FEBS Letters

FEBS PRESS science publishing by scientists

Induction of human hemogenesis in adult fibroblasts by defined factors and hematopoietic coculture

Michael G. Daniel[1,2,3], David Sachs[1,2,3], Jeffrey M. Bernitz[4], Yesai Fstkchyan[1,3], Katrina Rapp[1,2], Namita Satija[5,6], Kenneth Law[7], Foram Patel[5,6], Andreia M. Gomes[8], Huen-Suk Kim[1,2,3], Carlos-Filipe Pereira[9,10], Benjamin Chen[5,6], Ihor R. Lemischka[1,2,11] and Kateri A. Moore[1,2]

1 Department of Cell, Developmental and Regenerative Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
2 Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
3 The Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA
4 Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule Zurich, Basel, Switzerland
5 Division of Infectious Disease, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA
6 Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
7 Rocket Pharmaceuticals Ltd, New York, NY, USA
8 CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Cantanhede, Portugal
9 Division of Molecular Medicine and Gene Therapy, Lund University, Sweden
10 Wallenberg Center for Molecular Medicine, Lund University, Sweden
11 Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine, New York, NY, USA

Correspondence
K. A. Moore, Department of Cell, Developmental and Regenerative Biology, Icahn School of Medicine at Mount Sinai, Icahn (East) Building Floor 13 Room 20-E, 1425 Madison Ave, New York, NY 10029, USA
Tel: +1-212-659-8312
E-mail: kateri.moore@mssm.edu

(Received 25 May 2019, revised 20 September 2019, accepted 23 September 2019, available online 12 October 2019)

doi:10.1002/1873-3468.13621

Edited by Catherine Robin

Transcription factor (TF)-based reprogramming of somatic tissues holds great promise for regenerative medicine. Previously, we demonstrated that the TFs GATA2, GFI1B, and FOS convert mouse and human fibroblasts to hemogenic endothelial-like precursors that generate hematopoietic stem progenitor (HSPC)-like cells over time. This conversion is lacking in robustness both in yield and biological function. Herein, we show that inclusion of GFI1 to the reprogramming cocktail significantly expands the HSPC-like population. AFT024 coculture imparts functional potential to these cells and allows quantification of stem cell frequency. Altogether, we demonstrate an improved human hemogenic induction protocol that could provide a valuable human *in vitro* model of hematopoiesis for disease modeling and a platform for cell-based therapeutics.

Database
Gene expression data are available in the Gene Expression Omnibus (GEO) database under the accession number GSE130361.

Keywords: AFT024; coculture; hematopoietic stem cells; reprogramming; transcription factors

Abbreviations
3GF, GATA2, GFI1, GFI1B, FOS; ACE, angiotensin converting enzyme; AGM, aorta-gonad-mesonephros; CAFC, cobblestone area forming cell; CB, cord blood; CFU, colony forming unit; DOX, doxycycline; EHT, endothelial to hematopoietic transition; ELDA, extreme limiting dilution analysis; EPCR, endothelial protein C receptor; FGRS, FOSB, GFI1, RUNX1, SPI1; FL, fetal liver; GAG, glycosaminoglycan; GGF, GATA2, GFI1B, FOS; GO, gene ontology; HC, hydrocortisone; HDF, human dermal fibroblast; HE, hemogenic endothelium; HSC, hematopoietic stem cell; HSPC, hematopoietic stem progenitor cell; iPSC, induced pluripotent stem cell; LDA, limiting dilution analysis; LTC, long-term culture; LTC-IC, long-term culture-initiating cell; NSG, NOD-scid IL2Rg^{null}; PCA, principal component analysis; SD, standard deviation; SEM, standard error of the mean; tdT, tdTomato; TF, transcription factor; vWF, von Willebrand factor.

FEBS Letters **593** (2019) 3266–3287 © 2019 The Authors. *FEBS Letters* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

Hematopoietic stem cells (HSCs) are thought to generate all the cellular elements of the blood in a hierarchical manner [1], though recent work in this field suggests a more complicated process in the mouse system [2,3]. Multiple studies demonstrate an endothelial origin for multipotent HSCs, notably showing their emergence from a specific subset of cells called hemogenic endothelium (HE) through a process of cell budding termed the endothelial-to-hematopoietic transition (EHT) [4–8]. Due to their ability to repopulate the entire hematopoietic system upon transplantation in both mice and humans, HSCs represent the currently established standard for stem cell therapy. The source material required for these applications, however, remains in limited supply because HSCs notoriously die or differentiate *ex vivo* [9]. To this end, several studies exist that employ different methods to expand these cells *ex vivo* or generate them *de novo* [10–13]. This issue hinders their use for a multitude of *in vitro* applications, such as drug testing platforms and disease modeling systems. Allogeneic transplants involve an additional hurdle, in that they carry multiple risks of graft-versus-host disease and graft rejection due to poor HLA matching and a lack of ethnic diversity for sufficient matching material [14].

A paradigm shift in stem cell biology emerged once Yamanaka *et al*. demonstrated that overexpression of a defined set of transcription factors (TFs) could reprogram the differentiation of somatic cells to induced pluripotent stem cells (iPSCs) [15,16]. Ectopic TF overexpression to alter cell identity translated to the field of hematopoiesis, and multiple studies used different starting mouse or human cell populations, TF combinations, or culture conditions to obtain various types of *in vitro* derived blood products *de novo* [11,12]. Several of these studies focused on using either pluripotent [17–24] or somatic [25–30] cells with varying levels of success. Encouragingly, recent studies obtained transplantable cells using either human iPSCs [31] or human/mouse endothelial cells with a vascular niche coculture system [32,33].

Previously, we demonstrated that overexpression of TFs—Gata2, Gfi1B, and Fos (GGF)—was sufficient to induce a hemogenic program in mouse embryonic fibroblasts while the addition of Etv6 increased their efficiency. Together, these factors produce hematopoietic cells through a process that mimics developmental hematopoiesis. During reprogramming, the transduced cells appear to traverse from an endothelial precursor and then undergo an EHT to yield clonogenic progenitors [29]. The endothelial precursor cells bore a Prom1+, Sca1+, CD34+, and CD45− cell surface phenotype that we also found in cells of the aorta-gonad-mesonephros (AGM) and placenta. These cells gave rise to engraftable hematopoietic cells after maturation demonstrating that *in vitro* reprogramming can inform normal hematopoiesis [34]. Recent data demonstrate that GGF overexpression in iPSCs that are then used to form teratomas leads to an increased production of long-term HSCs [35]. We have transferred these findings to the human system and found that GGF reprogrammed fibroblasts appear to undergo the same type of process and can develop into short-term repopulating cells [30]. The yield and overall *in vitro*/*ex vivo* function of these human-derived cells, however, requires further improvement.

In this study, we sought to optimize our hemogenic induction process in adult human fibroblasts to produce greater yields of functional HSC-like cells. To this end, we identified an additional TF—GFI1—that expands the yield of functional hematopoietic progenitors when in concert with GGF reprogramming (the TF cocktail is now termed 3GF for GATA2, GFI1, GFI1B, and FOS). The cell surface phenotype of the derived cells correlates to that found on human HSCs. Gene expression profiles also show generation of hematopoietic cells through a hemogenic developmental process. AFT024 stromal cocultures permitted the derivation and quantification of functional HSC-like cells capable of forming colonies composed of various hematopoietic lineages. Additionally, cells reprogrammed with the new cocktail function at a much earlier time point as seen through our *in vitro* assays and short-term multilineage engraftment. Altogether, we demonstrate an enhanced human hemogenic induction strategy by using an additional TF as well as an *in vitro* maturation system.

Methods

Human dermal fibroblast, AFT024, and 293T cell culture

Human dermal fibroblasts (HDFs) used for all experiments were obtained from ScienCell (Carlsbad, CA, USA). Cells were plated in 10-cm tissue culture dishes in D10 media [Dulbecco's Modified Eagle Medium (DMEM); Thermo Fisher Scientific, Waltham, MA, USA] containing 10% Benchmark FBS (Gemini Bio-Products, West Sacramento, CA, USA), 1 mM L-glutamine, and penicillin/streptomycin (10 $\mu g \cdot mL^{-1}$; Thermo Fisher Scientific) at 37 °C. 293T cells for viral production were also cultured in standard D10 media at 37 °C. AFT024 cells used for long-term culture (LTC) and limiting dilution analysis (LDA) experiments were cultured in D10 media supplemented with 50 μM 2-ME at 32 °C for expansion. The day prior to the

experiments, AFT024 were mitotically inactivated *via* irradiation as previously described [36] and placed at 37 °C.

Molecular cloning, lentivirus production, and tdTomato-HDF generation

The coding regions of each candidate TF (Table S1A) were individually cloned into the pFUW-TetO vector where expression was controlled by the minimal CMV promoter and the tetracycline operator [37]. Lentiviral vectors carrying each of the chosen reprogramming factors were generated by calcium phosphate transfection into the 293T packaging cell line with a mixture of the viral plasmids of choice as well as the constructs that instruct viral packaging and the VSV-G protein (pMD2.G and psPAX2). For transgene activation, cells were cotransduced with lentiviral vectors containing the reverse tetracycline transactivator M2rtTA, which was controlled by the constitutively active human ubiquitin C promoter. Viral supernatants were collected 36, 48, and 60 h after 293T transfections, filtered (0.45 μm), and stored at −80 °C. Lentivirus carrying the pSin-tdTomato (tdT) vector (constitutively driven by the EF2 promoter) was generated as described above and used to transduce low-passage HDFs. The top 10% of tdT^+ cells were sorted and cultured to establish the tdT-HDF line in D10 media.

