

Reviewed Preprint

v2 • December 4, 2024

Revised by authors

Reviewed Preprint

v1 • April 29, 2024

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, Telma Ventura, Francesca Paola Luongo, Heather McClafferty, Alisha May, Helen Alice Taylor, Micheal J Shipston, Nicola Romanò, Lesley M Forrester, Pablo Menéndez, Antonella Fidanza 

Josep Carreras Leukemia Research Institute, Barcelona, Spain • Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK • Centre for Discovery Brain Sciences, Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK • Zhejiang University-University of Edinburgh Joint Institute, Zhejiang University School of Medicine, Zhejiang University, Haining, China • Red Española de Terapias Avanzadas (TERAV)-Instituto de Salud Carlos III (ISCIII), Madrid, Spain • CIBER-ONC, ISCIII, Barcelona, Spain • Institutió Catalana de Recerca i Estuis Avançats (ICREA), Barcelona, Spain • Department of Biomedicine, School of Medicine, University of Barcelona, Barcelona, Spain • Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

eLife Assessment

This study presents **useful** findings to inform and improve the in vitro differentiation of hematopoietic progenitor cells from human induced pluripotent stem cells. Relying on a well-characterised technical approach, the data analysis is overall **solid** and reasonably supports the main conclusions.

<https://doi.org/10.7554/eLife.94884.2.sa2>

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. This work has identified a key 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 tagged 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 our analysis indicated that IGFBP2 is involved in the remodeling of 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 and 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 IGFBP2 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

We, and others, have shown that *in vitro* differentiation of human iPSCs is a powerful tool to model intraembryonic hematopoiesis ^{14,16–18}. To understand the molecular basis underlying the challenges associated with the *in vitro* production of blood stem and progenitor cells *in vitro* from differentiating iPSCs, we compared our scRNAseq dataset from differentiating iPSCs ¹⁴ to that of cells derived *in vivo* from the human aorta-gonad-mesonephros (AGM) region ¹⁵. We integrated the transcriptomic data of *in vitro* derived endothelial (IVD_Endo) and hematopoietic cells (IVD_HPC) with that of arterial endothelial cells (aEC), arterial hemogenic cells (aHEC) and venous endothelial cells (vEC) derived from human embryos collected between Carnegie stage 12 and 14 (**Figure 1A-B**). To identify possible target genes that could be manipulated *in vitro* to improve iPSCs differentiation, we first determined the transcription

factors marking the aHEC *in vivo* (Table 1). Then, we filtered the transcription factor genes that were expressed at lower levels in the IVD_Endo and IVD_HPC. This strategy identified 9 target transcription factors *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*, *ZNF124*, *SOX6*, *ZNF33A*, *NFAT5* and *TFDP2* (Figure 1C), whose expression was detected across other AGM datasets in the hemogenic endothelium¹⁷ (Supplementary Figure 1).

Development of a DOX-inducible dCAS9-SAM activation system in human iPSCs

We previously developed an all-in-one Synergistic Activator Mediator system, UniSAM, that mediates the transcriptional activation of endogenous gene expression¹⁹. To activate the nine target genes identified in this study, we developed a novel doxycycline-inducible SAM (iSAM) cassette targeted into the *AAVS1* locus of human iPSCs (Figure 2A). We demonstrated this strategy could activate of *RUNX1C* expression in HeLa cells and that this was correlated with the doxycycline (DOX) concentration in a linear manner (Figure S2A-B). To verify gene activation in human PSCs at single-cell resolution, we employed a *RUNX1C*-GFP human embryonic stem cell (hESC) reporter cell line (Figure S2C-G). As predicted, the level of expression of the mCherry tag, the fluorescent mCherry reporter tag within the iSAM cassette, was proportional to the concentration of DOX (Figure S2D-E) and to the number of cells in which *RUNX1C*-GFP was activated (Figure S2 D-F). Furthermore, the *RUNX1C* expression level, measured by the mean fluorescence intensity (MFI) of the *RUNX1C*-GFP reporter, also correlated with the concentration of DOX (Figure S2G). We then tested the iSAM cassette in the iPSC line (SFCi55) (Figure 2B-D). Only when iPSCs were transfected with both iSAM and the gRNA directed to *RUNX1C* and treated with DOX, was the expression of the *RUNX1C* gene and *RUNX1* protein detected (Figure 2B-C).

We then targeted the iSAM cassette into the *AAVS1* locus using a Zinc Finger Nuclease (ZFN) strategy^{14,20,21}. iPSC clones that had specifically integrated the iSAM cassette into the *AAVS1* locus were validated by genomic PCR screening (Figure S2A-B) and sequencing. The *AAVS1* locus has been reported to be a “safe harbour” site that is resistant to epigenetic silencing and indeed, we had previously demonstrated that transgenes inserted into the *AAVS1* locus under the control of the constitutively active CAG promoter were efficiently expressed both in undifferentiated and in differentiated iPSCs^{20,22–24}. However, after the iSAM line had been established and cultured under self-renewal conditions, we noted a dramatic reduction in the number of mCherry+ cells in undifferentiated iSAM iPSCs upon DOX induction. We predicted this to be due to transgene-silencing of the rTTA DOX-inducible cassette (Figure S3C). To overcome this problem, we treated the iSAM iPSC line with an inhibitor of histone deacetylases (HDACs), sodium butyrate (SB), reported to have no adverse effect on iPSC maintenance^{25,26} and we also confirmed that the treatment with SB had no effect on viability and cell proliferation in our culture conditions (Figure S3F-G). A short 48-hour treatment significantly increased the number of mCherry+ cells upon DOX induction, proportional to the SB concentration (Figure S3D-E). We, therefore, maintained the iSAM iPSCs in the presence of SB and this fully restored the inducibility of the transgene with virtually all cells expressing mCherry in the presence of DOX (Figure S3G). We did not treat the cells with SB during the differentiation since we detected robust expression of the iSAM cassette during the differentiation (Figure S4F). Furthermore, since the HDACs are also involved in the chromatin remodelling during the EHT process²⁷ we predicted that SB treatment would negatively affect the differentiation.

To test the effect of activating the 9 target genes on the transcriptomes of differentiating iPSC cells, we engineered the gRNAs to allow their detection within the single cell RNA sequencing pipeline²⁸. We inserted a capture sequence prior to the termination signal to avoid any alteration in the secondary structure of the loops thus preserving the binding of the synergistic activators of the SAM system to the stem loops of the gRNAs. Of the two capture sequences available²⁸, we decided to use the one that was predicted to result in fewer secondary structure alterations and

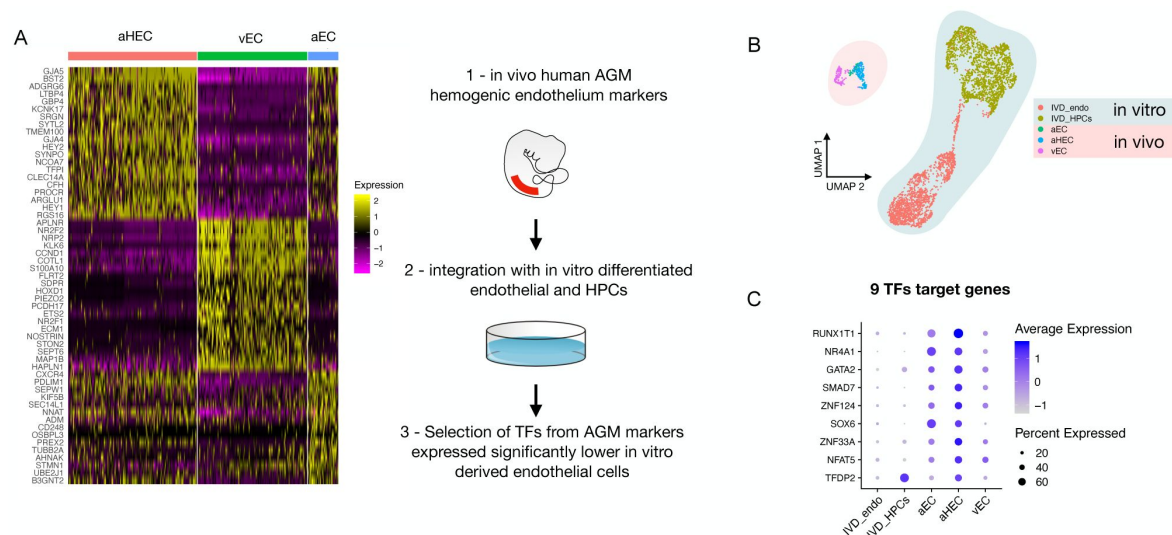


Figure 1.

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

A - Schematic of the analytic pipeline used to identify the target genes. **B** - 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. **C** - Target genes expression level showing higher expression in arterial hemogenic endothelium *in vivo* than *in vitro* derived cells.

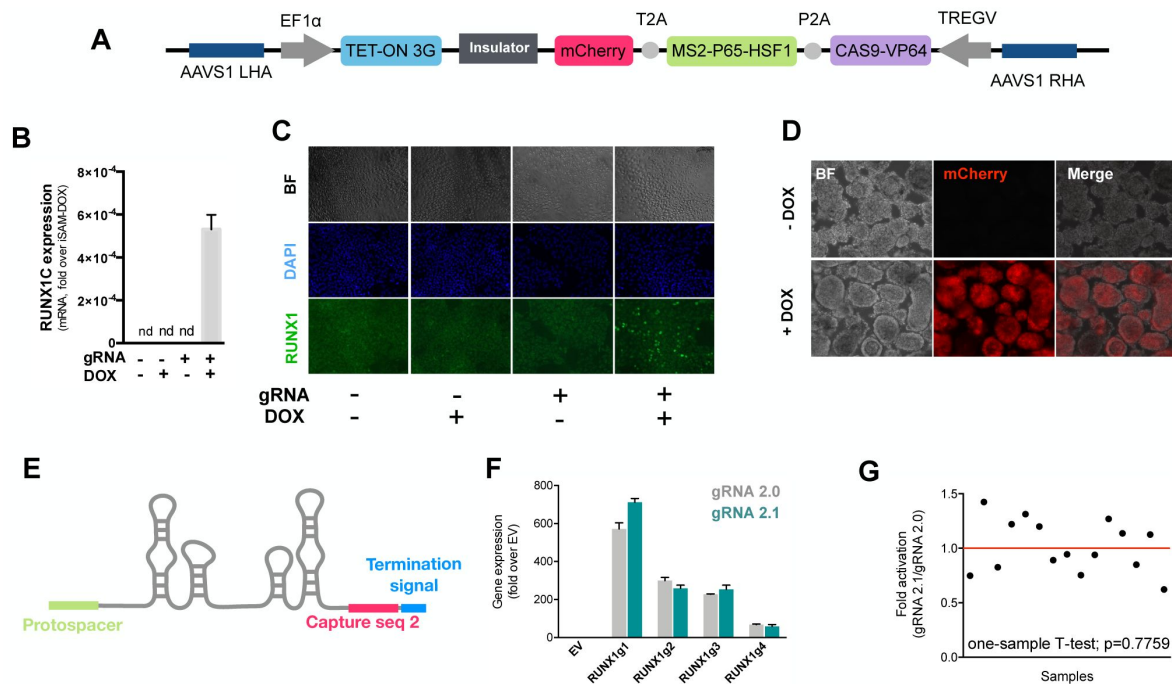


Figure 2.

