A novel human pluripotent stem cell-based gene activation system identifies IGFBP2 as a mediator in the production of hematopoietic progenitors in vitro.

Paolo Petazzi[1], Telma Ventura[6], Francesca Paola Luongo[6], Heather McClafferty[7], Alisha May[6], Helen Alice Taylor[6], Micheal J Shipston[7,8], Nicola Romanò[7,8], Lesley M. Forrester[6], Pablo Menéndez[1-5], Antonella Fidanza[6,9*].

1. Josep Carreras Leukemia Research Institute, Barcelona, Spain.
2. Red Española de Terapias Avanzadas (TERAV)-Instituto de Salud Carlos III (ISCIII), Madrid, Spain.
3. CIBER-ONC, ISCIII, Barcelona, Spain.
4. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.
5. Department of Biomedicine. School of Medicine, University of Barcelona, Barcelona, Spain.
6. Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK
7. Centre for Discovery Brain Sciences, Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK
8. Zhejiang University-University of Edinburgh Joint Institute, Zhejiang University School of Medicine, Zhejiang University, Haining 310058, China
9. Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK

* Correspondence to afidanza@ed.ac.uk

Abstract

A major challenge in the stem cell biology field is the ability to produce fully functional cells from induced pluripotent stem cells (iPSCs) that are a valuable resource for cell therapy, drug screening and disease modelling. Here we developed a novel inducible CRISPR-mediated activation strategy (iCRISPRa) to drive the expression of multiple endogenous transcription factors important for *in vitro* cell fate and differentiation of iPSCs to haematopoietic progenitor cells and used this to identify a key paracrine role for IGFBP2 in the development of hematopoietic progenitors. We first identified nine candidate transcription factors that we predicted to be involved in blood cell emergence during development, then generated novel gRNAs directed to the transcriptional start site of these transcription factors that could also be detected during scRNAseq. iCRISPRa activation of these endogenous transcription factors resulted in a significant expansion of arterial-fated endothelial cells expressing high levels of IGFBP2 and showed that IGFBP2 remodels the metabolic activity during *in vitro* endothelial to hematopoietic transition. As well as providing

fundamental new insights into the mechanisms of haematopoietic cell fate and differentiation, the broader applicability of iCRISPRa provides a valuable tool for studying dynamic processes controlling developmental events as well as for recapitulating abnormal phenotypes characterised by ectopic activation of specific endogenous gene expression in a wide range of systems.

Introduction

The haematopoietic system develops early during gestation through so called “waves” of hematopoietic progenitors that arise in different anatomical regions and result in the production of various progenitor and stem cells [1–5]. The precise signalling pathways leading to the production of hematopoietic stem and progenitor cells (HSPCs) during embryonic development are yet to be completely defined, posing a limitation on how to recapitulate the process *in vitro* from pluripotent stem cells.

During development, HSPCs are generated by a subset of endothelial cells, known as hemogenic endothelium [6–9], via the endothelial to hematopoietic transition (EHT) [10]. During the EHT, endothelial cells undergo profound transcriptional remodelling whereby the expression of endothelial genes is gradually downregulated, and the transcription of the hematopoietic program is initiated [11] while the cells round up and eventually detach to enter the circulation [12,13].

To explore the molecular control on the development of the hematopoietic system and to address the differences with the *in vitro* differentiation of induced pluripotent stem cells (iPSCs), we compared our single cell transcriptomics analysis of *in vitro* derived hemogenic endothelium and early progenitors [14] to that of the *in vivo* HSC-primed human hemogenic endothelium [15]. We then developed a novel DOX-inducible CRISPR gene activation system to assess the role of the genes that were expressed at a lower level within *in vitro*-derived cells compared to their *in vivo* counterparts. We employed single-cell RNA sequencing to track the presence of guide RNAs and we monitored the phenotypic effects of gene activation. With this experimental pipeline, we identified and functionally validated a novel role for IGFBP2, IGF Binding protein 2, in the development of *in vitro* progenitors and showed that IGBP2 remodels the metabolic activity during *in vitro* EHT.

Results

Comparison of iPSC-derived endothelial cells with AGM AGM endothelial cells dataset identifies 9 differentially expressed transcription factors.

76 We, and others, have shown that *in vitro* differentiation of human iPSCs is a powerful tool to
77 model intraembryonic hematopoiesis [14,16–18]. To understand the molecular basis underlying
78 the challenges associated with the *in vitro* production of blood stem and progenitor cells *in*
79 *vitro* from differentiating iPSCs, we compared our scRNAseq dataset from differentiating
80 iPSCs[14] to that of cells derived *in vivo* from the human aorta-gonad-mesonephros (AGM)
81 region [15]. We integrated the transcriptomic data of *in vitro* derived endothelial (IVD_Endo)
82 and hematopoietic cells (IVD_HPC) with that of arterial endothelial cells (aEC), arterial
83 hemogenic cells (aHEC) and venous endothelial cells (vEC) derived from human embryos
84 collected between Carnegie stage 12 and 14 (Figure 1A). To identify possible target genes
85 that could be manipulated *in vitro* to improve iPSCs differentiation, we first determined the
86 transcription factors marking the aHEC *in vivo* (Table 1). Then, we filtered the transcription
87 factor genes that were expressed at lower level in the IVD_Endo and IVD_HPC. This
88 strategy identified 9 target transcription factors *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*,
89 *ZNF124*, *SOX6*, *ZNF33A*, *NFAT5* and *TFDP2* (Figure 1B), whose expression was detected
90 across other AGM datasets in the hemogenic endothelium[17] (Supplementary Figure 1).
91
92 **Development of a DOX-inducible dCAS9-SAM activation system in human iPSCs**
93
94 We previously developed an all-in-one Synergistic Activator Mediator system, UniSAM, that
95 mediates the transcriptional activation of endogenous gene expression[19]. To activate the
96 nine target genes identified in this study, we developed a novel doxycycline-inducible SAM
97 (iSAM) cassette targeted into the *AAVS1* locus of human iPSCs (Figure 2A). We
98 demonstrated the activation of *RUNX1C* expression in HeLa cells and that this was
99 correlated with the doxycycline (DOX) concentration in a linear manner (Figure S2A-B). To
100 verify the activation in human PSCs at single-cell resolution, we employed a RUNX1C-GFP
101 human embryonic stem cell (hESC) reporter cell line (Figure S2C-G). As predicted, the level
102 of expression of the mCherry tag, the fluorescent reporter within the iSAM cassette, was
103 proportional to the concentration of DOX (Figure S2D-E) and to the number of RUNX1C-
104 GFP+ cells (Figure S2 D-F). Furthermore, the RUNX1C expression level, measured by the
105 mean fluorescence intensity (MFI) of the RUNX1C-GFP reporter, also correlated with the
106 concentration of DOX (Figure S2G). We then tested the iSAM cassette in the iPSC line
107 (SFCi55) (Figure 2B-D). Only when iPSCs were transfected with iSAM and the gRNA for
108 *RUNX1C*, in the presence of DOX, was the expression of the *RUNX1C* gene and RUNX1
109 protein detected (Figure 2B-C).
110 We then targeted the iSAM cassette into the *AAVS1* locus using a Zinc Finger Nuclease
111 (ZFN) strategy [14,20,21]. iPSC clones that had specifically integrated the iSAM cassette into the
112 *AAVS1* locus were validated by genomic PCR and sequencing (Figure S2A-B). The *AAVS1*