Viral transduction and reprogramming cell culture

Human dermal fibroblasts were transduced with a viral cocktail consisting of 33.33% D10 media, 33.33% viral supernatant containing M2rtTA, and the remaining 33.33% containing equal portions of each factor within the GGF or 3GF TF sets, to ensure equal multiplicities of infection of each individual viral particle, as well as 8 $\mu g \cdot mL^{-1}$ of polybrene. Additional factors were tested in combination with GGF in the same manner (Table S1B). Control transductions with mOrange in pFUW resulted in > 95% efficiency. HDFs on Day 1 were plated at a density of 1.5×10^5–3.0×10^5 on 0.1% gelatin-coated 10-cm dishes or six-well plates with D10 media. After 24 h, HDFs were transduced three times every 12 h over 24 h (at 0, 12, and 24 h). Twelve hours after the final transduction, the media was switched to D10 supplemented with 1 $\mu g \cdot mL^{-1}$ doxycycline (DOX) to begin transgene activation. On Day 4, transduced HDFs were dissociated with trypLE Express (Thermo Fisher Scientific) and split 1 : 2 onto 0.1% gelatin-coated six-well or 12-well plates with Myelocult media (H5100; Stem Cell Technologies, Vancouver, Canada) supplemented with hydrocortisone (HC; 10^{-6} M; Stem Cell Technologies), the cytokines SCF, FLT3L, and TPO (all R&D systems, 25 $ng \cdot mL^{-1}$ as previously described [38]), 1 $\mu g \cdot mL^{-1}$ DOX, and 50 μM 2-ME. Myelocult media was changed every 4 days for the duration of the cultures.

FACS analysis and sorting

Cells from standard reprogramming, CFU, or LTC experiments were first harvested using trypLE Express at specified days and washed with PBS supplemented with 5% FBS and 1 mM EDTA. Flow cytometry analysis was performed on a five-laser LSRII with DIVA software (BD Biosciences, San Jose, CA, USA) and analyzed with FCS Express 6 Flow Research Edition (Win64). Cells were stained with PE/CY7-hCD45 (2D1), FITC-hCD235a (GA-R2), APC-hCD41 (MReg30), BV421-hCD14 (M5E2), BV421-hCD34 (581), APC-hCD45 or FITC-hCD45 (2D1), PE-hEPCR (RCR-401), or APC-hCD49f (GoH3; all Biolegend, San Diego, CA, USA), as well as PE-hACE (BB9), FITC-hCD90 (5E10), PE-hCD49f (GoH3; BD Biosciences), APC-hCD90 (5E10; Thermo Fisher), or hACE-Biotin (BB9; R&D Scientific Corporation) and incubated with APC-Cy7 streptavidin (BioLegend). 4,6-diamidino-2-phenylindole (DAPI, 1 $\mu g \cdot mL^{-1}$; Sigma, St. Louis, MO, USA) or propidium iodide (PI, 1 $\mu g \cdot mL^{-1}$, R17755; Invitrogen, Carlsbad, CA, USA) was added prior to analysis to exclude dead cells. Sorting for RNA sequencing, transplantation, LTC, and CFU assays was performed with APC-CD49f alone, PE-CD49f alone, or BV421-CD34 and PE-CD49f using DAPI or PI to exclude dead cells..

Live imaging

Reprogrammed GGF and 3GF cells were analyzed by live staining at Days 14, 20, 28, and 35. Myelocult media was removed and 300 μL of 1 × PBS with 5% FBS was added and incubated with PE-hEPCR 1 : 20 for 15 min at 37 °C. The antibody mix was then aspirated, cells were washed with 1 × PBS with 5% FBS, their supplemented Myelocult media was replenished, and they were subsequently imaged on a Leica DMI 4000 B using Leica LAS AF software (Buffalo Grove, IL, USA). For CFU live stains, colonies were collected and washed with 1 × PBS with 5% FBS. They were then resuspended in 200 μL of trypLE Express, incubated at 37 °C for 5 min, triturated, and washed again with 1 × PBS with 5% FBS. Cells were then resuspended in 200 μL of 1 × PBS with 5% FBS, loaded into TC treated, sterile μ-Slide VI 0.4 ibiTreat chamber slides (Ibidi, Fitchburg, WI, USA, #80606), and imaged on a Leica DMI6000 inverted scope using Leica LAS AF software

Primary CFU, LTC-initiating cell (LTC-IC), and cobblestone area forming cell (CAFC) assays

Reprogrammed tdT-HDFs were harvested with trypLE Express at Days 15, 20, and 25 of reprogramming, washed in 1 × PBS, suspended in 500 μL of DMEM, and dispersed in 3 mL of cytokine-supplemented methylcellulose (h4435; Stem Cell Technologies). The cell suspension was then drawn into an 18G needle with a 5 mL syringe and plated 1ml per well of a non-TC-treated six-well plate

(Costar, Washington DC, USA, #3736). Empty spaces between the wells were filled with sterile H_2O and the plates were then incubated at 37 °C in 5% CO_2. For LTC assays, 12-well plates were first coated with 0.1% gelatin. Expanded AFT024 stromal cells were harvested and seeded at a density of 3×10^5–3.5×10^5 cells·mL^{-1} in D10 media supplemented with 50 μM 2-ME. About 1ml of the cell suspension was plated in each well of the gelatinized 12-well plates. Cells were then grown overnight at 32 °C with 5% CO_2. The next day cells were irradiated with 2000 rads. 20 000–30 000 Day 15 CD49f$^+$ sorted cells were then placed in each well with 4 mL of supplemented Myelocult media (with previously described concentrations of HC, SCF, FLT3L, TPO, DOX, and 2-ME) and incubated at 37 °C in 5% CO_2. Plates were then observed for colony growth and morphology, with weekly half-media changes for up to 5 weeks. For primary CFU assays using Lin$^-$CD34$^+$ cord blood (CB) HSCs, 250 cells were plated per 1ml of methylcellulose (h4435; Stem Cell Technologies), observed over 2 weeks, and colony types/total colony numbers were counted. Two thousand Lin$^-$CD34$^+$ CB HSCs were plated per well of a 12-well plate for AFT024 LTC cultures as well.

Colony imaging, LTC-IC CFU, and cytospins

Selected wells from the LTC assay for reprogrammed cells as well as Lin$^-$CD34$^+$ CB HSCs were harvested and plated in CFU assays as previously described. For LTC-initiating cell (IC) CFU plating for CB HSCs from LTC assays, 1 well from these cultures was taken and separated into 90% (therefore representative of 1800 of the initially seeded Lin$^-$CD34$^+$ CB HSCs) and 10% (200 initial cells) samples. Cobblestone and CFU colonies were imaged on a Leica DMI 4000 B automated inverted scope using Leica LAS AF software. After colony derivation in CFU from the LTC assays, colonies were collected in 1 × PBS supplemented with 5% FBS, washed, and resuspended in 200 μL of trypLE Express. Colonies were then incubated in trypLE at 37 °C for 5 min, titrated to make single-cell suspensions, washed, resuspended in 200 μL of 1 × PBS with 5% FBS, and loaded into cytospin-prepared slides. Samples were spun at 250 r.p.m. for 3 min, stained with Hematoxylin and Eosin, and then imaged on a Leica DM5500 upright scope using Leica LAS AF software.

AFT024 *in vitro* CAFC limiting dilution analysis (LDA)

AFT024 stroma was cultured as previously described [36] and harvested to a concentration of 350 000 cells·mL^{-1} in D10 supplemented with 50 μM 2-ME. In 0.1% gelatin-coated 96-well plates, 100 μL of this suspension was plated and allowed to grow overnight at 32 °C with 5% CO_2. The following day, cells were irradiated and the media was replaced with 100 μL of fresh supplemented Myelocult. On Day 15 of reprogramming, CD49f$^+$ tdT-HDF cells reprogrammed with either GGF or 3GF were sorted and seeded in all 12 wells of Row A of the prepared 96-well plate with 20 000 cells in 100 μL per well. Using a multichannel pipet, 100 μL of the 200 μL cell suspension was taken to Row B, mixed with the 100 μL already present in the wells, and then serially diluted down to Row H with the dilutions as follows: Row A: 10 000 cells per well; Row B: 5000 cells; Row C: 2500 cells; Row D: 1250 cells; Row E: 625 cells; Row F: 312.5 cells; Row G: 156.25 cells; and Row H: 78.125. Twenty hours after seeding, 100 μL of fresh supplemented Myelocult was added to each row. Half-media changes were performed weekly, and wells with emerging cobblestone colonies were counted after 5 weeks of LTC. Stem cell frequency was then calculated using Poisson statistics [36] and extreme LDA (ELDA) [39]. The same process was used for LDA analysis of Lin$^-$CD34$^+$ CB HSCs, but instead cell densities started at 1000 cells per well in Row A, 500 in Row B, 250 in Row C, 125 in Row D, 62.5 in Row E, 31.25 in Row F, 15.625 in Row G, and 7.8125 in Row H for primary LDA. For secondary LDA, cell densities were 250 cells per well in Row A, 125 in Row B, 62.5 in Row C, 31.25 in Row D, 15.625 in Row E, 7.8125 in Row F, 3.90625 in Row G, and 1.953125 in Row H. CB Lin$^-$CD34$^+$ cells were also grown in supplemented Myelocult media, but without added DOX.

mRNA, cDNA, and library sample preparation

3GF reprogrammed HDFs were reprogrammed to Day 15 and Day 25 and subsequently sorted in triplicate into two populations at both time points: CD49f+CD34− and CD49f+CD34+. About 10^5 cells for CD49f+CD34− replicates and at least 3×10^4 cells for CD49f+CD34+ replicates were sorted into 1 × PBS with 5% FBS, pelleted, and RNA was subsequently isolated using the NucleoSpin RNA XS extraction kit (Takara Bio USA, in Mountainv View, CA, USA, 740902.50). cDNA was synthesized and amplified using the SMART-seq v4 Ultra Low Input RNA Kit for Sequencing (Takara Bio, USA, 634889). Amplified cDNA was then purified using the Agencourt AMPure XP Kit (Beckman Coulter, Brea, CA, USA, A63880). The concentration of the derived cDNA was quantified using a Qubit fluorometer (Thermo Fisher Scientific). The quality of the derived cDNA samples was determined using the Agilent High Sensitivity DNA Kit (Agilent, Santa Clara, CA, USA, 5067-4626) and an Agilent 2100 Bioanalyzer. cDNA libraries were then created using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA, FC-131-1024) and the Nextera XT Index Kit (Illumina, FC-131-1001) and subsequently sequenced on an Illumina HiSeq 4000 with 25M 100-nt reads per sample.