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).

this new gRNA was named 2.1 (**Figure 2E** [↗](#)). We compared the activation level achieved with the new 2.1 gRNA to that of the original 2.0 backbone using various gRNAs targeting *RUNX1C* (**Figure 2F** [↗](#)). These results convincingly demonstrated that the addition of the capture sequence in the gRNA 2.1 does not alter the level of endogenous gene activation that could be achieved (**Figure 2G** [↗](#)). Altogether, these results show that the iSAM system is able to induce gene expression that predictably translates into an increased protein expression and thus provides a platform to steer phenotypical changes in cell identity.

To activate our target genes, we designed 5-7 gRNAs in the 200bp upstream of the transcriptional start sites of each of the 9 target genes. We subcloned a total of 49 gRNAs (Table S1) into the gRNA 2.1 backbone and packaged them into lentiviral particles, (herein referred to as the AGM library) as well as a non-targeting (NT) gRNA that was used as control. The iSAM iPSC line was transduced with the targeting gRNAs or the control non-targeting gRNA to generate the iSAM_AGM and iSAM_NT iPSCs line, respectively (Figure S4A). After puromycin selection, their integration in the genome was also confirmed by PCR and sequencing (Figure S3H-I).

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

To assess the transcriptional changes in response to the activation of the target genes, we differentiated the iSAM iPSCs, induced with DOX and subjected them to single-cell RNA sequencing using the 10X pipeline. After 10 days of differentiation, in the presence or absence of DOX, we FAC-sorted live CD34⁺ cells from iSAM_AGM and the iSAM_NT iPSCs (Figure S4B). Following data filtering, we selected cells in which the gRNAs expression was detected and showed that our approach activated all the target genes, except for *ZFN124*. A higher level of expression of *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*, *SOX6*, *ZNF33A*, *NFAT5*, *TFDP2* was observed following DOX treatment of cells of iSAM_AGM cells compared to the iSAM_NT cells (**Figures 3A** [↗](#), S3G). To study the effect of gene activation on transcriptional and cellular phenotype, we performed clustering analysis and detected a total of 7 clusters (**Figure 3B-E** [↗](#)). A high level of *GJA4* and *DLL4* expression was used to annotate the arterial-like cell cluster, while high levels of hemogenic markers, such as *RUNX1* and *CD44*, were used to annotate the hemogenic clusters, EHT_1 and EHT_2 (**Figure 3B** [↗](#), C-D). All the clusters expressed pan-endothelial markers such as *PECAM1* and *CDH5*, coding for CD31 and VECAD, respectively, which get progressively downregulated in EHT_1 and EHT_2, as expected for cells undergoing the EHT process. To assess the effect of the activation on cell identity, we analysed the proportion of the cell clusters between the libraries and detected a significant expansion of the arterial cluster in the DOX-induced iSAM-AGM compared to the iSAM_NT cells (**Figures 3E** [↗](#), S4D, S4I). To validate the effect of the activation on the expansion of arterial cell population, we analysed their prevalence by flow cytometry. Although DOX treatment resulted in an average 1.63 ± 0.13 fold increase of CD34⁺DLL4⁺ cells in the control iSAM_NT sample (Figure S3A), the increase observed in the iSAM-AGM cells was significantly greater, with an expansion of 3.33 ± 0.73 fold increase of CD34⁺DLL4⁺ immunophenotypic arterial cells identified by flow cytometry (**Figure 3F** [↗](#)). To assess the effect of the different cell composition in the CD34⁺ compartment upon activation on the emergence of colony forming progenitors, we isolated CD34⁺ cells using magnetic beads and cocultured 20,000 cells on OP9 supportive stromal cells for 7 days in the presence of hematopoietic differentiation cytokines. After one week, the progenitor cells were assessed by colony-forming unit (CFU) assays and scored 14 days later. We detected an increased number of CFU-E (colony-forming unit erythroid) and CFU-GM (colony-forming unit granulocyte/macrophage) and a reduction of CFU-M (colony-forming unit macrophage) in iSAM-AGM samples cultured in the presence of DOX compared to the absence of DOX but no significant effect of DOX in SAM_NT samples (**Figure 3G** [↗](#)). These data indicate that activation of target transcription factors using our novel CRISPR strategy results in transcriptional remodelling and a steer in cell identity that we detected as a functional difference in the hematopoietic progenitor profile.

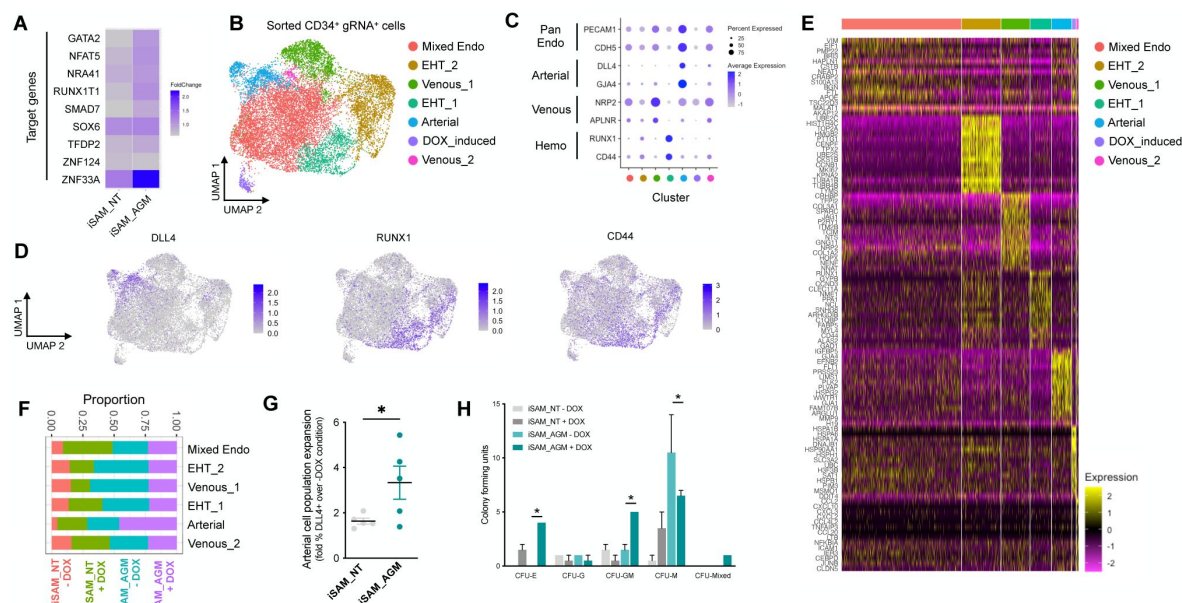


Figure 3.

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*, *APLN*) 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** - Heatmap of the top 15 marker genes for each of the clusters. **F** - 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. **G** - 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 = 0.0417 paired t-test). **H** - Colony forming potential of the suspension progenitor cells derived from the two lines treated with or without DOX following OP9 coculture activation, data show the colony obtained for 104 CD34+ input equivalent (n=3 from independent differentiations * p<0.05, Tukey's two-way ANOVA).

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

To better understand the molecular mechanism behind the increased progenitor development, we compared the expression profile of the arterial cells between the different activation libraries, and we obtained a list of genes upregulated upon activation of the targets (Table 1). The most upregulated gene, *IGFBP2*, was expressed at significantly higher levels in the iSAM_AGM library in the presence of DOX compared to the others (Figure 4A, Table 1). Although the arterial cells expressed *IGFBP2* at the highest level compared to other cell types, the activation of the gene was not cell-type specific and it was detected across the various clusters (Figure S4H). We then compared the gRNA distribution in these arterial cells from the iSAM_AGM treated with DOX to that of arterial cells without activation. We observed a significant enrichment of the *RUNX1T1*-specific gRNAs (Table 1), indicating that the increased *IGFBP2* expression could be downstream of *RUNX1T1* activation. IGF Binding Protein 2 is a member of the family of IGF binding proteins and is thought to be secreted from cells where it then binds IGF1, IGF2, and other extracellular matrix proteins, modulating their function. IGF1 and IGF2 are commonly used in differentiation protocols^{29–31}, including ours^{32,33}, due to their direct role in blood development. To test if the increased frequency of functional hematopoietic progenitors was due to IGFBP2 signalling, presumably derived from the arterial cells, we supplemented the media with IGFBP2 at 100ng/ml after the induction of endothelial cell differentiation. To explore the role of IGFBP2 we employed the parental iPSCs line, SFCi55 from which the iSAM line was derived. We isolated CD34⁺ cells at day 8 and co-cultured them on OP9 cells in presence of IGFBP2 for one week, then tested for their hematopoietic clonogenic potential using CFU assays (Figure 4B). Cells treated with IGFBP2 showed a significant increase in the total number of haematopoietic CFU colonies compared to cells in control cultures. To assess whether IGFBP2 also affected the production of arterial cells themselves, we analysed the proportion of DLL4⁺ cells, but no difference was detected in the presence of IGFBP2 (Figure 4C). This implies that the mechanism of action of IGFBP2 is different from that mediated by the gene activation that led to an expansion of the arterial population. To further understand the relationship between colonies' potential with the arterial identity, we sorted CD34⁺ into DLL4⁺ and DLL4[−] and cultured them on OP9 co-culture for a week prior to methylcellulose assay. CD34⁺DLL4⁺ and CD34⁺DLL4[−] plated on OP9 showed different capacity to generate suspension cells (Figure S5D), in line with the colonies' formation results showing that the DLL4[−] contains the largest progenitor activity. This aligns with the observation that the EHT process coincides with the downregulation of arterial markers such as *DLL4* and the upregulation of hemogenic markers such as *RUNX1* (Figure 3C).

This observation supports our hypothesis that the change in hematopoietic progenitor production following activation in the iSAM_AGM line is associated with both differential gene expression within the arterial cells rather than their expansion. We then focused on the characterization of the cells derived from the CD34⁺ cells after coculture with the OP9 in the presence of IGFBP2. Our results showed a significant expansion of the CD34⁺ and CD43⁺ cell populations, further supporting our hypothesis (Figure 4D). To address the potential role of *IGFBP2* *in vivo*, we analysed single cell sequencing data from the human AGM at Carnegie stages 14 and 15. *IGFBP2* is highly expressed within the AGM niche by stromal and epithelial cells and, most importantly, highest in endothelial cells (Figure 4E). Together these data show that IGFBP2 addition results in increased hematopoietic blood production *in vitro* and that the endothelial compartment is the most likely source of IGFBP2 both *in vitro* and *in vivo*.