113 locus has been reported to be a “safe harbour” site that is resistant to epigenetic silencing
114 and indeed, we had previously demonstrated that transgenes inserted into the *AAVS1* locus
115 under the control of the constitutively active CAG promoter were efficiently expressed both in
116 undifferentiated and in differentiated iPSCs [20,22–24]. However, after the iSAM line had been
117 established and cultured under self-renewal conditions, we noted a dramatic reduction in the
118 number of mCherry+ cells in undifferentiated iSAM iPSCs upon DOX induction, that we
119 predicted to be due to transgene-silencing of the rTTA DOX-inducible cassette (Figure S3C).
120 To overcome this problem, we treated the iSAM iPSC line with an inhibitor of histone
121 deacetylases (HDACs), sodium butyrate (SB), reported to have no adverse effect on iPSC
122 maintenance [25,26] and we also confirmed that the treatment with SB had no effect on viability
123 and cell proliferation in our culture conditions (Figure S3D-E). A short 48-hour treatment
124 significantly increased the number of mCherry+ cells upon DOX induction, proportional to
125 the SB concentration (Figure S3F). We, therefore, maintained the iSAM iPSCs in the
126 presence of SB and this fully restored the inducibility of the transgene with virtually all cells
127 expressing mCherry in the presence of DOX (Figure S3G).
128 To test the effect of activating the 9 target genes on the transcriptomes of differentiating
129 iPSC cells, we engineered the gRNAs to allow their detection within the single cell RNA
130 sequencing pipeline [27]. We inserted a capture sequence prior to the termination signal to
131 avoid any alteration in the secondary structure of the loops thus preserving the binding of the
132 synergistic activators of the SAM system to the stem loops of the gRNAs. Of the two capture
133 sequences available[27], we decided to use the one that was predicted to result in fewer
134 secondary structure alterations and this new gRNA was named 2.1 (Figure 2E). We
135 compared the activation level achieved with the new 2.1 gRNA to that of the original 2.0
136 backbone using various gRNAs targeting *RUNX1C* (Figure 2F). These results convincingly
137 demonstrated that the addition of the capture sequence in the gRNA 2.1 does not alter the
138 level of endogenous gene activation that could be achieved (Figure 2G). Altogether, these
139 results show that the iSAM system is able to induce gene expression that translates into an
140 increased protein expression and thus provides a platform to steer phenotypical changes in
141 cell identity.
142 To activate our target genes, we designed 5-7 gRNAs in the 200bp upstream of the
143 transcriptional start sites of each of the 9 target genes. We subcloned a total of 49 gRNAs
144 (Table S1) into the gRNA 2.1 backbone and packaged them into lentiviral particles, (herein
145 referred to as the AGM library) as well as a non-targeting (NT) gRNA that was used as
146 control. The iSAM iPSC line was transduced with the targeting gRNAs or the control non-
147 targeting gRNA to generate the iSAM_AGM and iSAM_NT iPSCs line, respectively. After
148 puromycin selection, their integration in the genome was also confirmed by PCR and
149 sequencing (Figure S3H-I).

150

151 **Single Cell RNA sequencing in combination with CRISPR activation identifies arterial**
152 **cell type expansion in association with higher hematopoietic progenitor potential**
153

154 To assess the transcriptional changes in response to the activation of the target genes, we
155 differentiated the iSAM iPSCs, induced with DOX and subjected them to single-cell RNA
156 sequencing using the 10X pipeline. After 10 days of differentiation, in the presence or
157 absence of DOX, we FAC-sorted live CD34+ cells from iSAM_AGM and the iSAM_NT
158 iPSCs. Following data filtering, we selected cells in which the gRNAs expression was
159 detected and showed that our approach activated all the target genes, except for ZFN124. A
160 higher level of expression of *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*, *SOX6*, *ZNF33A*, *NFAT5*,
161 *TFDP2* was observed following DOX treatment of cells of iSAM_AGM cells compared to the
162 iSAM_NT cells (Figure 3A). To study the effect of gene activation on transcriptional and
163 cellular phenotype, we performed clustering analysis and detected a total of 7 clusters
164 (Figure 3B). A high level of *GJA4* and *DLL4* expression was used to annotate the arterial-like
165 cell cluster, while high levels of hemogenic markers, such as *RUNX1* and *CD44*, were used
166 to annotate the hemogenic clusters, EHT_1 and EHT_2 (Figure 3B, C-D). To assess the
167 effect of the activation on cell identity, we analysed the proportion of the cell clusters
168 between the libraries and detected a significant expansion of the arterial cluster in the DOX-
169 induced iSAM-AGM compared to the iSAM_NT cells (Figure 3E). To validate the effect of
170 the activation on the expansion of arterial cell population, we analysed their prevalence by
171 flow cytometry. Although DOX treatment resulted in an average 1.63 ± 0.13 fold increase of
172 $CD34^+DLL4^+$ cells in the control iSAM_NT sample (Figure S3A), the increase observed in
173 the iSAM-AGM cells was significantly greater, with an expansion of 3.33 ± 0.73 fold increase
174 of $CD34^+DLL4^+$ immunophenotypic arterial cells identified by flow cytometry (Figure 3F). To
175 assess the effect of this predicted arterial cell expansion on the emergence of functional
176 progenitors, we isolated $CD34^+$ cells using magnetic beads and cocultured 20,000 cells on
177 OP9 supportive stromal cells for 7 days, in the presence of hematopoietic differentiation
178 cytokines. After one week, the progenitor cells were assessed by colony-forming unit (CFU)
179 assays and scored 14 days later. We detected an increased number of CFU-E (colony-
180 forming unit erythroid) and CFU-GM (colony-forming unit granulocyte/macrophage) and a
181 reduction of CFU-M (colony-forming unit macrophage) in iSAM-AGM samples cultured in the
182 presence of DOX compared to the absence of DOX but no significant effect of DOX in
183 SAM_NT samples (Figure 3G). These data indicate that activation of target transcription
184 factors using our novel CRISPR strategy results in transcriptional remodelling and a steer in
185 cell identity that we detected as a functional difference in the hematopoietic progenitor
186 profile.

187

188 **The addition of IGFBP2 to the *in vitro* differentiation leads to a higher number of**
189 **functional hematopoietic progenitor cells.**
190

191 To better understand the molecular mechanism behind the increased progenitor
192 development, we compared the expression profile of the arterial cells between the different
193 activation libraries, and we obtained a list of genes upregulated upon activation of the
194 targets (Table 1). The most upregulated gene, *IGFBP2,* was expressed at significantly
195 higher levels in the iSAM_AGM library in the presence of DOX compared to the others
196 (Figure 4A, Table 1). We then compared the gRNA distribution in these arterial cells from the
197 iSAM_AGM treated with DOX to that of arterial cells without activation. We observed a
198 significant enrichment of the *RUNX1T1-specific* gRNAs (Table1), indicating that the
199 increased *IGFBP2* expression could be downstream of *RUNX1T1* activation. IGF Binding
200 Protein 2 is a member of the family of IGF binding proteins and is thought to be secreted
201 from cells where it then binds IGF1, IGF2, and other extracellular matrix proteins, modulating
202 their function. To test if the increased frequency of functional hematopoietic progenitors was
203 due to IGFBP2 paracrine signalling, presumably derived from the arterial cells, we
204 supplemented the media with IGFBP2 at 100ng/ml after the induction of endothelial cell
205 differentiation. To explore the role of IGFBP2 we employed the parental iPSCs line, SFCi55
206 from which the iSAM line was derived. We isolated CD34+ cells at day 8 and co-cultured
207 them on OP9 cells in presence of IGFBP2 for one week, then tested for their hematopoietic
208 clonogenic potential using CFU assays (Figure 4B). Cells treated with IGFBP2 showed a
209 significant increase in the total number of haematopoietic CFU colonies compared to cells in
210 control cultures. To assess whether IGFBP2 also had a paracrine effect on the production of
211 arterial cells themselves, we analyzed the proportion of DLL4+ cells, but no difference was
212 detected in the presence of IGFBP2 (Figure 4C). This observation supports our hypothesis
213 that the change in hematopoietic progenitor production following activation in the
214 iSAM_AGM line is associated with differential gene expression within the arterial cells rather
215 than expansion of that cell population. We then focused on the characterization of the cells
216 derived from the CD34+ cells after coculture with the OP9 in the presence of IGFBP2. Our
217 results showed a significant expansion of the CD34+ and CD43+ cell populations further
218 supporting our hypothesis (Figure 4D). To address the potential role of *IGFBP2 in vivo*, we
219 analysed single cell sequencing data from the human AGM at Carnegie stages 14 and 15.
220 *IGFBP2* is highly expressed within the AGM niche by stromal and epithelial cells and, most
221 importantly, highest in endothelial cells (Figure 4E). Together these data show that IGFBP2
222 addition results in increased hematopoietic blood production *in vitro* and that the endothelial
223 compartment is the most likely source of IGFBP2 both *in vitro* and *in vivo*.