RNAseq analysis

Reads were mapped to the human genome (GRCh37.75) using STAR to avoid high mapping error rates, low

mapping speed, and read length limitation/mapping biases [40]. Reads mapping to annotated genes were counted using featureCounts [41]. Read count normalization and pairwise differential expression analyses between groups of samples were performed using DESeq2, which is used for differential analysis of count data (in our case normalized read counts) with shrinkage estimation for fold changes to improve the accuracy and stability of our results [42]. GAGE, a gene set analysis method that can manage datasets with different sample sizes or experimental designs, was used to perform gene set enrichment tests between pairwise groups of samples [43]. Enrichment tests were performed for custom gene sets extracted from prior literature and for gene sets associated with specific gene ontology (GO) terms [44,45]. Samples were additionally analyzed using principal component analysis (PCA) and hierarchical clustering in R. DESeq2. Normalized read counts for all relevant samples were plotted using GRAPHPAD PRISM 7 software. RNAseq data from Gomes *et al.* [30] was used for comparative purposes. The upregulation of TFs in 3GF cells and GGF cells was compared using BIOJUPIES [46].

NSG mouse transplants

After sorting CD49f$^+$ cells from 3GF reprogrammed HDFs cultured on 0.1% gelatin, cells were washed in 1 × PBS and then transplanted (3.0×10^5 cells) into NOD-scid IL2Rg^{null}(NSG) 0- to 2-day-old pups *via* intrahepatic injection. Mouse peripheral blood (PB) was analyzed at 4, 8, and 16 weeks postinjection for engraftment of human-derived cells. To distinguish levels of engraftment, cells were stained for mouse PacBlue-mCD45 (30-F11; Biolegend) and PE-Cy7-hCD45 (2D1; Thermo Fisher Scientific). Within the hCD45 compartment, we looked at cells stained with APC-eflour 780-hCD3 (UCHI1), PerCP-Cy5.5-hCD8 (RPA-T8), PE-hCD19 (HIB19), and Alexa488-hCD14 (all from Biolegend). Engraftment levels were compared to levels found in CB-derived CD34$^+$ hematopoietic progenitors isolated using Diamond CD34 Isolation Kits (Miltenyi, Bergisch Gladbach, Germany) using the same injection method as previously described (1.0×10^5 cells·pup^{-1}).

Statistical analysis

Data were analyzed with GRAPHPAD PRISM 7 software using the nonparametric Mann–Whitney test for samples not assuming a normally distributed dataset. Bars represent mean and error bars represent standard error of the mean (SEM). Statistically significant differences are as follows: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. Limiting dilution frequencies were determined using Poisson statistics and ELDA software [39].

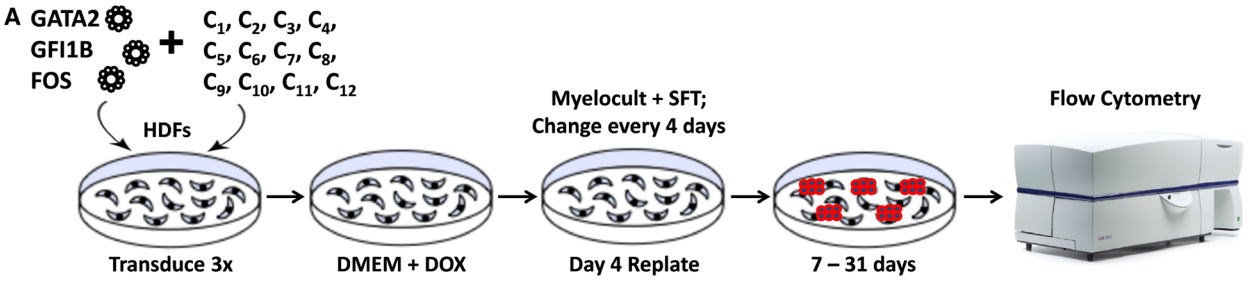
Results

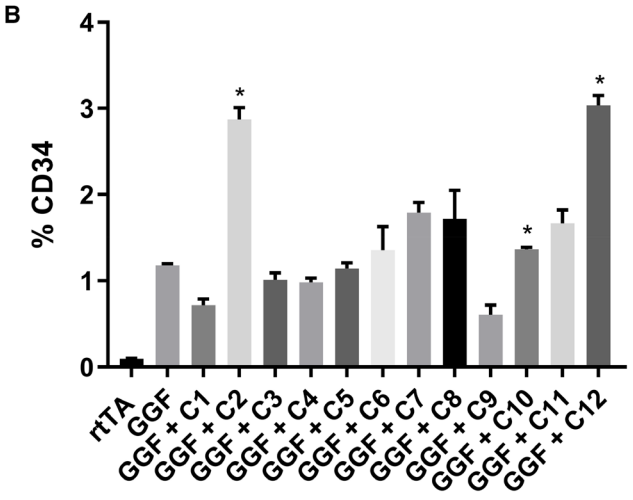
GFI1 added to the GGF cocktail expands the derived HSPC pool

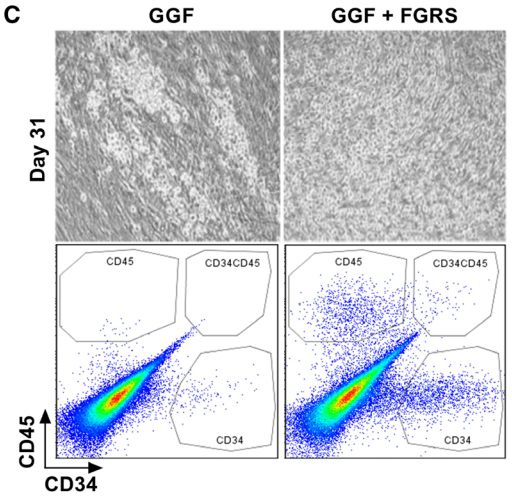
We recently demonstrated that three TFs, GATA2, GFI1B, and FOS, were sufficient to induce a hemogenic program in human fibroblasts [30]. Transplantation of CD49f$^+$ cells sorted at Day 25 of reprogramming in immunocompromised mice resulted in short-term multilineage engraftment. These results were encouraging but the overall yield and function of the reprogrammed cells needed improvement. We elected to ask if other TFs from previously published reprogramming strategies together with our factors would improve outcome. These factors were individually cloned into the DOX-inducible pFUW cassette used in our initial studies [37] (Table S1A) [17,18,23,26,29,32–34,47]. These TFs were arranged into 12 distinct combinations (Table S1B) and used in concert with GGF to induce hemogenesis in adult HDFs.

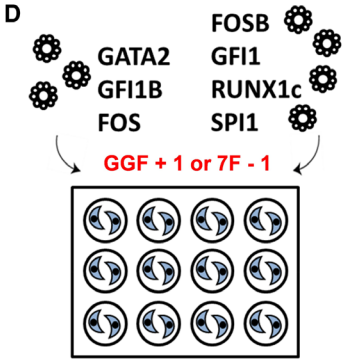
Reprogrammed cells were screened at Day 30 to determine levels of CD34 induction, a known HSC marker [29,48] (Fig. 1A). We found that three TF combinations significantly improved conversion to CD34$^+$ cells (Fig. 1B). C2 contains the polycistronic

