells (ECs), as a result of a number of internal and external signals, which drives these cells to adopting a hematopoietic fate (Ottersbach, 2019). These so-called hemogenic endothelial cells (HECs) undergo major morphological changes, resulting in the appearance of intra-aortic cell clusters in which HECs further mature from pro-HSCs to type I and type II pre-HSCs, before becoming fully mature HSCs (Rybtsov et al., 2014). While it is known that the two transcription factors RUNX1 and GATA2 are essential for the EHT (Chen et al., 2009; de Pater et al., 2013), it is clear that they do not act alone. For example, SOX17 is required upstream for arterial fate and HEC specification (Clarke et al., 2013; Corada et al., 2013), while GFI1 and GFI1B are downstream targets of RUNX1, necessary for EHT completion (Thambyrajah et al., 2016). How all of these transcription factors interact and whether they form multi-component complexes, however, remains unknown.

In this study, we reveal the transcription factor GATA3 as another important regulator of the EHT. Its expression is upregulated in HECs and early HSC precursors, thereby enriching for hemogenic potential, but is downregulated before fully mature HSCs are formed. Through RNA-Seq analysis we show co-expression of *Gata3* with many important EHT regulators and have been able to link GATA3 expression with a more quiescent cell state. Importantly, endothelial-specific deletion of *Gata3* significantly reduces hematopoietic stem and progenitor cell (HSPC) formation, which, together with its hematopoiesis-supportive role in the co-developing sympathetic nervous system (SNS) (Fitch et al., 2012) and the sub-aortic mesenchyme (Fitch et al., 2020), gives GATA3 a multi-faceted, central role in HSC generation.

## Results

### GATA3 is expressed in a subset of endothelial and hematopoietic cells

We had previously detected GATA3 expression in a subset of ECs (Fitch et al., 2012) (**Figure 1A, red arrowheads**), often in the vicinity of intra-aortic cell clusters (**Figure 1A, yellow arrowheads**), which suggested that GATA3 may be expressed in HECs. We carried out a more careful analysis using a *Gata3*-GFP reporter mouse line (Grote et al., 2006). Immunohistochemical co-staining for GATA3 with an anti-GATA3 and an anti-GFP antibody confirmed this construct to be a faithful reporter (**Figure S1A-F**), that recapitulated the expression of *Gata3* in individual ECs (**Figure 1B, arrows**) and at the base of intra-aortic clusters (**Figure 1B, yellow arrowhead**), but not within clusters. Flow cytometry analysis confirmed that a fraction (6.4%) of ECs (VE-CADHERIN [VEC]<sup>+</sup>) expresses *Gata3*-GFP at E10.5 when

HEC frequency is at its highest (**Figure 1C**). Interestingly, more than 40% of these *Gata3*-GFP+ ECs also express the hematopoietic markers CD45 (late hematopoietic marker) and/or CD41 (early marker), with 17% expressing both (**Figure 1D**). The percentage of CD45+ *Gata3*-GFP+ cells was much lower within the VEC- fraction, suggesting that the majority are VEC+ hematopoietic stem and progenitor cells (HSPCs) or HSC precursors. This percentage increases substantially at E11.5 (**Figure 1E-G**). To further investigate whether the *Gata3*-GFP+ fraction contains phenotypic HSC precursors, we employed the nomenclature developed by the Medvinsky lab (Rybtsov et al., 2014), with pro-HSCs defined as VEC+CD41+CD43-CD45-, pre-HSC type I as VEC+CD41+CD43+CD45- and pre-HSC type II as VEC+CD41+CD43+CD45+ (**Figure S1G**). All three populations were present in the *Gata3*-GFP+ fraction (**Figure 1H**).

To determine if the *Gata3*-GFP+ hematopoietic population at E10.5 contains HSPCs, we sorted them (**Figure S1H**) and plated them directly in colony-forming (CFU-C) assays alongside *Gata3*-GFP- hematopoietic cells (HCs) and *Gata3*-GFP+/- ECs. As expected, ECs do not produce hematopoietic colonies when plated directly in methylcellulose (**Figure 1I**). Interestingly, only the *Gata3*-GFP- HCs gave rise to colonies, but not the GFP+ fraction, suggesting that the latter consists either of mature hematopoietic cells or very early precursors that require further maturation towards the hematopoietic fate.

### **GATA3 expression enriches for hemogenic endothelial activity**

Inherent hemogenic potential can be revealed via a co-culture step on OP9 stromal cells (Swiers et al., 2013). *Gata3*-GFP+/- ECs were sorted (**Figure S1H**) and cultured with OP9 cells. Cells were then assessed for hematopoietic marker expression and progenitor potential (**Figure 2A**). Both EC populations were able to give rise to hematopoietic cells (**Figure 2B**), with a trend towards higher production of CD45+ cells from *Gata3*-GFP+ ECs (**Figure 2C**). In CFU-C assays, however, *Gata3*-GFP+ ECs had a noticeably higher colony output (**Figure 2D**), which was highly significant for all colony types.

### **GATA3 marks HSC precursors**

We also tested the potential of the *Gata3*-GFP+/- HC populations (**Figure 2A, S1H**) in co-cultures. Unsurprisingly, they gave an almost entirely CD45+ output (**Figure 2E**). Interestingly, while overall progenitor potential was higher in the *Gata3*-GFP- fraction (**Figure 2F**), splitting this into individual progenitor types revealed that *Gata3* expression enriches specifically for the most immature CFU-GEMM progenitor (**Figure 2G-I**), suggesting that *Gata3*-GFP may mark HSC precursors.

The intermediate steps involved in the maturation of HECs into transplantable HSCs have been dissected with the help of a culture system developed in the Medvinsky lab (Rybtsov et al., 2014;

Rybtsov et al., 2011). It involves aggregation of sorted cell populations with OP9 cells and culturing these co-aggregates in the presence of cytokines. The potential of the initial populations is then assessed by flow cytometry and transplantations (**Figure 2J**). Pro-HSCs emerge from E9 in the AGM region, while type I pre-HSCs are detected at E10-11. Pre-HSC type II emerge towards E11 (Rybtsov et al., 2014). *Gata3*-GFP +/- ECs and HCs were sorted (**Figure S2A**) and cultured as OP9 co-aggregates (**Figure 2J**). All four populations were able to generate CD45+ HCs, with the sorted HC fractions producing substantially more CD45+ cells than the EC populations (**Figure 2K**). ECs are unable to transdifferentiate into transplantable HSCs in these culture conditions (Rybtsov et al., 2014), and indeed neither *Gata3*+ nor *Gata3*- ECs produced any detectable chimerism in transplant recipients (**Figure 2L,M**). Intriguingly, repopulation activity at E9.5-E10.5 was restricted to the *Gata3*-GFP+ HC fraction, which completely shifted to the *Gata3*-GFP- fraction at E11.5. All of the results taken together with our previous data (Fitch et al., 2012) imply that *Gata3* expression is switched on in HECs, with *Gata3* expression continuing during the early stages of HSC maturation, but being switched off at the pre-HSC type II stage and remaining off in emerging HSCs (**Figure S2B**).

#### ***Gata3*-expressing cells show an enrichment for key regulators of EHT**

To get a better understanding of how GATA3 may be involved in the EHT, we performed RNA-Seq on small pools (20 cells/pool; 20 pools/population) of sorted *Gata3*-GFP+ ECs and HCs and compared their transcriptome to the *Gata3*-GFP- fractions (**Figure 3A, Figure S1H**). Principal Component Analysis showed a clear separation of ECs from HCs along component 1, with each population displaying a distinct subdivision according to *Gata3* expression along component 2 (**Figure 3B**). Differential expression analysis revealed a higher number of genes being upregulated in *Gata3*+ cells (EC: 557; HC: 1527) than were downregulated (EC: 253; HC: 232), with some overlap between the two cell types (**Figure 3C; Table S1**). Gene ontology analysis of the upregulated genes saw a significant enrichment of processes associated with stem cell development and differentiation and migration in both cell populations, which may be a reflection of these cells undergoing morphological changes during EHT (**Figure S3**). Genes that were downregulated in the *Gata3*+ HC fraction were largely associated with differentiated blood lineages (e.g. *Icos*, *Irf8*, *Klf4*, *Ccl3* and *Il6ra*), which reflects the HSC precursor status of the *Gata3*+ cells. Interestingly, *Gata3*+ ECs displayed a downregulation of genes linked to the positive regulation of cell cycle and proliferation.

To understand the position of GATA3 in the network of EHT regulators, we analyzed the expression of regulators that mark key stages of HSC generation (Ottersbach, 2019) (**Figure S2B**). *Sox17* is upregulated in arterial ECs and required for HEC specification, but subsequently needs to be downregulated for the EHT to proceed. This is confirmed in our dataset, where it is absent in HCs, but

upregulated in *Gata3*<sup>+</sup> ECs (**Figure 3D**). *Runx1*, *Gata2* and *Gfi1* are key transcription factors for initiating the hematopoietic transcriptional program in HEC, which is reflected in their upregulation in *Gata3*<sup>+</sup> ECs, with *Gata2* showing the most widespread expression. *Runx1* is much more highly expressed in HCs (*Gata3*<sup>+</sup> and – HCs), while *Gata2* and *Gfi1* become downregulated, which is consistent with the literature (Ottersbach, 2019). *Gfi1b* is involved in the final stages of the EHT, with expression restricted to intra-aortic clusters. Accordingly, its expression was confined to the *Gata3*<sup>+</sup> HC population.

Among the genes commonly upregulated in *Gata3*-GFP<sup>+</sup> ECs and HCs (**Table S1, Figure 3D**) were well-known HSC regulators such as *Hoxa10*, which was shown to drive the transition of HECs into HSCs derived from human iPSCs (Sugimura et al., 2017), *Meis1*, which promotes the reprogramming of hematopoietic progenitors to HSCs (Riddell et al., 2014), and *Mllt3*, which is essential for human HSC self-renewal (Calvanese et al., 2019) and was reported to be upregulated in cells undergoing EHT (Oatley et al., 2020). *Mecom*, a transcription factor essential for fetal and adult HSC function (Kumano and Kurokawa, 2010), that was recently, together with *Hoxa10* and *Meis1*, described as a marker for E9.5 HECs (Gao et al., 2018), was upregulated in *Gata3*-GFP<sup>+</sup> ECs. The Notch signaling pathway is essential for the EHT; however, its activity needs to be downregulated for HSCs to emerge (Gama-Norton et al., 2015), which is reminiscent of the downregulation of *Gata3*. Indeed, both *Notch1* and one of its important downstream targets, *Hey2*, were upregulated in *Gata3*-GFP<sup>+</sup> ECs.