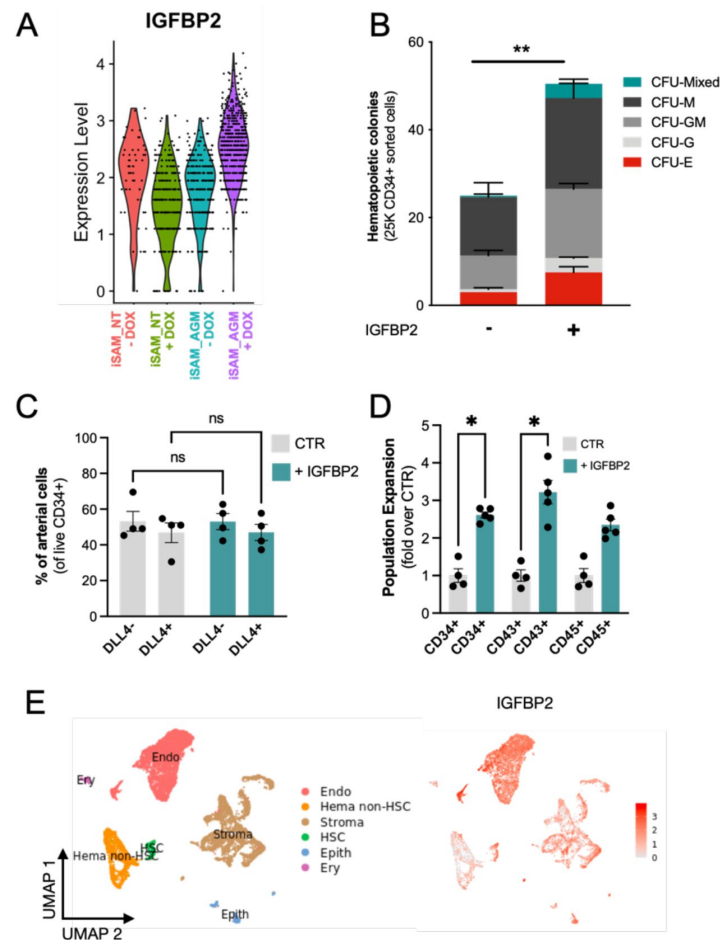


Figure 4.

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.

IGFBP2 enhances metabolic dependency on oxidative phosphorylation of differentiating endothelial cells

Following the observation that IGFBP2 supports haematopoietic progenitor differentiation *in vitro* from human iPSCs we performed a time course single-cell RNA experiment of SFCi55 iPSCs differentiated in its presence. We isolated CD34⁺ cells and plated them in EHT culture on laminin³⁴, rather than on OP9 support, to assess exclusively the specific effects of IGFBP2-on differentiating iPSCs. FAC-sorted single/live adherent cells from day 10 and 13 in the presence and absence of IGFBP2 were subjected to single-cell RNAseq (Figure S4C). Our time course transcriptomic analyses showed that IGFBP2 induced a change in the transcriptome, specifically on day 13 (**Figure 5A**). A cluster of endothelial cells enriched for genes associated with the KEGG pathway of growth factor binding, was detected almost exclusively at day 13 (Figure S5F); these cells displayed a different transcriptional signature (as indicated by a shift in their position in the UMAP embedding) in the presence of IGFBP2, and were enriched upon treatment (Figure S5F). Interestingly, this cluster showed expression of *GJA4* across the cells, indicating their broad arterial identity, with a specific reduction of *DLL4* expression level concomitantly to an upregulation of *RUNX1* in the region of the cluster induced by IGFBP2 (**Figure 5B**). This increase in hemogenic identity within the endothelial cell compartment is consistent with the observation of increased functional progenitor production. This suggests that IGFBP2 supports arterial cells in the acquisition of hemogenic capacity.

We compared the transcriptome of cells at day 13 in the presence and absence of IGFBP2 and, using KEGG enrichment analysis, we observed that the genes that were upregulated by IGFBP2 were highly enriched in the oxidative phosphorylation term (**Figure 5C-D**). Since the metabolic switch between glycolytic to mitochondrial metabolism has been previously reported in definitive hematopoiesis^{35,36}, we tested whether the addition of IGFBP2 could result in a different ATP production profile. We analysed the ATP production at day 13, when *RUNX1* expression was induced by IGFBP2 addition, by using specific inhibitors of the complex I and II and complex V to quantify the intracellular ATP and mitochondrial synthesis, respectively. We detected a reduction in glycolytic-derived ATP (**Figure 5E-F**), which translates to a higher contribution of mitochondria metabolism for IGFBP2-treated cells (**Figure 5G**). We then looked at the expression levels of genes encoding the enzymes of the glycolytic pathway and its checkpoints, glucose and lactate transporters, monocarboxylate transporters and the enzymes associated with hexokinase and phosphofructose reactions. We observed a general downregulation of the glycolytic enzymes and its checkpoints, with the exception of *PFKM* and *SLC16A1*, (Figures S5A-B), which were expressed at lower levels in the IGFBP2-treated cells at day 13 compared to the control on the same day. These differences were not observed on day 10 cells, which is consistent with our previous observations that the IGFBP2-induced remodelling of the endothelial cells transcriptome happens exclusively on day 13.

To assess whether the higher clonogenic potential and effect on the bioenergetic profile of the cells treated with IGFBP2 was the result of an increase in proliferation, we analysed the cell cycle profile of suspension hematopoietic cells obtained from the OP9 cocultured in the presence of IGFBP2. No differences in the cell cycle distribution of hematopoietic progenitors (Figure S3C) was observed, indicating that the detected increase in hematopoietic progenitors is not likely to be a consequence of their increased cycling. We, therefore, concluded the effect of IGFBP2 occurs prior to, or during their emergence via the EHT process.

In summary, our results show that the activation of the 9 target transcription factors, expressed at higher levels *in vivo* within the arterial cells in the AGM region, leads to the expansion of the arterial cells and, consequently, to an increase in progenitor activity (**Figure 6**). IGFBP2 was identified as the most upregulated gene in the arterial cell upon gene activation. The addition of

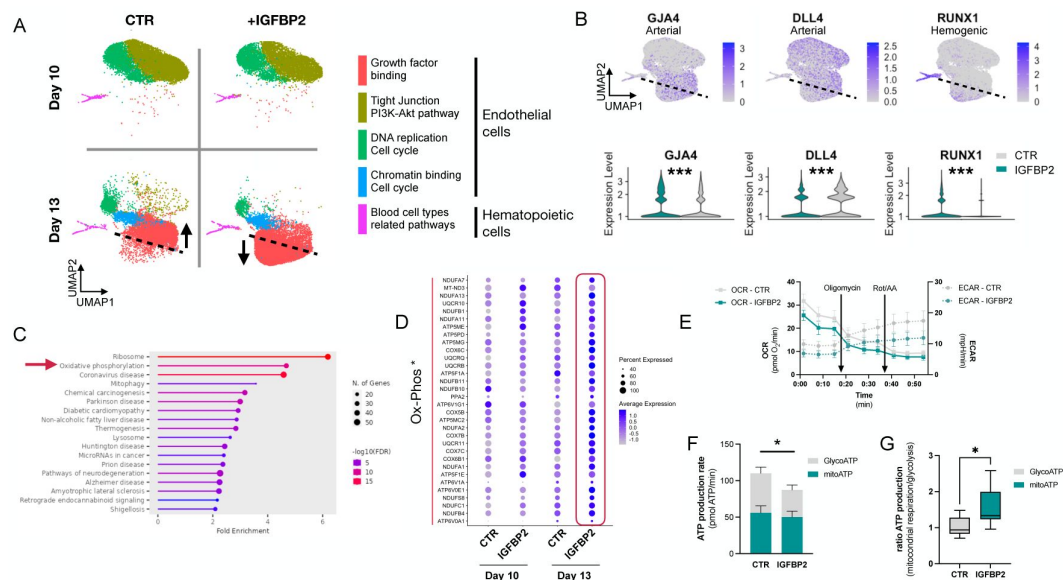


Figure 5.

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 (*GJA4* $p=1E^{-54}$, *DLL4* $p=1.2E^{-119}$, *RUNX1* $p=8.2E^{-163}$). **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.

IGFBP2 to the differentiation culture induced the upregulation of *RUNX1* and OxPhos genes and accelerated the metabolic dependency on OXPHOS in association with improved progenitor activity (**Figure 6** [↗](#)).

Discussion

The complexity and dynamism of developmental hematopoiesis *in vivo* have imposed challenges in accurately reproducing the process *in vitro*. We hypothesized that this could, in part, be due to the inability to recapitulate the appropriate transcriptional programme in cultured cells. Here, we developed a novel CRISPR-activation system to induce the expression of genes that are expressed at low levels in cells that are generated *in vitro*, to explore the downstream consequences of their activation and to assess the effects on emerging hematopoietic progenitor cells.

When we compared endothelial cells derived *in vitro* from hiPSCs to those in the AGM region of the human embryo, at the time point when early hematopoietic commitment takes place, we identified nine transcription factors that were expressed at a lower level in the cultured cells. Some of these had been associated with blood cell development, including *GATA2*³⁷[↗](#)–³⁹[↗](#), *SMAD7*⁴⁰[↗](#), *NR4A1*⁴⁰[↗](#), *SOX6*⁴¹[↗](#) and *RUNX1T1*¹⁵[↗](#),¹⁷[↗](#). Other genes such as,

ZNF124, ZNF33A, NFAT5 and TFDP2 had not been previously associated with hematopoiesis and could provide early evidence of their possible role in hematopoiesis that will require further studies. We were particularly interested in *RUNX1T1* (also known as ETO) which is associated with the leukemic fusion protein, AML1/ETO resulting from the t(8;21) chromosomal translocation⁴²[↗](#). Furthermore, *RUNX1T1* expression has been recently detected in transcriptomic analyses of the human AGM region¹⁵[↗](#),¹⁷[↗](#), but its precise role during the ontogeny of the blood system has not been elucidated. The addition of a capture sequence to the gRNA backbone enabled their detection coincidentally with the single-cell transcriptome, and this allowed us to demonstrate that *RUNX1T1* gRNAs were significantly enriched in cells within an expanded arterial cluster and their presence was associated with the highest expression of *IGFBP2* across clusters. *IGFBP2* KO mice show increased expression of cell-cycle inhibitors and HSC apoptosis, implicating *IGFBP2* as a modulator of HSCs cell cycle and survival⁴³[↗](#). More recently *IGFBP2* was reported as being highly expressed in the human AGM region at CS14 when HSCs are emerging¹⁷[↗](#), supporting a possible role during developmental hematopoiesis but the precise molecular process was unclear. In this study, we show that the addition of *IGFBP2* recombinant protein in our *in vitro* model of EHT results in the emergence of an increased number of functional hematopoietic progenitors, and we provide an early observation to suggest that *RUNX1T1* could be involved in the regulation its expression. Because *RUNX1T1* lacks a DNA binding domain, its direct involvement in the regulation of *IGFBP2* expression, or in the development of the hematopoietic system in general, must require association with other cofactors that are yet to be identified.