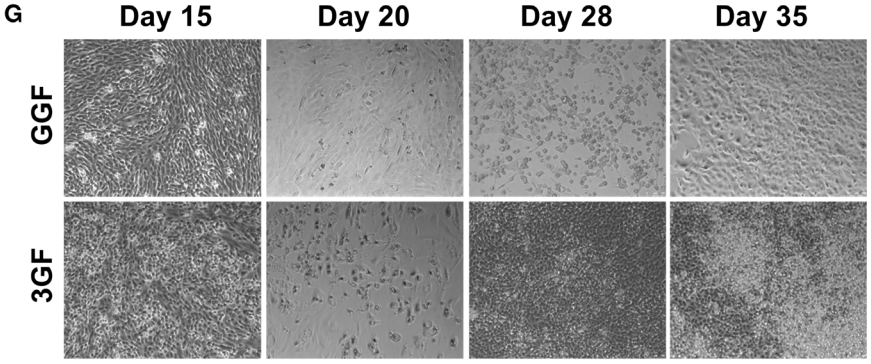
Fig. 1. GFI1 expands the CD34$^+$ progenitor pool induced by GGF reprogramming. (A) Scheme displaying the reprogramming process for the initial TF screen. SFT = SCF, Flt3L, and TPO (each at 25 ng·mL^{-1}). (B) Flow cytometric quantification of Day 30 reprogrammed cells identifies three TF combinations that significantly expand the CD34$^+$ progenitor population compared to GGF reprogrammed cells out of the 12 tested combinations (n = 2–3). (C) Day 31 cellular morphology during reprogramming culture reveals a massive proliferation of round hematopoietic-like cells when GGF reprogramming is supplemented with FGRS. Expansion of the CD34$^+$ and CD45$^+$ cell pools are also observed in the representative flow cytometry dot plots. (D) Scheme for $N + 1$ and $N - 1$ experiments to identify the TF or combination of TFs from the FGRS cocktail that permits the observed CD34$^+$ expansion. (E) Results from the $N + 1$ and $N - 1$ experiments reveal that GFI1 factor in the FGRS combination is responsible for the CD34$^+$ cell expansion (n = 6). (F) Representative flow plots of GGF and 3GF reprogrammed cells reveal the significant induction of CD34$^+$ cells (Auto represents an empty channel). (G) Representative reprogramming culture images throughout GGF and 3GF reprogramming recapitulate the massive proliferation of round hematopoietic-like cells present when GFI1 is used with GGF. Data are represented as mean ± SEM. Nonparametric Mann–Whitney *T*-test for samples not assuming a normally distributed dataset. $*P < 0.05$, $**P < 0.01$, $***P < .001$.

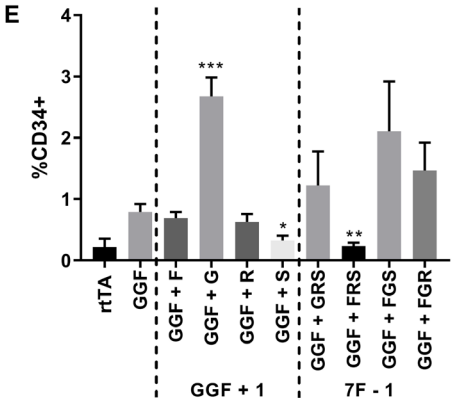

A
GATA2
GFI1B
FOS
C1, C2, C3, C4,
C5, C6, C7, C8,
C9, C10, C11, C12
HDFs
Myelocult + SFT;
Change every 4 days
Flow Cytometry
Transduce 3x
DMEM + DOX
Day 4 Replate
7 – 31 days

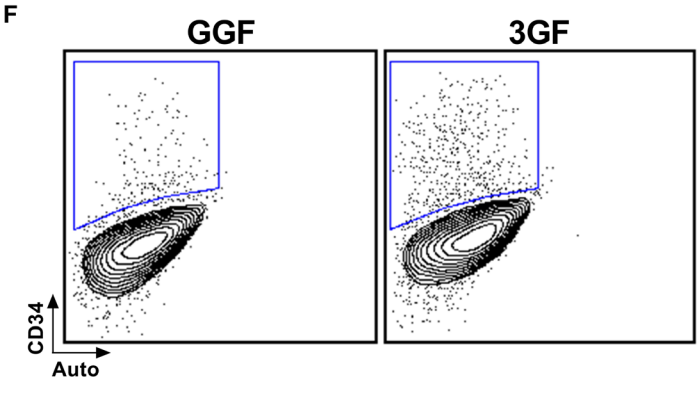

B
% CD34
rtTA
GGF
GGF + C1
GGF + C2
GGF + C3
GGF + C4
GGF + C5
GGF + C6
GGF + C7
GGF + C8
GGF + C9
GGF + C10
GGF + C11
GGF + C12


C
GGF
GGF + FGRS
Day 31
CD45
CD34CD45
CD34
CD45
CD34


D
GATA2
GFI1B
FOS
FOSB
GFI1
RUNX1c
SPI1
GGF + 1 or 7F - 1


G
Day 15
Day 20
Day 28
Day 35
GGF
3GF


E
%CD34+
rtTA
GGF
GGF + F
GGF + G
GGF + R
GGF + S
GGF + GRS
GGF + FRS
GGF + FGS
GGF + FGR
GGF + 1
7F - 1


F
GGF
3GF
CD34
Auto

STEMCCA pluripotency reprogramming cassette [49] and C10 contains a combination of shRNAs to p53, which has been shown to improve reprogramming efficiency upon repression of p53 [26,50]. To avoid both reprogramming to pluripotency and altering the p53 network of the reprogrammed cells, these cocktails were not used further. C12, however, contained a group of TFs: FOSB, GFI1, RUNX1c, and SPI1 (FGRS), which enhanced GGF reprogramming without induction of pluripotency (Fig. 1C). This cocktail of factors has been recently shown to successfully reprogram human and mouse endothelial cells to acquire a hematopoietic fate upon continuous coculture with a vascular niche [32,33].

$N - 1$ or $N + 1$ experiments using the GGF and FGRS cocktails (Fig. 1D) revealed GFI1 as the factor that improves the yield of $CD34^+$ cells as seen when GFI1 alone is added to the GGF cocktail, or when GFI1 is removed from the seven-factor combination (GGF + FGRS; Fig. 1E,F). Reprogramming with new reprogramming cocktail (now termed 3GF for GATA2, GFI1, GFI1B, and FOS) results in a significant expansion of rounded hematopoietic-like cells in culture as compared to GGF cells based on cellular morphology (Fig. 1G). In-house developed software 'GPSforGenes' [30] shows that the TF combinations for GGF and 3GF, as well as GATA2 + FOS or GATA2 + GFI1 + FOS, were all highly expressed both in $CD34^+$ hematopoietic stem progenitor cells (HSPCs) and placental tissue (Fig. S1A–D). This program takes gene expression data from human tissues in the GeneAtlas U133A database, and demonstrates that the TFs we selected for hemogenic induction are highly expressed in a combinatorial manner in multiple hematopoietic tissues and lineages. It is interesting that without the inclusion of GFI1B, expression in $CD71^+$ early erythroid lineages cells is lost (Fig. S1C,D).

With initial experiments suggesting that GFI1 improves reprogramming efficacy based on CD34 induction, we next characterized the HSC surface phenotype of the 3GF vs. GGF derived cells. To this end, cell surface marker expression of developmental human HSC markers CD34, CD49f, and angiotensin converting enzyme (ACE) was analyzed at several time points during reprogramming. ACE enriches for HSCs and also marks all cells in the developing embryo destined to adopt a hematopoietic fate [51–53]. Likewise, CD49f (also known as integrin α6) also enriches for HSCs [53,54]. 3GF reprogramming results in expanded yields of all populations of hematopoietic progenitors (in this case ACE^+, $CD49f^+$, or ACE^+CD49f^+ cells) as compared to GGF (Fig. 2A). 3GF also results in expanded populations of $CD34^+$ throughout reprogramming, as well as the $CD49f^+CD34^+$ and ACE^+CD34^+ populations (Fig. 2B–D). Likewise, 3GF reprogramming induces an expansion of $CD49f^+ACE^+CD34^+$ progenitors (Fig. 2E). Additionally, the bulk ACE^+ subset at Day 35 of GGF or 3GF reprogramming demonstrates robust populations of both $CD49^+CD34^-$ (red box) and $CD49f^+CD34^+$ (green box) cells (Fig. S2A).

Previous studies have identified endothelial protein C receptor (EPCR, CD201) as a marker of expanded $CD34^+$ progenitors from CB [55] as well as functional HSCs in murine BM [56]. Interestingly, both GGF and 3GF reprogrammed cells throughout the reprogramming process stain positive for EPCR, which by morphology appears to mark both endothelial-like cells as well as the rounded HSC-like cells that emerge in these prolonged cultures. In particular, the EPCR population in 3GF reprogramed cells significantly expands over time (Fig. 3A). As with the other HSC markers (Fig. 2), greater yields of $CD49f^+CD34^+$ $EPCR^+$ cells are obtained in the 3GF population, with the majority of $EPCR^+$ cells emerging at Day 27, which corresponds to a late-intermediate time point in reprogramming. Notably, virtually all of the $CD49f^+CD34^+$ cells are $EPCR^+$ (Fig. 3B,C). Within the bulk $EPCR^+$ subset of cells at Day 35 of GGF and 3GF reprogramming, distinct populations of $CD49f^+$ $CD34^-$ and $CD49f^+CD34^+$ cells can be observed as well (Fig. S2B).

Expression profiling of 3GF cells as they undergo reprogramming

Transcriptomic analyses *via* RNAseq of both 3GF and GGF reprogrammed cells were compared to determine if 3GF cells also underwent a developmental process as cells transition through endothelial intermediates to form hematopoietic cells while potentially remaining distinct from GGF cells [30]. To this end, four different populations of 3GF reprogrammed cells were sequenced in triplicate: (a) D15 3GF $CD49f^+CD34^-$; (b) D15 3GF $CD49f^+CD34^+$; (c) D25 3GF $CD49f^+CD34^-$; and (d) D25 3GF $CD49f^+CD34^+$.

