GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1

Roshana Thambyrajah[1,10], Milena Mazan[1,2,10], Rahima Patel[1], Victoria Moignard[3], Monika Stefanska[1], Elli Marinopoulou[1], Yaoyong Li[4], Christophe Lancrin[1,5], Thomas Clapes[6], Tarik Möröy[7], Catherine Robin[6], Crispin Miller[4], Shaun Cowley[8], Berthold Göttgens[3], Valerie Kouskoff[9,11] and Georges Lacaud[1,11]

In vertebrates, the first haematopoietic stem cells (HSCs) with multi-lineage and long-term repopulating potential arise in the AGM (aorta–gonad–mesonephros) region. These HSCs are generated from a rare and transient subset of endothelial cells, called haemogenic endothelium (HE), through an endothelial-to-haematopoietic transition (EHT). Here, we establish the absolute requirement of the transcriptional repressors GFI1 and GFI1B (growth factor independence 1 and 1B) in this unique trans-differentiation process. We first demonstrate that *Gfi1* expression specifically defines the rare population of HE that generates emerging HSCs. We further establish that in the absence of GFI1 proteins, HSCs and haematopoietic progenitor cells are not produced in the AGM, revealing the critical requirement for GFI1 proteins in intra-embryonic EHT. Finally, we demonstrate that GFI1 proteins recruit the chromatin-modifying protein LSD1, a member of the CoREST repressive complex, to epigenetically silence the endothelial program in HE and allow the emergence of blood cells.

The vertebrate haematopoietic system is characterized by several waves of blood progenitor generation[1]. The first two waves of blood cell emergence are initiated in the extra-embryonic yolk sac and generate mainly primitive erythrocytes and then erythroid–myeloid progenitors[2–5]. In the next wave, the intra-embryonic AGM region gives rise to the first haematopoietic stem cells (HSCs) with multi-lineage and long-term repopulating potential in adult recipients[6,7]. Recent studies have provided evidence for the existence of a subset of endothelial cells capable of generating blood cells; a haemogenic endothelium[8–12] (HE). Live imaging studies have revealed how during this process of endothelial-to-haematopoietic transition (EHT), spindle-shaped endothelial cells lose their epithelial morphology to adopt the typical round shape of haematopoietic cells[10,11,13–16]. Consistent with the generation of HSCs from endothelial cells, the first progenitors with HSC activity are detected in the major arteries[17] and cells with a HSC phenotype have been found in the intra-aortic haematopoietic clusters (IAHC) generated through EHT from the ventral wall of the dorsal aorta[13,18,19] (vDA).

The molecular mechanisms controlling EHT and the specification and emergence of HSCs are still poorly understood. The transcription factor RUNX1 (AML1) is a master regulator of definitive haematopoiesis that is expressed in the entire continuum of blood development from HE to mature haematopoietic cells[20–24]. *Runx1* deletion leads to mid-gestation embryonic lethality and a complete absence of definitive haematopoietic cells[20,22,25,26] as a result of a block in the generation of blood cells from HE (refs 11,15,27). More recently, the two highly homologous transcriptional repressors GFI1 and GFI1B were identified as downstream targets of RUNX1 and important players during the initial wave of haematopoietic commitment in the yolk sac[28].

In this study, we investigated the molecular mechanisms controlling the generation of HSCs during the intra-embryonic wave of haematopoiesis in the AGM and demonstrate a prominent role for the transcriptional repressors GFI1 and GFI1B in this process. By combining mouse reporter lines and functional assays, we established that *Gfi1* expression is initiated in a rare population of HE, present in the vDA during HSC emergence, whereas GFI1B expression is upregulated in emerging IAHC. Loss of GFI1 proteins impaired IAHC formation, indicating the pivotal role of these repressors during the emergence of haematopoietic cells within the AGM. Furthermore, we
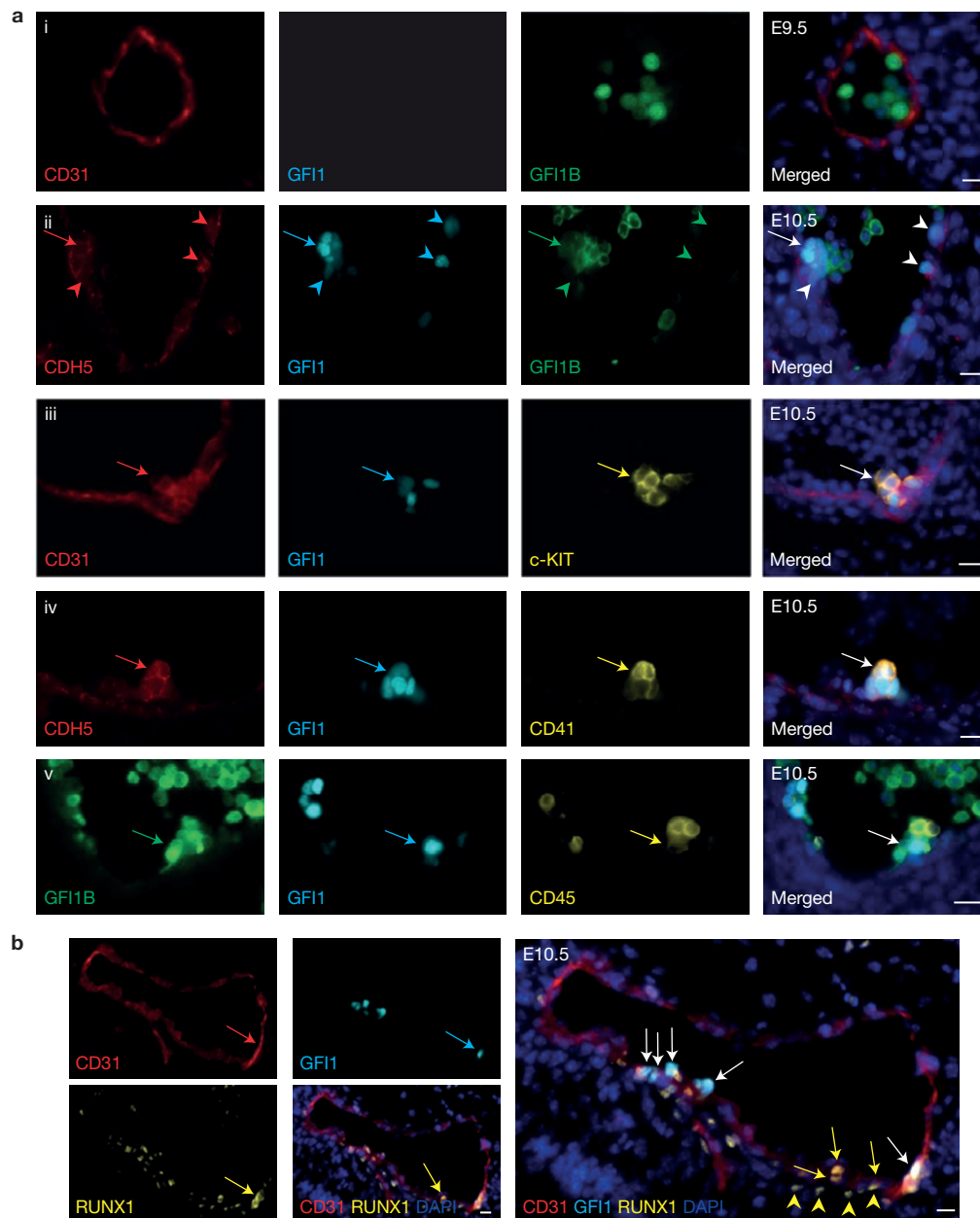
Figure 1 GFI1 and GFI1B define distinct cell populations in the E10.5 AGM. (**a**) Immunohistochemistry on E9.5 and E10.5 $Gfi1^{Tomato}/Gfi1b^{GFP}$ embryos for CD31 and CDH5 (both red), GFI1 (cyan), GFI1B (green) and c-KIT, CD41 and CD45 (all yellow). (i) E9.5 para-aortic splanchno-pleura region. (ii) GFI1$^+$/GFI1B$^-$ cells (arrowheads) and GFI1$^+$/GFI1B$^+$ cells (arrows) are indicated. (iii) GFI1 and c-KIT are co-expressed (white arrow). (iv) CD41 and GFI1 are co-expressed (white arrow). (v) CD45, GFI1 and GFI1B co-expression (white arrow). (**b**) Staining for RUNX1 (β-galactosidase; yellow), CD31 (red) and GFI1 (cyan). Co-expression of CD31, GFI1 and RUNX1 (white arrows), CD31 and RUNX1 (yellow arrows) or only RUNX1 expression (yellow arrowheads) are indicated. Representative images of AGM sections derived from more than 5 embryos from different litters are shown. Scale bars, 10 μm.

established that the critical function of the GFI1 proteins is mediated by the recruitment of LSD1, a member of the repressive CoREST complex. Deficiency in LSD1 prevents the morphological changes and loss of endothelial markers associated with the EHT. Accordingly, we observed that GFI1 and GFI1B directly bind to the regulatory regions of a large set of genes whose expression is maintained on LSD1 inhibition. This set of genes is closely associated with cardiovascular development, as well as maintenance and remodelling of blood vessels. Together, our results demonstrate that, in the AGM, the sequential

expression of GFI1 proteins orchestrates the generation of HSCs by recruiting the CoREST complex to silence the endothelial program.

RESULTS

GFI1$^+$/GFI1B$^-$ and GFI1$^+$/GFI1B$^+$ cells mark distinct populations within the AGM

To determine the specific expression and potential functions of GFI1 and GFI1B in the emergence of HSCs, we first analysed the spatiotemporal expression pattern of GFI1 and GFI1B in the

AGM. For this, a *Gfi1:H2B–Tomato* mouse reporter line (referred to as *Gfi1^{Tomato}* hereafter) was generated, validated (Supplementary Fig. 1a) and crossed with the *Gfi1b^{+/GFP}* knock-in line[29]. The expression of both *Gfi1* and *Gfi1b* was investigated in double transgenic embryos by fluorescence-activated cell sorting (FACS) and immunohistochemistry (IHC). At E9.5, when the dorsal aorta is developing within the para-aortic splanchno-pleura (P-Sp) region, *Gfi1* expression was not detected by IHC in sections of the dual reporter embryos. *Gfi1b* expression was restricted to circulating cells most likely representing primitive erythroid cells (Fig. 1a(i)).