224

225 **IGFBP2 remodels the metabolic profile of endothelial cells.**

226

227 Following the observation that IGFBP2 supports haematopoietic progenitor differentiation *in*
228 *vitro* from human iPSCs we performed a time course single-cell RNA experiment of SFCi55
229 iPSCs differentiated in its presence. We isolated $CD34^+$ cells and plated them in EHT culture
230 on laminin[28], rather than on OP9 support, to assess exclusively the specific effects of
231 IGFBP2- on differentiating iPSCs. FAC-sorted single/live cells from day 10 and 13 in the
232 presence and absence of IGFBP2 were subjected to single-cell RNAseq. Our time course
233 transcriptomic analyses showed that IGFBP2 induced a change in the transcriptome,
234 specifically on day 13 (Figure 5A). A cluster of endothelial cells enriched for genes
235 associated with the KEGG pathway of growth factor binding, was detected almost
236 exclusively at day 13; these cells displayed a different transcriptional signature (as indicated
237 by a shift in their position in the UMAP embedding) in the presence of IGFBP2..
238 Interestingly, this cluster showed expression of *GJA4* across the cells, indicating their broad
239 arterial identity, with a specific reduction of *DLL4* expression level concomitantly to an
240 upregulation of *RUNX1* in the region of the cluster induced by IGFBP2 (Figure 5B). This
241 increase in hemogenic identity within the endothelial cell compartment is consistent with the
242 observation of increased functional progenitor production. We compared the transcriptome
243 of cells at day 13 in the presence and absence of IGFBP2 and, using KEGG enrichment
244 analysis, we observed that the genes that were upregulated by IGFBP2 were highly
245 enriched in the oxidative phosphorylation term (Figure 5C-D). Since the metabolic switch
246 between glycolytic to mitochondrial metabolism has been previously reported in definitive
247 hematopoiesis[29,30], we tested whether the addition of IGFBP2 could result in a different ATP
248 production profile. We analysed the ATP production at day 13, when RUNX1 expression
249 was induced by IGFBP2 addition, by using specific inhibitors of the complex I and II and
250 complex V to quantify the intracellular ATP and mitochondrial synthesis, respectively. We
251 detected a reduction in glycolytic-derived ATP (Figure 5E-F), which translates to a higher
252 contribution of mitochondria metabolism for IGFBP2-treated cells (Figure 5G). We then
253 looked at the expression levels of genes encoding the enzymes of the glycolytic pathway
254 and its checkpoints, glucose and lactate transporters, monocarboxylate transporters and the
255 enzymes associated with hexokinase and phosphofructose reactions. We observed a
256 general downregulation of the glycolytic enzymes and its checkpoints, with the exception of
257 *PFKM* and *SLC16A1*, (Figure S4), which were expressed at lower levels in the IGFBP2-
258 treated cells at day 13 compared to the control on the same day. These differences were not
259 observed on day 10 cells, which is consistent with our previous observations that the
260 IGFBP2-induced remodelling of the transcriptome happens exclusively on day 13. These

261 results indicate that the presence of IGFBP2 induces the upregulation of *RUNX1* and
262 OxPhos genes and the metabolic remodelling of endothelial cells.
263
264 **Discussion**
265
266 The complexity and dynamism of developmental hematopoiesis *in vivo* have imposed
267 challenges in accurately reproducing the process *in vitro*. We hypothesized that this could, in
268 part, be due to the inability to recapitulate the appropriate transcriptional programme in
269 cultured cells. Here, we developed a novel CRISPR-activation system to induce the
270 expression of genes that are expressed at low levels in cells that are generated *in vitro*, to
271 explore the downstream consequences of their activation and to assess the effects on
272 emerging hematopoietic progenitor cells.
273 When we compared endothelial cells derived *in vitro* from hiPSCs to those in the AGM
274 region of the human embryo, at the time point when early hematopoietic commitment takes
275 place, we identified nine transcription factors that were expressed at a lower level in the
276 cultured cells. Some of these had been associated with blood cell development, including
277 *GATA2*[31–33], *SMAD7* [34], *NR4A1*[34], *SOX6*[35] and *RUNX1T1*[15,17]. Other genes such as,
278 ZNF124, ZNF33A, NFAT5 and TFDP2 had not been previously associated with
279 hematopoiesis and could provide early evidence of their possible role in hematopoiesis that
280 will require further studies. We were particularly interested in *RUNX1T1* (also known as
281 ETO) which is associated with the leukemic fusion protein, AML1/ETO resulting from the
282 t(8;21) chromosomal translocation[36]. Furthermore, *RUNX1T1* expression has been recently
283 detected in transcriptomic analyses of the human AGM region[15,17], but its precise role during
284 the ontogeny of the blood system has not been elucidated. The addition of a capture
285 sequence to the gRNA backbone enabled their detection coincidentally with the single-cell
286 transcriptome, and this allowed us to demonstrate that RUNX1T1 gRNAs were significantly
287 enriched in cells within an expanded arterial cluster and their presence was associated with
288 a high level of expression of *IGFBP2*. *IGFBP2* KO mice show increased expression of cell-
289 cycle inhibitors and HSC apoptosis, implicating IGFBP2 as a modulator of HSCs cell cycle
290 and survival [37]. More recently *IGFBP2* was reported as being highly expressed in the human
291 AGM region at CS14 when HSCs are emerging [17], supporting a possible role during
292 developmental hematopoiesis but the precise molecular process was unclear. In this study,
293 we show that the addition of IGFBP2 recombinant protein in our *in vitro* model of EHT results
294 in the emergence of an increased number of functional hematopoietic progenitors and we
295 provide evidence to support the hypothesis that RUNX1T1 could regulated its expression.
296 Because RUNX1T1 lacks a DNA binding domain, its direct involvement in the regulation of

297 IGFBP2 expression must require association with other cofactors that are yet to be
298 identified.
299 Our data, together with *in vivo* data[17], show that IGFBP2 is expressed predominantly by
300 endothelial cells. The supportive role of the endothelial niche in the development of HSCs
301 has been studied *in vivo* [38–40] and exploited *in vitro* to support HSCs emergence [39,41].
302 We describe here a novel role for IGFBP2 in the remodelling of the metabolism of iPSC-
303 derived endothelial cells and that this is associated with the induction of RUNX1, a hallmark
304 of hemogenic fate. The addition of IGFBP2 in the culture induces upregulation of the genes
305 of the oxidative phosphorylation pathway in association with an increased relative
306 mitochondrial contribution to cellular ATP production. This is due to reduced glycolysis in
307 association with the downregulation of genes coding for glycolytic enzymes and their
308 checkpoints. The switch between glycolytic to mitochondrial metabolism has been shown to
309 be essential for the EHT in definitive hematopoiesis[29,30]. This switch is induced *in vivo* by
310 mechanical cues downstream of the establishment of circulation, and it is required for
311 functional HSCs development[29], while *in vitro*, the switch is driven by pyruvate mitochondrial
312 catabolism, leading to definitive EHT as opposed to primitive[30]. Further *in vivo* studies are
313 required to characterize the role of IGFBP2 during AGM hematopoiesis to overcome the
314 limitation of using an *in vitro* model such as iPSCs differentiation.
315 In conclusion, we detected transcriptional differences between *in vivo* and *in vitro*
316 developmental hematopoiesis and developed a novel inducible gene activation system to
317 identify novel molecular players during the endothelial-to-hematopoietic transition. Our
318 multidisciplinary approach identified *RUNX1T1* and IGFBP2 as novel signalling molecules
319 that support human blood progenitor development *in vitro*, the latter inducing a metabolic
320 switch from cytoplasmic glycolysis to mitochondrial respiration. This study demonstrates that
321 the combination of CRISPR-mediated activation of target genes with single-cell
322 transcriptomic analysis in differentiating hPSCs can be a powerful approach to alter cell fate,
323 providing a tool for gene function studies during human development. The fine epigenetic
324 manipulation of the transcription can be readily applied to any cell lineage simply by adding
325 specific gRNAs and it will be instrumental in exploring other developmental processes that
326 can be, at least partially, be mimicked *in vitro* with human iPSCs.
327
328 Using this approach, we have identified the supportive role of paracrine IGFBP2,
329 predominantly produced by endothelial cells, which induces transcriptional and metabolic
330 remodelling in association with the induction of *RUNX1* expression.
331
332

Figures

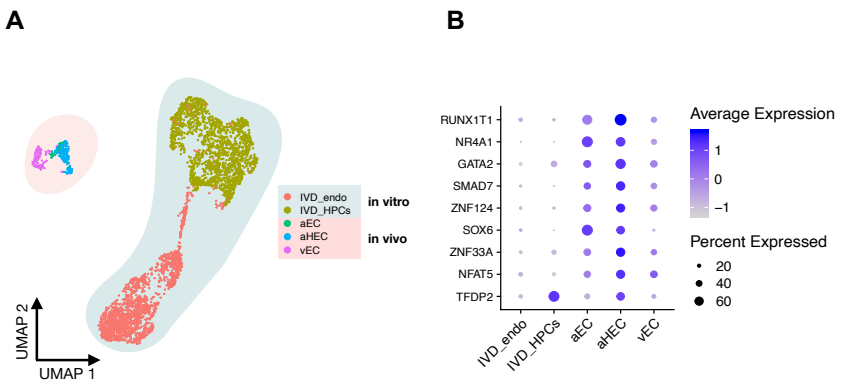

Figure 1 - Comparison of *in vitro* iPSC-derived and *in vivo* AGM-derived endothelial cells identifies 9 differentially expressed transcription factors.

A - Integrative analysis of single cell transcriptome of *in vitro* derived endothelial (IVD_Endo) and hematopoietic cells (IVD_HPCs) with *in vivo* developed endothelial cells (venous, vEC; arterial, aEC; arterial hemogenic, HECs) from human embryos (CS12-CS14) visualised on UMAP dimensions**. B** - Target genes expression level showing higher expression in arterial hemogenic endothelium *in vivo* compared to *in vitro* derived cells.