Molecular profiling of these four separate 3GF populations shows strong similarities to previously acquired datasets of GGF reprogrammed cells [30]. The first observation in this analysis was that dimension 1 in the PCA analysis likely accounted for technical variation between the GGF and 3GF experiments, and therefore in comparative PCA analyses and hierarchical clustering, only dimensions 2 and 3 were used (Fig. 4A,B). Comparison of dimensions 2 and 3 for GGF and 3GF reprogramming reveals a similar

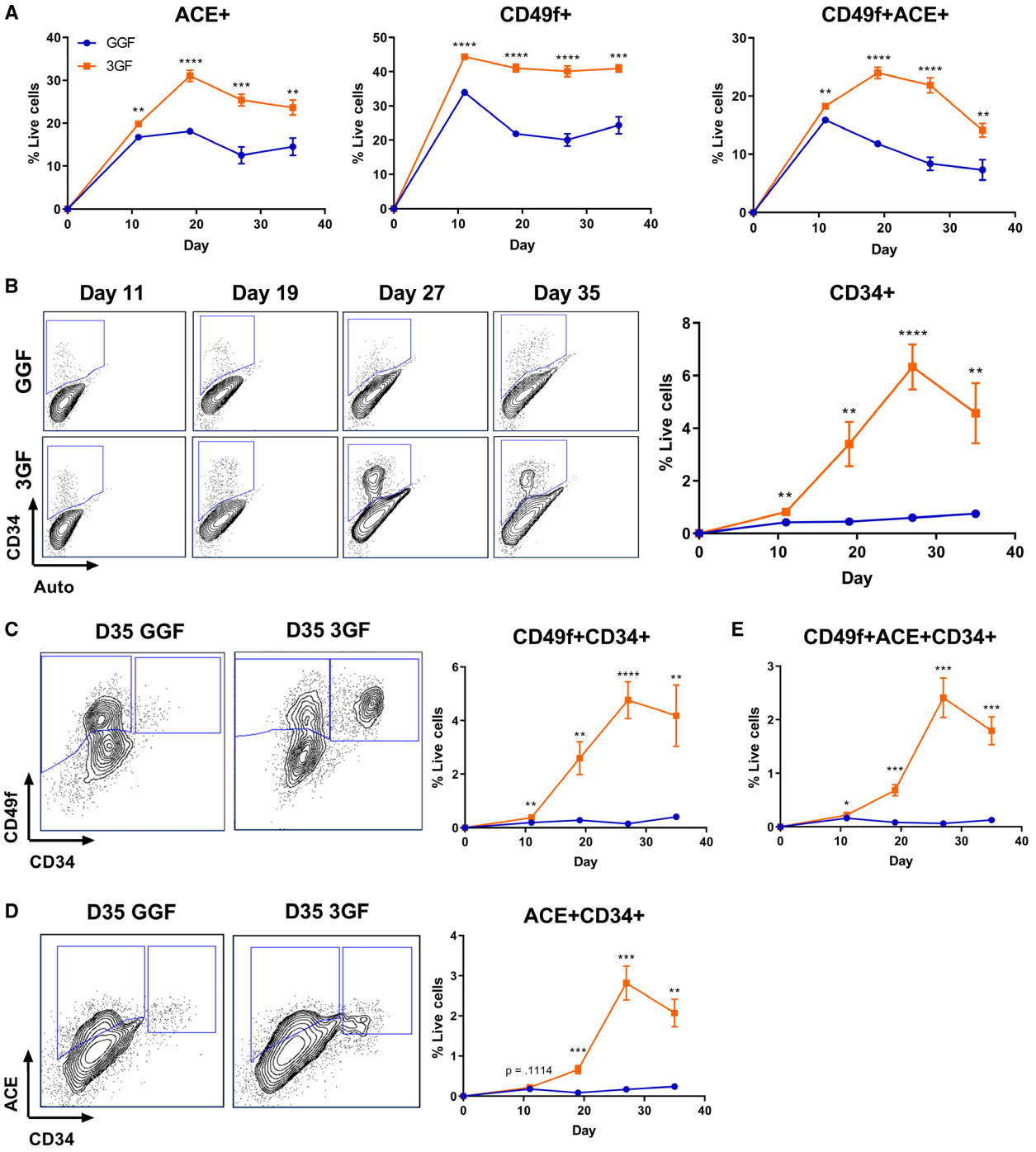


Fig. 2. 3GF reprogramming expands the hematopoietic progenitor pool based on cell-surface immunophenotype. (A) Flow cytometric quantification of ACE$^+$, CD49f$^+$, or CD49f$^+$ACE$^+$ cells (n = 7–12). (B) Representative flow plots of the CD34$^+$ population over the course of reprogramming showing a significant expansion of these cells with 3GF quantified on the right (n = 13–24). (C) Representative flow cytometry contour plots of CD49f$^+$CD34$^+$ cells in GGF and 3GF cells and their quantification (n = 14–24). (D) Representative flow plots of ACE$^+$CD34$^+$ cells in GGF and 3GF cells and their quantification (n = 7–12). (E) Quantification of flow data for CD49f$^+$ ACE$^+$CD34$^+$ cells in GGF and 3GF reprogramming (n = 7–12). Data are represented as mean ± SEM. Nonparametric Mann–Whitney T-test for samples not assuming a normally distributed dataset. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

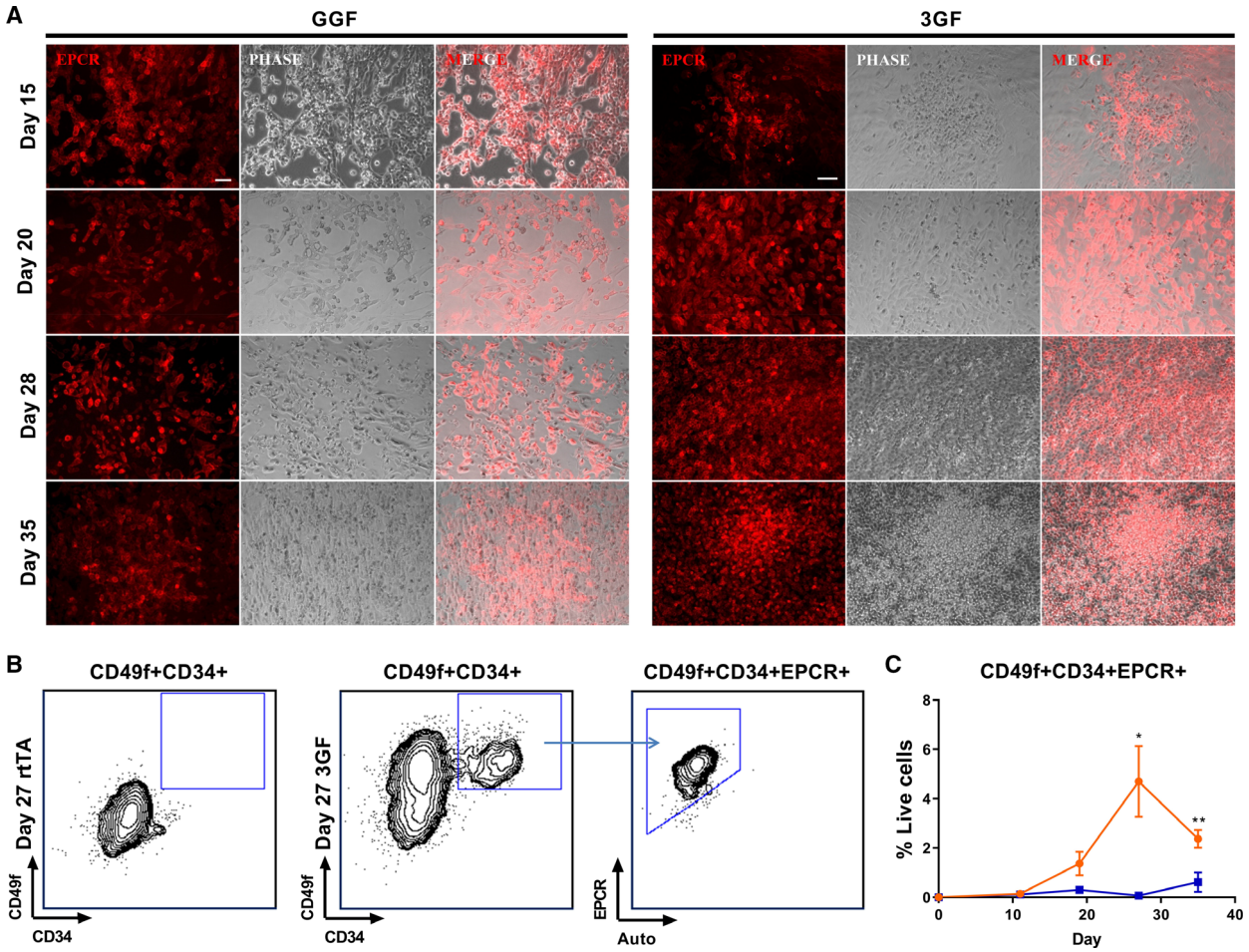


Fig. 3. GGF- and 3GF-induced cells express the stem cell marker EPCR. (A) Live staining of EPCR in GGF and 3GF cells throughout reprogramming. (B) Representative flow plots from D27 of rtTA control and 3GF cells stained for CD49f, CD34, and EPCR. (C) Quantification of $CD49^+CD34^+EPCR^+$ cells (n = 4–12). Data are represented as mean ± SEM. Nonparametric Mann–Whitney T-test for samples not assuming a normally distributed dataset. $^*P < 0.05$, $^{**}P < 0.01$.

developmental trajectory taken by 3GF cells and a strong separation of HDF negative controls from the reprogrammed populations (Fig. 4A). Hierarchical clustering of GGF and 3GF cells, again after removal of dimension 1 to correct for batch effects, shows a close relationship between D25 $CD49f^+CD34^+$ cells and, interestingly, clustering of GGF D25 $CD49f^+CD34^-$ and 3GF D15 $CD49f^+CD34^+$ cells (Fig. 4B). Previous data show that the GGF D25 $CD49f^+CD34^-$ population clusters most closely resemble known HSCs from published work [54], suggesting that the 3GF D15 $CD49f^+CD34^+$ population may most closely resemble endogenous HSCs in the 3GF reprogramming system.