Our data, together with *in vivo* data¹⁷[↗](#), show that *IGFBP2* is expressed predominantly by endothelial cells. The supportive role of the endothelial niche in the development of HSCs has been studied *in vivo*⁴⁴[↗](#)–⁴⁶[↗](#) and exploited *in vitro* to support HSCs emergence⁴⁵[↗](#),⁴⁷[↗](#). We describe here a novel role for *IGFBP2* in the remodelling of the metabolism of iPSC-derived endothelial cells and that this is associated with the induction of *RUNX1*, a hallmark of hemogenic fate. The addition of *IGFBP2* in the culture induces upregulation of the genes of the oxidative phosphorylation pathway in association with an increased relative mitochondrial contribution to cellular ATP production. This is due to reduced glycolysis in association with the downregulation of genes coding for glycolytic enzymes and their checkpoints. The switch between glycolytic to mitochondrial metabolism has been shown to be essential for the EHT in definitive hematopoiesis³⁶[↗](#),⁴⁸[↗](#). This switch is induced *in vivo* by mechanical cues downstream of the establishment of circulation, and it is required for functional HSCs development³⁵[↗](#), while *in vitro*, the switch is driven by pyruvate mitochondrial catabolism, leading to definitive EHT as

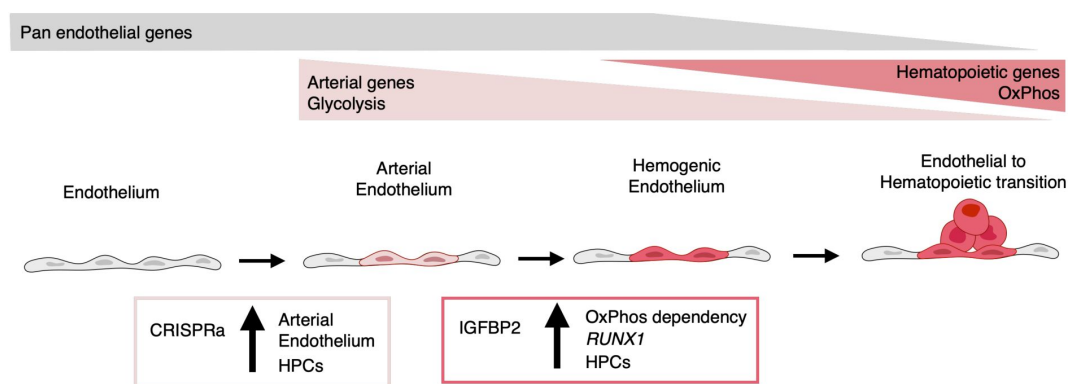


Figure 6.

Model summarising the results.

During development, some endothelial cells undergo arterialisatation, as identified by their arterial genes' expression profile (e.g. *DLL4*). Arterial cells, characterised by high dependency on glycolysis, are expanded by our CRISPR activation approach, resulting in more blood production since these cells are the cell-of-origin of the hemogenic endothelium. Once arterial cells commit to hemogenic endothelium fate, they start to express hematopoietic genes (e.g. *RUNX1*); this process is enhanced by IGFBP2 via induction of both *RUNX1* expression and increased dependency on Oxidative Phosphorylation, known to be important for the progression of the endothelial to hematopoietic transition.

opposed to primitive³⁶. Further *in vivo* studies are required to characterize the role of IGFBP2 during AGM hematopoiesis to overcome the limitation of using an *in vitro* model such as iPSCs differentiation.

In conclusion, we detected transcriptional differences between *in vivo* and *in vitro* developmental hematopoiesis and developed a novel inducible gene activation system to identify novel molecular players during the endothelial-to-hematopoietic transition. Our multidisciplinary approach identified IGFBP2 as novel signalling molecules that support human blood progenitor development *in vitro*, inducing a metabolic switch from cytoplasmic glycolysis to mitochondrial respiration. This study demonstrates that combining CRISPR-mediated activation of target genes with single-cell transcriptomic analysis in differentiating hPSCs can be a powerful approach to alter cell fate, providing a tool for gene function studies during human development. The fine epigenetic manipulation of the transcription can be readily applied to any cell lineage simply by adding specific gRNAs and it will be instrumental in exploring other developmental processes that can be, at least partially, be mimicked *in vitro* with human iPSCs. Some limitations remain in applying CRISPR-mediated gene activation strategies in long differentiation protocols due to the challenge of detecting high copy numbers of the gRNAs, limiting the possibility of providing the fine statistical correlation needed to predict downstream target genes. Testing different promoters driving the gRNA expression or structural modifications of the gRNA scaffold could result in a more robust expression and allow for correlation analysis.

Finally, using this approach, we have identified the supportive role of IGFBP2, predominantly produced by endothelial cells, which induces transcriptional and metabolic remodelling in association with the induction of *RUNX1* expression, and result in higher hematopoietic progenitors' activity.

Methods

Resource availability

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

Pluripotent Stem Cells maintenance

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

Transfection

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

Immunocytochemistry

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

Gene expression analysis

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

Pluripotent Stem Cells differentiation to hematopoietic progenitors

hPSCs were differentiated in a xeno-free composition of SFD medium ¹⁴[\[14\]](#), BSA was substituted with human serum albumin, HSA (Irvine-Scientific). Day 0 differentiation medium, containing 10 ng/ml BMP4 was added to the colonies prior cutting. Cut colonies were transferred to a Cell Repellent 6 wells Plates (Greniner) to form embryoid bodies and cultured for two days. At day 2 media was changed and supplemented with 3 μ M CHIR (StemMacs). At day 3, EBs were transferred into fresh media supplemented with 5 ng/ml bFGF and 15 ng/ml VEGF. At day 6 media was changed for final haematopoietic induction in SFD medium supplemented with 5 ng/ml bFGF, 15 ng/ml VEGF, 30 ng/ml IL3, 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 25 ng/ml IGF1. From day 6 onward, cytokines were replaced every two days.

CD34 isolation

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

OP9 coculture and colony assay

OP9 cells were maintained in α -MEM supplemented with 20% serum (Gibco) and sodium bicarbonate (Gibco) and passaged with Trypsin every 3-4 days. The day before the co-culture, 45.000 OP9 cells were plated for each 12 well plates' well in SFD media. The day of the co-culture the 20.000 iSAM cells or 25.000 SFCi55 or H9 were plated in each well and culture in SFD media supplemented with 5 ng/ml bFGF, 15 ng/ml VEGF, 30 ng/ml IL3, 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 25 ng/ml IGF1 and 100ng/ml IGFBP2. Cytokines were replaced twice during one week of coculture. At the end of the coculture, cells were collected by Trypsin and half of the well equivalent was plated in 2 ml of methylcellulose medium (Human enriched H4435, Stemcell Technologies). Cells were incubated in the assay for 14 days and then scored.

Laminin EHT culture

Laminin EHT culture was performed as previously described³⁴. Briefly, 24 well plates were coated for at least 2 hours with recombinant human Laminin-521 (Thermo-Fisher). Following CD34⁺ isolation at day 8, 400,000 CD34⁺ cells were seeded in each well of a pre-coated 24-well plate in SFD media supplemented with 5 ng/ml bFGF, 15 ng/ml VEGF, 30 ng/ml IL3, 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 25 ng/ml IGF1 and 100ng/ml IGFBP2.

Flow cytometry staining and cell sorting

Embryoid bodies were dissociated using Accutase (Life Technologies) at 37°C for 30'. Cells were centrifuged and resuspended in PBS + 0.5% BSA + 2mM EDTA, counted and stained at 10⁵ cells for a single tube. Cells were stained with antibodies for 30' at room temperature gently shaking. Flow cytometry data were collected using DIVA software (BD). For the sorting experiments, the cells were prepared as above and stained at 10⁷ cells/ml in presence of the specific antibodies. Sorting was performed using FACSARIA Fusion (BD) and cells were collected in PBS + 1% BSA. Data were analysed using FlowJo version 10.4.2.

Flow cytometry antibodies

For flow cytometry 10⁵ cells per test were stained in 50 µl of staining solution with the following antibodies: CD34 Percp-Efluor710 (4H11 eBioscience, 1:100), CD34 Pe (4H11 eBioscience, 1:200), CD43 APC (eBio84-3C1, 1:100), CD45 FITC (2D1 ebioscience, 1:100), DLL4 Pe (MHD4-46 Biolegend, 1:200), CD41 PE (HIP8 Biolegend, 1:200), CD144 APC (16B1 eBioscience, 1:100), CD235a FITC (HIR2 BD Bioscience, 1:250).

iSAM plasmid generation

The iSAM plasmid was obtained by Gibson assembly of four fragments. The first fragment, the backbone, was a DOX-inducible AAVS1 targeted plasmid expressing an E6-E7-IRES-ZsGreen which was excised by BstBI and NdeI. The second fragment, one of the adapters, was derived from the UniSAM plasmid that we previously generated (Addgene #99866) by PCR with the following primer sets FW_aggggaccgggttcgagaaggggctctcatcactaggccgctagctctagagcgctcgaatt, RV_ttcgggtcccaattgccgtcgtctggcggtcttccaccttctcttcttggggtcatgttgcc. The UniSAM cassette was obtained also from the UniSAM plasmid via digestion with BsrGI and BsiWI. Finally, the last fragment consisting of another adapter for the Gibson was custom synthesised and contained overlapping sequences flanking a chicken b-globin insulator that we inserted to prevent cross-activation of the EF1α-promoter and the TRE-GV promoter driving the iSAM. Correct assembly was verified by Sanger sequencing.

iSAM cell lines derivation

The iSAM plasmid was used together with ZNFs specific for the AAVS1 locus to mediate specific integration in SFCi55 human iPSCs line^{14,20}. Briefly, 10 µg of AAVS1-iSAM with 2.5 µg of each ZNFs, left and right, using Xfect (Takara) according to the manufacturer protocol. Cells were selected using Neomycin. Single clones were picked, amplified, and initially screened by mCherry expression upon DOX addition. Clones that expressed the fluorescent tag were screened for specific integration using PCR followed by Sanger sequencing for the correctly integrated clones. 100 ng of genomic DNA was amplified using the EmeraldAmp® MAX HS Takara and specific primer sets (Table 1). For the specific AAVS1 integration site, Sigma_AAVS1 - CGG AAC TCT GCC CTC TAA CG and NeoR -GAT ATT GCT GAA GAG CTT GGC GG were used with the PCR conditions of 95 °C for 7 min, 32 cycles of 95 °C for 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 the Wild type locus screening, Sigma_AAVS1 - CGG AAC TCT GCC CTC

TAA CG and AAVS1_EXT3_RV – ACA CCC AGA CCT GAC CCA AA were used with the PCR cycling 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, with the final elongation step at 72 °C for 6 min.