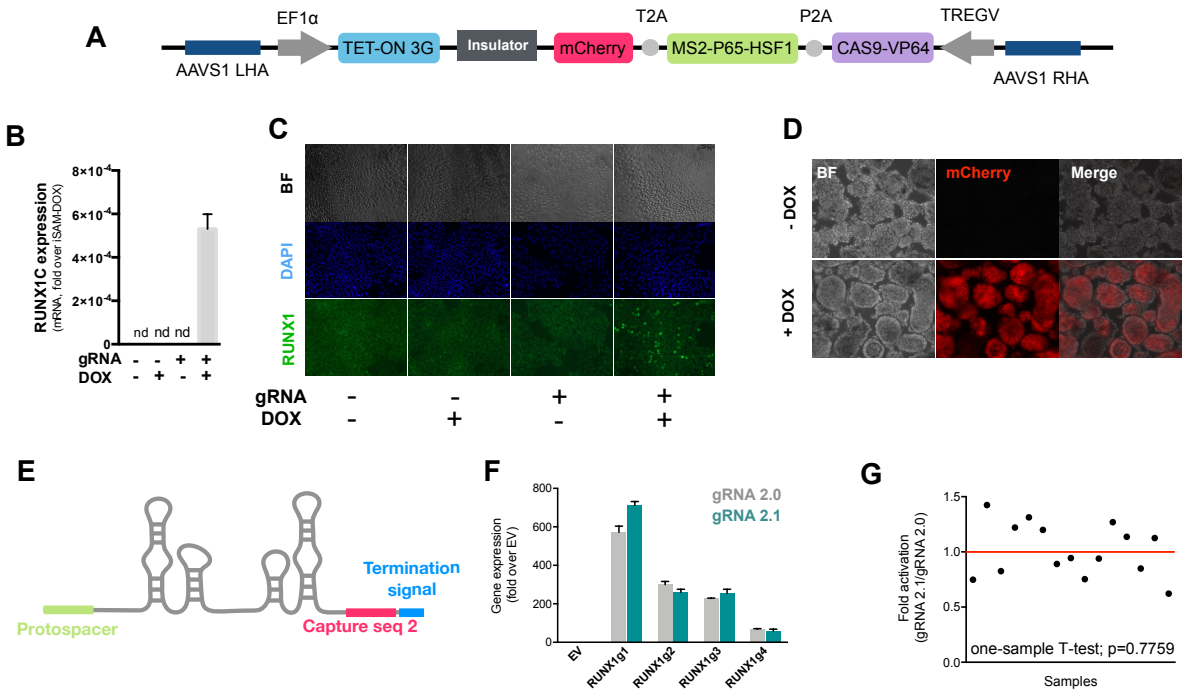

Figure 2 - The inducible iSAM cassette successfully mediates activation of endogenous gene expression upon DOX induction.

A - Schematic of the iSAM cassette containing the TET-on system under the control of EF1α and dCAS9-P2A-MS2-p65-HSF1-T2A-mCherry under the rTTA responsive elements, separated by genetic silencer and flanked by AAVS1 specific homology arms. **B** - *RUNX1C* gene expression activation after transient transfection of the iSAM plasmid and gRNAs in presence or absence of DOX in human iPSC line (n=3 from independent transfections). **C** – RUNX1 protein expression upon iSAM activation after transient transfection of the iSAM plasmid and gRNAs with DOX in human iPSC line detected by immunostaining. **D** Expression of the iSAM cassette reported by mCherry tag during the differentiation protocol, the representative images (bright field – BF, and fluorescence) show embryoid bodies at day 3 of differentiation. **E** - Schematic of the gRNA 2.1 containing the capture sequence for detection during the scRNAseq pipeline. **F** - *RUNX1C* gene activation level obtained using either the gRNA 2.0 or 2.1 backbone (n=3 from independent transfections of the 4 different gRNAs). **G** - Statistical analysis of the gRNAs activation level showing no significant variation following addition of the capture sequence (n=3 for each of the 4 different gRNAs).

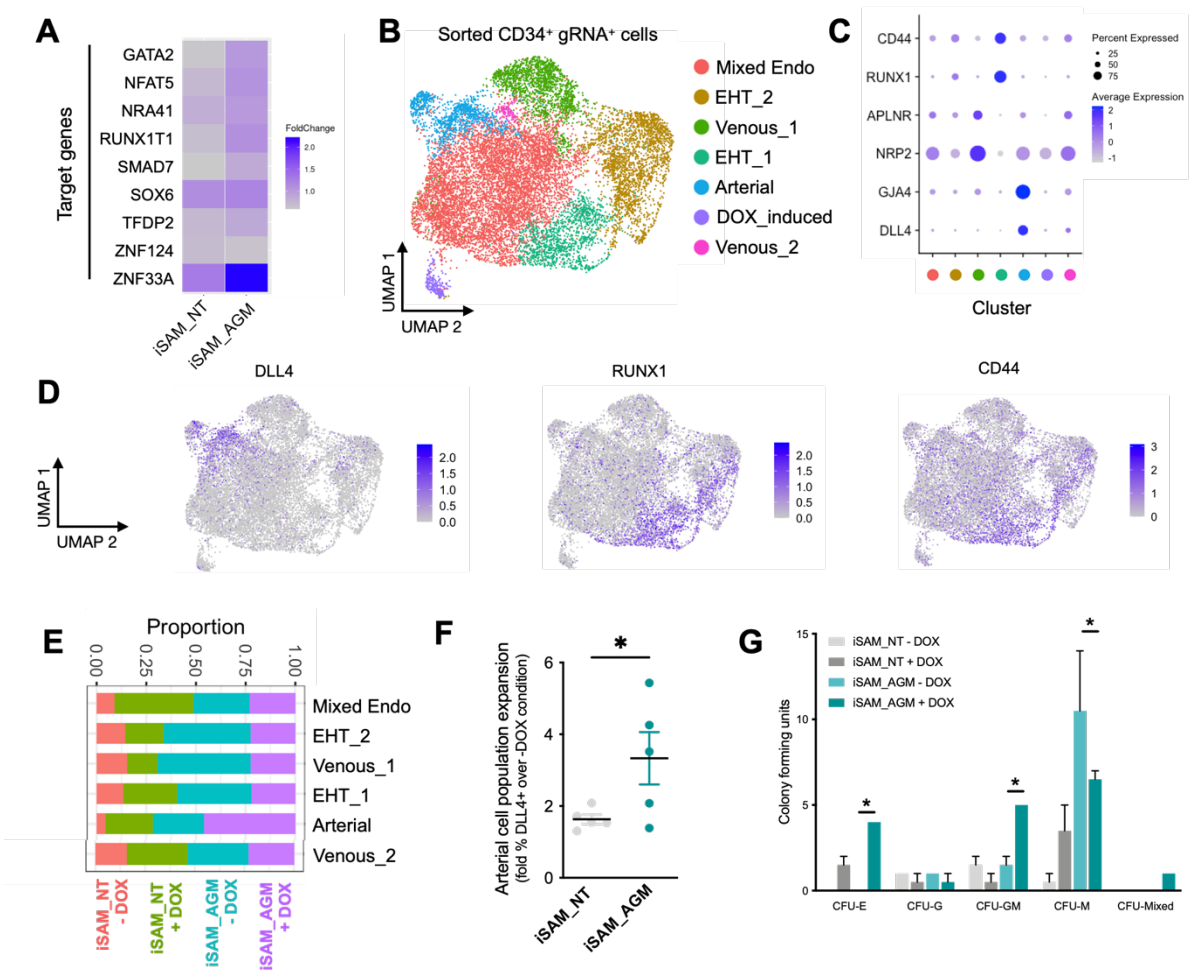

Figure 3 - Single Cell RNA sequencing in combination with CRISPR activation identifies arterial cell type and functional hematopoietic expansion in association with activation of the 9 target genes.