Interactive notebooks containing bioinformatics data from this RNAseq data were created using BIOJUPIES software [46]. We focused on comparative analyses of upregulated TFs between GGF and 3GF populations on the same day cell phenotype. We observed that a variety of TFs were upregulated specifically in the 3GF populations in each comparison. These TFs (labeled in blue, red, green, or purple) are found in the literature and support acquisition of a more hematopoietic TF landscape [6,57,58] (Fig. 4C). Notably, this analysis demonstrates clear upregulation of MAZ, RCOR1, and ZKSCAN1 in each 3GF population. MAZ is known to be involved in the activity of the SCL 1b promoter in $CD34^+$ primitive myeloid cells [59]. RCOR1, another TF enriched in HSCs, is suggested to play an important role in HSC specification from the AGM and was also found to be upregulated in 3GF cells [60]. ZKSCAN1 was found in some hPSC-derived HE cells that engrafted immunocompromised mice [31], as well as in other transcriptomic

FEBS Letters **593** (2019) 3266–3287 © 2019 The Authors. *FEBS Letters* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

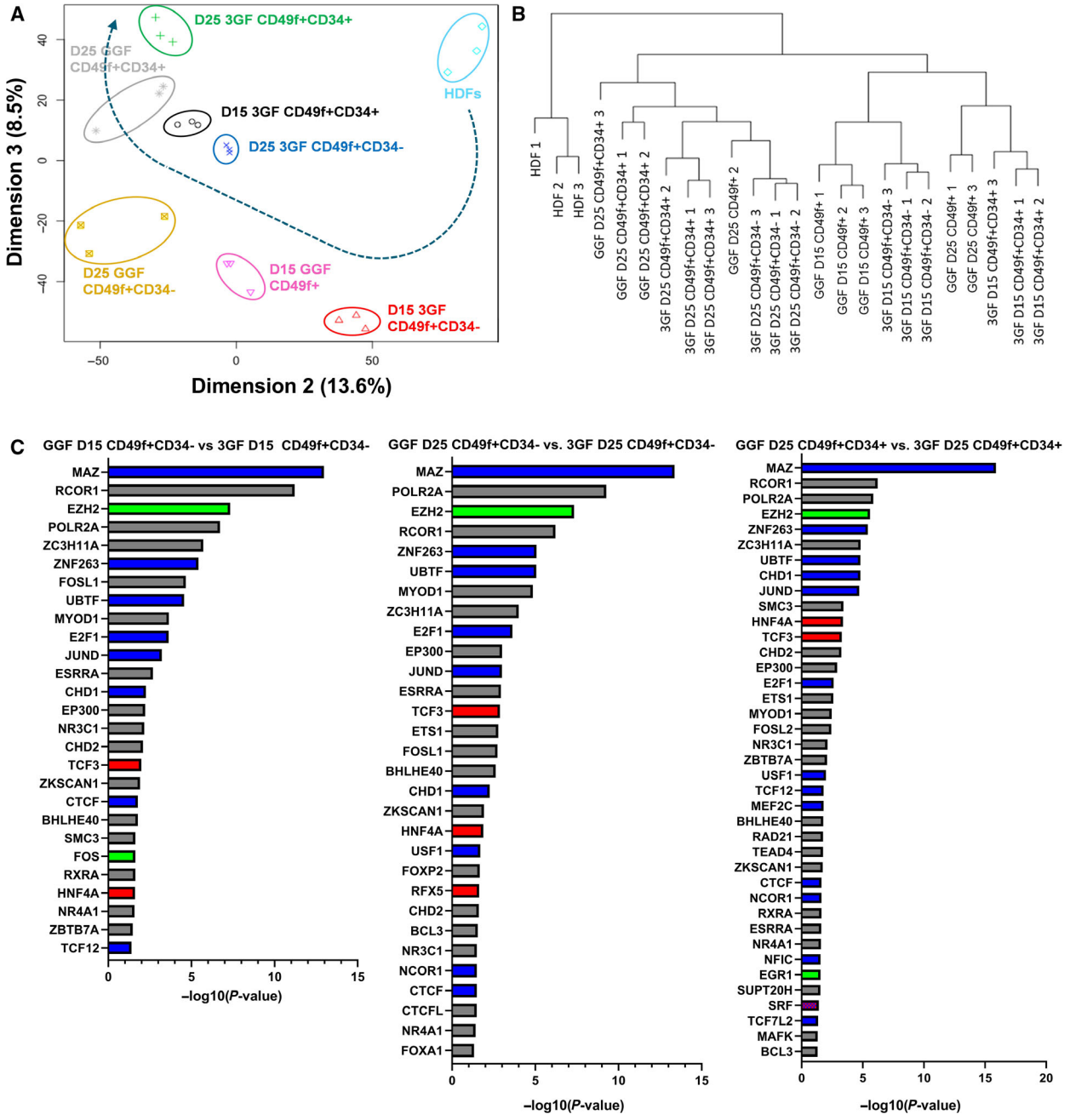


Fig. 4. Comparative RNA sequencing analyses of GGF vs. 3GF cells. (A) PCA of dimension 2 (13.6%) and dimension 3 (8.5%) between GGF and 3GF reprogrammed cells. The blue curved arrow represents the hypothetical trajectory of cells as they traverse from endothelial intermediates, defined by the data as $CD49f^{+}CD34^{-}$, toward hematopoietic cells, defined as $CD49f^{+}CD34^{+}$. (B) Hierarchical clustering analysis of GGF and 3GF cells after depletion of dimension 1. (C) Upregulated TFs between GGF and 3GF populations at different time points and surface phenotype determined using BIOJUPIES [46] and plotted to reflect statistically significant differential expression. Known TFs identified from the human $CD34^{+}$ hematopoietic cell database are labeled by blue bars [57]. TFs seen to be upregulated in E9.5 HE and E9.5 artery HE are shown in red [58], and hematopoietic TFs upregulated as hPSC-derived $CD34^{+}$ cells undergo mesodermal specification, EHT, and eventual blood production are shown in green [6]. The patterned purple bar for SRF reflects TFs identified in both Gomes *et al.* [57] and Gao *et al.* [58].