At E10.5, GFI1^+/GFI1B^− cells were found either embedded within the endothelium (CD31 or VE-cadherin/CDH5; arrowheads, Fig. 1a(ii) and Supplementary Fig. 1b), and/or at the base of IAHCs (arrows in Fig. 1a(ii)). Some GFI1^+ cells in IAHCs were also positive for GFI1B (Fig. 1a(ii)). We next analysed the expression of *Gfi1* and *Gfi1b* in the context of previously characterized markers of HE and IAHC in the AGM: RUNX1, c-KIT, CD41 and CD45. Staining for β-galactosidase on *Runx1:nls-LacZ/Gfi1^{GFP/+}* E10.5 AGMs showed that *Gfi1* expression was restricted to a subset of RUNX1^+ cells within the endothelial layer (white arrows, Fig. 1b) and did not extend to the sub-aortic mesenchyme and endothelium (yellow arrows and arrowheads, Fig. 1b). Previous studies have shown that CDH5^+ cells in the vDA generate IAHCs that express the blood markers c-KIT and CD41 before acquiring CD45. These CDH5^+/CD41^+ and CDH5^+/CD45^+ intermediates represent pre-HSCs that mature into directly transplantable HSCs (refs 30–32). We observed that GFI1^+ cells in the clusters co-stained for c-KIT (arrow, Fig. 1a(iii)). CD41 (arrow, Fig. 1a(iv)) and CD45 (arrow, Fig. 1a(v)) at E10.5. Similarly at E11.5, GFI1^+GFI1B^− and GFI1^+GFI1B^+ cells were still closely associated with endothelial lining and IAHCs at E11.5 (Fig. 2a).

In agreement with the IHC data, only a very low frequency of GFI1^+/GFI1B^− or GFI1^+/GFI1B^+ cells were found by FACS in E10.5 and E11.5 AGMs (Fig. 2b). A large fraction of GFI1^+/GFI1B^− cells were in the CDH5^+/CD45^− endothelial subset or in the CDH5^+/CD45^+ double-positive cell population (Fig. 2c). In contrast, GFI1^+/GFI1B^+ cells were mostly associated with the more mature CDH5^+/CD45^+ or CDH5^−/CD45^+ populations (Fig. 2c). Within the CDH5^+ endothelial compartment, the frequency of GFI1^+/GFI1B^− was similar at E10.5 and E11.5, whereas GFI1^+/GFI1B^+ tended to increase (Supplementary Fig. 1c). Within this compartment, we also found a small population of cells that expressed only *Gfi1b* at E11.5 (Supplementary Fig. 1d). Finally, to visualize the acquisition of *Gfi1b* expression by GFI1^+ cells, we performed *in vitro* culture of FACS-sorted CDH5^+/GFI1^+/GFI1B^− cells and observed the emergence of GFI1B-expressing cells (Supplementary Fig. 2). Similarly, time-lapse imaging of *Gfi1^{Tomato}Gfi1b^{GFP}* embryonic sections showed the acquisition of *Gfi1b* expression by CDH5^+/GFI1^+/GFI1B^− cells (Fig. 2d and Supplementary Video 1).

Overall, these data revealed a dynamic and distinct expression pattern for *Gfi1* and *Gfi1b*. *Gfi1* expression is initiated at E10.5 in RUNX1^+ endothelial cells within the vDA and is sustained in the first emerging cells within the IAHCs, suggesting that *Gfi1* expression marks HE undergoing the transition to haematopoiesis. *Gfi1b*-expressing cells, in contrast, are mainly localized in IAHCs as c-KIT^+ cells and coincide with further commitment to CD41 and CD45 expression.

GFI1^+/GFI1B^− cells mark HE whereas HSC activity is mostly associated with GFI1^+/GFI1B^+ expression

To demonstrate that the endothelial GFI1^+/GFI1B^− cells correspond to HE that subsequently become GFI1^+/GFI1B^+ blood precursors, we analysed the expression of a set of 70 endothelial and haematopoietic genes in single cells of the different CDH5^+ populations defined by *Gfi1* and *Gfi1b* expression. These included endothelial CDH5^+ cells, the CDH5^+ cells expressing *Gfi1* found in the endothelial lining (CDH5^+/GFI1^+/GFI1B^−/cKIT^−), or emerging in the lumen (CDH5^+/GFI1^+/GFI1B^−/c-KIT^+), the CDH5^+/GFI1^+/GFI1B^+ cells present in the clusters and the rare population of CDH5^+/GFI1^−/GFI1B^+ cells (Fig. 3a and Supplementary Fig. 3). Consistent with a developmental continuum between the three GFI1^+ populations, principal component analyses indicated that CDH5^+/GFI1^+/GFI1B^−/c-KIT^− and CDH5^+/GFI1^+/GFI1B^+ cells formed separate clusters but were linked by the CDH5/GFI1^+/GFI1B^−/c-KIT^+ population (Fig. 3b). In contrast, endothelial CDH5^+ or CDH5^+/GFI1^−/GFI1B^+ cells clustered as separate entities with the latter exhibiting an expression profile consistent with megakaryocytes. *Gfi1* expression closely coincided with the expression of endothelial genes (*Dll4*, *Procr*, *Kdr*, *Egfl7*, *Notch1*) whereas the expression of *Gfi1b* was linked to haematopoietic genes such as *c-kit*, *Itga2b*, *CD45*, *Mpo* and *c-myb* (Fig. 3b). Together, these analyses suggest that CDH5^+/GFI1^+GFI1B^−/c-KIT^−, CDH5^+/GFI1^+/GFI1B^−/c-KIT^+ and CDH5^+/GFI1^+/GFI1B^+ are distinct but sequentially related cell populations corresponding to haemogenic endothelium, cells undergoing EHT and emerging haematopoietic cells, respectively.

By definition, HE cells require a further maturation step to generate haematopoietic progenitors. To functionally evaluate the HE potential of the GFI1^+ fraction, we isolated CDH5^+/GFI1^− and CDH5^+/GFI1^+ cells from E10.5 AGM while excluding blood precursors (CD41^+ and/or CD45^+). Both populations were plated into haematopoietic assays, either directly, or following a 4-day maturation step of co-culture with OP9 cells. Direct replating into haematopoietic assays yielded very few colonies (Fig. 3d) in contrast to control sorted GFI1^+ and/or GFI1B^+ cells containing CD41^+ and CD45^+ cells (Supplementary Fig. 4a). Following maturation, the CDH5^+/GFI1^−/CD41^−/CD45^− fraction did not gain any haematopoietic potential, whereas the CDH5^+/GFI1^+/CD41^−/CD45^− cells generated haematopoietic colonies at a high frequency, indicating that at E10.5, HE cells were present within the CDH5^+/GFI1^+ population (Fig. 3d).

Finally, we tested whether GFI1^+ cells in the vDA contain HSCs. Single-cell suspensions of *Gfi1^{Tomato/+}* E11.5 AGMs were stained for CDH5 and separated into CDH5^+/GFI1^+ (that are either expressing *Gfi1b* or not) and CDH5^+/GFI1^− cells. Both fractions were transplanted into lethally irradiated mice. Long-term primary and secondary contributions were mainly detected in recipients transplanted with the CDH5^+/GFI1^+ fraction (Fig. 3e(i) and Supplementary Fig. 4b,c), indicating that HSC activity is restricted to the CDH5^+/GFI1^+ population at E11.5. We further separated the CDH5^+/GFI1^+ cell population into CDH5^+/GFI1^+/GFI1B^− and CDH5/GFI1^+/GFI1B^+ cells to test whether there were significant differences between these two populations. Consistent with the less mature state of the
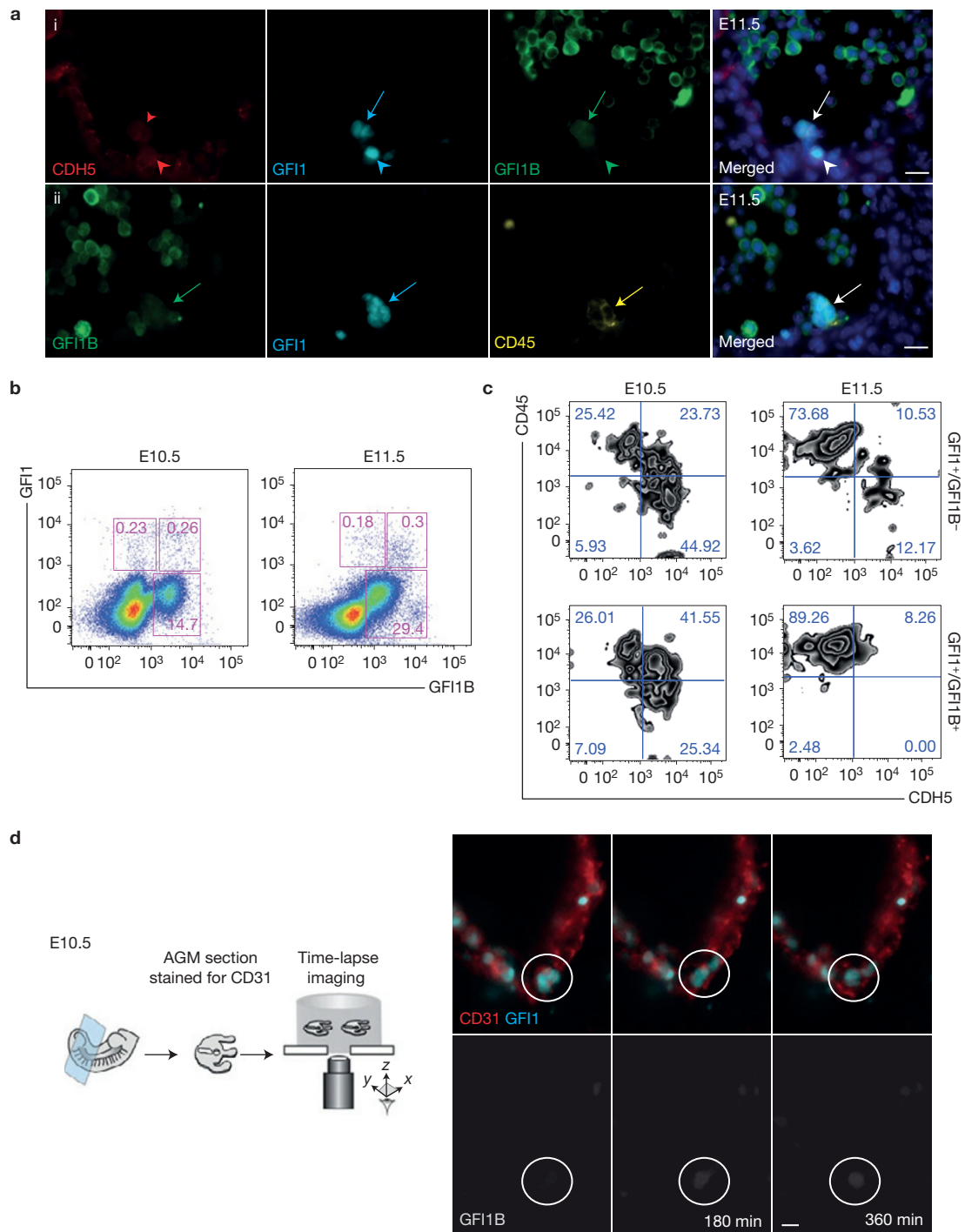
Figure 2 GFI1 marks haemogenic endothelial cells in the E10.5 AGM. (**a**) Immunohistochemistry for CDH5 (red), GFI1 (cyan), GFI1B (green) and CD45 (yellow) at E11.5. (i) GFI1$^+$/GFI1B$^-$ (arrowheads) and GFI1$^+$/GFI1B$^+$ cells (arrows) are indicated. (ii) GFI1$^+$/GFI1B$^+$ cells co-expressing CD45 (arrow) are indicated. Representative images of AGM sections derived from more than 3 embryos from different litters are shown. (**b**) FACS analysis of GFI1 (Tomato) and GFI1B (GFP) expression in E10.5 and E11.5 AGM (representative FACS data from more than 2 different litters are shown). (**c**) FACS analysis of CDH5 and CD45 expression by GFI1$^+$/GFI1B$^-$ and GFI1$^+$/GFI1B$^+$ cells in E10.5 and E11.5 AGMs (representative FACS data from more than 2 different litters are shown). (**d**) Time-lapse imaging performed on $Gfi1^{Tomato}Gfi1b^{GFP}$ E10.5 sections stained with CD31 ($n = 3$ independent experiments with AGMs from 2 or more embryos, one representative sequence of imaging is shown). Scale bars, 10 μm.