A - Gene expression profile of target genes following target genes' activation, heatmap shows the expression level of the target genes in the iSAM_NT and iSAM_AGM treated with DOX following normalisation on the -DOX control. **B** - Dimension reduction and clustering analysis of the scRNAseq data following activation, filtered on cells where the gRNA expression was detected. **C** - Arterial (*GJA4*, *DLL4*), venous (*NRP2*, *APLNR*) and hemogenic marker (*CD44*, *RUNX1*) expression distribution in the clusters indicated by the colour. **D** - Expression distribution visualised on the UMAP plot showing the location of arterial cells marked by *DLL4*, and hemogenic endothelium marked by *CD44* and *RUNX1*. **E** - Contribution of the different libraries to the clusters showing that arterial cell cluster is overrepresented in the iSAM_AGM treated with DOX, compared to the other libraries. **F** - Expansion of the arterial population assessed by the membrane marker expression of DLL4+ following targets' activation, quantified by flow cytometry at day 8 of differentiation (Data are normalised on the iSAM_NT + DOX sample, n=5 independent differentiations, * p

397 = 0.0417 paired t-test). **G** - Colony forming potential of the suspension progenitor cells
398 derived from the two lines treated with or without DOX following OP9 coculture activation,
399 data show the colony obtained for 104 CD34+ input equivalent (n=3 from independent
400 differentiations * $p<0.05$, Tukey's two-way ANOVA).
401

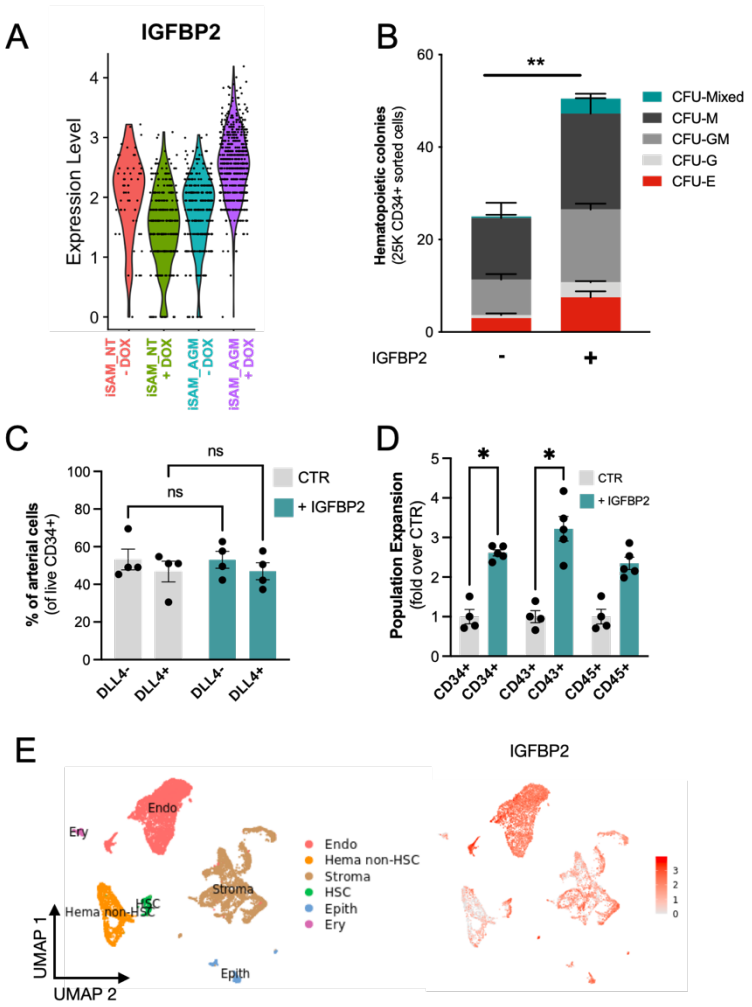

Figure 4 - IGFBP2 addition to the *in vitro* differentiation leads to a higher number of functional hematopoietic progenitor cells. A - Violin plot of *IGFBP2* expression profile in the arterial cells obtained from the different conditions, in the presence or absence of gRNAs and DOX. **B** -Number of hematopoietic colonies obtained after coculture on OP9 in presence or absence of IGFBP2 (n=3-4 from independent differentiations, ** p=0.0080, Sidak's Two way ANOVA). **C** – Percentage of DLL4$^+$ arterial cells differentiation within the CD34$^+$ compartment analysed by flow cytometry in day 8 EBs (n=4 from independent differentiations, Two-way Anova, ns =p>0.99). **D** – Expansion of hematopoietic progenitors analysed using markers' expression on suspension progenitors derived after coculture of CD34+ cells onto OP9 support (data are expressed as fold over the CTR in the absence of IGFBP2 (n=4 from independent differentiations, * p<0.02, Sidak's Two way ANOVA. **E** - Single-cell transcriptomic analysis of developing AGM collected from human embryos at Carnegie Stages14 and 15 enriched for CD31$^+$ and CD34$^+$ showing the *IGFBP2* expression profile *in vivo* in the AGM.

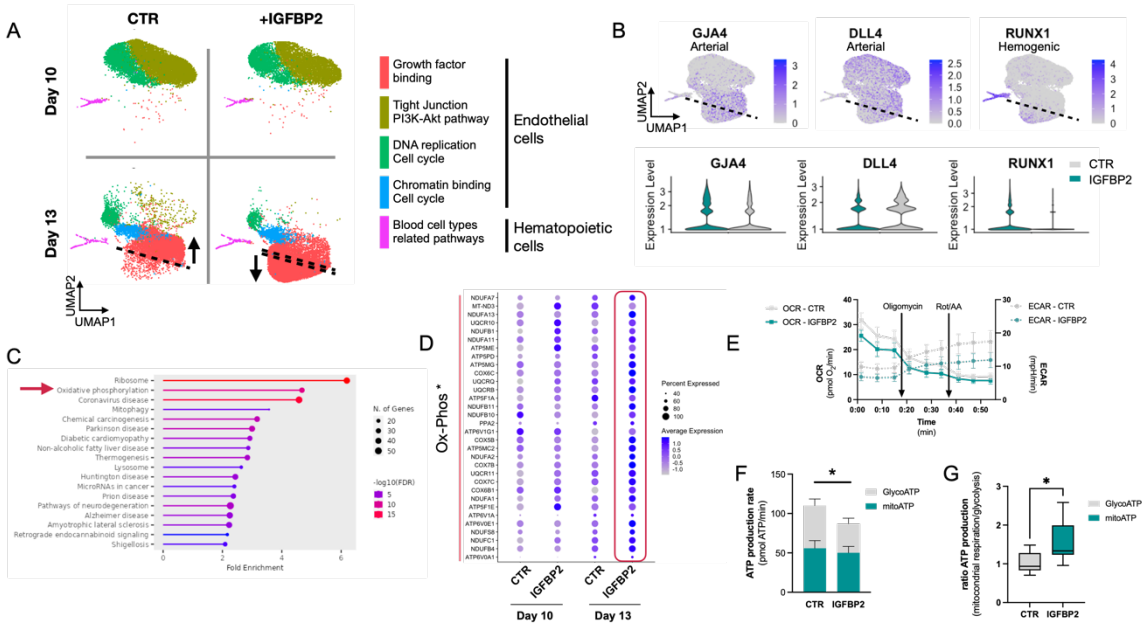

Figure 5 - IGFBP2 alters cell metabolism by inducing a reduction in glycolytic ATP production.

A – Clustering analysis of the single cell transcriptomic time course analysis of differentiating cells at day 10 and day 13 in the absence (CTR) or presence of IGFBP2. Arrows indicate the difference in the clustering due to the addition of IGFBP2 compared to control. **B** – Expression profile of arterial markers, *GJA4* and *DLL4*, and hemogenic marker *RUNX1* (top - the dashed line shows the location of the shift in gene expression of cells treated with IGFBP2) and their expression profile in the endothelial cells cluster marked by Growth factor binding in absence (CTR) and in presence of IGFBP2. **C –** KEGG enrichment analysis of the genes upregulated at day 13 upon IGFBP2 treatment. The arrow shows the ranking of the Oxidative Phosphorylation pathway. **D** - Dot Plot showing the expression profile of the genes coding for the enzyme of the Oxidative Phosphorylation pathway. **E** -Oxygen Consumption Rate (OCR) and Extracellular Acidification Rate (ECAR) profile in cells at day 13 of differentiation reporting mitochondrial respiration and glycolysis, respectively. **F** – ATP production rate divided by that deriving from glycolysis and from mitochondrial respiration, in cells treated with IGFBP2 and controls at day 13. **G** – Ratio of the ATP production between glycolysis and mitochondrial respiration in cells treated with IGFBP2 and controls at day 13.

437
438 **Methods**

439

440 **Resource availability**

441 R code is available at https://github.com/afidanza/CRISPRa. Raw and processed data have
442 been deposited to ArrayExpress (E-MTAB-12748 – currently on hold). The AAVS1-iSAM are
443 available on Addgene (Addgene #211495) and the gRNA 2.1 plasmids (Addgene #211496)
444 (both currently on hold). Further information and requests for resources and reagents should
445 be directed to and will be fulfilled by the corresponding author.

446

447 **Pluripotent Stem Cells maintenance**

448 hPSCs were maintained *in vitro* in StemPro hESC SFM (Gibco) with bFGF (R&D) at 20
449 ng/ml. Wells were coated with Vitronectin (ThermoFisher Scientific) at least 1 hour before
450 plating and cells were passaged using the StemPro EZPassage tool (ThermoFisher
451 Scientific). Media change was performed every day and cells passaged every 3–4 days at a
452 ratio of 1:4.