Capture sequencing addition to the gRNA backbone

The Capture sequence 2 was added to the gRNA_Puro_Backbone (Addgene #73797) by PCR. Briefly, the capture sequence was added before the termination signal of the gRNA followed by a BamHI site using the following PCR primers:

gRNA_FW gagggcctatttcccatgattcct,

gRNA_Cap_RV aaaaaaggatccaaaaaCCTTAGCCGCTAATAGGTGAGCgcaccgactcggtgcc. The gRNA backbone was replaced from the original plasmid via NdeI and BamHI digestion, followed by ligation of the PCR produced following the same digestion. Correct integration of the insert was verified by Sanger sequencing.

Target genes identification

Candidate genes were identified by comparing *in vivo* hemogenic endothelium ¹⁵ (GSE135202) and *in vitro* iPSCs-derived endothelial cells that we previously generated ¹⁴ (E-MTAB-9295). Briefly, the two datasets were merged and normalised using the R package Seurat. Specific markers for hemogenic endothelium were identified and transcription factors were sorted based on their GO annotation. Within those genes we filtered those detected in more than 50% of the *in vivo* hemogenic endothelium and expressed in less than 25% of the *in vitro* derived endothelial cells. This pipeline identified 9 target genes *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*, *ZNF124*, *SOX6*, *ZNF33A*, *NFAT5*, *TFDP*.

AGM gRNAs library preparation

sgRNA design was performed by selecting the top candidates for on-target and off-target score. Between 5 and 7 guides per gene were designed for *RUNX1T1*, *NR4A1*, *GATA2*, *SMAD7*, *ZNF124*, *SOX6*, *ZNF33A*, *NFAT5*, *TFDP2* using the CRISPRpick tool from the Broad Institute (<https://portals.broadinstitute.org/gppx/crispick/public>) (Table S1). All the guide variants were Golden Gate cloned with the gRNA 2.1 backbone according to the established protocol ⁴⁹. The 49 plasmids were pooled together in an equimolar ratio and the library prep was subsequently used to produce lentiviral particles with a second-generation production system. Briefly, the psPAX2 packaging plasmid, pMD2.G envelope, and the AGM vector library were co-transfected using polyethyleneimine (PEI) (Polysciences, Warrington, PA, USA) as previously detailed ⁵⁰. Lentiviral particles-containing supernatants were harvested 48–72 h post-transfection, concentrated by ultracentrifugation and titered in hiPSCs cells.

iSAM_AGM and iSAM_NT cell line derivation

The selected iSAM clone (3.13 internal coding) was infected with viral particles containing either the AGM library or the non targeting gRNA (NT) at a MOI of 10. The iSAM cells were plated the afternoon before at 7×10^6 cells into a T125 in the presence of 10 μ M Rock Inhibitor (Merk) which was maintained until the day following the infection. Cells were infected in presence of 8 μ g/ml of Polybrene (Merk). Puromycin selection was initiated 36 hours post-infection and maintained during their culture until the beginning of the differentiation. Both lines were tested for integration of the gRNAs. Briefly, 100 ng of isolated gDNA was amplified using the PrimeSTAR MAX PCR mix (Takara) using the primers gRNA_screening_FW and gRNA_screening_RV (Table1). Purified amplicons were subjected to Sanger sequencing.

Single Cell RNA sequencing

For the iSAM_AGM and iSAM_NT single cell RNA sequencing experiment, embryoid bodies obtained from day 10 of differentiation were dissociated using Accutase (Life Technologies') at 37°C for 30'. For the IGFBP2 experiment, day 10 and day 13 cells from the Laminin EHT culture were detached from the adherent layer using Accutase (Life Technologies') at 37°C for 5'. From both experiments, cells were centrifuged and resuspended in CD34-Pe staining solution at a density of 10^7 /ml. CD34+/live/single cells were FAC-sorted in PBS + 0.1 % BSA. Cell viability was also confirmed by Trypan blue stain for an accurate count. Around 15000 cells per sample were loaded into the 10X Chromium Controller, and single cell libraries were obtained using the Chromium single cell 3' Reagent Kits v3 (10XGenomics) according to manufacturer protocol. The four libraries were indexed using SI PCR primers with different i7 indexes to allow for demultiplexing of the sequencing data. RNA concentration was obtained using Qubit RNA HS (Thermo-Fisher). Quality of the obtained libraries were verified using LabChip GX (PerkinElmer). Libraries were sequenced using NextSeq 2000 technology (Illumina) at 50.000 reads/cell. Data were aligned to GRCh38 using the Cell Ranger dedicated pipeline (10XGenomics). Data filtering, dimension reduction, clustering analysis, differentially expressed genes and cell cycle analysis were obtained using Seurat R package (version 4.1.0)⁵¹. Cells were subjected to QC and filtering using both the number of genes (1000-7500) and the percentage of mitochondrial genes detected (1-15%), resulting in 9025, 10942, 9468, 13073 cells respectively for the samples iSAM_NT, iSAM_NT+DOX, iSAM_AGM and iSAM_AGM+DOX. Pseudotemporal ordering was performed using Monocle 3 R package⁵². KEGG pathways was performed using ShinyGo⁵³. The gRNAs' expression matrix was used to select cells in which the expression of the gRNAs was detected. Briefly, for the libraries derived from the iSAM_NT control containing only the non-targeting gRNA, the filter was set for cells expressing one gRNA, while for the iSAM_AGM libraries was set to more than one. The code is available on GitHub at <https://github.com/afidanza/CRISPRa>, the raw data have been submitted to Array Express (E-MTAB-12748), and the browsable processed data will be made available at the time of publication on our website containing previous sequencing data at <https://lab.antonellafidanza.com>.

ATP production analysis

At day 8 of differentiation, 20.000 CD34+ were plated for each precoated well of the Seahorse XFp mini Cell Culture Plates (Agilent) precoated with rhLaminin-521 (Thermo-Fisher). On day 13, the cells' ATP production rate was analysed with the Seahorse XF Real-Time ATP Rate Assay kit (Agilent) according to manufacturer protocol, following confirmation of comparable cell densities across replicates and conditions. Briefly, the media was changed prior to the assay for the XF DMEM medium, pH 7.4, supplemented with 10 mM Seahorse XF glucose, 1mM Seahorse XF pyruvate and 2mM Seahorse XF Glutamine (Agilent) and incubated for 45-60 minutes in a non-CO₂ incubator. Oligomycin and Rotenone/AA solutions were prepared and added in the cartridge and finally loaded together with the cells in the Seahorse XF Mini Analyzer using the dedicated software. Data were collected at the end of the run, and the values of ATP production were calculated according to the ATP Production Rate Calculation provided by Agilent. Briefly, OCR ATP was calculated as $OCR_{\text{basal}} - OCR_{\text{oligo}}$ averaged across the three reads for each well. Mitochondrial ATP was calculated as OCR_{ATP} multiplied by the molecular oxygen consumption rate of 2 and by the P/O value of 2.75. For the glycolytic ATP production, we calculated the MitoPER as the $OCR_{\text{basal}} - OCR_{\text{rot}}$ times the CO₂ contribution Factor of 0.5. The PER was then calculated as the ECAR times the Buffer Factor of 2.6, the volume of reaction of 2.28 µl and the Kvol value of 1.1. The GlycoATP Production Rate was obtained by removing the MitoPER from the PER, and the TotalATP was calculated by adding the GlycoATP and the MytoATP. Mito/Glyco ratio was obtained by dividing their ATP production value.

Cell cycle analysis

DAPI staining and flow cytometry analysis were performed to verify the proliferation rate of the cells. Briefly, cells were collected from the supernatant by aspiration using a pastette, washed in PBS + 0.5% BSA + 2mM EDTA, counted and stained at 10^5 cells for a single tube. Cells were stained 1:1 v:v with a solution of 1% NP40 and 5 µg/ml DAPI for 2 minutes and acquired using the DIVA software (BD), and analysed using FlowJo version 10.4.2.

Author contribution

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.

Acknowledgements

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)

References

1. Medvinsky A., Dzierzak E (1996) **Definitive hematopoiesis is autonomously initiated by the AGM region** *Cell* **86**:897–906
2. Palis J., Robertson S., Kennedy M., Wall C., Keller G (1999) **Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse** *Development* **126**:5073–84
3. Patel S. H., et al. (2022) **Lifelong multilineage contribution by embryonic-born blood progenitors** *Nature* **606**:747–753
4. Böiers C., et al. (2013) **Lymphomyeloid Contribution of an Immune-Restricted Progenitor Emerging Prior to Definitive Hematopoietic Stem Cells** *Cell Stem Cell* **13**:535–548
5. Hoeffel G., et al. (2015) **C-Myb(+) erythro-myeloid progenitor-derived fetal monocytes give rise to adult tissue-resident macrophages** *Immunity* **42**:665–78
6. Jaffredo T., Gautier R., Eichmann A., Dieterlen-Lièvre F (1998) **Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny** *Development* **125**:4575–83
7. Zovein A. C., et al. (2008) **Fate Tracing Reveals the Endothelial Origin of Hematopoietic Stem Cells** *Cell Stem Cell* **3**:625–636
8. Bertrand J. Y., et al. (2010) **Haematopoietic stem cells derive directly from aortic endothelium during development** *Nature* **464**:108–111
9. Boisset J. C., et al. (2010) **In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium** *Nature* **464**:116–120
10. Ottersbach K. (2019) **Endothelial-to-haematopoietic transition: an update on the process of making blood** *Biochem Soc Trans* **47**:591–601 <https://doi.org/10.1042/BST20180320>
11. Swiers G., et al. (2013) **Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level** *Nature Communications* **4**:1–10
12. Kissa K., Herbomel P (2010) **Blood stem cells emerge from aortic endothelium by a novel type of cell transition** *Nature* **464**:112–115
13. Eilken H. M., Nishikawa S. I., Schroeder T (2009) **Continuous single-cell imaging of blood generation from haemogenic endothelium** *Nature* **457**:896–900
14. Fidanza A., et al. (2020) **Single cell analyses and machine learning define hematopoietic progenitor and HSC-like cells derived from human PSCs** *Blood* **136**:2893–2904
15. Zeng Y., et al. (2019) **Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing** *Cell Res* **29**:881–894
16. Ng E. S., et al. (2016) **Differentiation of human embryonic stem cells to HOXA+ hemogenic vasculature that resembles the aorta-gonad-mesonephros** *Nat Biotechnol* **34**:1168–1179