CDH5$^+$/GFI1$^+$/GFI1B$^-$ cells, the repopulation potential of this fraction was lower than that of CDH5$^+$/GFI1$^+$/GFI1B$^+$ cells (Fig. 3e(ii)).

GFI1 and GFI1B are required for EHT in the AGM

To investigate whether GFI1 proteins have a functional importance in HSC emergence, we examined the presence of IAHC in E10.5

Figure 3 GFI1+ cells acquire GFI1B expression and harbour HSC activity. (**a**) Schematic representation of the sorting strategy for the single-cell PCRs. (**b**) Left: principal component analysis for the cells and genes analysed. Right: principal component analysis for endothelial and haematopoietic genes expressed by CDH5+/GFI1+/c-KIT− and CDH5+/GFI1+/GFI1B+ cells. (**c**) Vi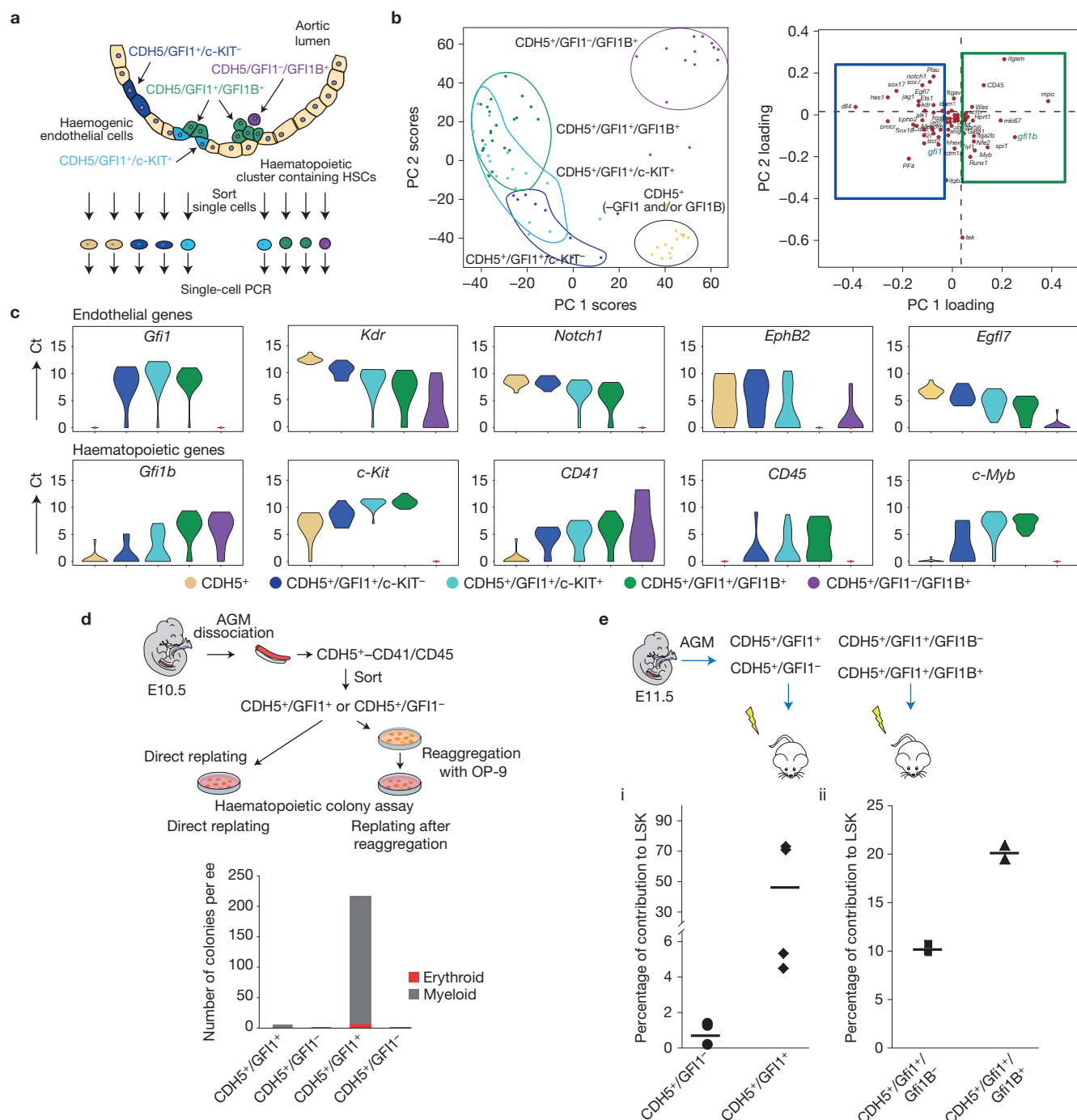olin plots of single-cell PCR data for the indicated genes (data are derived from one single Fluidigm single-cell PCR run performed with n = 96 AGM cells from 3 different litters). (**d**) Haemogenic potential of E10.5 CDH5+/GFI1+/CD41−/CD45− and CDH5+/GFI1−/CD41−/CD45− cells. The CDH5+/GFI1+/CD41−/CD45− and CDH5+/GFI1−/CD41−/CD45− populations were either directly replated or replated after a co-aggregation step with OP-9 cells for 4 days (2 independent experiments were performed with AGMs from 2 and 9 embryos; data from one representative experiment are shown; the source data for the second experiment can

be found in Supplementary Table 3). ee, embryo equivalent. (**e**) (i) Long-term repopulation potential of E11.5 CDH5+/GFI1+ and CDH5+/GFI1− cells from *Gfi1^{Tomato}* AGMs. Donor contributions to the Lin−Sca-1+c-Kit+ (LSK) compartment at 17 weeks after transplantation are shown (one independent experiment with AGM cells sorted from one litter, n = 3 NSG mice transplanted with CDH5+/GFI1− and n = 4 NSG mice transplanted with CDH5+/GFI1+ cells). (ii) E11.5 *Gfi1^{Tomato}Gfi1b^{GFP}* AGMs were dissociated and the CDH5+/GFI1+/GFI1B− and CDH5+/GFI1+/GFI1B+ cells were injected intravenously into irradiated NSG mice. Donor contributions to the LSK compartment at 16 weeks after transplantation are shown (one independent experiment with AGM cells sorted from one litter, n = 2 NSG mice transplanted with CDH5+/GFI1+/GFI1B− and n = 2 NSG mice transplanted with CDH5+/GFI1+/GFI1B− cells). Mean donor contribution is depicted with a line.
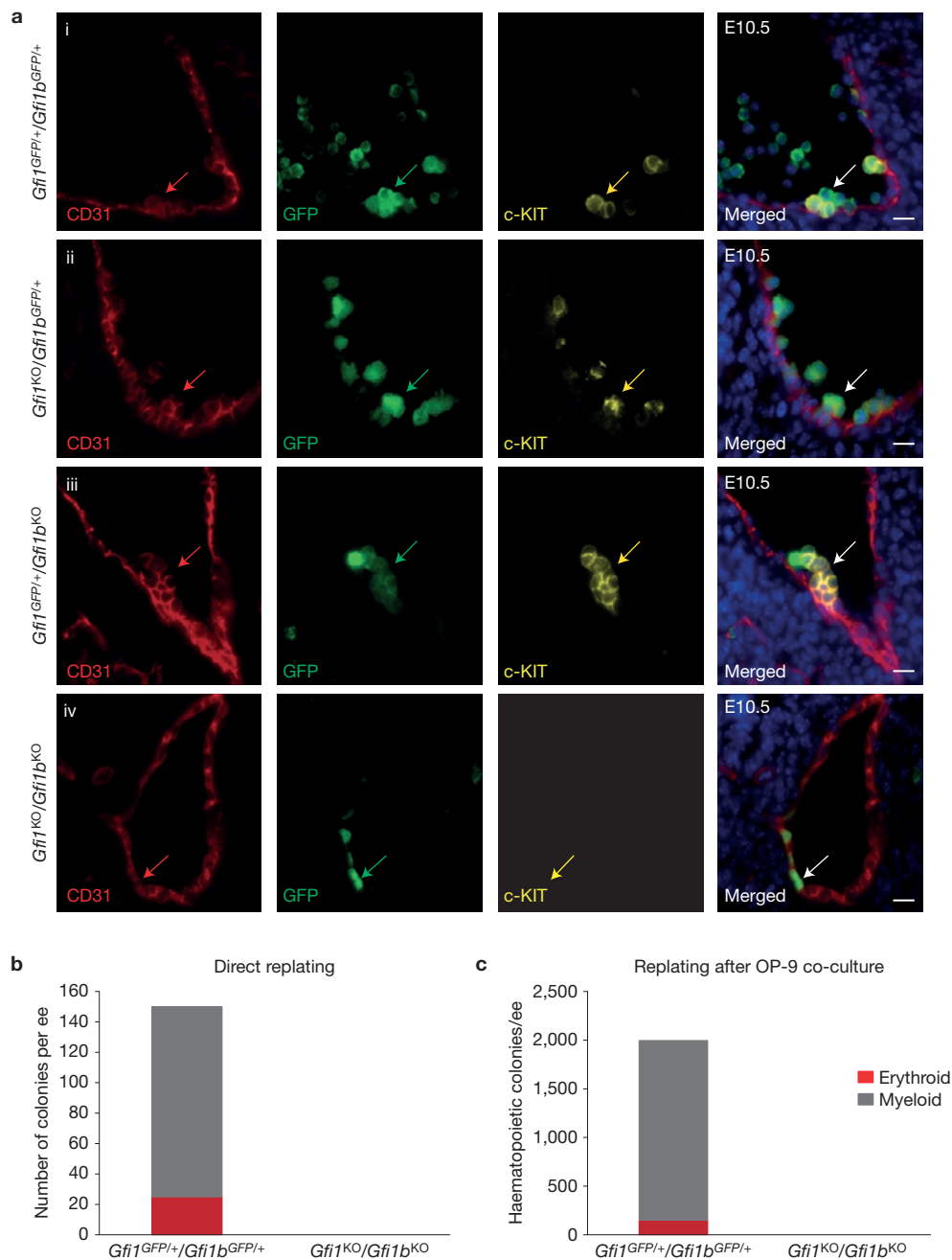
Figure 4 GFI1 and GFI1B are essential for intra-aortic cluster formation. (**a**) Immunohistochemistry on *Gfi1*het/*Gfi1b*het (*Gfi1*GFP/+/*Gfi1b*GFP/+), *Gfi1*KO/*Gfi1b*het (*Gfi1*GFP/GFP/*Gfi1b*GFP/+), *Gfi1*het/*Gfi1b*KO (*Gfi1*GFP/+/*Gfi1b*GFP/GFP) and *Gfi1*KO/*Gfi1b*KO (*Gfi1*GFP/GFP/*Gfi1b*GFP/GFP) E10.5 AGMs for CD31 (red), GFI1 and GFI1B (green) and c-KIT (yellow). Representative images of AGM sections derived from more than 4 embryos from different litters are shown.