453

454 **Transfection**

455 iPSCs SFCi55 and hESCs RUNX1-GFP were plated at 3×10^5 cells per a well of a 6 well
456 plate and reverse transfected with 2 µg of DNA using the Xfect Transfection reagent
457 (Clontech) and analyzed 2 days later.
458 HeLa cells were cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12
459 (DMEM/F12) with Glutmax and 5% FCS (Gibco) and passaged every few days, at a ratio of
460 1:6. HEL were cultured in Iscove's Modified Dulbecco's Medium (IMDM) with 10% FCS
461 (Gibco) and passaged every few days, at a ratio of 1:4. 2×10^5 cells were plated, transfected
462 at 6–8 hours with 0.75 µg of DNA using Xfect Transfection reagent (Clontech) and then
463 analysed 2 days after.

464

465 **Immunocytochemistry**

466 Cells were fixed in 4% PFA in PBS at room temperature for 10', permeabilized in PBS-T
467 (0.4% Triton-X100) for 20' and blocked in PBS-T with 1% BSA and 3% goat serum for
468 1 hour. Primary antibodies were incubated in blocking solution over night at 4 °C (RUNX1
469 1:200 - ab92336, Abcam). Cells were then washed in PBS-T and incubated with secondary
470 antibodies for 1 hour at room temperature (donkey α-rabbit 1:200 - A-11008 - Thermo
471 Scientific). Cells were washed in PBS-T and counterstained with DAPI. Images were
472 generated using the Zeiss Observer microscope.

473

474 **Gene expression analysis**

475 Total RNA was purified using the RNAeasy Mini Kit (Qiagen) and cDNA synthesized from
476 500 ng of total RNA using the High Capacity cDNA synthesis Kit (Applied Biosystem). 2 ng of
477 cDNA were amplified per reaction and each reaction was performed in triplicate using the
478 LightCycler 384 (Roche) with SYBR Green Master Mix II (Roche). A melting curve was
479 performed and analyzed for each gene to ensure the specificity of the amplification. *β-*
480 *Actin* was used as reference genes to normalize the data [19].

481

482 **Pluripotent Stem Cells differentiation to hematopoietic progenitors**

483 hPSCs were differentiated in a xeno-free composition of SFD medium [14], BSA was
484 substituted with human serum albumin, HSA (Irvine-Scientific). Day 0 differentiation medium,
485 containing 10 ng/ml BMP4 was added to the colonies prior cutting. Cut colonies were
486 transferred to a Cell Repellent 6 wells Plates (Greniner) to form embryoid bodies and
487 cultured for two days. At day 2 media was changed and supplemented with 3 μM CHIR
488 (StemMacs). At day 3, EBs were transferred into fresh media supplemented with 5 ng/ml
489 bFGF and 15 ng/ml VEGF. At day 6 media was changed for final haematopoietic induction in
490 SFD medium supplemented with 5 ng/ml bFGF, 15 ng/ml VEGF, 30 ng/ml IL3, 10 ng/ml IL6,
491 5 ng/ml IL11, 50 ng/ml SCF, 2 U/ml EPO, 30 ng/ml TPO, 10 ng/ml FLT3L and 25 ng/ml
492 IGF1. From day 6 onward, cytokines were replaced every two days.

493

494 **CD34 isolation**

495 CD34+ cells were isolated using CD34 Magnetic Microbeads from Miltenyi Biotec, according
496 to their manufacturing protocol. Briefly, Embryoid bodies were dissociated using Accutase
497 (Life Techonologies) at 37°C for 30'. Cells were centrifuged and resuspended in 150 μl of
498 PBS + 0.5% BSA + 2mM EDTA with 50 μl Fcr blocker and 50 μl of magnetic anti-CD34 at
499 4°C for 30'. Cells were washed using the same buffer and transferred to pre-equilibrated
500 columns, washed three times and eluted. After centrifugation, cells were resuspend in SFD
501 media, counted and plated for OP9 coculture.

502

503 **OP9 coculture and colony assay**

504 OP9 cells were maintained in α-MEM supplemented with 20% serum (Gibco) and sodium
505 bicarbonate (Gibco) and passaged with Trypsin every 3-4 days. The day before the co-
506 culture, 45.000 OP9 cells were plated for each 12 well plates' well in SFD media. The day of
507 the co-culture the 20.000 iSAM cells or 25.000 SFCi55 or H9 were plated in each well and
508 culture in SFD media supplemented with 5 ng/ml bFGF, 15 ng/ml VEGF, 30 ng/ml IL3, 10
509 ng/ml IL6, 5 ng/ml IL11, 50 ng/ml SCF, 2 U/ml EPO, 30 ng/ml TPO, 10 ng/ml FLT3L and 25

510 ng/ml IGF1 and 100ng/ml IGFBP2. Cytokines were replaced twice during one week of
511 coculture. At the end of the coculture, cells were collected by Trypsin and half of the well
512 equivalent was plated in 2 ml of methylcellulose medium (Human enriched H4435, Stemcell
513 Technologies). Cells were incubated in the assay for 14 days and then scored.
514
515 **Laminin EHT culture**
516 Laminin EHT culture was performed as previously described[28]. Briefly, 24 well plates were
517 coated for at least 2 hours with recombinant human Laminin-521 (Thermo-Fisher). Following
518 CD34+ isolation at day 8, 400.000 CD34+ cells were seeded in each well of a pre-coated
519 24-well plate in SFD media supplemented with 5 ng/ml bFGF, 15 ng/ml VEGF, 30 ng/ml IL3,
520 10 ng/ml IL6, 5 ng/ml IL11, 50 ng/ml SCF, 2 U/ml EPO, 30 ng/ml TPO, 10 ng/ml FLT3L and
521 25 ng/ml IGF1 and 100ng/ml IGFBP2.
522
523 **Flow cytometry staining and cell sorting**
524 Embryoid bodies were dissociated using Accutase (Life Techonologies) at 37°C for 30'.
525 Cells were centrifuged and resuspended in PBS + 0.5% BSA + 2mM EDTA, counted and
526 stained at 10^5 cells for a single tube. Cells were stained with antibodies for 30' at room
527 temperature gently shaking. Flow cytometry data were collected using DIVA software (BD).
528 For the sorting experiments, cells were stained at 10^7cells/ml in presence of the specific
529 antibodies. Sorting was performed using FACSAria Fusion (BD) and cells were collected in
530 PBS + 1% BSA. Data were analysed using FlowJo version 10.4.2.
531
532 **Flow cytometry antibodies**
533 For flow cytometry 10^5 cells per test were stained in 50 μl of staining solution with the
534 following antibodies: CD34 Percp-Efluor710 (4H11 eBioscience, 1:100), CD34 Pe (4H11
535 eBioscience, 1:200), CD43 APC (eBio84-3C1, 1:100), CD45 FITC (2D1 ebioscience, 1:100),
536 DLL4 Pe (MHD4-46 Biolegend, 1:200)
537
538 **iSAM plasmid generation**
539 The iSAM plasmid was obtained by Gibson assembly of four fragments. The first fragment,
540 the backbone, was a DOX-inducible AAVS1 targeted plasmid expressing an E6-E7-IRES-
541 ZsGreen which was excised by BstBI and NdeI. The second fragment, one of the adapters,
542 was derived from the UniSAM plasmid that we previously generated (Addgene #99866) by
543 PCR with the following primer sets
544 FW_aggggacccggttcgagaaggggctcttcatcactagggccgctagctctagagagcgtcgaatt,
545 RV_ttcgggtcccaattgccgtcgtgctggcggctcttcccacctttctcttcttcttggggctcatggtggcc. The UniSAM
546 cassette was obtained also from the UniSAM plasmid via digestion with BsrGI and BsiWI.