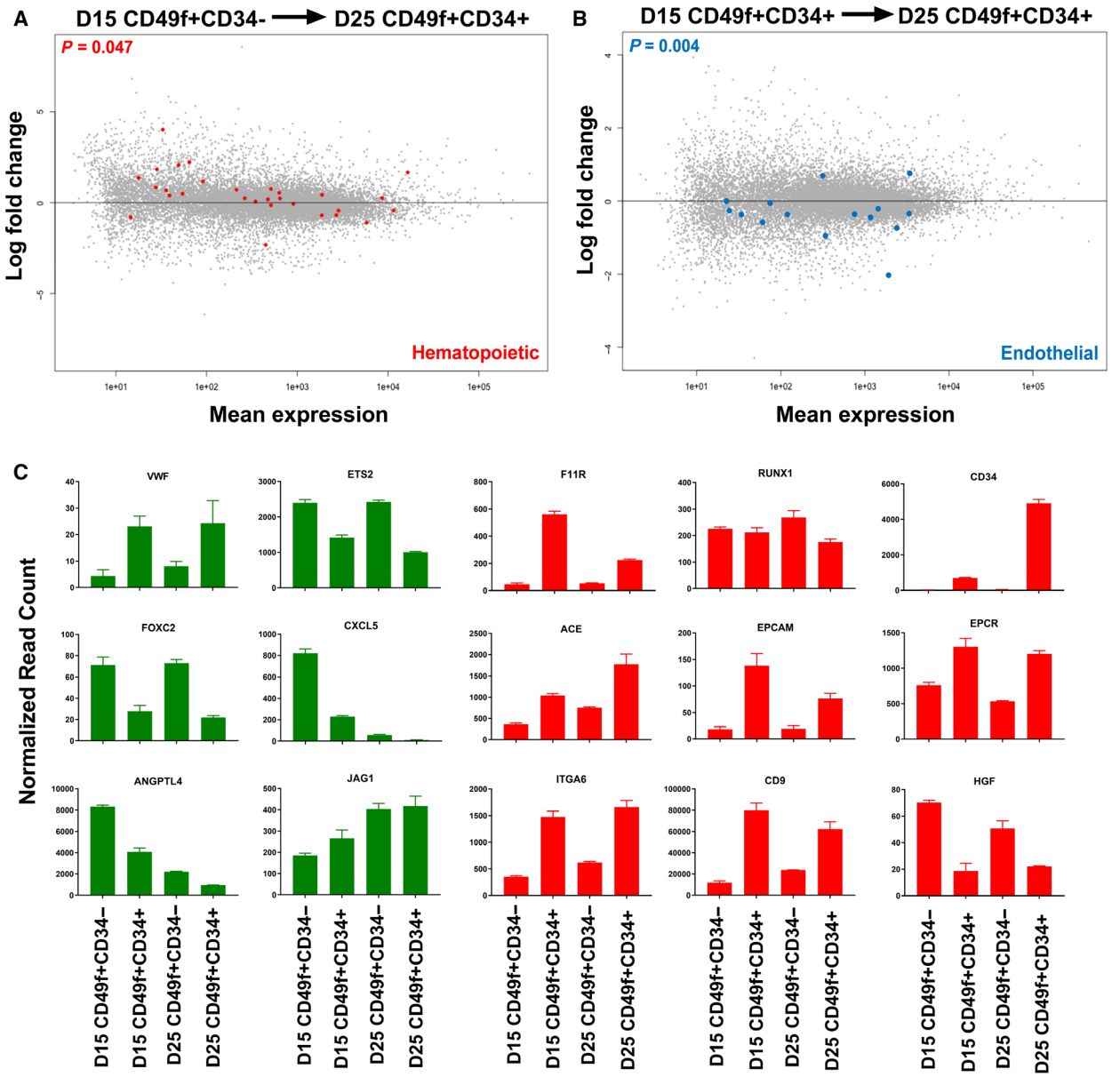


Fig. 5. Differential gene expression analysis of 3GF cells. (A) DESeq2 MA plot showing comparative analysis of the 3GF D15 CD49f$^+$CD34$^-$ to D25 CD49f$^+$CD34$^+$ populations using a CD34$^+$CD38$^-$CD45RA$^-$Thy1$^+$CD49f$^+$ HSC gene list from Notta *et al.* (red dots; [54]). (B) DESeq2 MA plot showing comparative analysis of the 3GF D15 CD49f$^+$CD34$^+$ to D25 CD49f$^+$CD34$^+$ populations using an endothelial gene list from F. Pereira (blue dots; [30]). (C) Quantification of normalized read counts for select endothelial and hematopoietic genes across the four RNA sequenced 3GF populations. Bar graphs in green represent genes commonly associated with endothelial identity. Bar graphs in red represent genes commonly associated with hematopoietic identity. Data are represented as mean ± SEM (n = 3).

analyses of HSCs [61,62]. These data suggest that the TF profile of 3GF cells more closely resembles native HSCs than GGF cells.

Using a list of upregulated genes in CD34$^+$CD38$^-$CD45RA$^-$CD90$^+$CD49f$^+$ HSCs as compared to CD34$^+$CD38$^-$CD45RA$^-$CD90$^-$CD49f$^-$ MPPs [54], there is a statistically significant upregulation of these genes in the 3GF dataset as D15 CD49f$^+$CD34$^-$ cells mature to D25 CD49f$^+$CD34$^+$ cells, $P = 0.047$ (Fig. 5A, red dots are genes from Notta *et al.* [54], elaborated in Table S2A). Interestingly, comparative analysis of D15 CD49f$^+$CD34$^+$ and D25 CD49f$^+$CD34$^+$ with highlighted endothelial genes [30] shows a significant downregulation of this gene list, $P = 0.004$

FEBS Letters **593** (2019) 3266–3287 © 2019 The Authors. *FEBS Letters* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

(Fig. 5B, blue dots are endothelial genes elaborated in Table S2B). Comparative analyses of these previously published gene lists highlight and support the theorized developmental trajectory initiated by 3GF reprogramming, where hematopoietic cells emerge from endothelial intermediates as they mature throughout the induction process either through time or with activation of CD34.

To globally assess expression differences between the maturation stages captured in our 3GF reprogramming, D15 $CD49f^+CD34^-$ cells were used as a baseline for subsequent GO term analysis. From this analysis, a consistent downregulation of key pathways pertaining to the cell cycle, including M phase, mitotic cell cycle, DNA packaging, and microtubule cytoskeleton organization GO terms, is observed (Fig. S3A). This suggests that the cells generated later in the reprogramming (as well as D15 $CD49f^+CD34^+$ cells) shut down the machinery required for the cell cycle, which is consistent with known inactive cell cycle machinery in endogenous quiescent HSCs that usually remain in the G_0 phase in the BM [63]. Significantly upregulated GO terms in this analysis include skeletal system/vasculature development, acute inflammatory response, activation of the immune response, polysaccharide metabolic process, aminoglycan metabolic process, and glycoprotein catabolic process. Upregulation of the skeletal system/vasculature development terms could indicate enrichment for a response to HSC-support factors typically found in the endogenous HSC niche [64]. Unsurprisingly, the terms for the inflammatory and immune response indicate an activation of hematopoietic-type genes, as HSCs are involved in these pathways [65]. Perhaps most interestingly, several terms involved with metabolism and catabolism of aminoglycans, polysaccharides, and glycoproteins are seen (Fig. S3A). This suggests that 3GF reprogrammed cells are primed to utilize inductive signals from a stromal niche layer to mature further.

Key individual endothelial and hematopoietic genes possess interesting expression patterns throughout 3GF reprogramming. Regarding endothelial genes, we see von Willebrand factor (vWF), a factor known to be expressed in both endothelial cells and HSCs, is expressed in each analyzed population, with a predominance of vWF in the $CD49f^+CD34^+$ hematopoietic cells. Conversely, there is a predominance of ETS2 and FOXC2 in the $CD49f^+CD34^-$ cells, and both genes are known to play key roles in HE. Interestingly, there is a downregulation of CXCL5 and ANGPTL4 across time and throughout each population. JAG1, however, appears to increase as time goes on and as 3GF cells acquire CD34 (Fig. 5C, green bars represent endothelial genes). This further indicates the endothelial identity these cells take on and shows that the cells may be constructing an intrinsic endothelial niche during the reprogramming process. There is also an induction of key hematopoietic genes throughout the reprogramming, several of which confirm what is observed in the flow data (Figs 2 and 3). CD34 expression limited to the $CD34^+$ sorted populations validates the accuracy of sorting and sequencing. Interestingly, although expression of our other markers is observed in all populations (as expected), there is a predominance of ACE, EPCR, and ITGA6 (CD49f) in the $CD49f^+CD34^+$ populations. These markers are known to purify for functional human HSCs [51,52,54,55,66], further confirming the HSC-like identity that is induced in 3GF reprogramming. Interestingly, there is an increased expression of F11R, EPCAM, and CD9 in the D15 $CD49f^+CD34^+$ cells as compared to the D25 $CD49f^+CD34^+$ cells, suggesting that the 3GF D15 $CD49f^+CD34^+$ population possesses greater functional potential than the more mature counterparts. CD9 complexes with c-kit in CB $CD34^+$ cells, and may regulate hematopoietic progenitor proliferation and differentiation [67]. Unsurprisingly, each population expresses RUNX1, further supporting the thought that the derived cells are hemogenic and theoretically undergoing EHT. Interestingly, HGF has been found to act as a mobilizer of HSCs to the PB [68] (Fig. 5C, red bars represent hematopoietic genes). What it does in endothelial populations (where it appears to be primarily expressed), however, remains unknown.

AFT024 LTC imparts functional potential to 3GF reprogrammed cells

tdT-3GF cells reprogrammed to Days 15, 20, and 25 displayed hematopoietic morphology in reprogramming cultures, but when harvested and plated in CFU assays no colonies formed (Fig. 6A). Interestingly, although the derived 3GF cells clearly displayed a cell surface phenotype highly similar to endogenous human HSCs (Figs 2 and 3), their *in vitro* functional potential required further optimization. We theorized that the reprogramming process required a separate maturation step in the form of a coculture system as has been shown to be crucial for other reprogramming strategies [29,32–34,69].

As a positive control, 250 Lin^-CD34^+ CB HSCs·mL^{-1} were plated directly for CFU assays and CFU-GEMM, BFU-E, and CFU-GM cells were counted after 2 weeks (Fig. S4A). FACS analysis of these colonies revealed that the majority population