17. Calvanese V., et al. (2022) **Mapping human haematopoietic stem cells from haemogenic endothelium to birth** *Nature* **604**:534–540
18. Sturgeon C. M., Ditadi A., Awong G., Kennedy M., Keller G (2014) **Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells** *Nat Biotechnol* **32**:554–561
19. Fidanza A., et al. (2017) **An all-in-one UniSam vector system for efficient gene activation** *Sci Rep* **7**
20. Yang C.-T., et al. (2017) **Activation of KLF1 Enhances the Differentiation and Maturation of Red Blood Cells from Human Pluripotent Stem Cells** *Stem Cells* **35**:886–897
21. Lopez-Yrigoyen M., et al. (2018) **A human iPSC line capable of differentiating into functional macrophages expressing ZsGreen: A tool for the study and in vivo tracking of therapeutic cells** *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**
22. Lopez-Yrigoyen M., et al. (2019) **Genetic programming of macrophages generates an in vitro model for the human erythroid island niche** *Nat Commun* **10**
23. Lopez-Yrigoyen M., et al. (2018) **A human iPSC line capable of differentiating into functional macrophages expressing ZsGreen: A tool for the study and in vivo tracking of therapeutic cells** *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**
24. May A., et al. (2023) **Modelling the erythroblastic island niche of dyserythropoietic anaemia type IV patients using induced pluripotent stem cells** *Front Cell Dev Biol* **11**
25. Kang S. J., Park Y. Il, So B., Kang H. G. (2014) **Sodium butyrate efficiently converts fully reprogrammed induced pluripotent stem cells from mouse partially reprogrammed cells** *Cell Reprogram* **16**:345–354
26. Zhang Z., Xiang D., Wu W. S (2014) **Sodium Butyrate Facilitates Reprogramming by Derepressing OCT4 Transactivity at the Promoter of Embryonic Stem Cell-Specific miR-302/367 Cluster** *Cell Reprogram* **16**
27. Thambyrajah R., et al. (2018) **HDAC1 and HDAC2 Modulate TGF- β Signaling during Endothelial-to-Hematopoietic Transition** *Stem Cell Reports* **10**:1369–1383
28. Replogle J. M., et al. (2020) **Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing** *Nat Biotechnol* <https://doi.org/10.1038/s41587-020-0470-y>
29. Ng E. S., et al. (2016) **Differentiation of human embryonic stem cells to HOXA⁺ hemogenic vasculature that resembles the aorta-gonad-mesonephros** *Nat Biotechnol* **34**:1168–1179
30. Sturgeon C. M., Ditadi A., Awong G., Kennedy M., Keller G (2014) **Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells** *Nat Biotechnol* **32**:554–561
31. Ditadi A., et al. (2015) **Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages** *Nat Cell Biol* **17**:580–591
32. Ventura T., Egan E. J., Romanò N., Fidanza A (2023) **In Vitro Model of Fetal Human Vessel On-chip to Study Developmental Mechanobiology** *J Vis Exp* **2023**

33. Fidanza A., et al. (2020) **Single cell analyses and machine learning define hematopoietic progenitor and HSC-like cells derived from human PSCs** *Blood* **136**:2893–2904
34. Ventura T., Egan E. J., Romanò N., Fidanza A (2023) **In Vitro Model of Fetal Human Vessel On-chip to Study Developmental Mechanobiology** *J Vis Exp* **2023**
35. Azzoni E., et al. (2021) **The onset of circulation triggers a metabolic switch required for endothelial to hematopoietic transition** *Cell Rep* **37**
36. Oburoglu L., et al. (2022) **Pyruvate metabolism guides definitive lineage specification during hematopoietic emergence** *EMBO Rep* **23**
37. Castaño J., et al. (2019) **GATA2 Promotes Hematopoietic Development and Represses Cardiac Differentiation of Human Mesoderm** *Stem Cell Reports* **13**:515–529
38. de Pater E., et al. (2013) **Gata2 is required for HSC generation and survival** *Journal of Experimental Medicine* **210**:2843–2850
39. Ling K. W., et al. (2004) **GATA-2 Plays Two Functionally Distinct Roles during the Ontogeny of Hematopoietic Stem Cells** *J Exp Med* **200**
40. McGarvey A. C., et al. (2017) **A molecular roadmap of the AGM region reveals BMPER as a novel regulator of HSC maturation** *J Exp Med* **214**:3731–3751
41. McGrath K. E., et al. (2011) **A transient definitive erythroid lineage with unique regulation of the β -globin locus in the mammalian embryo** *Blood* **117**
42. Rejeski K., Duque-Afonso J., Lübbert M (2021) **AML1/ETO and its function as a regulator of gene transcription via epigenetic mechanisms** *Oncogene* **40**:5665–5676
43. Huynh H. D., et al. (2011) **IGF binding protein 2 supports the survival and cycling of hematopoietic stem cells** *Blood* **118**
44. Crosse E. I., et al. (2020) **Multi-layered Spatial Transcriptomics Identify Secretory Factors Promoting Human Hematopoietic Stem Cell Development** *Cell Stem Cell* **27**
45. Hadland B., et al. (2022) **Engineering a niche supporting hematopoietic stem cell development using integrated single-cell transcriptomics** *Nature Communications* **13**:1–17
46. Hadland B. K., et al. (2015) **Endothelium and NOTCH specify and amplify aorta-gonad-mesonephros-derived hematopoietic stem cells** *J Clin Invest* **125**:2032–2045
47. Sandler V. M., et al. (2014) **Reprogramming human endothelial cells to haematopoietic cells requires vascular induction** *Nature* **511**:312–318
48. Azzoni E., et al. (2021) **The onset of circulation triggers a metabolic switch required for endothelial to hematopoietic transition** *Cell Rep* **37**
49. Konermann S., et al. (2015) **Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex** *Nature* **517**:583–588
50. Petazzi P., et al. (2020) **Robustness of Catalytically Dead Cas9 Activators in Human Pluripotent and Mesenchymal Stem Cells** *Mol Ther Nucleic Acids* **20**:196–204

51. Hao Y., et al. (2021) **Integrated analysis of multimodal single-cell data** *Cell* **184**:3573–3587
52. Cao J., et al. (2019) **The single-cell transcriptional landscape of mammalian organogenesis** *Nature* **566**:496–502
53. Ge S. X., Jung D., Yao R (2020) **ShinyGO: a graphical gene-set enrichment tool for animals and plants** *Bioinformatics* **36**:2628–2629

Author information

Paolo Petazzi

Josep Carreras Leukemia Research Institute, Barcelona, Spain

Telma Ventura

Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK

Francesca Paola Luongo

Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK

Heather McClafferty

Centre for Discovery Brain Sciences, Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK

Alisha May

Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK

Helen Alice Taylor

Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK

Micheal J Shipston

Centre for Discovery Brain Sciences, Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK, Zhejiang University-University of Edinburgh Joint Institute, Zhejiang University School of Medicine, Zhejiang University, Haining, China

Nicola Romanò

Centre for Discovery Brain Sciences, Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK, Zhejiang University-University of Edinburgh Joint Institute, Zhejiang University School of Medicine, Zhejiang University, Haining, China

Lesley M Forrester

Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK
ORCID iD: [0000-0002-9717-8299](https://orcid.org/0000-0002-9717-8299)

Pablo Menéndez

Josep Carreras Leukemia Research Institute, Barcelona, Spain, Red Española de Terapias Avanzadas (TERAV)-Instituto de Salud Carlos III (ISCIII), Madrid, Spain, CIBER-ONC, ISCIII, Barcelona, Spain, Institució Catalana de Recerca i Estuis Avançats (ICREA), Barcelona, Spain, Department of Biomedicine, School of Medicine, University of Barcelona, Barcelona, Spain

Antonella Fidanza

Centre for Regenerative Medicine, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, UK, Edinburgh Medical School, Biomedical Sciences, University of Edinburgh, Edinburgh, UK
ORCID iD: [0000-0002-1266-6454](https://orcid.org/0000-0002-1266-6454)

For correspondence: afidanza@ed.ac.uk

Editors

Reviewing Editor

K VijayRaghavan

National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India

Senior Editor

K VijayRaghavan

National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore, India

Reviewer #1 (Public Review):**Summary:**

The work from Petazzi et al. aimed at identifying novel factors supporting the differentiation of human hematopoietic progenitors from induced-pluripotent stem cells (iPSCs). The authors developed an inducible CRISPR-mediated activation strategy (iCRISPRa) to test the impact of newly identified candidate factors on the generation of hematopoietic progenitors in vitro. They first compared previously published transcriptomic data of iPSC-derived hematopoietic populations with cells isolated ex vivo from the aorta-gonad-mesonephros (AGM) region of the human embryo and they identified 9 transcription factors expressed in the aortic hemogenic endothelium that were poorly expressed in the in vitro differentiated cells. They then tested the activation of these candidate factors in an iPSC-based culture system supporting the differentiation of hematopoietic progenitors in vitro. They found that the IGF binding protein 2 (IGFBP2) was the most upregulated gene in arterial endothelium after activation and they demonstrated that IGFBP2 promotes the generation of functional hematopoietic progenitors in vitro.

Strengths:

The authors developed a very useful doxycycline-inducible system to activate the expression of specific candidate genes in human iPSC. This approach allows us to simultaneously test the impact of 9 different transcription factors on in vitro differentiation of hematopoietic cells, and the system appears to be very versatile and applicable to a broad variety of studies. Using this approach, the authors exhaustively demonstrated the role of IGFBP2 in promoting the generation of functional hematopoietic progenitors in vitro from iPSCs.

Weaknesses:

The authors performed a very thorough characterization of the system in proof-of-principle experiments activating a single transcription factor. However, when 9 independent factors were used, it is not always clear whether the observed results were the consequence of the simultaneous activation of all 9 TF or just a subset of them.

<https://doi.org/10.7554/eLife.94884.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public Review):

Summary:

The work from Petazzi et al. aimed at identifying novel factors supporting the differentiation of human hematopoietic progenitors from induced pluripotent stem cells (iPSCs). The authors developed an inducible CRISPR-mediated activation strategy (iCRISPRa) to test the impact of newly identified candidate factors on the generation of hematopoietic progenitors in vitro. They first compared previously published transcriptomic data of iPSCderived hemato-endothelial populations with cells isolated ex vivo from the aorta-gonadmesonephros (AGM) region of the human embryo and they identified 9 transcription factors expressed in the aortic hemogenic endothelium that were poorly expressed in the in vitro differentiated cells. They then tested the activation of these candidate factors in an iPSCbased culture system supporting the differentiation of hematopoietic progenitors in vitro. They found that the IGF binding protein 2 (IGFBP2) was the most upregulated gene in arterial endothelium after activation and they demonstrated that IGFBP2 promotes the generation of functional hematopoietic progenitors in vitro.