547 Finally, the last fragment consisting of another adapter for the Gibson was custom
548 synthetised and contained overlapping sequences flanking a chicken b-globin insulator that
549 we inserted to prevent cross-activation of the EF1α-promoter and the TRE-GV promoter
550 driving the iSAM. Correct assembly was verified by Sanger sequencing.
551
552 **iSAM cell lines derivation**
553 The iSAM plasmid was used together with ZNFs specific for the AAVS1 locus to mediate
554 specific integration in SFCi55 human iPSCs line [14,20]. Briefly, 10 μg of AAVS1-iSAM with 2.5
555 μg of each ZNFs, left and right, using Xfect (Takara) according to the manufacturer protocol.
556 Cells were selected using Neomycin. Single clones were picked, amplified, and initially
557 screened by mCherry expression upon DOX addition. Clones that expressed the fluorescent
558 tag were screened for specific integration using PCR followed by Sanger sequencing for the
559 correctly integrated clones. 100 ng of genomic DNA was amplified using the EmeraldAmp®
560 MAX HS Takara and specific primer sets (Table 1). For the specific AAVS1 integration site,
561 Sigma_AAVS1 - CGG AAC TCT GCC CTC TAA CG and NeoR -GAT ATT GCT GAA GAG
562 CTT GGC GG were used with the PCR conditions of 95 °C for 7 min, 32 cycles of 95 °C for
563 15 s, 57 °C for 30 s, and 72 °C for 1 min, with the final elongation step at 72 °C for 7 min. For
564 the Wild type locus screening, Sigma_AAVS1 - CGG AAC TCT GCC CTC TAA CG and
565 AAVS1_EXT3_RV – ACA CCC AGA CCT GAC CCA AA were used with the PCR cycling
566 conditions of 95 °C for 7 min, 30 cycles of 95 °C for 15 s, 57 °C for 30 s, and 72 °C for 2 min,
567 with the final elongation step at 72 °C for 6 min.
568
569
570 **Capture sequencing addition to the gRNA backbone**
571 The Capture sequence 2 was added to the gRNA_Puro_Backbone (Addgene #73797) by
572 PCR. Briefly, the capture sequence was added before the termination signal of the gRNA
573 followed by a BamHI site using the following PCR primers:
574 gRNA_FW gagggcctatttcccatgattcct,
575 gRNA_Cap_RV aaaaaaggatccaaaaaaCCTTAGCCGCTAATAGGTGAGCgcaccgactcggtgcc.
576 The gRNA backobone was replaced from the original plasmid via NdeI and BamHI digestion,
577 followed by ligation of the PCR produced following the same digestion. Correct integration of
578 the insert was verified by Sanger sequencing.
579
580 **Target genes identification**
581 Candidate genes were identified by comparing *in vivo* hemogenic endothelium [15]
582 (GSE135202) and *in vitro* iPSCs-derived endothelial cells that we previously generated [14]

583 (E-MTAB-9295). Briefly, the two datasets were merged and normalised using the R
584 package Seurat. Specific markers for hemogenic endothelium were identified and
585 transcription factors were sorted based on their GO annotation. Within those genes we
586 filtered those detected in more than 50% of the *in vivo* hemogenic endothelium and
587 expressed in less than 25% of the *in vitro* derived endothelial cells. This pipeline identified 9
588 target genes *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*, *ZNF124*, *SOX6*, *ZNF33A*, *NFAT5*, *TFDP.*
589
590 **AGM gRNAs library preparation**
591 sgRNA design was performed by selecting the top candidates for on-target and off-target
592 score. Between 5 and 7 guides per gene were designed for *RUNX1T1*, *NR4A1*, *GATA2*,
593 *SMAD7*, *ZNF124*, *SOX6*, *ZNF33A*, *NFAT5*, *TFDP2* using the CRISPRpick tool from the
594 Broad Institute (https://portals.broadinstitute.org/gppx/crispick/public) (Table S1). All the
595 guide variants were Golden Gate cloned with the gRNA 2.1 backbone according to the
596 established protocol [42]. The 49 plasmids were pooled together in an equimolar ratio and the
597 library prep was subsequently used to produce lentiviral particles with a second-generation
598 production system. Briefly, the psPAX2 packaging plasmid, pMD2.G envelope, and the AGM
599 vector library were co-transfected using polyethyleneimine (PEI) (Polysciences, Warrington,
600 PA, USA) as previously detailed [43], Lentiviral particles-containing supernatants were
601 harvested 48–72 h post-transfection, concentrated by ultracentrifugation and titered in
602 hiPSCs cells.
603
604 **iSAM_AGM and iSAM_NT cell line derivation**
605 The selected iSAM clone (3.13 internal coding) was infected with viral particles containing
606 either the AGM library or the non targeting gRNA (NT) at a MOI of 10. The iSAM cells were
607 plated the afternoon before at $7x10^6$ cells into a T125 in the presence of 10 μM Rock
608 Inhibitor (Merk) which was maintained until the day following the infection. Cells were
609 infected in presence of 8μg/ml of Polybrene (Merk). Puromycin selection was initiated 36
610 hours post-infection and maintained during their culture until the beginning of the
611 differentiation. Both lines were tested for integration of the gRNAs. Briefly, 100 ng of isolated
612 gDNA was amplified using the PrimeSTAR MAX PCR mix (Takara) using the primers
613 gRNA_screening_FW and gRNA_screening_RV (Table1). Purified amplicons were
614 subjected to Sanger sequencing.
615
616 **Single Cell RNA sequencing**
617 For the iSAM_AGM and iSAM_NT single cell RNA sequencing experiment, embryoid bodies
618 obtained from day 10 of differentiation were dissociated using Accutase (Life Technologies')

619 at 37°C for 30'. For the IGFBP2 experiment, day 10 and day 13 cells from the Laminin EHT
620 culture were detached from the adherent layer using Accutase (Life Technologies') at 37°C
621 for 5'. From both experiments, cells were centrifuged and resuspended in CD34-Pe staining
622 solution at a density of 10^7/ml. CD34+/live/single cells were FAC-sorted in PBS + 0.1 %
623 BSA. Cell viability was also confirmed by Trypan blue stain for an accurate count. Around
624 15000 cells per sample were loaded into the 10X Chromium Controller, and single cell
625 libraries were obtained using the Chromium single cell 3' Reagent Kits v3 (10XGenomics)
626 according to manufacturer protocol. The four libraries were indexed using SI PCR primers
627 with different i7 indexes to allow for demultiplexing of the sequencing data. RNA
628 concentration was obtained using Quibit RNA HS (Thermo-Fisher). Quality of the obtained
629 libraries were verified using LabChip GX (PerkinElmer). Libraries were sequenced using
630 NextSeq 2000 technology (Illumina) at 50.000 reads/cell. Data were aligned to GRCh38
631 using the Cell Ranger dedicated pipeline (10XGenomics). Data filtering, dimension
632 reduction, clustering analysis, differentially expressed genes and cell cycle analysis were
633 obtained using Seurat R package (version 4.1.0)[44]. Pseudotemporal ordering was performed
634 using Monocle 3 R package [45]. KEGG pathways was performed using ShinyGo[46]. The code
635 is available on Github at https://github.com/afidanza/CRISPRa, the raw data have been
636 submitted to Array Express (E-MTAB-12748), and the browsable processed data will be
637 made available at the time of publication on our website containing previous sequencing
638 data at https://lab.antonellafidanza.com.
639
640 **ATP production analysis**
641
642 At day 8 of differentiation, 20.000 CD34+ were plated for each precoated well of the
643 Seahorse XFp mini Cell Culture Plates (Agilent) precoated with rhLaminin-521 (Thermo-
644 Fisher). At day 13, the cells' ATP production rate was analysed with the Seahorse XF Real-
645 Time ATP Rate Assay kit (Agilent) according to manufacturer protocol. Briefly, the media
646 was changed prior to the assay for the XF DMEM medium, pH 7.4, supplemented with 10
647 mM Seahorse XF glucose, 1mM Seahorse XF pyruvate and 2mM Seahorse XF Glutamine
648 (Agilent) and incubated for 45-60 minutes in a non-CO_2 incubator. Oligomycin and
649 Rotenone/AA solutions were prepared and added in the cartridge and finally loaded together
650 with the cells in the Seahorse XF Mini Analyzer using the dedicated software. Data were
651 collected at the end of the run, and the values of ATP production were calculated according
652 to the ATP Production Rate Calculation provided by Agilent.
653
654 **Author contribution**
655

AF designed the study, performed experiments and bioinformatic analysis, wrote the paper and led the research. PP, TV, FPL, HM performed experiments. AM, HT, MS provided support to the experiments. NR performed bioinformatic analysis. LF and PM helped the design of the study and the research. All authors provided essential feedback on the experiments and the manuscript.