(**b**,**c**) Haematopoietic potential of *Gfi1*het/*Gfi1b*het and *Gfi1*KO/*Gfi1b*KO AGMs. Cells were replated either directly (**b**) or after a co-aggregation step with OP-9 cells for 4 days (one independent experiment with AGM cells sorted from one litter) (**c**) into haematopoietic colony assays (one independent experiment with AGM cells sorted from one litter). ee, embryo equivalent. Scale bars, 10 μm.

embryos. IAHC were readily observed in control het/het and knockout (KO) embryos for either GFI1 or GFI1B (Fig. 4a(i)–(iii) and Supplementary Fig. 5a). In contrast, no c-KIT⁺ IAHCs were found in the double *Gfi1s* KO (*Gfi1*GFP/GFP/*Gfi1b*GFP/GFP) AGMs (Fig. 4a(iv)). Instead, we detected GFP⁺ cells, which were fully embedded within the lining of the vDA, suggesting a critical impairment in the EHT and IAHC generation. The presence of clusters in *Gfi1*-deficient

embryos might be explained by the upregulation of *Gfi1b* that could result in a functional compensation for the loss of GFI1. Indeed, it has been previously demonstrated that GFI1B could functionally replace and rescue the haematopoietic defects observed in *Gfi1*-deficient animals[28,33]. Accordingly, we observed the appearance of *Gfi1b* transcripts in the endothelial lining of the dorsal aorta of *Gfi1*-deficient embryos by *in situ* staining (Supplementary Fig. 5b).
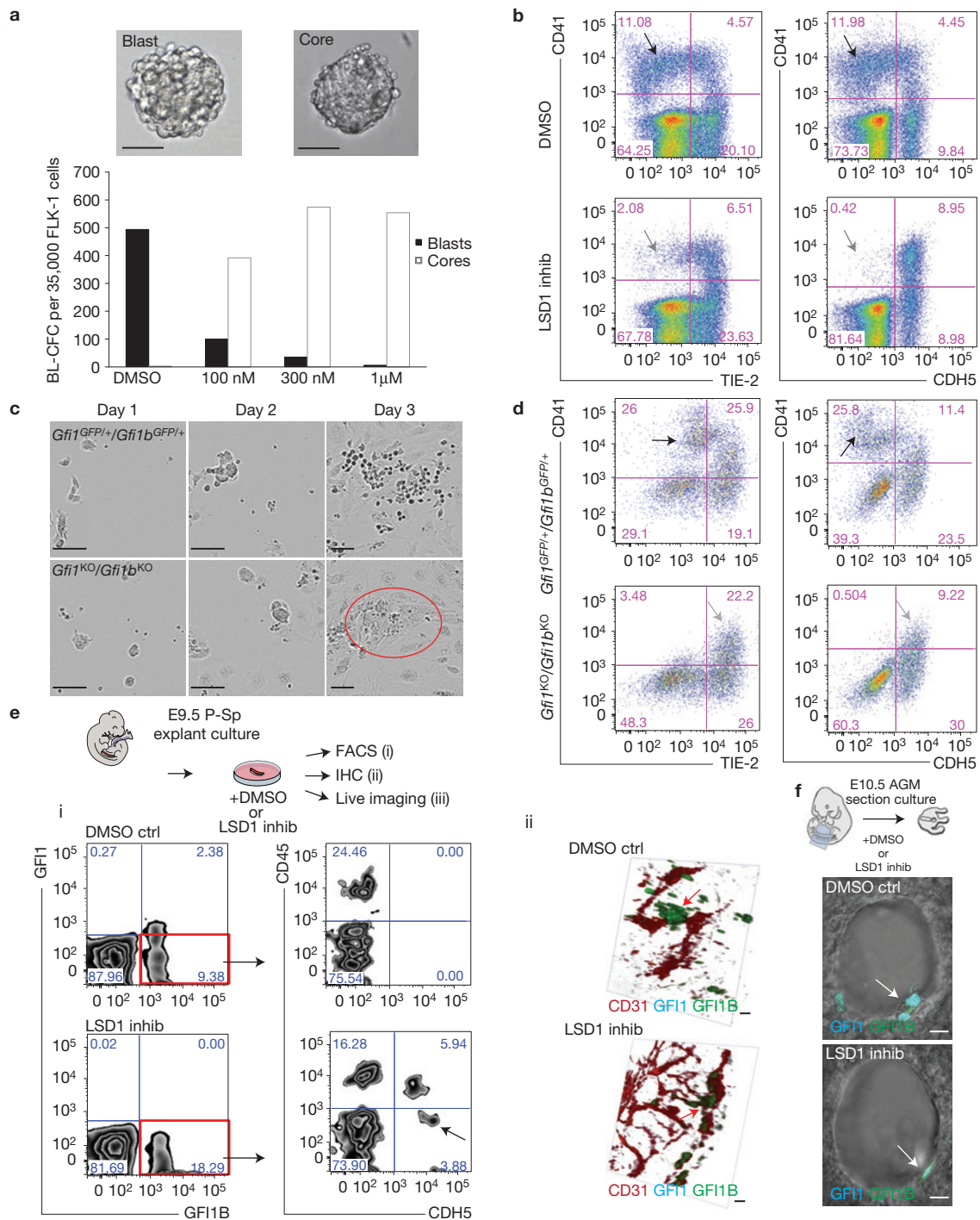
Figure 5 LSD1 is critical for EHT. (**a**) ESC-derived FLK1[+] cells were either cultured with dimethylsulphoxide (DMSO) or different concentrations of the LSD1 inhibitor in blast colony-forming units (BL-CFC; one representative experiment out of two independent experiments is shown, data from the second experiment can be found in Supplementary Table 3). (**b**) FACS on day 3 monolayer cultures of ESC-derived FLK1[+] cells treated with DMSO or LSD1 inhibitor (300 nM; representative FACS plots of more than 6 independent experiments). CD41[+] single (black arrow), TIE-2[+]/CD41[+] and CDH5[+]/CD41[+] double-positive (grey arrow) cells are indicated. (**c**) Time-lapse images of monolayer cultures (Day 3) of $Gfi1^{KO}/Gfi1b^{KO}$ or $Gfi1^{het}/Gfi1b^{het}$ ESC-derived FLK1[+] cells (representative images of 2 independent experiments). (**d**) FACS on cells isolated from day 5 EBs of $Gfi1^{KO}/Gfi1b^{KO}$ or $Gfi1^{het}/Gfi1b^{het}$ (representative FACS plots of more than 6 independent experiments). CD41[+] single

(black arrow), TIE-2[+]/CD41[+] and CDH5[+]/CD41[+] double-positive (grey arrow) cells are indicated. (**e**) E9.5 P-Sp $Gfi1^{Tomato}/Gfi1b^{GFP}$ explant cultures treated with DMSO (DMSO ctrl) or LSD1 inhibitor (300–500 nM). The explants were analysed either by FACS for the presence of GFI1B[+] cells in the CDH5[+] endothelial compartment (representative data of more than 3 independent experiments) (i) or by imaging to generate a three-dimensional reconstruction of E9.5 AGM explants cultured with DMSO (DMSO ctrl) or 500 nM of the LSD1 inhibitor (LSD1 inhib; representative data of 3 independent experiments) (ii). IHC for CD31 (red), GFI1 (cyan) and GFI1B (green) was performed before imaging. (**f**) Imaging of E10.5 $Gfi1^{Tomato}/Gfi1b^{GFP}$ AGM sections after overnight *ex vivo* culture with DMSO (DMSO ctrl) or 500 nM of the LSD1 inhibitor (LSD1 inhib; representative data of 3 independent experiments). Scale bars, 50 μm (**a**,**c**), 10 μm (**e**) and 20 μm (**f**).

By quantitative PCR, we also detected an upregulation of *Gfi1b* expression in the whole AGM lysate and more specifically in HE cells (CDH5$^+$/GFP$^+$/c-KIT$^-$) of *Gfi1*-deficient embryos (*Gfi1*^{KO}/*Gfi1b*^{het}; Supplementary Fig. 5c). Together, these data suggest that GFI1 is critical for the initiation of the EHT, but that its absence can be compensated by induction of *Gfi1b* expression.

Finally, to test whether *Gfi1*^{KO}/*Gfi1b*^{KO} AGM cells possessed any haematopoietic potential, we plated single-cell suspensions of *Gfi1*^{KO}/*Gfi1b*^{KO} AGM in haematopoietic assays either directly, or after a maturation step on OP9 cells (Fig. 4b,c). None of the conditions led to haematopoietic colony formation, indicating an absolute requirement for GFI1 and GFI1B in the generation, and/or the subsequent maturation, of blood precursors in the AGM.

LSD1 is essential for the EHT

Both GFI1 and GFI1B have been shown to epigenetically repress transcription in MEL (murine erythroleukaemia) cells by recruiting the chromatin regulatory CoREST complex[34]. The CoREST complex includes the histone demethylase LSD1 (KDM1A) and the histone deacetylases HDAC1 and HDAC2. On recruitment to target sites, HDACs and LSD1 remove the activating acetylation and methylation marks from histones within their proximity[34]. Stable repression of the target loci is then further induced by the recruitment of G9 or SUV39H1, and subsequent methylation of repressive sites.