Strengths:

The authors developed an extremely useful doxycycline-inducible system to activate the expression of specific candidate genes in human iPSC. This approach allows us to simultaneously test the impact of 9 different transcription factors on in vitro differentiation of hematopoietic cells, and the system appears to be very versatile and applicable to a broad variety of studies.

The system was extensively validated for the expression of 1 transcription factor (RUNX1) in both HeLa cells and human iPSC, and a detailed characterization of this test experiment was provided.

The authors exhaustively demonstrated the role of IGFBP2 in promoting the generation of functional hematopoietic progenitors in vitro from iPSCs. Even though the use of IGFBP2interacting proteins IGF1 and IGF2 have been previously reported in human iPSC-derived hematopoietic differentiation in vitro (Ditadi and Sturgeon, Methods 2016; Ng et al., Nature Biotechnology 2016), and IGFBP-2 itself has been shown to promote adult HSC expansion ex vivo (Zhang et al., Blood 2008), its role on supporting in vitro hematopoiesis was demonstrated here for the first time.

Weaknesses:

Although the authors performed a very thorough characterization of the system in proof-of-principle experiments activating a single transcription factor, the data provided when 9

independent factors were used is not sufficient to fully validate the experimental strategy. Indeed, in the current version of the manuscript, it is not clear whether the results presented in both the scRNAseq analysis and the functional assays are the consequence of the simultaneous activation of all 9 TF or just a subset of them. This is essential to establish whether all the proposed factors play a role during embryonic hematopoiesis, and a more complete analysis of the scRNAseq dataset could help clarify this aspect.

Similarly, the data presented in the manuscript are not sufficient to clarify at what stage of the endothelial-to-hematopoietic transition (EHT) the TF activation has an impact. Indeed, even though the overall increase of functional hematopoietic progenitors is fully demonstrated, the assays proposed in the manuscript do not clarify whether this is due to a specific effect at the endothelial level or to an increased proliferation rate of the generated hematopoietic progenitors. Similar conclusions can be applied to the functional validation of IGFBP2 in vitro.

The overall conclusions are sometimes vague and not always supported by the data. For instance, the authors state that the CRISPR activation strategy resulted in transcriptional remodeling and a steer in cell identity, but they do not specify which cell types are involved and at what level of the EHT process this is happening. In the discussion, the authors also claim that they provided evidence to support that RUNX1T1 could regulate IGFBP2 expression. However, this is exclusively based on the enrichment of RUNX1T1 gRNA in cells expressing higher levels of IGFBP2 and it does not demonstrate any direct or indirect association of the two factors.

We thank the reviewer for the positive comments about the importance of our work and have now addressed the points raised as weaknesses by performing additional analysis and experiments, adding a new schematic of the mechanism, and rewording our claims.

We have clarified the different effects mediated by the activation and the IGFBP2 addition in a summary section at the end of the results and added Figure 6, showing this in visual form. We have also clearly stated the limitations related to the correlation between RUNX1T1 and IGFBP2 in the discussion and toned down our claims regarding this throughout the entire paper. We have also reworded the text to clarify the specific cell types identified in the sequencing data that we refer to.

Reviewer #2 (Public Review):

To enable robust production of hematopoietic progenitors in-vitro, Petazzi et al examined the role of transcription factors in the arterial hemogenic endothelium. They use IGFBP2 as a candidate gene to increase the directed differentiation of iPSCs into hematopoietic progenitors. They have established a novel induced-CRISPR mediated activation strategy to drive the expression of multiple endogenous transcription factors and show enhanced production of hematopoietic progenitors through expansion of the arterial endothelial cells. Further, upregulation of IGFBP2 in the arterial cells facilitates the metabolic switch from glycolysis to oxidative phosphorylation, inducing hematopoietic differentiation. While the overall study and resources generated are good, assertions in the manuscript are not entirely supported by the experimental data and some claims need further experimental validation.

We thank the reviewer for the positive comments, and we have provided new data and analysis to make sure that all our assertions are clearly supported and also reworded those where limitations were identified by the reviewers.

Recommendations for the authors:

Reviewing Editor (Recommendations For The Authors):

The assessment could change from "incomplete" to "solid" if the authors: i) improve data analysis (for both scRNAseq and functional assays) by providing additional information that could strengthen their conclusions, as suggested in the specific comments by both reviewers; ii) either provide new functional evidence supporting their mechanistic conclusion or alternatively tone down the claims that are not fully supported by data and acknowledge the limitations raised by reviewers in the discussion; (iii) the issue of paracrine signaling to expand only hematopoietic progenitors needs to be addressed.

We have now improved the data analysis and provided additional functional tests to strengthen our conclusions and toned down those that were identified by the reviewers as not supported enough and included a discussion on these limitations. We have also reworded the section about the paracrine signaling throughout the paper.

Reviewer #1 (Recommendations For The Authors):

Figure 1 contains exclusively published data. It might be more appropriate to use it as a supplementary figure or as part of a more exhaustive figure (maybe combining Figures 1 and 2 together?).

Figure 1 contained novel bioinformatic analyses that represent the base of our research and it has a different content and focus to figure 2, which is already a large figure. We therefore believe it is better to keep it as a separate figure, containing a new panel now too.

It seems there is an issue with Figure S3 labelling:

- *In line 112, Figure S2A-B does not display genomic PCR and sequencing results;*
- *In line 123, Figure S3D-E does not show viability and proliferation data;*
- *In line 127, Figure S3G does not show mCherry expression in response to DOX;*

We apologies for the confusion with the numbers, we have now correctly labelled the figures.

It would be more informative to include gates and frequency on flow cytometry plots in Figure S3, to be able to evaluate the extent of the reduction in mCherry expression.

We have now included the gating and frequency of mCherry-expressing cells in Supplementary Figure 3D.

It is not clear from the text and figures whether the SB treatment was maintained throughout the hematopoietic differentiation protocol (line 122):

- *If so, it would be important to confirm that HDAC treatment does not affect EHT cultures*
- *If not, can the authors provide some evidence that transgene silencing is not occurring during hematopoietic differentiation?*

We have clarified that we decided to treat the cells with SB exclusively in maintenance conditions because HDACs have been shown to be essential for the EHT (lines 138-142). We have now also included additional data showing the high expression of the mCherry tag reporting the iSAM expression on day 8 (Supplementary Figure 4F).

Can the authors provide a simple diagram summarizing the experimental strategy for each differentiation experiment in the respective supplementary figure? For instance, at what stage of the protocol was DOX added in Figure 3? Or at what stage IGFBP2 was added in Figure 5? It would be a very useful addition to the interpretation of the results.

We have now included three schematics for all the experiments in the manuscript in supplementary figure 4 A-C.

In Figure 3, the authors should provide more detailed information about the data filtering of the scRNAseq experiment, and more specifically:

- *How many cells were included in the analysis for each library after QC and filtering?*
- *How "cells in which the gRNAs expression was detected" were selected? Do they include only cells showing expression of gRNAs for all 9 TF?*

This information is now included in the method section lines 773-781; the detailed code is available on the GitHub link provided in the same section. We have filtered the cells expressing one gRNA for the non-targeting gRNA (iSAM_NT) control and more than one for the iSAM_AGM sample.

In Figure 3A, it is not clear whether the expression of the 9 factors is consistently detected in all cells or just a subset of them, and the heatmap in Figure 3A does not provide this information. It would be more accurate to provide expression on a per-cell basis, for instance, as a violin plot displaying single dots representing each cell.

We have now included this violin plot in Supplementary Figure 4G as requested. However, this visualisation is difficult to interpret because some of the target genes' expression seems variable in both experimental and control conditions. We had envisaged that this could have been the case and so this is why we had included the three different controls. For this reason we chose to show the normalised expression which takes all the different variables into account (Figure 3A).

In Figure 3B-C, it seems that clusters EHT1 and EHT2 do not express endothelial markers anymore. Are these fully differentiated hematopoietic cells rather than cells undergoing EHT? In general, it would be quite important to provide evidence of expressed marker genes characterizing each cluster (eg. heatmap summarizing top DEG in the supplementary figure?).

We have now provided a spreadsheet containing the clusters' markers that we used in

Supplementary Table 1) a heatmap in Figure 3E. Furthermore, we have now edited Figure 3C to include Pan Endothelial markers (PECAM1 and CDH5). These data show that the EHT1 and EHT2 cluster both express endothelial markers but are progressively downregulated as expected during endothelial to hematopoietic transition. We have also included and discussed this in the manuscript lines 192-195 and a schematic for the mechanism in Figure 6.

In Figure 3E, displaying the proportion of clusters within each sample/library would be a more accurate way of comparing the cell types present in each library (removing potential bias introduced by loading different numbers of cells in each sample).

We have now included the requested data in Supplementary Figure 4I and it confirms again the expansion of arterial cells in the activated cells.

In Figure 3G, by plating 20,000 total CD34+, the assay does not account for potential differences in sample composition. It is then hard to discriminate between the increased number of progenitors in the input or an enhanced ability of HE to undergo EHT. This is an important aspect to consider to precisely identify at what level the activation of the 9 factors is acting. A proper quantification of flow cytometry data summarizing the % of progenitors, arterial cells, etc. would be useful to interpret these results.

Lines 204-205 reworded. We are very much aware of the fact that the CD34+ cell population consists of a range of cells across the EHT process and this is precisely why we carried out this single cell sequencing analyses. We purposely tested the effect of the observed changes in composition by colony assays

In Figure 3G, it seems that NT cells w/o DOX have very little CFU potential (if any). Can the authors provide an explanation for this?

We think that the limited CFU potential is due to the extensive genetic manipulation and selection that the cells underwent for the derivation of all the iSAM lines but this did not impede us from observing an effect of gene activation on CFU numbers. This is one of the primary reasons that we then validated our overall findings using the parental iPSC line in control condition and with the addition of IGFBP2. We show that the parental iPSC line gives rise to hematopoietic progenitor, both immunophenotypically (Figure 4D) and functionally, at expected levels (Figure 4B left column).

Figure 4A shows an upregulation of IGFBP2 in arterial cells as a result of TF activation. However, from the data presented here, it is not possible to evaluate whether this is specific to the arterial cluster, or it is a common effect shared by all cell types regardless of their identity.