### **Gata3 expression correlates with a more quiescent cell state**

Among the most prominent gene ontology terms associated with the genes downregulated in *Gata3*-GFP<sup>+</sup> ECs were proliferation and positive regulation of the cell cycle (**Figure S3**). We interrogated the differentially expressed genes specifically for cell cycle regulators and uncovered as a general trend an upregulation of cell cycle inhibitors and a downregulation of cell cycle promoters in *Gata3*-GFP<sup>+</sup> cells (**Figure 3E**). Furthermore, we observed a significant enrichment of quiescent cells in the *Gata3*-GFP<sup>+</sup> subsets (**Figure 3F**). Since the cell cycle inhibitor *Cdkn1c* was upregulated in both *Gata3*-GFP<sup>+</sup> cell types (**Figure 3E,G; Table S1**), we hypothesized that it was responsible for the quiescent phenotype. Indeed, there was a decrease in *Gata3*-GFP<sup>+</sup> ECs and HCs in the G0/G1 cell cycle phase from *Cdkn1c*-deficient embryos (**Figure 3H**). We also noticed that *Gata3*-GFP<sup>+</sup> ECs were generally more quiescent than HCs, suggesting that HECs may exit the cell cycle to undergo the massive morphological changes required for their transition into HCs, with emerging HCs then starting to proliferate to expand the pre-HSC pool, as reported by others (Batsivari et al., 2017; Oatley et al., 2020; Rybtsov et al., 2016). Interestingly, an upregulation of *RUNX1C*, which is normally upregulated towards the end of EHT, has been associated with an exit from the cell cycle in undifferentiated human

pluripotent stem cells (Fidanza et al., 2017). These cells, in which *RUNX1* was activated via dCAS9 promoter targeting (UniSam system) also had higher levels of *GATA3* (**Figure 3I**). This not only confirms the association of *Gata3* expression with a more quiescent state in different models and species, but may also point to a direct link between these two EHT-associated genes.

### **GATA3 function in ECs is required for normal HSPC numbers**

To confirm a functional role for GATA3 in the EHT, we crossed a conditional *Gata3* knockout line (Zhu et al., 2004) with a VEC-Cre line (Chen et al., 2009) to delete *Gata3* specifically within the endothelial lineage and its derivatives and assessed how this affected HSPC numbers in the AGM (**Figure 4A**). Heterozygous and homozygous deletion of *Gata3* significantly reduced progenitor numbers at both E10.5 and E11.5, demonstrating that EC-specific *Gata3* expression is required for HSPC formation and that this is dose dependent (**Figure 4B,C**). A complete knockout of *Gata3* resulted in a very similar reduction of progenitors, showing that it is the expression of *Gata3* in ECs that is important in this context (**Figure S4A,B**). We had previously reported that explant-culturing of AGMs rescued the HSC defect in *Gata3*<sup>+/-</sup> embryos (Fitch et al., 2012). To see if that could also rescue the progenitor defect, we added a 3-day explant culture step. Progenitor numbers in VEC-Cre+ *Gata3*<sup>+/-</sup> AGMs recovered slightly, as there was now a statistically significant difference between heterozygous and homozygous knockout AGMs (**Figure 4D,E**). This was also the case with germline-deleted embryos (**Figure S4C,D**). A substantial number of progenitors remain in AGMs where both copies of *Gata3* were deleted from ECs (**Figure 4B-E**); however, genotyping of individual colonies revealed that 7 out of 20 progenitors had escaped deletion (**Figure S4E**). This suggests that the effect on progenitors shown here is likely an underestimate, although germline-deleted embryos also retained some GATA3-independent progenitors (**Figure S4A-D**).

AGMs were also transplanted to determine the effect of endothelial-specific *Gata3* deletion on HSC numbers. Repopulation activity was significantly reduced in homozygous knockout AGMs, demonstrating that GATA3 expression in the aortic endothelium is required for HSC generation (**Figure 4F**).

### **Discussion**

This newly described cell-intrinsic role of GATA3 in the EHT adds another facet to the complex way in which GATA3 promotes HSC production in the AGM. We have previously demonstrated that GATA3 is expressed in two AGM niche compartments, the SNS (Fitch et al., 2012) and the sub-aortic mesenchyme (Fitch et al., 2020), from where it also supports emerging HSCs. The relative contribution to AGM hematopoiesis of GATA3 in these three different compartments requires further dissection,

although its role in HECs is likely to be dominant as the effect of the EC-specific knockout on HSPC numbers closely mirrors that of the germline knockout ((Fitch et al., 2012) and this manuscript). Yet, this defect was rescued through the administration of catecholamine derivatives (Fitch et al., 2012), which could be explained by the fact that HSC production in *Gata3*-null AGMs is not entirely disrupted, with the remaining HSC activity being amplified through catecholamine addition. The function of GATA3 in HECs is also likely to temporally precede its role in supporting HSCs via catecholamine production. As we show here, *Gata3* is already expressed at E9.5 in pro-HSCs, while the maturation of neural crest cells into catecholamine-secreting SNS cells only commences from E10.5. Furthermore, we have previously unveiled a reduction in HECs and intra-aortic cluster formation in the absence of GATA3 (Fitch et al., 2012), suggesting a role for GATA3 at the early stages of EHT initiation.

Another interesting question is the relationship of GATA3 with the other two major regulators of HSC emergence, GATA2 and RUNX1. We showed that GATA3 is required for *Gata2* expression in the SNS, but not in the endothelium, and the expression of *Runx1* in the AGM is reduced by half in *Gata3*-null embryos (Fitch et al., 2020; Fitch et al., 2012); however, deletion of RUNX1 and GATA2 from ECs has a much more profound negative impact on HSPC numbers (Chen et al., 2009; de Pater et al., 2013). Data from a recent scRNA-Seq analysis of aortic EC subpopulations show an earlier upregulation of *Gata3* in arterial ECs as compared with *Runx1*, *Gata2* and *Gfi1* (Hou et al., 2020). Instead, it seems to closely mirror the expression pattern of *Hey2*, *Sox7* and *Sox17*, which need to be downregulated for HSCs to mature (Lizama et al., 2015). This expression pattern of *Gata3* is conserved in humans (Calvanese et al., 2022).

*Gata3* has been shown to be expressed in adult quiescent LT-HSCs where it regulates their entry into the cell cycle (Frelin et al., 2013; Ku et al., 2012). scRNA-Seq datasets from mouse embryos and human PSCs revealed that EC and HEC are more quiescent, but enter the cell cycle at the end of the EHT process (Canu et al., 2020; Fadlullah et al., 2022; Oatley et al., 2020; Zeng et al., 2019). The subsequent increase in cycling as HSC precursors mature is also supported by data from a cell cycle reporter mouse (Batsivari et al., 2017) and coincides with the stage at which *Gata3* becomes downregulated. Importantly, when cell cycle progression of PSC-derived ECs was chemically blocked, these cells could no longer generate HSPCs, implying that cell cycle entry is required to complete EHT. Taken all of these data into account, GATA3 may thus keep ECs/HECs quiescent while they undergo the major morphological changes required for the transdifferentiation into HCs, and then, in analogy to its reported role in adult HSCs, may promote their re-entry into the cell cycle for completion of the EHT.

## **Experimental procedures**

### **Mice**

All animal work was carried out under a UK Home Office-approved licence and following local ethical approval. Males and females of *Gata3-LacZ* knock-in mice (van Doorninck et al., 1999), *Gata3-GFP* knock-in mice (Grote et al., 2006), conditional *Gata3* knockout mice (Zhu et al., 2004), *p57Kip2* knockout mice (Zhang et al., 1997), VEC-Cre transgenic mice (Chen et al., 2009) and C57BL/6J mice were crossed to obtain embryos of the desired stage and genotype. The day of vaginal plug detection was considered as E0.5. The developmental stage of embryos was specified by either counting somite pairs (E9.5-E10.5) or by eye pigmentation (E11.5). Embryos smaller than their littermates or lacking a heartbeat were excluded.

### **AGM explant cultures**

AGMs were dissected and cultured on Durapore filters (Millipore) on M5300 long-term culture medium (Stem Cell Technologies) supplemented with  $10^{-6}$ M hydrocortisone (Sigma). After 3 days, AGMs were dissociated with 0.125% collagenase (Alfa Aesar).

### **OP9 co-cultures**

These were carried out according to Swiers et al., 2013, with further details provided in the Supplemental Experimental Procedures.

### **Co-aggregation cultures**

AGM cells from E9.5-11.5 embryos were sorted and co-aggregated at 1 embryo equivalent (e.e.) with OP9 cells according to Rybtsov et al (Rybtsov et al., 2011). Cultured aggregates were dissociated using collagenase and either analyzed by flow cytometry or transplanted into irradiated mouse recipients.

### **Colony-forming assays**

Dissociated AGM cells were plated in triplicates (65.000 cells/plate) in methylcellulose (M3434; Stem cell technologies), incubated at 37°C and colonies scored 7 days later.

To detect *Gata3* deletion by VEC-Cre, individual colonies were picked and analysed by PCR, using the following primers: forward, CAGTCTCTGGTATTGATCTGCTTCTT; and reverse, GTGCAGCAGAGCAGGAAACTCTCAC.

### **Transplantations**

Single-cell suspensions were injected intravenously into irradiated (split dose of 460-475 rad, Caesium source) recipients together with  $2 \times 10^5$  spleen cells (for direct transplantations) or  $2 \times 10^4$  bone marrow cells (for co-aggregates). Recipients were CD45.1/.2 or CD45.1/.1 and the donors were

CD45.2/2 on a C57BL6J background. Recipient blood was analyzed at 1 and 4 months post-transplantation.

### **Flow cytometry**

Antibody stainings were performed 30min on ice in the dark, with further details on antibodies, controls and equipment provided in the Supplemental Experimental Procedures.

For cell cycle analysis, sorted cells were incubated for 1min at room temperature 1:1 with DAPI staining solution (5ug/ml DAPI (Sigma) and 1% (v/v) Nonidet P40 (Sigma) in dH<sub>2</sub>O) and analyzed on an LSRFortessa (BD Bioscience).

### **Immunohistochemistry**

Cryosections were prepared and stained as described previously (Fitch et al., 2012) with further details provided in the Supplemental Experimental Procedures.

### **X-gal staining**

*Gata3*<sup>l<sup>z</sup>/+</sup> embryos were fixed for 1h with PBS/10% Formal Saline/0.2% Glutaraldehyde/2mM MgCl<sub>2</sub>/5mM EGTA/0.02% NP40 at 4°C and stained overnight at room temperature with 5mM K<sub>3</sub>Fe(CN)<sub>6</sub>/5 mM K<sub>4</sub>Fe(CN)<sub>6</sub>/2 mM MgCl<sub>2</sub>/0.01% Nadeoxycholate/0.02% NP40/0.1% X-gal (all in PBS). Stained embryos were cryopreserved and sectioned as above and sections counterstained with Neutral Red.

### **Endogenous gene activation in human iPSCs**

Human iPSCs were cultured and transfected with the *RUNX1C*-activating UniSAM system as described previously (Fidanza et al., 2017), with experimental details and qPCR primers provided in the Supplemental Experimental Procedures.