To investigate whether the EHT is regulated by a similar epigenetic mechanism, we examined the consequences of pharmacological LSD1 inactivation[35] on blood emergence during the *in vitro* differentiation of embryonic stem cells (ESCs). In this system, FLK1$^+$ mesodermal haemangioblasts plated in methylcellulose form tight HE cores from which round haematopoietic cells emerge to produce blast colonies[11,36]. We first tested the LSD1 inhibitor in this blast colony assay and observed that increasing amounts of the inhibitor markedly decreased the number of blast colonies. In parallel, we detected a directly correlated increase in the number of cores, suggesting that the LSD1 inhibition blocks the emergence of round blood cells from HE cores (Fig. 5a). We next assessed the impact of LSD1 inhibition in monolayer cultures by time-lapse imaging and FACS. In these cultures, HE cells aggregate as tight and adherent clusters that are characterized by the co-expression of endothelial (CDH5 and TIE2/TEK) and haematopoietic markers (c-KIT and CD41). These HE clusters then generate free-floating haematopoietic cells concomitantly with progressive loss of endothelial markers[9]. We observed that control cells transited through a HE stage and generated round blood cells (Supplementary Video 2). Strikingly, cultures treated with the LSD1 inhibitor generated HE clusters, but did not progress to produce free-floating blood cells (Supplementary Video 3). FACS analysis of the control cultures confirmed the presence of HE cells (CDH5$^+$/CD41$^+$ or TIE2$^+$/CD41$^+$) and of the subsequent CD41$^+$ single-positive blood cells (Fig. 5b). In contrast, in the presence of the inhibitor, very few CD41$^+$ single-positive blood cells were detected, suggesting a block in EHT. We confirmed this striking phenotype in ESCs carrying floxed *Lsd1* alleles[37]. On induction of deletion, we observed a similar defect in the generation of round blood cells (Supplementary Videos 4 and 5 and Supplementary Fig. 6a). Finally, consistent with previous reports[38,39], we observed a decrease in cell proliferation

and an increase in apoptosis on *Lsd1* deletion during *in vitro* differentiation (Supplementary Fig. 6b,c).

To correlate this observation with the activity of GFI1 proteins, we assessed the *in vitro* phenotype resulting from the combined loss of *Gfi1* and *Gfi1b*. As observed for LSD1 deficiency, differentiating *Gfi1*^{KO}/*Gfi1b*^{KO} ESCs formed HE clusters, but the emergence of round haematopoietic cells (Fig. 5c), and the downregulation of CDH5 and TIE2, were not observed (Fig. 5d). To investigate whether LSD1 activity was also required in the AGM for the generation of blood cells, we analysed *Lsd1* expression and performed *ex vivo* LSD1 inhibition in E9.5 P-Sp explants and E10.5 embryonic sections of *Gfi1*$^{Tomato/+}$/*Gfi1b*$^{GFP/+}$ embryos. We observed that *Lsd1* is ubiquitously expressed in the mouse embryo, consistent with a previous report[37] (Supplementary Fig. 7a). Following culture of control E9.5 P-Sp explants, GFP$^+$ (*Gfi1b* expressing) cells were found to be negative for the endothelial markers CD31 or CDH5 by FACS and were detected as clusters of cells by confocal microscopy (Fig. 5e and Supplementary Video 6). In contrast, following LSD1 inhibition, some of the GFI1B$^+$ cells were still CDH5$^+$ (Fig. 5e). In addition, some *Gfi1b*-expressing cells were still found embedded within the CD31$^+$ endothelial lining of the vDA (Supplementary Video 7) but expressed the haematopoietic marker CD45 (Fig. 5e(ii) and Supplementary Fig. 7b). Similarly, in cultures of E10.5 sections treated with the LSD1 inhibitor, *Gfi1*- and *Gfi1b*-expressing cells remained embedded in the endothelial lining of the vDA (Fig. 5f and Supplementary Fig. 7c). Together, these results indicate the critical requirement of LSD1 for the emergence of blood cells, both in differentiating ESCs and in the AGM.

GFI1 and GFI1B shut down the HE program by recruiting LSD1 to target genes

To identify the genome-wide molecular changes induced by GFI1 and GFI1B through the recruitment of LSD1 during EHT, we compared global gene expression profiles of LSD1-inhibited ESC-derived HE cells (CDH5$^+$/CD41$^+$) and the corresponding CD41 compartment of control cells (CDH5$^+$/CD41$^+$ HE and CD41$^+$ haematopoietic cells). We found 244 protein-coding transcripts present at higher, and 165 at lower, levels in LSD1-inhibited cells (Fig. 6a). Transcripts with higher levels in LSD1-inhibited cells were enriched for genes implicated in cardiovascular system development, whereas transcripts with lower levels of expression were associated with haematological system development/function and cell morphology (Fig. 6b). In parallel, we also mapped GFI1- and GFI1B-binding sites in ES-derived HE cells using the DamID (DNA adenine methyltransferase Identification) strategy[40,41] (Fig. 6c and Supplementary Fig. 7d). Consistent with their potential function in HE, the genes bound by GFI1 and/or GFI1B were highly enriched for transcripts involved in integrin, actin cytoskeleton, and gap junction signalling (Fig. 6d); pathways that need to be modulated for EHT. Next, we overlapped the lists of genes bound by GFI1 and/or GFI1B with the list of genes expressed at higher levels when LSD1 activity is blocked. The resulting list of 78 candidate genes (Fig. 7a) contained genes with well-established roles in stem cells (*Lgr5, Lin28A, Sall1*), as well as in cardiovascular development, blood vessels maintenance and remodelling (*Egfl7, Nos3, Pcdh12, Ptprb, Sox7, Cdh5, Esam, Hey2, Hand2, ID1, ID3*). Accordingly, this list was significantly enriched for genes involved in development and, in particular, in cardiovascular system development
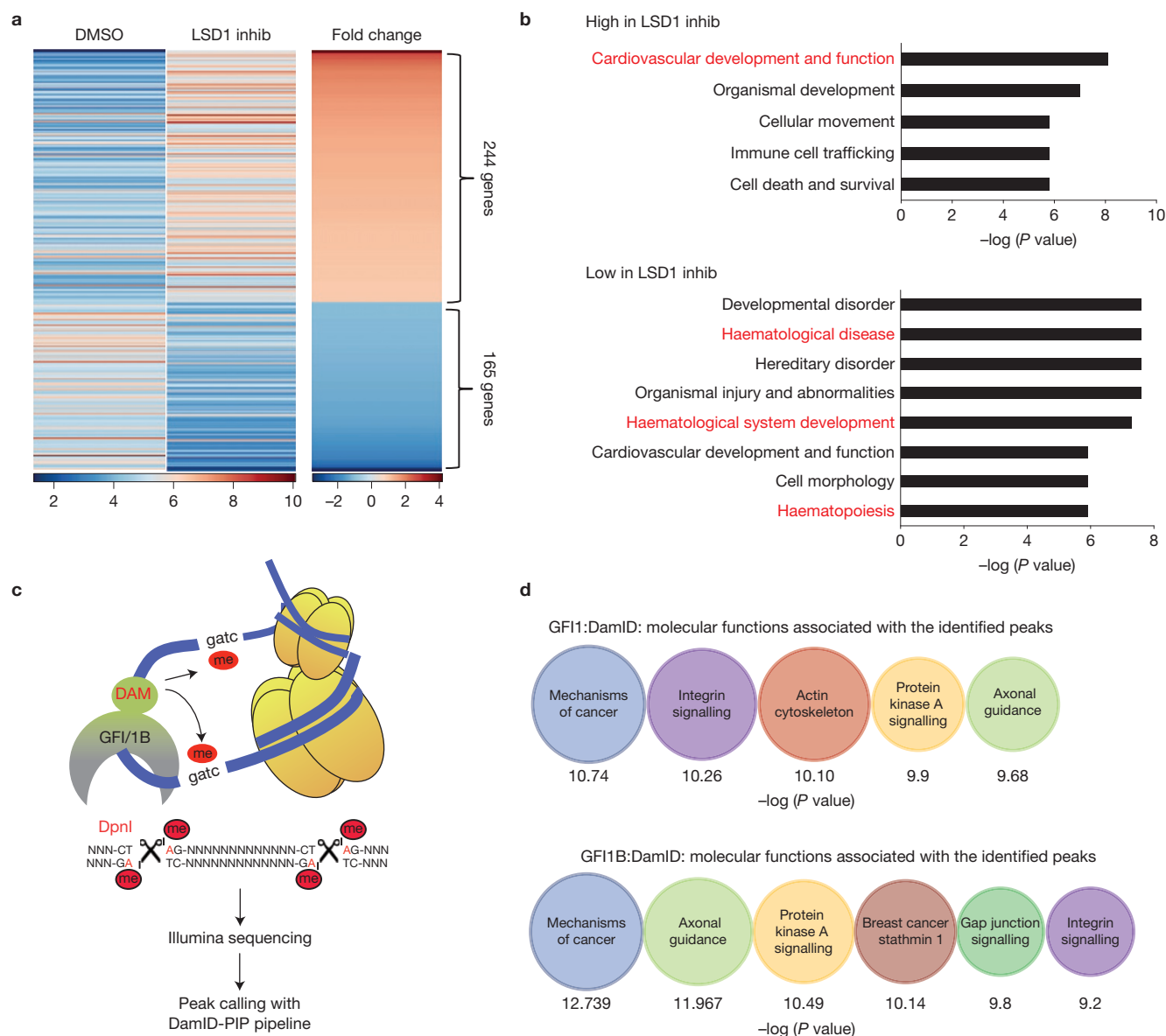
Figure 6 Molecular program governed by LSD1 and GFI1/1B. (**a**) Heat map of differentially expressed genes between ESC-derived CD41+ cells in day 2 liquid blast culture in the presence of LSD1 inhibitor (300 nM) or DMSO. (**b**) Molecular function of the genes expressed at higher and lower levels in the LSD1-inhibited cells compared with control cells. (**c**) Schematic representation of the DamID technique. (**d**) GO terms of genes bound by GFI1 and GFI1B in day 2 liquid blast-derived HE cells.

(Fig. 7b). Interestingly, 33 out of the 78 genes have been previously shown to be bound by RUNX1 in ChIP-seq experiments during EHT (ref. 42; Fig. 7c and genes highlighted in red in Fig. 7a). *In silico* analysis also indicated a frequent occurrence of RBPJ-binding motifs within, or close to the GFI1 and GFI1B peaks (Fig. 7d and genes indicated by an asterisk in Fig. 7a), suggesting that both RUNX1 and *Notch* signalling could additionally regulate the transcription of these genes. Finally, the *S1P* signalling pathway, implicated in vascular permeability[43], was found to be significantly associated with genes present in this list (Fig. 7b, bottom panel). Together, our molecular data strongly suggest that during EHT, GFI1 and GFI1B bind to genes associated with the global maintenance of endothelial identity, and this epigenetically silences this program through the recruitment of the CoREST complex to orchestrate the generation of HSCs and blood progenitors (Fig. 7e).

DISCUSSION

Recent studies have conclusively demonstrated the existence of HE cells that have the unique ability to give rise to HSCs and blood precursors. However, the precise nature of this transient and rare subpopulation of endothelial cells, and how they generate different types of blood cell through EHT, remains poorly understood.