Acknowledgment

AF and LF acknowledge financial support from the Biotechnology and Biological Sciences Research Council; Grant S002219/1. AF was supported by a European Hematology Association Advanced Research Grant (EHA RAG 2021), and by the American Society for Hematology (Research Global Award). TV and AM were supported by PhD studentships from the Medical Research Council (Precision Medicine) and the College of Medicine and Veterinary Medicine, respectively. FPL was supported by an Erasmus+ Traineeship Program 2016/2017. PM acknowledges financial support from a PERIS program from the Catalan Government and a Retos collaboration project from the MINECO (RTC-2018-4603-1)

677

Bibliography

678 1. Medvinsky, A. & Dzierzak, E. Definitive hematopoiesis is autonomously initiated by
679 the AGM region. *Cell* **86**, 897–906 (1996).
680 2. Palis, J., Robertson, S., Kennedy, M., Wall, C. & Keller, G. Development of erythroid
681 and myeloid progenitors in the yolk sac and embryo proper of the mouse.
682 *Development* **126**, 5073–84 (1999).
683 3. Patel, S. H. *et al.* Lifelong multilineage contribution by embryonic-born blood
684 progenitors. *Nat. 2022 6067915* **606**, 747–753 (2022).
685 4. Böiers, C. *et al.* Lymphomyeloid Contribution of an Immune-Restricted Progenitor
686 Emerging Prior to Definitive Hematopoietic Stem Cells. *Cell Stem Cell* **13**, 535–548
687 (2013).
688 5. Hoeffel, G. *et al.* C-Myb(+) erythro-myeloid progenitor-derived fetal monocytes give
689 rise to adult tissue-resident macrophages. *Immunity* **42**, 665–78 (2015).
690 6. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lièvre, F. Intraaortic hemopoietic
691 cells are derived from endothelial cells during ontogeny. *Development* **125**, 4575–83
692 (1998).
693 7. Zovein, A. C. *et al.* Fate Tracing Reveals the Endothelial Origin of Hematopoietic
694 Stem Cells. *Cell Stem Cell* **3**, 625–636 (2008).
695 8. Bertrand, J. Y. *et al.* Haematopoietic stem cells derive directly from aortic endothelium
696 during development. *Nature* **464**, 108–111 (2010).
697 9. Boisset, J. C. *et al.* In vivo imaging of haematopoietic cells emerging from the mouse
698 aortic endothelium. *Nature* **464**, 116–120 (2010).
699 10. Ottersbach, K. Endothelial-to-haematopoietic transition: an update on the process of
700 making blood. (2019). doi:10.1042/BST20180320
701 11. Swiers, G. *et al.* Early dynamic fate changes in haemogenic endothelium
702 characterized at the single-cell level. *Nat. Commun. 2013 41* **4**, 1–10 (2013).
703 12. Kissa, K. & Herbomel, P. Blood stem cells emerge from aortic endothelium by a novel
704 type of cell transition. *Nature* **464**, 112–115 (2010).
705 13. Eilken, H. M., Nishikawa, S. I. & Schroeder, T. Continuous single-cell imaging of blood
706 generation from haemogenic endothelium. *Nature* **457**, 896–900 (2009).
707 14. Fidanza, A. *et al.* Single cell analyses and machine learning define hematopoietic
708 progenitor and HSC-like cells derived from human PSCs. *Blood* **136**, 2893–2904
709 (2020).
710 15. Zeng, Y. *et al.* Tracing the first hematopoietic stem cell generation in human embryo
711 by single-cell RNA sequencing. *Cell Res.* **29**, 881–894 (2019).
712 16. Ng, E. S. *et al.* Differentiation of human embryonic stem cells to HOXA+ hemogenic

713 vasculature that resembles the aorta-gonad-mesonephros. *Nat. Biotechnol.* **34**, 1168–
714 1179 (2016).
715 17. Calvanese, V. *et al.* Mapping human haematopoietic stem cells from haemogenic
716 endothelium to birth. *Nat. 2022 6047906* **604**, 534–540 (2022).
717 18. Sturgeon, C. M., Ditadi, A., Awong, G., Kennedy, M. & Keller, G. Wnt signaling
718 controls the specification of definitive and primitive hematopoiesis from human
719 pluripotent stem cells. *Nat. Biotechnol.* **32**, 554–561 (2014).
720 19. Fidanza, A. *et al.* An all-in-one UniSam vector system for efficient gene activation. *Sci.*
721 *Rep.* **7**, 6394 (2017).
722 20. Yang, C.-T. *et al.* Activation of KLF1 Enhances the Differentiation and Maturation of
723 Red Blood Cells from Human Pluripotent Stem Cells. *Stem Cells* **35**, 886–897 (2017).
724 21. Lopez-Yrigoyen, M. *et al.* A human iPSC line capable of differentiating into functional
725 macrophages expressing ZsGreen: A tool for the study and in vivo tracking of
726 therapeutic cells. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, (2018).
727 22. Lopez-Yrigoyen, M. *et al.* Genetic programming of macrophages generates an in vitro
728 model for the human erythroid island niche. *Nat. Commun.* **10**, 881 (2019).
729 23. Lopez-Yrigoyen, M. *et al.* A human iPSC line capable of differentiating into functional
730 macrophages expressing ZsGreen: A tool for the study and in vivo tracking of
731 therapeutic cells. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, (2018).
732 24. May, A. *et al.* Modelling the erythroblastic island niche of dyserythropoietic anaemia
733 type IV patients using induced pluripotent stem cells. *Front. cell Dev. Biol.* **11**, (2023).
734 25. Kang, S. J., Park, Y. Il, So, B. & Kang, H. G. Sodium butyrate efficiently converts fully
735 reprogrammed induced pluripotent stem cells from mouse partially reprogrammed
736 cells. *Cell. Reprogram.* **16**, 345–354 (2014).
737 26. Zhang, Z., Xiang, D. & Wu, W. S. Sodium Butyrate Facilitates Reprogramming by
738 Derepressing OCT4 Transactivity at the Promoter of Embryonic Stem Cell–Specific
739 miR-302/367 Cluster. *Cell. Reprogram.* **16**, 130 (2014).
740 27. Replogle, J. M. *et al.* Combinatorial single-cell CRISPR screens by direct guide RNA
741 capture and targeted sequencing. *Nat. Biotechnol.* (2020). doi:10.1038/s41587-020-
742 0470-y
743 28. Ventura, T., Egan, E. J., Romanò, N. & Fidanza, A. In Vitro Model of Fetal Human
744 Vessel On-chip to Study Developmental Mechanobiology. *J. Vis. Exp.* **2023**, (2023).
745 29. Azzoni, E. *et al.* The onset of circulation triggers a metabolic switch required for
746 endothelial to hematopoietic transition. *Cell Rep.* **37**, (2021).
747 30. Oburoglu, L. *et al.* Pyruvate metabolism guides definitive lineage specification during
748 hematopoietic emergence. *EMBO Rep.* **23**, (2022).
749 31. Castaño, J. *et al.* GATA2 Promotes Hematopoietic Development and Represses

750 Cardiac Differentiation of Human Mesoderm. *Stem Cell Reports* **13**, 515–529 (2019).
751 32. de Pater, E. *et al.* Gata2 is required for HSC generation and survival. *J. Exp. Med.*
752 **210**, 2843–2850 (2013).
753 33. Ling, K. W. *et al.* GATA-2 Plays Two Functionally Distinct Roles during the Ontogeny
754 of Hematopoietic Stem Cells. *J. Exp. Med.* **200**, 871 (2004).
755 34. McGarvey, A. C. *et al.* A molecular roadmap of the AGM region reveals BMPER as a
756 novel regulator of HSC maturation. *J. Exp. Med.* **214**, 3731–3751 (2017).
757 35. McGrath, K. E. *et al.* A transient definitive erythroid lineage with unique regulation of
758 the β-globin locus in the mammalian embryo. *Blood* **117**, 4600 (2011).
759 36. Rejeski, K., Duque-Afonso, J. & Lübbert, M. AML1/ETO and its function as a regulator
760 of gene transcription via epigenetic mechanisms. *Oncogene* **40**, 5665–5676 (2021).
761 37. Huynh, H. D. *et al.* IGF binding protein 2 supports the survival and cycling of
762 hematopoietic stem cells. *Blood* **118**, 3236 (2011).
763 38. Crosse, E. I. *et al.* Multi-layered Spatial Transcriptomics Identify Secretory Factors
764 Promoting Human Hematopoietic Stem Cell Development. *Cell Stem Cell* **27**, 822
765 (2020).
766 39. Hadland, B. *et al.* Engineering a niche supporting hematopoietic stem cell
767 development using integrated single-cell transcriptomics. *Nat. Commun. 2022 131* **13**,
768 1–17 (2022).
769 40. Hadland, B. K. *et al.* Endothelium and NOTCH specify and amplify aorta-gonad-
770 mesonephros-derived hematopoietic stem cells. *J. Clin. Invest.* **125**, 2032–2045
771 (2015).
772 41. Sandler, V. M. *et al.* Reprogramming human endothelial cells to haematopoietic cells
773 requires vascular induction. *Nature* **511**, 312–318 (2014).
774 42. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered
775 CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).
776 43. Petazzi, P. *et al.* Robustness of Catalytically Dead Cas9 Activators in Human
777 Pluripotent and Mesenchymal Stem Cells. *Mol. Ther. - Nucleic Acids* **20**, 196–204
778 (2020).
779 44. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-
780 3587.e29 (2021).
781 45. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis.
782 *Nat. 2019 5667745* **566**, 496–502 (2019).
783 46. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for
784 animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
785