### **RNA sequencing**

Libraries for RNA sequencing were prepared according to the protocol by Picelli et al. (Picelli et al., 2014). Cells were initially sorted from 3-5 biological replicates into tubes based on their populations: *Gata3*-GFP+ EC, *Gata3*-GFP- EC, *Gata3*-GFP+ HC and *Gata3*-GFP- HC. Each population was then sorted again into 96 wells plate, 20 cells/well (20 pools per population), containing 2.3μl of lysis buffer (0.2% RNase inhibitor (Ambion, Thermo Fisher Scientific) in Triton X-100 (Sigma)). Details on the reverse transcription, PCR pre-amplification, library preparation steps and data analysis are provided in the Supplemental Experimental Procedures.

### **Statistical Analysis**

Graph preparations and statistical analysis were performed using GraphPad Prism. The Mann-Whitney test was used for transplantation experiments, paired t-test for colony forming assay following OP9 co-culture and co-aggregates, and two-way ANOVA test for colony forming assay.

### **Data availability**

The data have been deposited in NCBI's Gene Expression Omnibus (GEO) repository under accession number GSE114926.

### **Acknowledgments**

The authors are very grateful to the staff of the animal facilities both at the Cambridge Institute for Medical Research and the Centre for Regenerative Medicine for their support with animal experiments and to the flow cytometry teams at both of these institutes, Dr. Reiner Schulte, Dr. Chiara Cossetti, and Michal Maj in Cambridge and Fiona Rossi and Dr Claire Cryer in Edinburgh, as well as the Cambridge NIHR BRC Cell Phenotyping Hub, for excellent cell sorting services and help with flow cytometry analyses. We would further like to acknowledge the assistance of Matthew Gratian and Mark Bowen (Cambridge Institute for Medical Research) and Bertrand Vernay (Centre for Regenerative Medicine) with microscopy, and Drs Céline Souilhoul and Stanislav Rybtsov with setting up the co-aggregate assays. We are also indebted to Jingfang Zhu and Meinrad Busslinger for providing the conditional *Gata3* KO and the *Gata3*-GFP mice, respectively. Core facilities at the Edinburgh Centre for Regenerative Medicine were supported by centre grant MR/K017047/1. This work was funded by an Intermediate Fellowship from the Kay Kendall Leukaemia Fund (to K.O.), a Blood Cancer UK Bennett Senior Fellowship (10015 to K.O.), a fellowship from the King Abdullah International Medical Research Centre (KAIMRC), Ministry of National Guard Health Affairs (to N.Z.) and a Wellcome Trust studentship (to L.N.). This research was also funded in part by the Wellcome Trust and the UKRI Medical Research Council. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

### **Author contribution**

N.Z. performed and designed the majority of experiments; E.D. and R.H. performed bioinformatics analyses; N.K.W. and B.G. provided advice and assistance with scRNA-Seq experiment; L.N. and A.F. performed experiments; L.M.F. provided important reagents; K.O. conceived and supervised the study and wrote the manuscript.

## Declaration of interest

The authors have no conflicts of interest to declare.

## References

- Batsivari, A., Rybtsov, S., Souilhol, C., Binagui-Casas, A., Hills, D., Zhao, S., Travers, P., and Medvinsky, A. (2017). Understanding Hematopoietic Stem Cell Development through Functional Correlation of Their Proliferative Status with the Intra-aortic Cluster Architecture. *Stem Cell Reports* 8, 1549-1562.
- Calvanese, V., Capellera-Garcia, S., Ma, F., Fares, I., Liebscher, S., Ng, E.S., Ekstrand, S., Aguade-Gorgorio, J., Vavilina, A., Lefaudeux, D., *et al.* (2022). Mapping human haematopoietic stem cells from haemogenic endothelium to birth. *Nature* 604, 534-540.
- Calvanese, V., Nguyen, A.T., Bolan, T.J., Vavilina, A., Su, T., Lee, L.K., Wang, Y., Lay, F.D., Magnusson, M., Crooks, G.M., *et al.* (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. *Nature* 576, 281-286.
- Canu, G., Athanasiadis, E., Grandy, R.A., Garcia-Bernardo, J., Strzelecka, P.M., Vallier, L., Ortmann, D., and Cvejic, A. (2020). Analysis of endothelial-to-haematopoietic transition at the single cell level identifies cell cycle regulation as a driver of differentiation. *Genome Biol* 21, 157.
- 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.
- Clarke, R.L., Yzaguirre, A.D., Yashiro-Ohtani, Y., Bondue, A., Blanpain, C., Pear, W.S., Speck, N.A., and Keller, G. (2013). The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol* 15, 502-510.
- Corada, M., Orsenigo, F., Morini, M.F., Pitulescu, M.E., Bhat, G., Nyqvist, D., Breviario, F., Conti, V., Briot, A., Iruela-Arispe, M.L., *et al.* (2013). Sox17 is indispensable for acquisition and maintenance of arterial identity. *Nat Commun* 4, 2609.
- de Bruijn, M.F., Speck, N.A., Peeters, M.C., and Dzierzak, E. (2000). Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J* 19, 2465-2474.
- 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.
- Fadlullah, M.Z.H., Neo, W.H., Lie, A.L.M., Thambyrajah, R., Patel, R., Mevel, R., Aksoy, I., Do Khoa, N., Savatier, P., Fontenille, L., *et al.* (2022). Murine AGM single-cell profiling identifies a continuum of hemogenic endothelium differentiation marked by ACE. *Blood* 139, 343-356.
- Fidanza, A., Lopez-Yrigoyen, M., Romano, N., Jones, R., Taylor, A.H., and Forrester, L.M. (2017). An all-in-one UniSam vector system for efficient gene activation. *Sci Rep* 7, 6394.
- Fitch, S.R., Kapeni, C., Tsitsopoulou, A., Wilson, N.K., Gottgens, B., de Bruijn, M.F., and Ottersbach, K. (2020). Gata3 targets Runx1 in the embryonic haematopoietic stem cell niche. *IUBMB Life* 72, 45-52.
- Fitch, S.R., Kimber, G.M., Wilson, N.K., Parker, A., Mirshekar-Syahkal, B., Gottgens, B., Medvinsky, A., Dzierzak, E., and Ottersbach, K. (2012). Signaling from the sympathetic nervous system regulates hematopoietic stem cell emergence during embryogenesis. *Cell Stem Cell* 11, 554-566.
- Frelin, C., Herrington, R., Janmohamed, S., Barbara, M., Tran, G., Paige, C.J., Benveniste, P., Zuniga-Pflucker, J.C., Souabni, A., Busslinger, M., *et al.* (2013). GATA-3 regulates the self-renewal of long-term hematopoietic stem cells. *Nat Immunol* 14, 1037-1044.
- Gama-Norton, L., Ferrando, E., Ruiz-Herguido, C., Liu, Z., Guiu, J., Islam, A.B., Lee, S.U., Yan, M., Guidos, C.J., Lopez-Bigas, N., *et al.* (2015). Notch signal strength controls cell fate in the haemogenic endothelium. *Nat Commun* 6, 8510.
- Gao, L., Tober, J., Gao, P., Chen, C., Tan, K., and Speck, N.A. (2018). RUNX1 and the endothelial origin of blood. *Exp Hematol* 68, 2-9.

Grote, D., Souabni, A., Busslinger, M., and Bouchard, M. (2006). Pax 2/8-regulated Gata 3 expression is necessary for morphogenesis and guidance of the nephric duct in the developing kidney. *Development* *133*, 53-61.

Hou, S., Li, Z., Zheng, X., Gao, Y., Dong, J., Ni, Y., Wang, X., Li, Y., Ding, X., Chang, Z., *et al.* (2020). Embryonic endothelial evolution towards first hematopoietic stem cells revealed by single-cell transcriptomic and functional analyses. *Cell Res* *30*, 376-392.

Ku, C.J., Hosoya, T., Maillard, I., and Engel, J.D. (2012). GATA-3 regulates hematopoietic stem cell maintenance and cell-cycle entry. *Blood* *119*, 2242-2251.

Kumano, K., and Kurokawa, M. (2010). The role of Runx1/AML1 and Evi-1 in the regulation of hematopoietic stem cells. *J Cell Physiol* *222*, 282-285.

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

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

Oatley, M., Bolukbasi, O.V., Svensson, V., Shvartsman, M., Ganter, K., Zirngibl, K., Pavlovich, P.V., Milchevskaya, V., Foteva, V., Natarajan, K.N., *et al.* (2020). Single-cell transcriptomics identifies CD44 as a marker and regulator of endothelial to haematopoietic transition. *Nat Commun* *11*, 586.

Ottersbach, K. (2019). Endothelial-to-haematopoietic transition: an update on the process of making blood. *Biochem Soc Trans* *47*, 591-601.

Picelli, S., Faridani, O.R., Bjorklund, 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.

Riddell, J., Gazit, R., Garrison, B.S., Guo, G., Saadatpour, A., Mandal, P.K., Ebina, W., Volchkov, P., Yuan, G.C., Orkin, S.H., *et al.* (2014). Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell* *157*, 549-564.

Rybtsov, S., Batsivari, A., Bilotkach, K., Paruzina, D., Senserrich, J., Nerushev, O., and Medvinsky, A. (2014). Tracing the origin of the HSC hierarchy reveals an SCF-dependent, IL-3-independent CD43(-) embryonic precursor. *Stem Cell Reports* *3*, 489-501.

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

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

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

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.

Thambyrajah, R., Mazan, M., Patel, R., Moignard, V., Stefanska, M., Marinopoulou, E., Li, Y., Lancrin, C., Clapes, T., Moroy, T., *et al.* (2016). GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nat Cell Biol* *18*, 21-32.

van Doorninck, J.H., van Der Wees, J., Karis, A., Goedknegt, E., Engel, J.D., Coesmans, M., Rutteman, M., Grosveld, F., and De Zeeuw, C.I. (1999). GATA-3 is involved in the development of serotonergic neurons in the caudal raphe nuclei. *J Neurosci* *19*, RC12.

Zeng, Y., He, J., Bai, Z., Li, Z., Gong, Y., Liu, C., Ni, Y., Du, J., Ma, C., Bian, L., *et al.* (2019). Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. *Cell Res* *29*, 881-894.

Zhang, P., Liegeois, N.J., Wong, C., Finegold, M., Hou, H., Thompson, J.C., Silverman, A., Harper, J.W., DePinho, R.A., and Elledge, S.J. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome. *Nature* *387*, 151-158.

Zhu, J., Min, B., Hu-Li, J., Watson, C.J., Grinberg, A., Wang, Q., Killeen, N., Urban, J.F., Jr., Guo, L., and Paul, W.E. (2004). Conditional deletion of *Gata3* shows its essential function in T(H)1-T(H)2 responses. *Nat Immunol* 5, 1157-1165.