In this study, we revealed an unrecognized requirement for the two transcriptional repressors GFI1 and GFI1B for HSC emergence in the AGM. Indeed, we did not detect any IAHCs or any haematopoietic potential in E10.5 *Gfi1*[KO]/*Gfi1b*[KO] AGMs, even following further

a

GFI1:DAM   GFI1B:DAM

1,335   2,115   1,664

31

17   30

149

LSD1 inhib (high)

AIM2	MDFIC
AS3MT	NOS3
ASAH2	NPL
ASB4	PCDH12
✳ CD93	PDYN
CTTNBP2NL	PTPRB
CYP26A1	✳ RGL1
DPYSL3	SCN11A
✳ EGFL7	✳ SGPP1
EPS8	✳ SHE
✳ FBN1	✳ SOX7
✳ GATA2	STC1
HAND2	✳ SYNE2
ID3	TBX3
LPAR6	WIPF3
MCTS1	

1700066B19Rik
ARHGEF15
CHST7
Ctla2a/Ctla2b
DUSP4
ENC1
GREB1L
HDX
HSD3B2
LGR5
LIN28A
NKAIN2
OIT3
PARP8
PRR16
TCF15
ZNF761

1600029D21Rik
2700046A07Rik
ACER2
ARAP3
BAALC
BC043934
BCL6B
C1orf54
C8orf4
CDH5
CYB5D1
✳ EDN1
ELTD1
EMX2
ESAM
FBN2
✳ GIMAP4
Gm266
GPM6A
HELT
HEY2
ICAM2
ID1
NPR1
POMC
RASGRP3
SALL1
SEC24D
TDGF1
THSD1

b Functional classification of the 78 candidates

Cardiovascular system development
Embryonic development
Organ development
Organismal development
Tissue development
Gene expression

P value (–log)

Enriched pathways amongst the 78 candidates

Spingosine and sphingosine-1-phosphate metabolism
Ceramide degradation
Corticotropin releasing hormone signalling

P value (–log)

c Co-regulated by GFI1 and/or GFI1B and LSD1 in HE

RUNX1-binding 33/78   No RUNX1-binding 44/78

d Binding motifs 100 bp up/down stream of GFI1:DAM and/or GFI1B:DAM peaks in the 78 final candidates:

RBPJ binding motif

P-value: 10^{-10}

e

Aortic lumen

CDH5[+]/RUNX1[+]/GFI1[+] (c-KIT[−])

Haemogenic endothelial cells

CDH5[+]/RUNX1[+]   GFI1/GFI1B LSD1/CoREST

Silencing of HE programme

CDH5[+]/RUNX1[+]/GFI1[+] (c-KIT[−])

CDH5[+]/RUNX1[+]/ GFI1[+]/GFI1B[−] (c-KIT[+])

CDH5[+]/RUNX1[+]

Haemogenic endothelial cells

CDH5[+]/RUNX1[+]/GFI1[+] (c-KIT[−])

CDH5[+]/RUNX1[+]/GFI1[+]/ GFI1B[+] (CD45[+])

CDH5[+]/RUNX1[+]/ GFI1[+]/GFI1B[+] (CD45[−])

CDH5[+]/RUNX1[+]/ GFI1[−]/GFI1B[+] (CD45[−])

CDH5[+]/RUNX1[+]

Haemogenic endothelial cells

CDH5[+]/RUNX1[+]/GFI1[+]/GFI1B[+] (c-KIT[+])
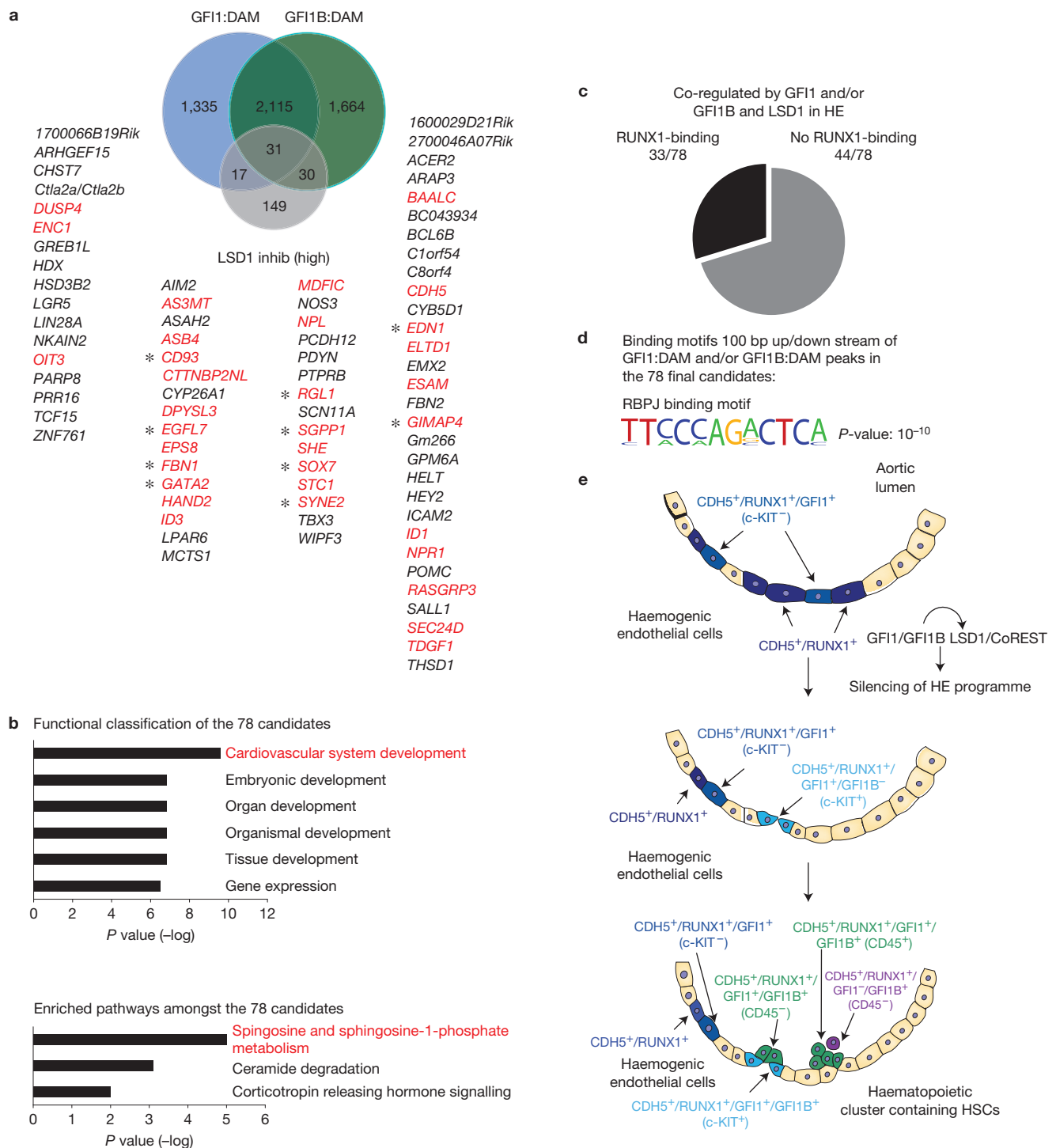
Haematopoietic cluster containing HSCs

Figure 7 Identification of genes repressed by GFI1 and GF1B in HE. (**a**) Overlap of genes bound by *Gfi1* and *Gfi1b* and expressed at higher levels of transcription on LSD1 inhibition. Genes previously shown to harbour RUNX1 ChIP-seq peaks are indicated in red and genes with RBPJ-binding motifs within 100 bp of GFI1 or GFI1B peaks are indicated by an asterisk. (**b**) Functional classification and pathway enrichment analysis for the 78 candidate genes. (**c**) Pie chart representing the number of genes with/out RUNX1 ChIP-seq binding peaks. (**d**) RBPJ-binding motif identified by HOMER in 100 bp proximity of GFI1 and/or GFI1B peaks. (**e**) Model of the expression patterns and function of GFI1 and GFI1B during HE specification and EHT in the AGM.

maturation, indicating an absolute requirement for GFI1 and GFI1B for the commitment and generation of haematopoietic cells. Interestingly, this developmental block is not observed in *Gfi1* single

KO suggesting that the critical function of GFI1 in HE cells could be rescued by the compensatory expression of *Gfi1b* that we observed in these cells. In the double-KO embryos, discrete GFP[+] cells were

observed embedded within the endothelium of the vDA, indicating that HE cells are specified in the absence of GFI1 and GFI1B proteins, but are unable to progress through EHT. This is reminiscent of the situation in the yolk sac where the double-KO GFP$^+$ cells remain clustered together and are not disseminated in the vasculature[28]. In contrast, the absence of blood potential in *Gfi1*^{KO}/*Gfi1b*^{KO} AGMs is distinct from the situation in the yolk sac where GFP$^+$ cells were shown to exhibit haematopoietic potential when dissociated and replated[28]. These results suggest differences in the molecular regulation during the generation of blood cells in the yolk sac and in the AGM. This could also reflect the existence of different types of HE at these two sites, as previously suggested[44]. Interestingly, this dichotomy mirrors the different requirements for Notch signalling that is dispensable for yolk sac haematopoiesis but essential for AGM haematopoiesis[45–48].

Detailed analysis of the EHT has so far been challenging given the limited number of genetic markers specifically defining HE. Our results suggest that *Gfi1* expression identifies the specific subset of cells in the endothelial lining that are actively undergoing the transition to the haematopoietic fate. As such, GFI1 delineates a more refined cell population than previously described markers such as the endothelial markers SOX17 (ref. 49) and SOX7 (ref. 50) that are widely expressed in endothelial cells, or the haematopoietic markers SCA-1 (ref. 51) and RUNX1 (refs 20,52) that are extensively expressed in other AGM cell populations.

We also revealed here that LSD1, a member of the CoREST repressive complex, is critical for the EHT process. *Lsd1* deficiency phenocopies the developmental block observed in *Gfi1*^{KO}/*Gfi1b*^{KO}, strongly suggesting that the repressive function of the *Gfi1s* during EHT is mediated by the CoREST complex. Our molecular analysis identified a specific set of 78 genes that are bound and repressed by GFI1 and GFI1B during EHT. This list contains not only genes that have not been previously identified as GFI or GFI1B targets[53,54], but also some genes that have known, or suspected functions in stem cell and endothelial cells. This list also harbours a wealth of new candidates potentially implicated in the EHT. In addition, a large proportion of these genes has been shown to be bound by RUNX1 (ref. 42) and contains a RBPJ-binding motif suggesting that GFI1 repression, RUNX1 binding and Notch signalling overlap and cooperate to orchestrate the EHT process in the AGM.

Together, our study provides mechanistic insights and valuable new candidates and leads for future studies of the molecular program governing HE specification and the generation of blood cells and HSCs in the AGM. Deciphering the regulatory pathways that specifically control this EHT process is essential for precisely recapitulating this program during ESC differentiation[55–58] or during reprogramming[59–61] to ultimately generate HSCs *in vitro*.   □

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary Information is available in the online version of the paper

AUTHOR CONTRIBUTIONS
R.T. designed and performed most of the experiments, analysed the data and wrote the manuscript, M.M. initiated the project, designed performed experiments and analysed the data. R.P., V.M., M.S. and C.L. designed and performed experiments. E.M. and Y.L. performed bioinformatics analysis on the sequencing data and microarray. T.C., T.M., C.R., C.M., S.C. and B.G. contributed valuable tools and protocols. V.K. and G.L. designed and supervised the research project, analysed the data, and wrote the manuscript.