Data has now been included in Supplementary Figure 4H, which shows that all the cells show an increase in IGFBP2, but arterial cells show the highest increase. We have now edited the text to reflect this, in lines 228-230.

In Figure 5A-B only a minority of arterial cells express RUNX1 in response to IGFBP2 treatment. Is this sufficient to explain the very significant increase in the generation of functional hematopoietic progenitors described in Figure 4? Quantification and statistical analysis of RUNX1 upregulation would strengthen this conclusion.

We have now provided the statistical analysis showing significant upregulation of RUNX1 upon IGFBP2 addition. The p values are now provided in the figure 5 legend.

In Figure 5 the authors conclude that IGFBP2 remodels the metabolic profile of endothelial cells. However, it is not clear which cell types and clusters were included in the analysis of Figure 5C-G. Is the switch from Glycolysis to Oxidative Phosphorylation specific to endothelial cells? Or it is a more general effect on the entire culture, including hematopoietic cells?

We based this conclusion on the fact that the single-cell RNAseq allows to verify that the metabolic differences are obtained in the endothelial cells. Given that we sorted the adherent cells, the majority of these are endothelial cells as shown in Figure 5A. The Seahorse pipeline includes a number of washing steps resulting in the analyses being performed on the adherent compartment which we know consists primarily of endothelial cells. We cannot exclude some contamination from non-endothelial cells but we highlight to this reviewer that the initial observation of the metabolic changes was identified in endothelial cells in the

single cell sequencing data. Taken together, we believe that this implies that metabolic changes are specific to this population. We have clarified this in the line 317.

In the discussion, the authors conclude that they "provide evidence to support the hypothesis that RUNX1T1 could regulate IGFBP2 expression". To further support this conclusion, the authors could provide a correlation analysis of the expression of the two genes in the cell type of interest.

Following the observation of the IGFBP2 high expression across clusters, we have now reworded this sentence in lines 382-385 We have tried to perform the correlation analysis but we believe this not to be appropriate due to the detection level of the gRNA, we have now included this as a limitation point in the discussion lines 416-427, and also toned down the conclusion we did draw about RUNX1T1 throughout the whole manuscript.

As mentioned by the authors, IGFBP2 binds IGF1 and IGF2 modulating their function. Both IGF1 (<http://dx.doi.org/10.1016/j.ymeth.2015.10.001>) and IGF2 (doi:10.1038/nbt.3702) have been used in iPSC differentiation into definitive hematopoietic cells. It would be relevant to discuss/reference this in the discussion.

We have now included the suggested reference in the section where we discuss the role of IGFBP2 in binding IGF1 and IGF2.

Reviewer #2 (Recommendations For The Authors):

(1) Figure 1 compares the transcriptome of human AGM and in-vitro derived hemogenic endothelial cells (HECs). It is not clear why only the genes downregulated in the latter were chosen. Are there any significantly upregulated genes, knockdown/knockout which could also serve a similar purpose? Single-cell transcriptome database analysis is very preliminary. A detailed panel with differences in cluster properties of HECs between the two systems should be provided. A heatmap of all differentially expressed genes between the two samples must be generated, along with a logical explanation for choosing the given set of genes.

We have now included another panel in figure 1 to better clarify the logic behind the strategy used to identify our target genes (Figure 1A).

(2) Figure 2 - a panel describing the workflow of gRNA design and targeting for the 9 candidate genes, along with lentiviral packaging and transduction would make it easier to follow.

We have now included three schematics for all the experiments in the manuscript in supplementary figure 4 A-C.

(3) Figure 3- to assess the effect of arterial cell expansion on the emergence of hematopoietic progenitors, CD34+ Dll4+ cells should be sorted for OP9 co-culture assay. Using only CD34+ cells does not answer the question raised. Also, the CFU assay performed does not fully support the claim of enhanced hematopoietic differentiation since only CFU-E and CFU-GM colonies are increased in Dox-treated samples, with no effect on other colony types. OP9 co-culture assay with these cells would be required to strengthen this claim.

We wanted to clarify that the effect on the methylcellulose coming from the activated cells was not limited to CFU-E, as the reviewer reported; instead, it also affected CFU-GM and CFU-M.

We have now performed additional experiments where we sorted the CD34+ compartment into DLL4- and DLL4+ in Supplementary Figure 5D-E, which we discussed in lines 250-258.

(4) In Figure 3F, there appears to be a lot of variation in the DLL4% fold change values for DOX treated iSAM_AGM sample, which weakens the claim of increased arterial expansion.

Can the authors explain the probable reason? It is suggested that the two other controls (iSAM_+DOX and iSAM_-DOX) should be included in this analysis. It is imperative to also show % populations rather than just fold change to gain confidence.

We agree that there is a lot of variability. That is because differentiation happens in 3D in embryoid bodies, which contain many different cell types that differentiate in different proportions across independent experiments. We have now included the raw data in Supplementary Figure 4 D, with additional statistical analysis to show the expansion of arterial cells including also the suggested additional controls.

(5) How does activation of these target genes cause increased arterialization? Is the emergence of non-HE populations suppressed? Or is it specific to the HE? The data on this should be clarified and also discussed. ANTO/Lesley text

We have provided additional data clarifying the connection between increased arterialisation and hemogenic potential. We showed that the activation induces increased arterialisation and that IGFBP2 acts by supporting the acquisition of hemogenic potential. We have discussed this in lines 326-348 and provided a new figure to explain this in detail (figure 6)

(6) Considering that IGFBP2 was chosen from the activated target gene(s) cluster, can the authors explain why the reduced CFU-M phenomenon observed in Figure 3G does not appear in the MethoCult assay for IGFBP2 treated cells (Figure 4B)?

The difference could be explained by the fact that in Figure 3G, the cells underwent activation of multiple genes, while in Figure 4B, they were only exposed to IGFBP2. Our results show that IGFBP2 could at least partially explain the phenotype that we see with the activation, but we believe that during the activation experiments, there might be other signals available that might not be induced by IGFBP2 alone. We have also added a summary section and a figure to clarify the different mechanisms of action of the gene activation and IGFBP2.

(7) Figure 4- while the experiments conducted support the role of IGFBP2 in increasing hematopoietic output, there is no experimental evidence to prove its function through paracrine signalling in HECs. The authors need to provide some evidence of how IGFBP2 supplementation specifically expands only the hematopoietic progenitors. Experimental strategies involving specifically targeting IGFBP2 in hemogenic/arterial endothelial cells are required to prove its cell type specific function. Additionally, assessing the in vivo functional potential of the hematopoietic cells generated in the presence of IGFBP2, by bone-marrow transplantation of CD34+ CD43+ cells, is essential.

The role of IGFBP2 in the context of HSC production and expansion was not the topic of our research, and we have not claimed that IGFBP2 affects the long-term repopulating capacity of HSPCs. Therefore, we believe that the requested experiments are not required to support the specific claims that we do make. We have now provided more experiments and bioinformatic analysis that support the role of IGFBP2 in inducing the progression of EHT from arterial cells to hemogenic endothelium, and to avoid misunderstandings, we have toned down our claims by editing the text regarding its paracrine effect s.

(8) Figure 4C-D -It is recommended to plot % populations along with fold change value. As this is a key finding, it is important to perform flow cytometry for additional hematopoietic markers- CD144, CD235a and CD41a to demonstrate whether this strategy can also expand erythroid-megakaryocyte progenitors. Telma

Figure 4C already shows the percentage values; we have now added the percentage for Figure 4D in SF5C. We have also performed additional analysis as requested and added the data obtained to Supplementary Figure 5D.

(9) In Figure 5, analysis showing the frequency of cells constituting different clusters, between untreated and IGFBP2-treated samples in the single-cell transcriptome analysis is essential. Additional experiments are required to validate the function of IGFBP2 through modulation of metabolic activity. Inhibition of oxidative phosphorylation in the IGFBP2-treated cells should reduce the hematopoietic output. Authors should consider doing these experiments to provide a stronger mechanistic insight into IGFBP2-mediated regulation of hematopoietic emergence.

We have now included the requested cluster composition in Supplementary Figure 5F. We decided not to include further tests on the metabolic profile of IGFBP2 as we already discussed in other papers that showed, using selective inhibitors, that the EHT coincides with a glycol to OxPhos switch.

(10) It is very striking to see that IGFBP2 supplementation changes the transcriptional profile of developing hematopoietic cells by increasing transcription of OXPHOS-related genes with concomitant reduction of glycolytic signatures, particularly at Day 13. However, the mitochondrial ATP rate measurements do not seem convincing. The bioenergetic profiles show that when mitochondrial inhibitors are added, both groups exhibit decreased OCR values and, on the other hand, higher ECAR. This indicates that both groups have the capability to utilize OXPHOS or glycolysis and may only differ in their basal respiration rates.

Differences in proliferation rate can cause basal respiration to change. There is no information on how the bioenergetic profile was normalized (cell no./protein amount). Given that IGFBP2 has been shown to increase proliferation, it is very likely that the cells treated with IGFBP2 proliferated faster and therefore have higher OCR. The data needs to be normalized appropriately to negate this possibility.

We have previously tested whether IGFBP2 causes an increase in proliferation by analysing the cell cycle of cells treated with it, as we initially thought this could be a mechanism of action. We have now provided the quantification of the cell cycle in the cells treated with IGFBP2, showing no effect was observed in cell cycle Supplementary Figure 4E. Following this analysis, we decided to plate the same number of cells and test their density under the microscope before running the experiment; each experiment was done in triplicate for each condition. We have now added this info to the method sections lines 806-813. We did not comment on the basal difference, which we agree might be due to several factors, but we only compared the difference in response to the inhibitors, which isn't affected by the basal level but exclusively by their D values. We have also included the formulas used to calculate the ATP production rate.

Overall, it appears that IGFBP2 does not seem to primarily cause metabolic changes, but simply accelerates the metabolic dependency on OXPHOS. Hence, the term 'metabolic remodelling' must be avoided unless IGFBP2 depletion/loss of function analysis is shown.

We thank the reviewer for suggesting how to interpret the data about the dependency on OXPHOS. We have now changed the conclusions and claims about the effect of IGFBP2. We have also included a cell cycle analysis of the hematopoietic cells derived upon IGFBP2 addition to show that they don't show differences in proliferation that could cause the increase in colony formation we observed. Regarding the assay, we have plated the same number of cells for each group to make sure we were comparing the same number of cells, which we also assessed in the microscope before the test, and we eliminated the suspension cells during the washes that preceded the measurement. The review is correct in indicating that there is a basal difference in the value of OCR and ECAR where the IGFBP2 is lower at the start and not higher, which would not conceal higher proliferation. Finally, the ATP production rate is calculated on the variation of OCR and ECAR upon the addition of inhibitors, which normalizes for the basal differences.

<https://doi.org/10.7554/eLife.94884.2.sa0>