## Figure legends

### Figure 1 *Gata3* is expressed in a subset of endothelial and hematopoietic cells.

(A) Cryosections of E10.5 *Gata3<sup>lox/+</sup>* with *Gata3*-LacZ expression in blue and counterstained with Neutral Red. Yellow arrowheads point to *Gata3*<sup>+</sup> cells near intra-aortic clusters, red arrowheads highlight *Gata3*<sup>+</sup> ECs. (B) Section from a *Gata3*-GFP<sup>+</sup> E11.5 embryo stained with CD34 (green), *Gata3*-GFP (magenta) and TH (yellow). Orange box indicates area shown at higher magnification on the right. The yellow arrowhead points to a *Gata3*<sup>+</sup> cell at the base of an intra-aortic cluster. The orange arrows in the higher magnification image highlight *Gata3*<sup>+</sup> ECs. AG: adrenal anlage, DA: dorsal aorta, SG: sympathetic ganglia, MD: mesonephric duct. Flow cytometry analysis of *Gata3*-GFP<sup>+</sup> E10.5 AGMs, stained for GFP, VEC, CD41 and CD45. Representative flow plots shown in (C) and results from 3 biological replicates shown in (D). Flow cytometry analysis of *Gata3*-GFP<sup>+</sup> E11.5 AGMs, stained for GFP, VEC, CD41 and CD45. Representative flow plots shown in (E) and results from 4 biological replicates shown in (F). 1 biological replicate = 1 embryo. (G) Increase of VEC+GFP+CD41+CD45<sup>+</sup> cells from E10.5 to E11.5. (H) Percentage of pro-HSCs (VEC+CD41+CD43-CD45<sup>-</sup>), pre-HSC type I (VEC+CD41+CD43+CD45<sup>-</sup>) and pre-HSC type II (VEC+CD41+CD43+CD45<sup>+</sup>) within the *Gata3*-GFP<sup>+</sup> fraction of E10.5 AGMs. n=3; 3 embryos used. An unpaired t-test was performed; \* p<0.05; \*\* p<0.01; \*\*\* p<0.001. (I) CFU-C assay of freshly sorted and directly plated G3-EC: CD41/45-GFP-VEC<sup>+</sup>, G3+EC: CD41/45-GFP+VEC<sup>+</sup>, G3-HC: CD41/45+GFP<sup>-</sup>, and G3+HC: CD41/45+GFP<sup>+</sup>; 3 technical replicates of 1 biological replicate.

See also Figure S1

### Figure 2 *Gata3* enriches for HECs and marks early HSC precursors.

(A) Outline of co-culture experiments. (B) Flow cytometry analysis of hematopoietic output from *Gata3*<sup>+/-</sup> ECs. Total CD45<sup>+</sup> cells (flow cytometry) (C) and hematopoietic progenitors (CFU-C) (D) produced by *Gata3*<sup>+/-</sup> ECs; n=4, with pooled and sorted cells from 8-12 embryos used per experiment. (E) Flow cytometry analysis of CD45<sup>+</sup> cells produced from *Gata3*<sup>+/-</sup> HCs from 3 independent experiments (n=3). Total progenitors (F), CFU-M (G), CFU-GM (H) and CFU-GEMM (I) produced by *Gata3*<sup>+/-</sup> HCs; n=3, with pooled and sorted cells from 8-12 embryos used per experiment. \*\*\* p<0.001; \*\*\*\* p<0.0001; Mann-Whitney test. (J) Outline of co-aggregate experiments. (K) Total CD45<sup>+</sup> cell output (flow cytometry) from E9.5-10.5 *Gata3*<sup>+/-</sup> ECs and HCs. Percent donor contribution after 4 months in recipients of co-aggregates from E9.5-10.5 (L) and E11.5 (M) *Gata3*<sup>+/-</sup> ECs and HCs. Arrow

highlights data point from analysis after 1 month as this recipient died unexpectedly before the 4 months analysis. n=7 for E9.5-10.5. Cells were sorted from 7-13 embryos into the 4 populations, which were co-aggregated as 1 embryo equivalent (ee), with 1 co-aggregate transplanted per recipient (28 recipients in total; 7 per cell population); n=4-5 for E11.5. Cells were sorted from 6-10 embryos into the indicated populations, which were co-aggregated as 1 ee (4 co-aggregates per EC population; 5 co-aggregates per HC population), with 1 co-aggregate transplanted per recipient (18 recipients in total).

**See also Figure S1 and S2**

**Figure 3 *Gata3* expression correlates with other EHT regulators and marks quiescent cells**

(A) Cell populations analyzed by RNA-Seq. (B) Principal Component Analysis of individual samples of the 4 populations. Cells were from 3-5 biological replicates, with cells from 2-6 GFP+ embryos pooled per replicate. (C) Venn diagrams of genes upregulated and downregulated in *Gata3*-GFP+ ECs and HCs. (D) tSNE plots of the 4 cell populations colored for the expression of the indicated genes. (E) Cell cycle regulators differentially expressed in *Gata3*-GFP+ ECs and HCs. (F) Percentage of quiescent cells in the *Gata3*-GFP+/- EC and HC populations, with representative flow cytometry plots. n=4, with cells from 2-10 GFP+ embryos pooled per independent experiment. \* p<0.05; \*\* p<0.01; paired t-test. (G) tSNE plot colored for *Cdkn1c* expression. (H) Percentage of *Cdkn1c* wild-type and knockout ECs and HCs in the different cell cycle phases. n=3, using a total of 11 WT (4, 4, 3) and 11 KO (4, 4, 3) embryos. \* p<0.05; paired t-test. (I) Expression of *RUNX1C* and *GATA3* by qPCR in hiPSCs with (+) and without (-) *RUNX1C*-activating gRNA as part of the UniSam endogenous gene activation system. n=3. \* p<0.05; \*\*\* p<0.001; Mann-Whitney test.

**See also Figure S2 and S3**

**Figure 4 *GATA3* function in endothelial cells is required for normal HSPC numbers.**

(A) Outline of experiments performed with EC-specific (VEC-Cre) *Gata3* knockout embryos. Colony numbers obtained with E10.5 (B) and E11.5 (C) uncultured AGMs of the indicated genotypes. Colony numbers obtained with E10.5 (D) and E11.5 (E) AGMs of the indicated genotypes after explant culture. n=3-4, 1 embryo per replicate. \*\* p<0.01; \*\*\* p<0.001; \*\*\*\* p<0.0001; Two-way ANOVA. (F) Percent donor contribution 4 months post-transplant with E11.5 AGM cells of the indicated genotypes. f/f or f/+ n=13, f/+;+/Cre n=10, f/f;+/Cre n= 11, 1 ee per recipient. \* p<0.05; Mann-Whitney test.

**See also Figure S4**

## Highlights

- *Gata3* marks hemogenic endothelial cells and early HSC precursors
- *Gata3* expression correlates with early EHT drivers
- *Gata3* is associated with a quiescent cell state
- GATA3 is essential for the endothelial-to-hematopoietic transition

## eTOC blurb

Ottersbach and colleagues demonstrate that *Gata3* marks hemogenic endothelial cells and early hematopoietic stem cell precursors where its expression correlates with other early drivers of the endothelial-to-hematopoietic transition and a quiescent cell cycle state. Importantly, they show that GATA3 facilitates the transition of hemogenic endothelial cells into fully functional hematopoietic stem cells.