1. Costa, G., Kouskoff, V. & Lacaud, G. Origin of blood cells and HSC production in the embryo. *Trends Immunol.* **33,** 215–223 (2012).
2. Palis, J., Robertson, S., Kennedy, M., Wall, C. & Keller, G. M. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126,** 5073–5084 (1999).
3. Frame, J. M., McGrath, K. E. & Palis, J. Erythro-myeloid progenitors: 'definitive' hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells Mol. Dis.* **51,** 220–225 (2013).
4. McGrath, K. E. *et al.* Distinct sources of hematopoietic progenitors emerge before HSCs and provide functional blood cells in the mammalian embryo. *Cell Rep.* **11,** 1892–1904 (2015).
5. Frame, J. M., Fegan, K. H., Conway, S. J., McGrath, K. E. & Palis, J. Definitive hematopoiesis in the yolk sac emerges from Wnt-responsive hemogenic endothelium independently of circulation and arterial identity. *Stem Cells* http://dx.doi.org/10.1002/stem.2213 (2015).
6. Medvinsky, A. J. & Dzierzak, E. A. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* **86,** 897–906 (1996).
7. Müller, A. M., Medvinsky, A. J., Strouboulis, J., Grosveld, F. & Dzierzak, E. A. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* **1,** 291–301 (1994).
8. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lièvre, F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125,** 4575–4583 (1998).
9. Zovein, A. C. *et al.* Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* **3,** 625–636 (2008).
10. Eilken, H. M., Nishikawa, S., Nishikawa, S. & Schroeder, T. Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457,** 896–900 (2009).
11. Lancrin, C. *et al.* The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* **457,** 892–895 (2009).
12. Nishikawa, S. *et al. In vitro* generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. *Immunity* **8,** 761–769 (1998).
13. Boisset, J.-C. *et al. In vivo* imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* **464,** 116–120 (2010).
14. Bertrand, J. Y. *et al.* Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* **464,** 108–111 (2010).
15. Kissa, K. & Herbomel, P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* **464,** 112–115 (2010).
16. Lam, E. Y. N., Hall, C. J., Crosier, P. S., Crosier, K. E. & Flores, M. V. Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. *Blood* **116,** 909–914 (2010).
17. de Bruijn, M. F. T. R., Speck, N. A., Peeters, M. C. & Dzierzak, E. A. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J.* **19,** 2465–2474 (2000).
18. Taoudi, S. & Medvinsky, A. J. 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).
19. Boisset, J.-C. *et al.* Progressive maturation toward hematopoietic stem cells in the mouse embryo aorta. *Blood* **125,** 465–469 (2015).
20. North, T. *et al.* Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126,** 2563–2575 (1999).
21. Sroczynska, P., Lancrin, C., Kouskoff, V. & Lacaud, G. The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. *Blood* **114,** 5279–5289 (2009).
22. Lacaud, G. *et al.* Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development *in vitro*. *Blood* **100,** 458–466 (2002).

23. North, T. E. *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16**, 661–672 (2002).

24. North, T. E., Stacy, T., Matheny, C. J., Speck, N. A. & de Bruijn, M. F. T. R. Runx1 is expressed in adult mouse hematopoietic stem cells and differentiating myeloid and lymphoid cells, but not in maturing erythroid cells. *Stem Cells* **22**, 158–168 (2004).

25. Okuda, T., Hiebert, S. W. & Downing, J. R. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84**, 321–330 (1996).

26. Wang, Q. *et al.* 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 (1996).

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

28. Lancrin, C. *et al.* GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood* **120**, 314–322 (2012).

29. Vassen, L., Okayama, T. & Möröy, T. Gfi1b: green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. *Blood* **109**, 2356–2364 (2007).

30. Rybtsov, S. *et al.* Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. *J. Exp. Med.* **208**, 1305–1315 (2011).

31. Taoudi, S. *et al.* Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin+CD45+ pre-definitive HSCs. *Cell Stem Cell* **3**, 99–108 (2008).

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

33. Fiolka, K. *et al.* Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. *EMBO Rep.* **7**, 326–333 (2006).

34. Saleque, S., Orkin, S., Kim, J. & Rooke, H. M. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Mol. Cell* **27**, 562–572 (2007).

35. Harris, W. J. *et al.* The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. *Cancer Cell* **21**, 473–487 (2012).

36. Lancrin, C. *et al.* Blood cell generation from the hemangioblast. *J. Mol. Med.* **88**, 167–172 (2010).

37. Foster, C. T. *et al.* Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol. Cell. Biol.* **30**, 4851–4863 (2010).

38. Wang, J. *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat. Genet.* **41**, 125–129 (2009).

39. Whyte, W. A. *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* **482**, 221–225 (2012).

40. Vogel, M. J., Peric-Hupkes, D. & van Steensel, B. Detection of *in vivo* protein-DNA interactions using DamID in mammalian cells. *Nat. Protoc.* **2**, 1467–1478 (2007).

41. Lie-A-Ling, M. *et al.* RUNX1 positively regulates a cell adhesion and migration program in murine hemogenic endothelium prior to blood emergence. *Blood* **124**, e11–e20 (2014).

42. Lichtinger, M. *et al.* RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J.* **31**, 4318–4333 (2012).

43. Wang, L. & Dudek, S. M. Regulation of vascular permeability by sphingosine 1-phosphate. *Microvasc. Res.* **77**, 39–45 (2009).

44. Chen, M. J. *et al.* Erythroid/myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells. *Stem Cell* **9**, 541–552 (2011).

45. Hadland, B. K. *et al.* A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. *Blood* **104**, 3097–3105 (2004).

46. Kumano, K. *et al.* Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. *Immunity* **18**, 699–711 (2003).

47. Bigas, A. & Robert-Moreno, L. The Notch pathway in the developing hematopoietic system. *Int. J. Dev. Biol.* **54**, 1175–1188 (2010).

48. Guiu, J. *et al.* Hes repressors are essential regulators of hematopoietic stem cell development downstream of Notch signaling. *J. Exp. Med.* **210**, 71–84 (2013).

49. Costa, G. *et al.* SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. *Development* **139**, 1587–1598 (2012).

50. Clarke, R. L. *et al.* The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat. Cell Biol.* **15**, 1–10 (2013).

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

52. Swiers, G. *et al.* Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nat. Commun.* **4**, 2924 (2013).

53. Kim, W., Klarmann, K. D. & Keller, J. R. Gfi-1 regulates the erythroid transcription factor network through Id2 repression in murine hematopoietic progenitor cells. *Blood* **124**, 1586–1596 (2014).

54. Möröy, T. & Khandanpour, C. Growth factor independence 1 (Gfi1) as a regulator of lymphocyte development and activation. *Semin. Immunol.* **23**, 368–378 (2011).

55. Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G. & Kouskoff, V. *In vivo* repopulating activity emerges at the onset of hematopoietic specification during embryonic stem cell differentiation. *Stem Cell Rep.* **4**, 431–444 (2015).

56. Ledran, M. H. *et al.* Efficient hematopoietic differentiation of human embryonic stem cells on stromal cells derived from hematopoietic niches. *Cell Stem Cell* **3**, 85–98 (2008).

57. Wang, L. *et al.* Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic HOXB4 expression. *J. Exp. Med.* **201**, 1603–1614 (2005).

58. Kyba, M., Perlingeiro, R. C. R. & Daley, G. Q. HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29–37 (2002).

59. Pereira, C. F. *et al.* Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* **13**, 205–218 (2013).

60. Sandler, V. M. *et al.* Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature* **511**, 312–318 (2014).

61. Batta, K., Florkowska, M., Kouskoff, V. & Lacaud, G. Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *Cell Rep.* **9**, 1871–1884 (2014).

METHODS

Mouse lines and embryos. The *Gfi1^Tomato* line was generated as previously described[62,63]. Briefly, the cDNA encoding the fluorescent H2B–(Td)Tomato protein was introduced into a BAC containing the entire *Gfi1* locus by recombineering. We used the experimental conditions previously described in ref. 62 but replaced the LacZ cDNA with the H2B–(Td)Tomato cDNA. The modified BAC was used to create a transgenic ESC line, and subsequently a *Gfi1:H2B–Tomato* mouse line. *Tomato* expression defined similar haematopoietic compartments as *GFP* expression in the previously established *Gfi1:GFP* knock-in line[64] (Supplementary Fig. 1a). The *Gfi1^+/GFP* and *Gfi1b^+/GFP* mouse lines have been previously described[29,64]. For timed matings, wild-type ICR females (aged 6–12 weeks) were mated to *Gfi1^Tomato*/*Gfi1b^GFP* males, or *Gfi1^+/GFP*/*Gfi1b^+/GFP* females (C57/Bl6, aged 6–20 weeks) were mated to *Gfi1^+/GFP*/*Gfi1b^+/GFP* males (C57/Bl6, aged 6–30 weeks). Vaginal plug detection was considered as day 0.5. All animal work was performed under regulation in accordance with the United Kingdom Animal Scientific Procedures Act (ASPA) 1986. Animal experiments were approved by the Animal Welfare and Ethics Review Body (AWERB) of the Cancer Research UK Manchester Institute.

Immunohistochemistry. IHC staining was described previously[28,32]. For IHC, embryos were fixed for 2 h in 4% PFA and 10 μm cryostat sections were prepared from frozen embryos. The sections were washed 3 times in PBS + 1% Triton for 30 min at room temperature before a blocking step. The sections were incubated for 1 h in blocking solution (PBS 1% Triton + 10% FCS + 0.2% sodium azide) at room temperature twice and then incubated in peroxidase block (3% H_2O_2 diluted in blocking buffer) overnight at 4 °C. This was followed by two washes for 15 min each in blocking buffer. Subsequently, a streptavidin/biotin block (Streptavidin/Biotin blocking kit SP-2002 by Vector laboratories) was performed according to the manufacturer's instructions. The samples were washed 2 × 15 min in blocking buffer and the primary antibody was added at the required dilution/concentration overnight (list of antibodies in Supplementary Table 1). The samples were washed 3 × 1 h in PBS 1% Triton + 10% FCS 0.2% sodium azide, followed by a 3 × 10 min wash in PBS 1% Triton. The secondary antibody was added in blocking buffer (no sodium azide) overnight. Finally, the samples were washed 3 × 1 h in PBS. The stained sections were mounted using Prolong Gold anti-fade medium with DAPI (Life Technologies) and images were acquired with a low-light microscope (Leica) and the Metamorph imaging software and processed with ImageJ. IHC was performed on several AGM sections from independent litters and representative images were chosen for publication.

***In situ* hybridization.** E10.5 embryos were fixed in 4% PFA overnight and dehydrated into methanol using a graded methanol/PBT series. After re-hydration, the embryos were cut into 5-μm-thick paraffin sections. *In situ* hybridization for *Gfi1b* was performed using an RNAview probe from Affymetrix (*Gfi1b* probe raised against NM_008114) at a dilution of 1:120 and the samples were run on a Leica Bond Rx according to the manufacturer's instructions with the following modification: before the hybridization step, antigene retrieval was obtained with a 5 min incubation step at −20° in acetone. The sections were washed in ddH₂O before and after the acetone step.