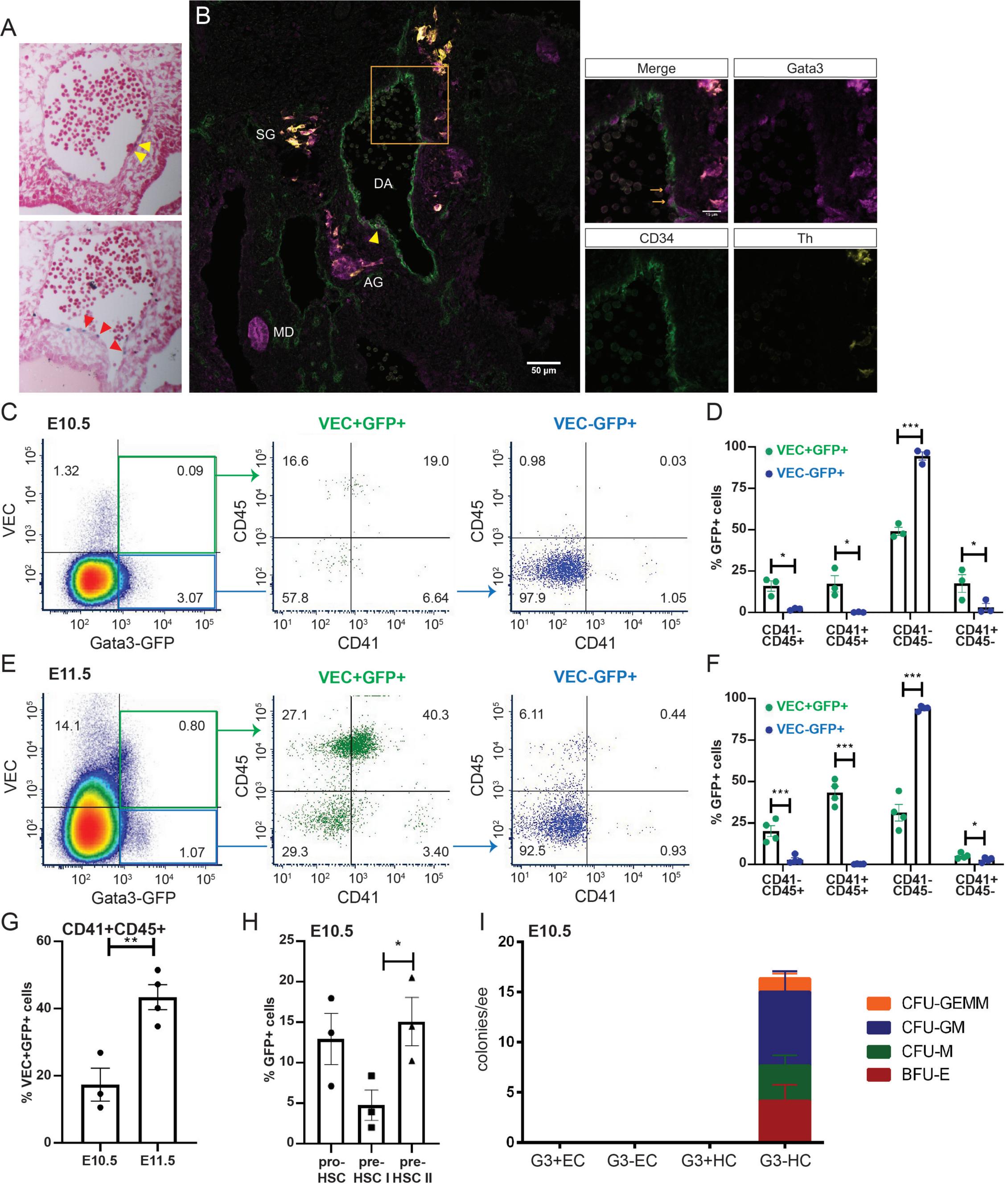


Figure 1

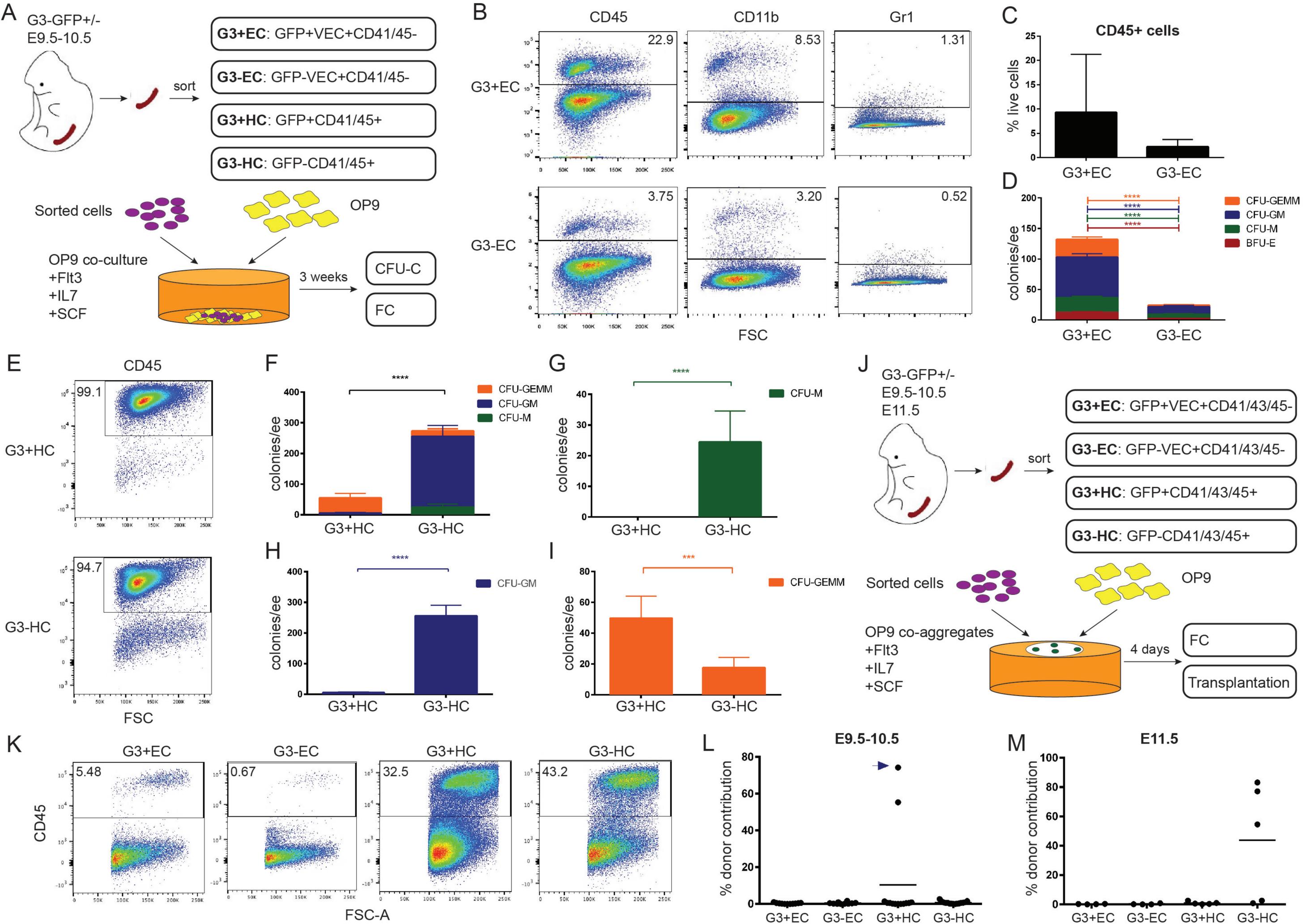


Figure 2

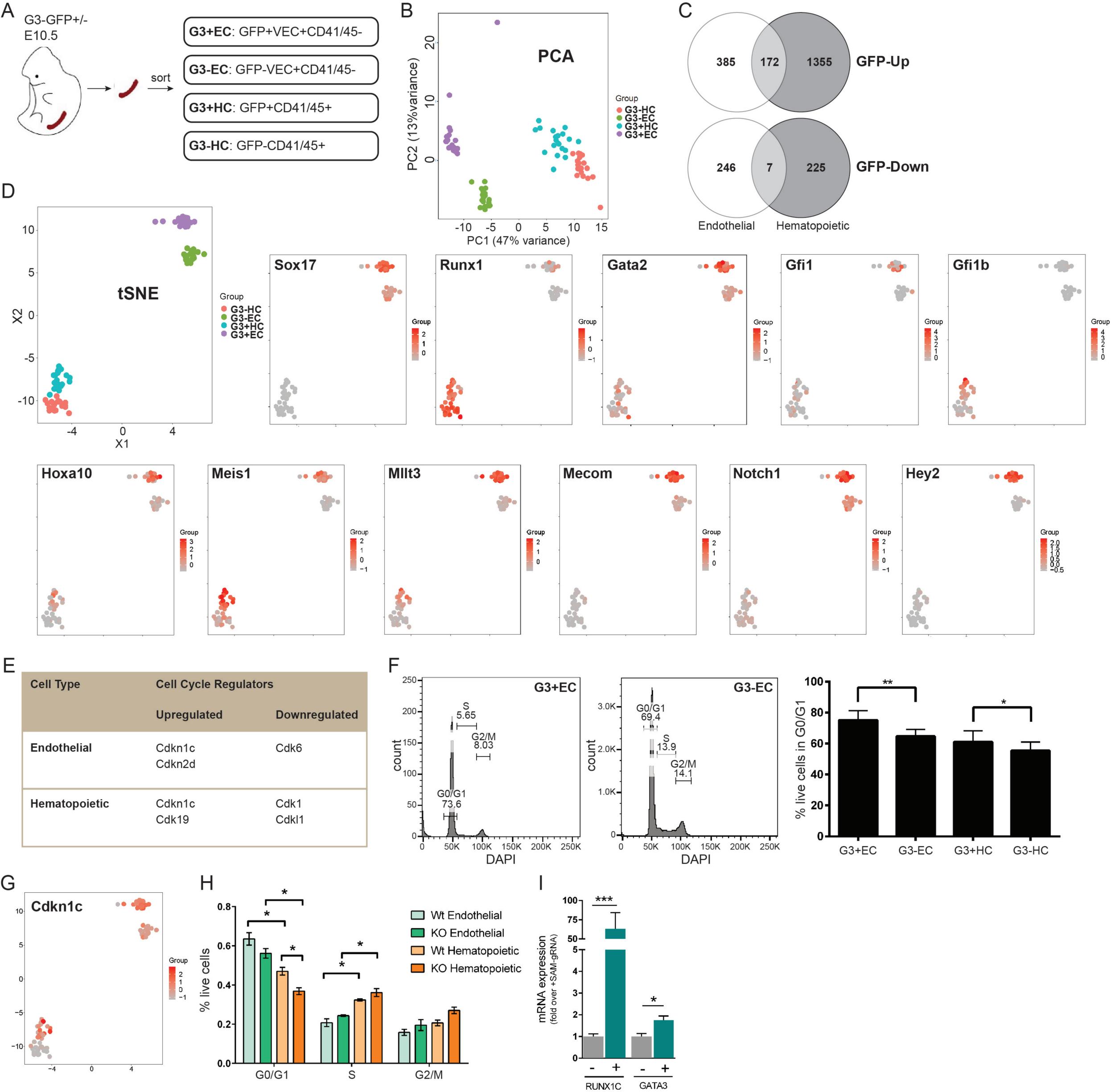
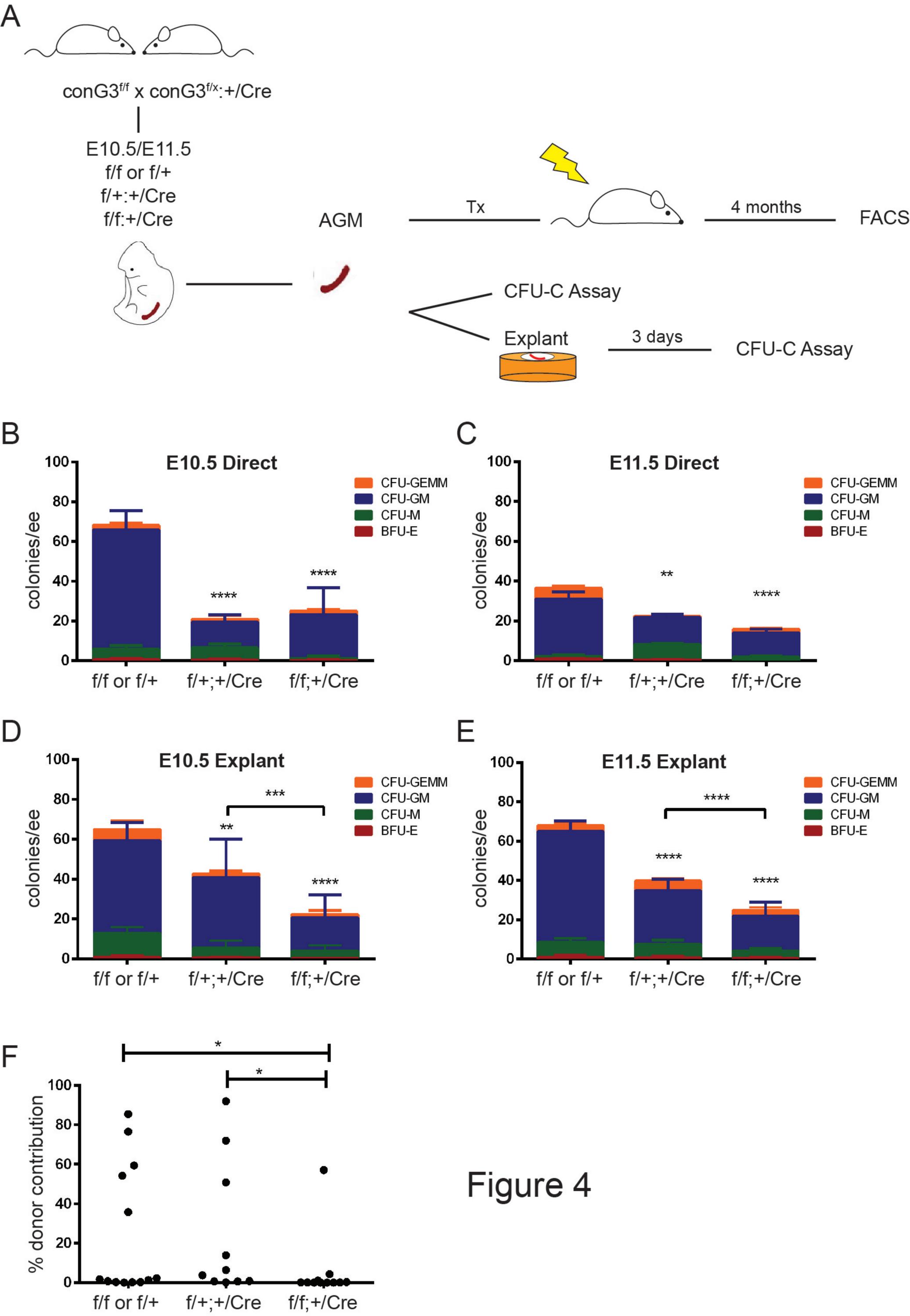
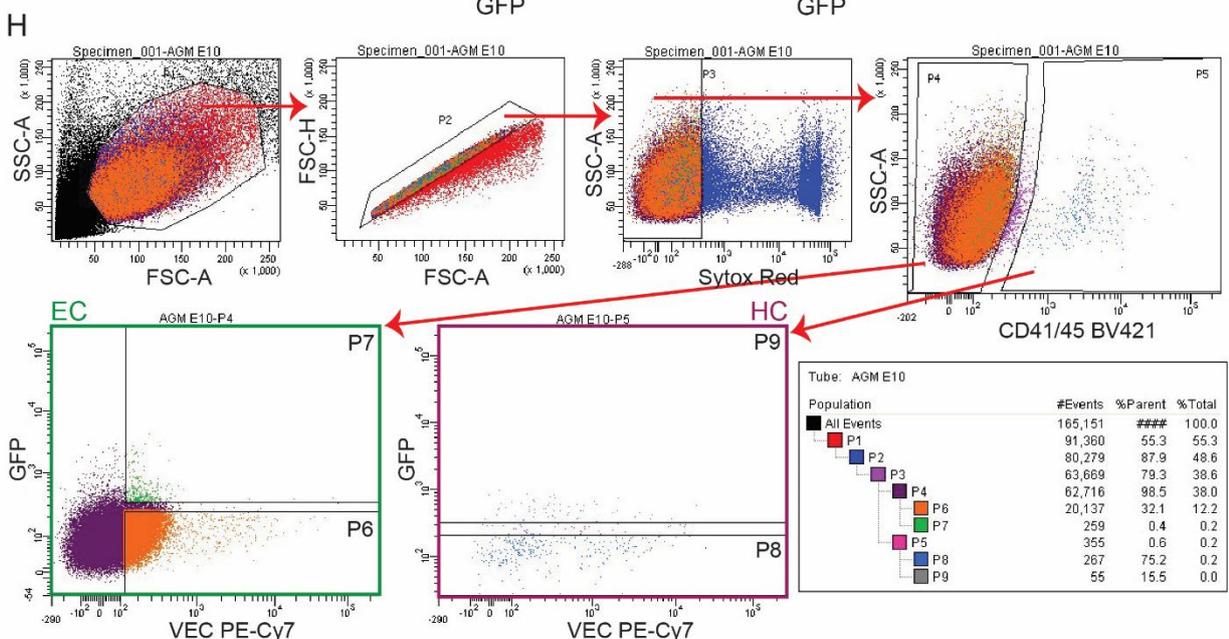
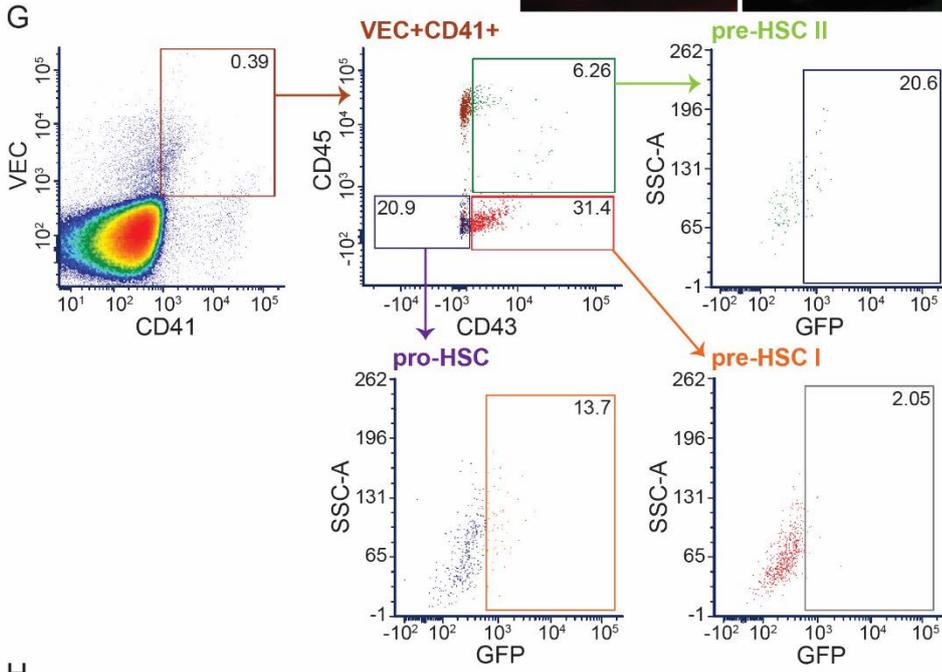
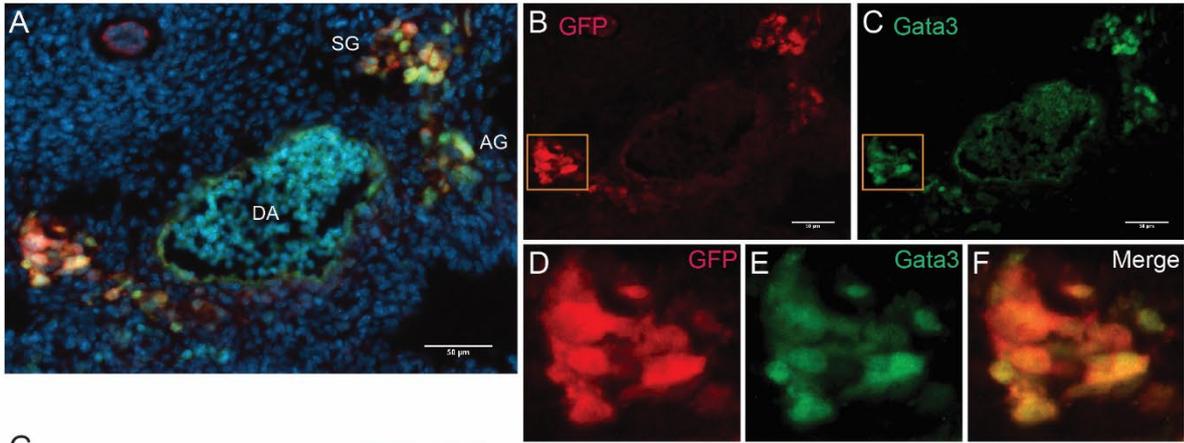


Figure 3



**Supplemental Figures**

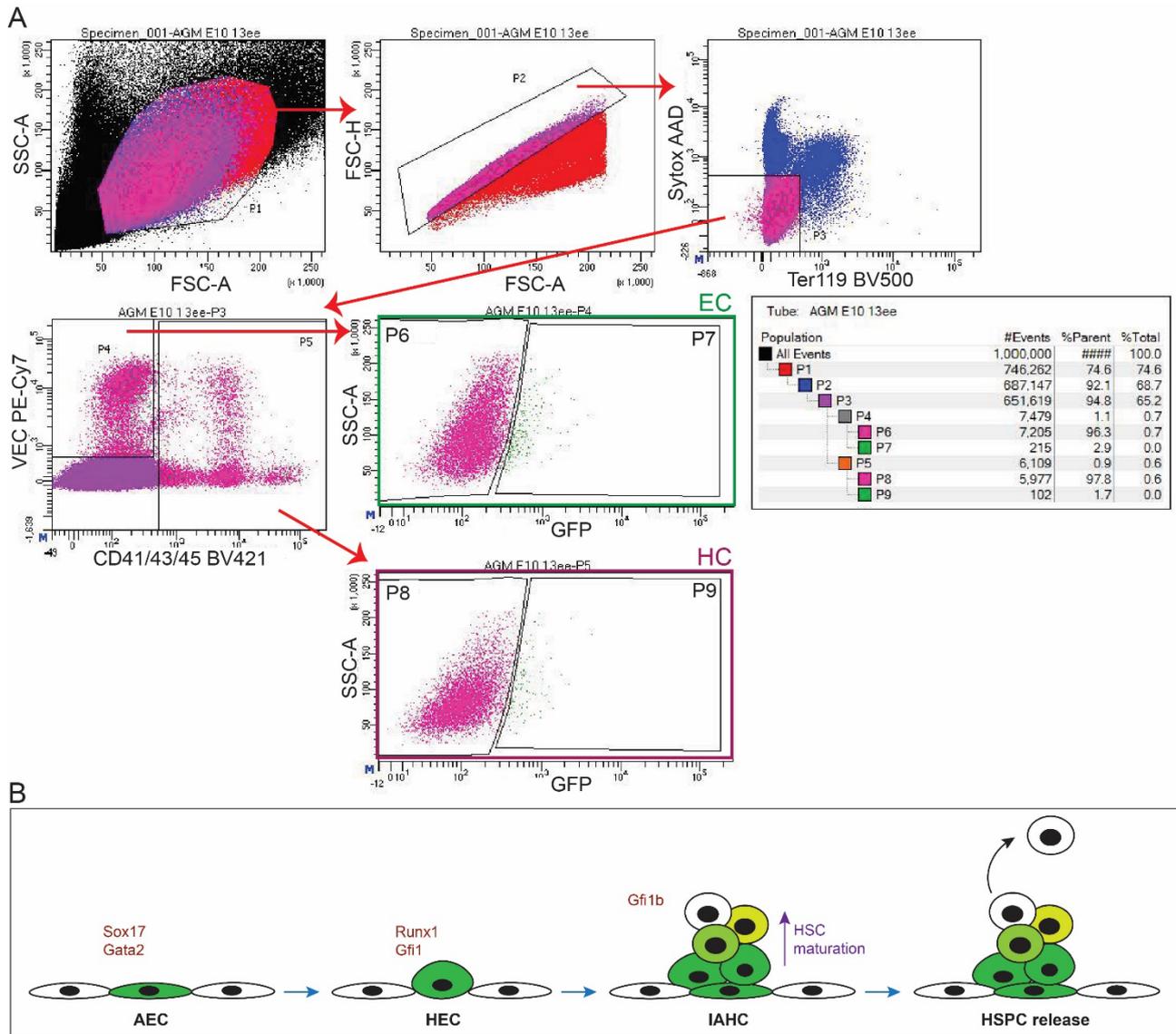


Average cell number per embryo

<b>EC</b>		<b>HC</b>	
Gata3-GFP+	60.8	Gata3-GFP+	38.7
Gata3-GFP-	1507.7	Gata3-GFP-	1168.7

**Figure S1 *Gata3*-GFP expression mirrors that of endogenous GATA3 protein and is expressed in HSC precursors – related to Figure 1 and 2**