Explant and re-aggregation cultures. P-Sp and AGM regions were dissected and dissociated as previously described[65]. In brief, the AGMs of E10.5 or E11.5 embryos were dissected in dissecting media (PBS with 7% fetal calf serum (FCS) and penicillin/streptomycin (100 U ml⁻¹)). For dissociation of the AGM into a single-cell suspension, AGMs were incubated for 20–30 min in 500 μl of 1 mg ml⁻¹ of collagenase/dispase (Roche cat. no. 10269638001) before mechanical dissociation with a syringe and needle. The resulting single-cell suspension was used for FACS sorting or re-aggregation with OP-9 cells. For explant cultures, whole AGMs were cultured on top of a 0.65 μm Durapore filter (Millipore, cat. no. DVPP02500) supported by a steel mesh in MyeloCult media (Stem Cell cat. no. M5300) supplemented with 10⁻⁵ M hydrocortisone at 37 °C, 5% CO_2 and a gas–liquid interface for 2 days[6]. Re-aggregation culture was performed as described previously[26]. Single AGM cells were mixed with OP-9 cells and cultured as a hanging drop (Terasaki plates, greiner bio cat. no. CPL-147) overnight in re-aggregation medium (IMDM (Invitrogen), 20% fetal calf serum, L-glutamine (4 mM), penicillin/streptomycin (50 units ml⁻¹), mercaptoethanol (0.1 mM), IL-3 (100 ng ml⁻¹), SCF (100 ng ml⁻¹) and Flt3L (100 ng ml⁻¹); growth factors were purchased from Peprotech) before being moved on top of a 0.65 μm Durapore filter (Millipore, cat. no. DVPP02500) supported by a steel mesh. Tissues were maintained in 5% CO_2 at 37 °C in a humidified incubator.

***In vitro* culture of AGM cells on OP-9 stromal cells.** FACS-sorted *Gfi1^Tomato* and *Gfi1b^+/GFP* cells were cultured on an OP-9 stromal cell-coated plate in

re-aggregation media (IMDM (Invitrogen), 20% fetal calf serum, L-glutamine (4 mM), penicillin/streptomycin (50 units ml⁻¹), mercaptoethanol (0.1 mM), IL-3 (100 ng ml⁻¹), SCF (100 ng ml⁻¹) and Flt3L (100 ng ml⁻¹)).

Live *ex vivo* imaging of embryonic sections. E10.5 *Gfi1^Tomato Gfi1b^GFP* AGM sections (50 μm thick) were cut using a tissue chopper (McIlwain) and stained with CD31–Alexa 647 in re-aggregation media for 20–30 min as previously published[13]. The sections were then washed 2 times for 5 min with re-aggregation media and mounted in low-melting-point agarose in glass-bottomed 35 mm Petri dishes (Matek, cat no. P35G-1.5-14-C). Time-lapse images were performed on a Zeiss axiovert 200M. The microscope is equipped with a 300 w xenon bulb light source and an automated *xy* stage in a temperature and gas controlled cabinet. A Zeiss ×20.0 neofluor dry lens objective was used.

EDU and AnnexinV staining. EDU and AnnexinV/7AAD staining were performed on day 2 of blast cultures according to the manufacturer's instructions (Click-it EDU Alexa647, Invitrogen cat. no. C10419 and AnnexinV apoptosis detection kit from eBioscience cat. no. 88-8007-72). In brief, day 2 liquid blast cultures were incubated at 37 °C in medium with/out EdU (final concentration 10 μM) for 2 h. The cells were transferred into tubes and washed with 1 ml 1% BSA/PBS and stained according to the manufacturer's instructions and analysed on a LSRII FACS Analyzer.

Repopulation assay. AGM tissues from CD45.1/CD45.2 embryos were dissected, disassociated into a single-cell suspension and stained for flow cytometry. Sorted cells were intravenously injected into 6–8-week-old sub-lethally irradiated NSG mice (CD45.1) at 4.5 ee (embryo equivalent) along with 2 × 10⁴ BM cells from C57/Bl6 mice (CD45.2). Donor chimaerism in the recipients was determined at 5, 9, 12 and at 16 or 17 (terminal samples) weeks post transplantation by FACS analysis of CD45.1 and CD45.2 expression.

ESC line derivation, growth, differentiation and time-lapse imaging. ESCs were derived as previously described[66] from blastocysts generated by crossing *Gfi1^+/GFP*/*Gfi1b^+/GFP* mice. *Bry–GFP* (ref. 67), *LSD1^Δ/lox* (ref. 37), *Gfi1^KO*/*Gfi1b^KO*, *Gfi1:Dam*, *Gfi1b:Dam* and *Dam* ESC lines were used. Culture and differentiation of ESC lines were performed as described before[68]. ESC lines were authenticated by PCR and tested monthly for mycoplasma contamination. For time-lapse imaging, day 3 FLK1⁺ cells were plated at 6.5 × 10⁴ in 6-well plates and imaged using time-lapse imaging in the IncuCyte FLR device (Essen Instruments). The videos were cropped and labelled with Final Cut Pro.

BL-CFC and CFU assays. Blast and CFU colony assays were performed as previously described[50].

Flow cytometry and cell sorting. Stainings were analysed on a Fortessa (BD Biosciences). Sorts were performed on a FACSAriaIII, a BD Influx (BD Biosciences), or by using magnetic sorting (Miltenyi Biotec). Monoclonal antibodies and streptavidin used were FLK-1–bio, TIE2–PE and c-KIT–APC–eFluor780 (2B8; eBioscience), CD41–PE (MWReg30; BD Biosciences), CDH5–APC (eBioBV13; eBioscience) and Strep–PECy7 (eBioscience), CD41–PE–Cy7 (MWReg30; eBioscience), CD45–PerCP Cy5.5(30-F11; Biolegend). FACS data were analysed using FlowJo software (TreeStar).

Fluidigm single-cell PCR. Single-cell gene expression analysis was performed as previously described[69]. Individual CDH5⁺/GFI1⁻/GFI1B⁻ (CDH5⁺, *n* = 14), CDH5⁺/GFI1⁺/GFI1B⁻/c-KIT⁻ (*n*=12), CDH5⁺/GFI1⁺/GFI1B⁻/c-KIT⁺ (*n*=22), CDH5⁺/GFI1⁺/GFI1B⁺ (*n*=23) and CDH5⁺/GFI1⁻/GFI1B⁺ (*n*=17) from E10.5 and E11.5 AGM (E11.5 for CDH5⁺/GFI1⁻/GFI1B⁺ cells only) were analysed. Single-cell quantitative RT–PCR was carried out using the Fluidigm BioMark platform as described, with a detection limit of Ct 25 (Supplementary Table 2). Gene expression was subtracted from the detection limit and normalized on a cell-wise basis to the mean expression of the four housekeeping genes (*Eif2b1, Mrpl19, Polr2a* and *Ubc*) in each cell. Only cells that expressed all four housekeeping genes, or for which the mean of the four housekeepers was +/− 3 s.d. from the mean of all cells were considered for subsequent analysis. A dCt value of −14 was then set as a limit where a gene was not detected. In the final analysis, 14 CDH5⁺, 10 CDH5⁺/GFI1⁺/GFI1B⁻/c-KIT⁻, 20 CDH5⁺/GFI1⁺/GFI1B⁻/c-KIT⁺, 21 CDH5⁺/GFI1⁺/GFI1B⁺ and 14 CDH5⁺/GFI1⁻/GFI1B⁺ cells were included. Hierarchical clustering was performed in R (www.r-project.org) using the hclust package and heatmap.2 from the gplots package using Spearman rank correlation and complete linkage. PCA was performed using the prcomp function.

DamID ESC lines, samples preparation and sequencing. The cDNAs encoding GFI1:DAM and GFI1B:DAM fusion proteins or a DAM control were

inserted in a *Dox*-inducible HPRT docking site in ESCs (ref. 58). Following differentiation, HE cell populations were isolated, genomic DNA prepared and digested with DpnI. The methylated DNA fragments were amplified as previously described[40,41] and sequenced on an Illumina HiSeq platform. The DamID peak identification was performed using the DamID-PIP pipeline previously described[41].

Peak validation, ontology analysis and motif discovery. GFI1 and GFI1B DamID peak motif discovery and binding motif analysis was performed using HOMER and Opossum software (default settings). Gene ontology analysis was performed with IPA software (Ingenuity Systems, www.ingenuity.com).

Accession codes. The microarray data and the GFI1:DAM and GFI1B:DAM data were submitted to GEO (accession number GSE57251).

Reproducibility of experiments. Sample sizes were determined by the nature of the experiment and variability of the output, not by a statistical method. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. Numbers of embryos are indicated in the figure legends. All representative IHC, time-lapse or *in situ* images (Figs 1, 2a and 4a, and Supplementary Figs 1b, 2a and 5a,b) were reproduced at least once. Transplantation experiments and haematopoietic colony formation assays with sorted cells from *Gfi1*^GFP/GFP^/*Gfi1b*^GFP/GFP^ mice were performed once. Time-lapse experiments of E10.5 AGM sections with/out LSD1 inhibition were repeated 5 times. All of the FACS plots are representatives of two or more analyses with embryos from different litters. Fluidigm single-cell PCR runs were performed twice. *In vitro* culture differentiation of the different ESC lines and blast colony assays were performed at least twice. The blast formation assay with/out LSD1 inhibitor (Fig. 5a) was repeated once. The source data are available in the source data file.

62. Wilson, N. K. *et al.* Gfi1 expression is controlled by five distinct regulatory regions spread over 100 kilobases, with Scl/Tal1, Gata2, PU.1, Erg, Meis1, and Runx1 acting as upstream regulators in early hematopoietic cells. *Mol. Cell. Biol.* **30,** 3853–3863 (2010).
63. Stefanska, M., Costa, G., Lie-A-Ling, M., Kouskoff, V. & Lacaud, G. Smooth muscle cells largely develop independently of functional hemogenic endothelium. *Stem Cell Res.* **12,** 222–232 (2014).
64. Yücel, R., Kosan, C., Heyd, F. & Möröy, T. Gfi1:green fluorescent protein knock-in mutant reveals differential expression and autoregulation of the growth factor independence 1 (Gfi1) gene during lymphocyte development. *J. Biol. Chem.* **279,** 40906–40917 (2004).
65. Kumaravelu, P. *et al.* 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 (2002).
66. Nichols, J. *et al.* Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. *Nat. Med.* **15,** 814–818 (2009).
67. Fehling, H. J. *et al.* Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* **130,** 4217–4227 (2003).
68. Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V. & Lacaud, G. *In vitro* differentiation of mouse embryonic stem cells as a model of early hematopoietic development. *Methods Mol. Biol.* **538,** 317–334 (2009).
69. Moignard, V. *et al.* Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat. Cell Biol.* **15,** 363–372 (2013).