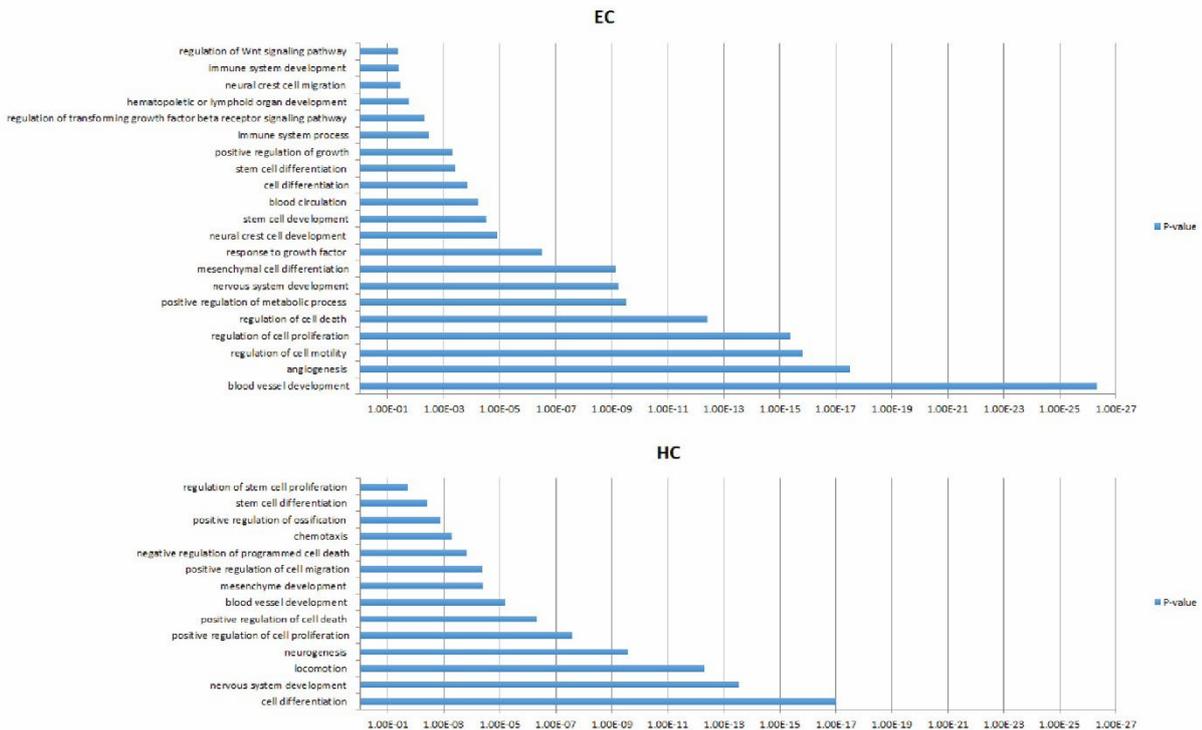
Cryosection from a *Gata3*-GFP+ E11.5 embryo showing the co-staining between GATA3 (green) and GFP (red) with DAPI as nuclear stain in blue. **(A)** merged image, **(B)** antibody to GFP, **(C)** antibody to GATA3. **(D-F)** Magnification of area indicated by yellow box in B and C demonstrating co-staining between GATA3 and GFP. DA: dorsal aorta, AG: adrenal anlage, SG: sympathetic ganglia. **(G)** Flow cytometry gating strategy for detecting *Gata3*-GFP expression in pro-HSCs (VEC+CD41+CD43-CD45-), pre-HSC I (VEC+CD41+CD43+CD45-) and pre-HSC II (VEC+CD41+CD43+CD45+). **(H)** Gating strategy for flow cytometry sorting of *Gata3*-GFP+/- ECs (VEC+ CD41/45-) and HCs (CD41/45+) for the CFU-C assays (Fig.1E) and the co-culture experiments (Fig. 2A). Average number of cells per embryo are stated underneath.



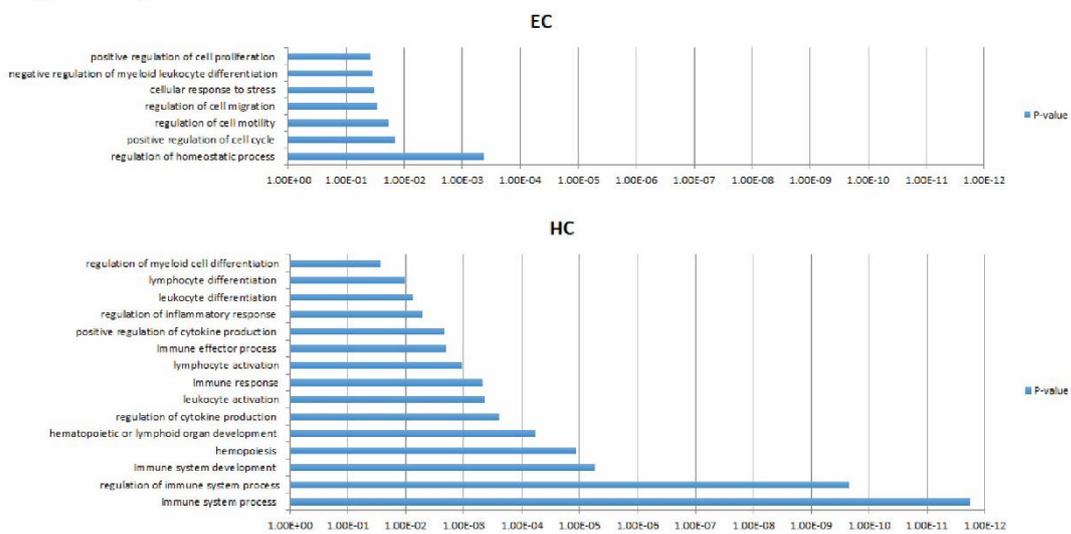
**Figure S2 *Gata3*-GFP<sup>±</sup> EC and HC sorting strategy – related to Figure 2 and 3**

(A) Gating strategy for flow cytometry sorting of *Gata3*-GFP<sup>±</sup> ECs (VEC<sup>+</sup> CD41/43/45<sup>-</sup> Ter119<sup>-</sup>) and HCs (CD41/43/45<sup>+</sup> Ter119<sup>-</sup>) for the co-aggregate experiments in Figure 2J. (B) Schematic diagram of *Gata3* expression (green) during the endothelial-to-hematopoietic transition, with upregulation of other key transcription factors indicated in red. *Gata3* is expressed in arterial endothelial cells (AEC), in which a hematopoietic transcriptional program is then switched on as they become hemogenic endothelial cells (HEC). These then mature into hematopoietic stem and progenitor cells (HSPC) inside intra-aortic hematopoietic clusters (IAHC) via pro-HSC, pre-HSC I and pre-HSC II stages, during which *Gata3* is downregulated until it is no longer expressed in HSPCs released into the circulation.

## Upregulated genes

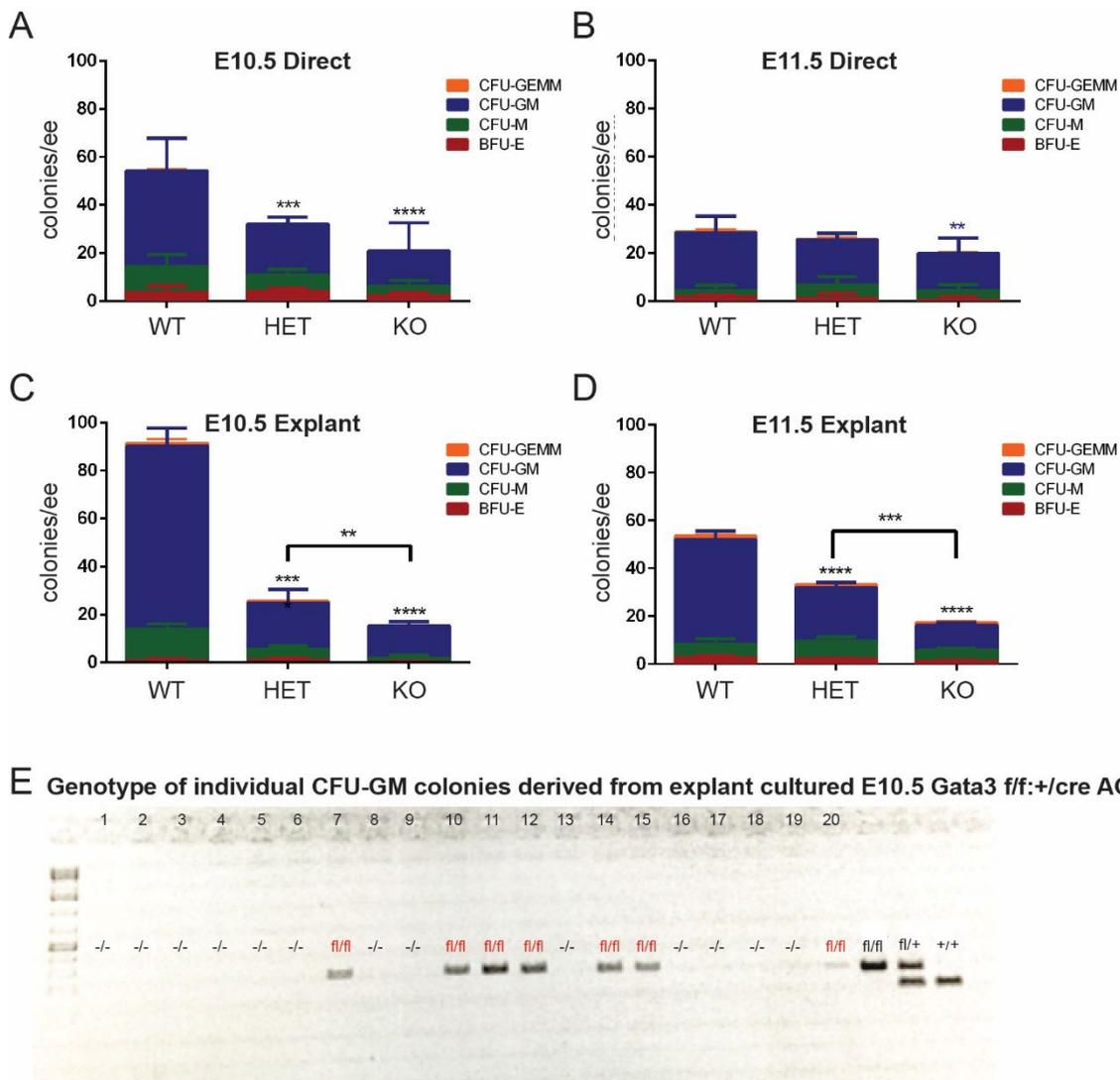


## Downregulated genes



**Figure S3 Gene ontology terms enriched in differentially expressed genes – related to Figure 3**

GO terms significantly enriched amongst the genes upregulated and downregulated in *Gata3*-GFP+ ECs and HCs.



**Figure S4 Colony-forming assays with *Gata3* germline-deleted AGMs – related to Figure 4**

CFU-C assays of *Gata3*<sup>+/+</sup> (WT), *Gata3*<sup>+/-</sup> (HET) and *Gata3*<sup>-/-</sup> (KO) uncultured E10.5 (A) and E11.5 (B) AGM cells and following 3 days of explant culture of E10.5 (C) and E11.5 (D) AGMs. Results are from 3-4 independent experiments (n=3-4, 1 embryo per replicate). \*\* p<0.01; \*\*\* p<0.001; \*\*\*\* p<0.0001; Two-way ANOVA. (E) Gel electrophoresis result from the genotyping of 20 individual CFU-GM colonies picked from methylcellulose plates seeded with cells from *Gata3*-fl/fl; *VEC-Cre*<sup>+</sup> AGMs cultured as explants, demonstrating that some cells retained the floxed allele (highlighted in red) and therefore had escaped recombination and deletion of *Gata3*. n=20.

**Supplemental Table S1**

Differentially expressed genes in *Gata3*-GFP<sup>+/-</sup> hematopoietic (HC) and endothelial (EC) cells. Supplied as a separate Excel spreadsheet.

## Supplemental Experimental Procedures

### OP9 co-cultures

These were carried out according to Swiers et al., 2013, with further details provided in the Supplemental Experimental Procedures. OP9 cells were maintained in  $\alpha$ MEM (Gibco) with 20% heat deactivated FCS (Hyclone) and 0.22% sodium bicarbonate (Gibco) at 37°C with 5% CO<sub>2</sub>. 24h prior to the start of the co-cultures, cells were plated in  $\alpha$ MEM with 10% FCS and 0.01% of 2-mercaptoethanol. Sorted hematopoietic and endothelial cell populations were plated on confluent OP9 stroma. Cultures were supplemented with SCF, FLT3 ligand and IL7 at 10ng/ml (all from Pepro-Tech) and incubated for 8–10 days at 37°C, 5% CO<sub>2</sub>. Hematopoietic output was assessed by CFU-C and flow cytometry.

### Flow cytometry

Antibody stainings were performed 30min on ice in the dark, using the following antibodies: CD41-BV421 (1:100; Biolegend; cat# 133911), CD45-BV421 (1:100; Biolegend; cat# 103133), CD45-APC-Cy7 (1:100; BD Bioscience; cat# 561037), CD45-A700 (1:50; Biolegend; cat# 103128), CD45.1-PE (1:200; eBioscience; cat# 12-0453-82), CD45.2-A700 (1:200; Biolegend; cat# 109821), CD43-BV421 (1:100; BD Bioscience; cat# 752957), Ter119-V500 (1:100; BD Bioscience; cat# 562120), VEC-PE-Cy7 (1:100; Biolegend; cat# 138016), VEC-AF647 (1:100; BD Bioscience; cat# 562242), CD11b-PB (1:200; Biolegend; cat# 101224) and Gr1-PB (1:200; Biolegend; cat# 108429). Peripheral blood samples from transplant recipients were pre-treated with Red Cell Lysis buffer (BD Bioscience). Dead cells were excluded via 7-aminoactinomycin D staining 1:1000 (7AAD, Invitrogen) or Sytox AAD (1:5000) (Invitrogen). All experiments included the following controls: unstained cells, single stained samples and fluorescent minus one (FMO) controls. Cells were analyzed using LSRFortessa (BD Bioscience) or sorted using MoFlo (Beckman Coulter), ARIA (BD Bioscience), or Fusion (BD Bioscience) and data analyzed with the FlowJo software (BD Bioscience).

### Immunohistochemistry

Embryos were fixed in 2% paraformaldehyde (Sigma) in PBS for 1.5h at 4°C, cryoprotected overnight in 30% sucrose/PBS at 4°C and embedded in OCT TissueTek. 10 $\mu$ m sections were prepared on a cryostat (Leica, CM3050 S).

For antibody staining, either an Avidin/biotin system or fluorescent-labelled secondary antibodies were used. Cryosections were blocked with 200 $\mu$ l of PBS/0.05% Tween/1% BSA, containing Avidin/biotin block where appropriate, and incubated with primary antibody for 24 hours at 4°C in the dark. The next day, the slides were incubated with secondary antibody or fluorescently labelled streptavidin for 45min at room temperature in the dark and then mounted with Vectashield containing DAPI (Vectorlabs). The following

antibodies were used: TH (mouse; 1:300; Millipore; cat# MAB318), GFP (chicken; 1:400; ThermoFisher; cat# 600-901-215), GFP (rabbit; 1:500; Life Technologies; cat# A11122), CD34-FITC (rat; BD Bioscience; 1:100; cat# 560238), Gata3 (goat; 1:300; BD Bioscience; cat# AF2605), anti-chicken-Alexa647 (1:500; Jackson ImmunoResearch; cat# 703-605-155), anti-mouse-Alexa546 (1:200; Life Technologies; cat# A10036), anti-rabbit-Alexa647 (1:200; Life Technologies, cat# A31573), anti-rabbit-Alexa555 (1:500; Life Technologies; cat# A31572), and anti-goat CF 633 (1:200; Sigma; cat# SAB4600128). Images were acquired on a Leica SP8 confocal microscope and analyzed with Leica Las X software.

### **Endogenous gene activation in human iPSCs**

Human iPSCs were cultured in StemPro hESC SFM (Gibco) supplemented with 20ng/ml bFGF (R&D) on CELLstart (Gibco) coated wells. Single cell suspension was obtained using Accutase (Gibco) and  $3 \times 10^5$  cells were reverse transfected with 2 $\mu$ g of UniSAM DNA using the Xfect Transfection reagent (Clontech) and plated into a coated 6 well plate. Each well was transfected with the PB-UniSAM plasmid (Addgene 99866 (Fidanza et al., 2017) containing either one of the four gRNAs against *RUNX1C* promoter (Fidanza et al., 2017)), which results in activation and upregulation of *RUNX1C*, or no guide (Empty vector control). Total RNA was extracted using the RNeasy Mini Kit (Qiagen) 2 days post transfection and cDNA synthesized using the High-Capacity cDNA synthesis Kit (Applied Biosystem). Gene expression analysis was performed in triplicate using the LightCycler 384 (Roche) with SYBR Green Master Mix II (Roche),  $\beta$ -Actin was used as reference genes. Gene activation values were calculated as fold change relative to the empty vector control group. The following primers were used: *RUNX1C\_fw* agcctggcagtgtcagaagt, *RUNX1C\_rv* gggactcaatgattcttttacc, *GATA3\_fw* gctcttcgctaccaggtg, *GATA3\_rv* gtaaaaagggcgacgactc, *ACTB\_fw* ccaaccgcgagaagatga, *ACTB\_rv* ccagaggcgtacagggatag.

### **RNA Sequencing**

#### *Reverse transcription*

For the reverse transcription step, 2 $\mu$ l of annealing mix (5% ERCC RNA spike-In Mix (pre-diluted at 1:25,000; Invitrogen), 5% Oligo-dT (5'-AAGCAGTGGTATCAACGCAGAGTACT30VN-3'; 100 $\mu$ M; biomers.net), 50% dNTP 10mM (Fermentas) and 40% distilled water) were added to each well and the plates incubated at 72°C for 3min and immediately placed on ice. 5.7 $\mu$ l of reverse transcription mixture (0.5 $\mu$ l Superscript II RT (200 U/ $\mu$ l; Invitrogen), 0.25 $\mu$ l RNase inhibitor (20 U/ $\mu$ l), 2 $\mu$ l 5x Superscript II First Strand Buffer (Invitrogen), 0.5 $\mu$ l DTT (Invitrogen), 2 $\mu$ l 100 $\mu$ M Betaine (Sigma), 0.06 $\mu$ l 1M MgCl<sub>2</sub> (Ambion), 0.1 $\mu$ l TSO Oligo (5'-AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3'; 100 $\mu$ M; Exiqon) and 0.29 $\mu$ l distilled water) were added and the plate placed in a PCR cycler: 42°C for 90min, 10 cycles of: 50°C for 2min, 42°C for 2min, then at the end 70°C for 15min.

### *PCR Pre-amplification*

For the PCR Amplification, 15µl of the PCR mixture were added per well, consisting of 12.5µl KAPA HiFi Hotstart ReadyMix (2x; KAPA Biosystems), 0.25µl ISPCR primer (5'-AAGCAGTGGTATCAACGCAGAGT-3'; 10µM; biomers.net) and 2.25µl distilled water. The reactions were run at 98°C for 3min, 21 cycles of: 98°C for 20sec, 67°C for 15sec, 72°C for 6min, and 72°C for 5min at the end.

Ampure XP beads (Beckman Coulter) were used for PCR clean up. The size distribution of the cDNA library was checked on an Agilent high-sensitivity DNA chip (Agilent Technologies), according to the manufacturer instructions.

### *Sequencing library preparation*

Tagmentation was carried out using the Illumina Nextera XT DNA sample preparation kit (Illumina) according to an optimized Tagmentation protocol (Fluidigm). Index Primers 1 (N701-N712) and 2 (S501-S508) at a ratio of 12.5% (each) were combined that each well was uniquely labelled and dual-indexing metadata could be obtained (Nextera XT 96-Index kit; Illumina). Following PCR amplification, libraries were pooled and cleaned up with Ampure XP beads. The library size distribution was checked on an Agilent high-sensitivity DNA chip and the library quantified using the KAPA library quantification kit (KAPA Biosystems). Pooled libraries were sequenced on an Illumina Hi-Seq 4000 (Sanger, Cambridge), as single-end 125 base pair reads.

### *Sequencing data analysis*

The data were aligned using STAR (Dobin et al., 2013) to Ensembl genome build 81 (Zerbino et al., 2018), with gene counts obtained using HT-Seq (Anders et al., 2015). Quality control filtering and normalization was performed in R. More than 500,000 reads uniquely mapped (either to ERCC spike-ins or endogenous mRNA), with more than 20% of total reads mapped to mRNA, less than 20% of mapped reads allocated to mitochondrial genes, less than 20% of reads mapped to ERCC spike-ins and more than 8000 high coverage genes. Cells were normalized with Scran (Lun et al., 2016) and highly variable genes identified estimating technical variance with the ERCC spike-ins (Brennecke et al., 2013). In-house programs in R were used for PCA and t-distributed Stochastic Neighbor Embedding (t-SNE) dimensionality reduction. Genes differentially expressed between cell types and Gata3 expression groups were identified using the rank\_genes\_groups function with the t-test\_overestim\_var method. P-values were adjusted using the benjamini-hochberg procedure, and genes with adjusted p-value < 0.01 considered as significant.

Gene Ontology (GO) analysis was performed using the Gene Ontology Consortium Enrichment Analysis, which utilizes PANTHER Classification System for biological processes in *mus musculus*.

### Supplemental References

- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166-169.
- Brennecke, P., Anders, S., Kim, J.K., Kolodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., *et al.* (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 10, 1093-1095.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.
- Fidanza, A., Lopez-Yrigoyen, M., Romano, N., Jones, R., Taylor, A.H., and Forrester, L.M. (2017). An all-in-one UniSam vector system for efficient gene activation. *Sci Rep* 7, 6394.
- Lun, A.T., Bach, K., and Marioni, J.C. (2016). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol* 17, 75.
- Picelli, S., Faridani, O.R., Bjorklund, 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.
- 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.
- Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C.G., *et al.* (2018). Ensembl 2018. *Nucleic Acids Res* 46, D754-D761.