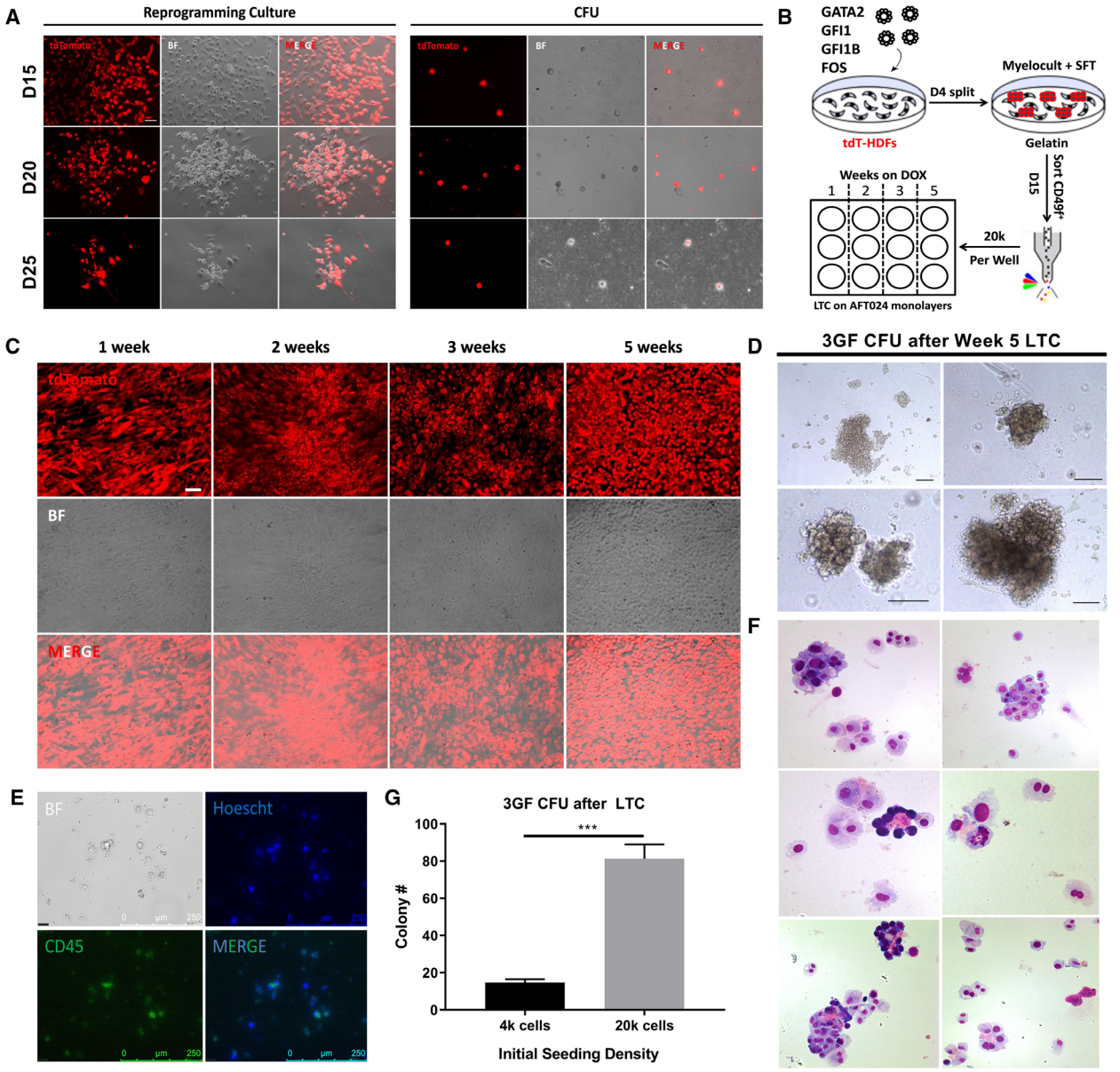


Fig. 6. *In vitro* maturation on AFT024 stroma imparts functional potential on 3GF reprogrammed cells. (A) Images of 3GF reprogrammed tdT-HDFs at Days 15, 20, and 25 of culture on gelatin with DOX and after an additional 3 weeks in CFU culture. CFUs failed to form. (B) Experimental scheme for LTC experiments on AFT024 stroma. Initially transduced cells were cultured on gelatin with DOX. After 15 days, $CD49f^+$ sorted cells were placed on AFT024 DOX, and continued culturing for varying lengths of time: 1 week = 1 week of DOX and 4 weeks NO DOX, 2 weeks = 2 weeks DOX and 3 weeks NO DOX, etc. (C) Images of 3GF reprogrammed tdT-HDFs after 5 weeks on AFT024 with the designated lengths of DOX exposure. (D) Resultant CFU from 5 weeks of LTC on AFT024 with DOX. (E) CD45 staining of CFU. (F) Cytospins of CFU. (G) Number of colonies derived from 4000 or 20 000 initially seeded $CD49f^+$ 3GF reprogrammed cells after 5 weeks of LTC on AFT024 with continuous DOX (n = 3–6). Error bars represent SEM. Nonparametric Mann–Whitney T-test for samples not assuming a normally distributed dataset. $^{***}P < 0.001$.

was $CD235a^+CD14^-$. A population of $CD235a^+CD14^+$ cells also emerged, which may indicate the presence of erythro-myeloid precursors in these cultures. Populations of $CD235a^-CD14^+$ and $CD41^+$ cells also existed within these colonies (Fig. S4B,C). Quantification of these colonies showed an average of 71.33 [standard deviation (SD) of 19.442] colonies per 250 initially seeded cells, signifying a colony forming potential of roughly 1 in 3.5 Lin^-CD34^+ cells. From these counts, there is an average of 9.333 CFU-GM (SD of 5.339),

6.44 BFU-E (SD of 2.506), and 55.556 CFU-GEMM (SD of 14.293; Fig. S4D).

Extensive early experiments to test *in vitro* maturation systems using OP9-DL1 [34] or E4EC [32,33] monolayers resulted in a loss of reprogrammed cells that did not display functional potential (data not shown). We elected to try coculture with the mouse fetal liver (FL) cell line AFT024 that we had previously shown to support both mouse and human HSCs [36,70,71]. On Day 15, 3GF reprogrammed tdT-HDFs were sorted for $CD49f^+$ cells and plated onto confluent monolayers of AFT024 with varying lengths of DOX exposure for varied activation of the transgenes (Fig. 6B). Strikingly, when Day 15 $CD49f^+$ sorted cells were cultured on AFT024 for 5 weeks with continuous DOX exposure, cobblestone-like colonies emerged (Fig. 6C and Fig. S5A). When these cells were harvested and plated in CFU assays, hematopoietic colonies formed that possessed myeloid and erythroid cells and stained positive for human CD45 (Fig. 6D–F). The frequency of CFU formation was ~ 1 in 250 cells that initially seeded the AFT024 LTC (Fig. 6G).

Cobblestone-like colonies with GGF reprogrammed cells also emerged in AFT024 LTCs (Fig. S5B), but the quality and quantity of the subsequent CFU were negligible (Fig. S5C,D). This establishes AFT024 as a stromal coculture niche that imparts the required signals to permit 3GF reprogrammed cells to mature and adopt hematopoietic functional potential.

AFT024 LTC permits quantification of stem cell frequency

The above results prompted the use of LDA and the AFT024 *in vitro* maturation system to determine the stem cell frequency of both GGF and 3GF reprogrammed cells, with the cobblestone area forming cell (CAFC) capacity correlating with repopulation activity [36,72] (Fig. 7A). This system allowed for distinct identification of positive cobblestone colonies as compared to tdT-HDFs that did not form colonies (Fig. S5A). Quantification of cobblestone-like colony formation demonstrated a significantly greater CAFC frequency of 1/4020 in 3GF reprogrammed cells as compared to 1/7465 in GGF cells ($P = 0.0270$; Fig. 7B). Using ELDA software, the likelihood ratio test of a single-hit model, a score test of heterogeneity, was statistically significant ($P = 0.00106$ and 0.0015, respectively). Positive control experiments using Lin^-CD34^+ CB HSCs also shows the formation of large cobblestone-like colonies after 5 weeks of LTC on AFT024. Cytospin of these cells reveals a majority of myeloid cells as well as some erythroid cells (Fig. S4E). Primary LDA of CB HSCs shows a steady decrease in HSC frequency over 5 weeks, suggesting that short-term progenitors in these purified cells may expand and exhaust, revealing a CAFC frequency that may correlate with long-term repopulating HSCs as previously shown [70] (Frequencies: W1, 1/11.8; W2, 1/19.5; W3, 1/23.3; W4, 1/55.3; and W5, 1/85.4; Fig. S4F and Table S3).

To determine if hematopoietic populations could be isolated from LTC on AFT024, FACS analysis was performed for several of the aforementioned markers on Day 15 $CD49f^+$ cells sorted onto either 0.1% gelatin- or AFT024-coated plates for 5 weeks with DOX supplementation. Some populations show a significant decrease in cell yield ($CD49f^+$, ACE^+, and $CD49f^+CD34^+$) while others remain the same ($CD34^+$, ACE^+CD34^+, $CD49f^+ACE^+$, and $CD49f^+ACE^+CD34^+$). The more mature $CD38^+$ population significantly decreases in cells grown on AFT024 (Fig. S5E). This suggests that some progenitor populations are maintained in AFT024 cocultures and that these populations can be isolated and sorted for downstream applications. Additional positive control experiments using Lin^-CD34^+ CB HSCs after AFT024 LTC show the continued derivation of multilineage colonies in secondary CFU assays after 5 weeks of LTC on AFT024 (Fig. S4G), with a majority of $CD45^+$ cells composed primarily of $CD14^+$ myeloid cells. Interestingly, the $CD235a^+CD14^+$ population is more abundant in these cells after AFT024 maturation, potentially indicating that this *in vitro* system may also support their expansion (Fig. S4H,I). From these counts, there is an average of 35.333 CFU-GM (SD of 5.164), 2.00 BFU-E (SD of 2.098), and 20.667 CFU-GEMM (SD of 4.227; Fig. S4J). Secondary LDA after 5 weeks of AFT024 coculture shows a sustained, higher stem cell frequency as compared to primary LDA assays, signifying the maintenance of true HSCs *in vitro* with this AFT024 cell line (Frequencies: W1, 1/9.04; W2, 1/7.0; W3, 1/12.6; W4, 1/24.61; and W5, 1/31.31; Fig. S4K and Table S3).

3GF cells engraft with short-term multilineage potential

To determine the engraftment potential of 3GF cells, derived $CD49f^+$ cells grown on gelatin were sorted at three different time points (Days 12, 15, and 18) of the reprogramming process and subsequently transplanted into the liver of newborn immunocompromised mice (P0–P2; Fig. 7C). Similar to what AFT024 supplies for the reprogrammed cells *in vitro*, the mouse microenvironment was theorized to provide necessary *in vivo*

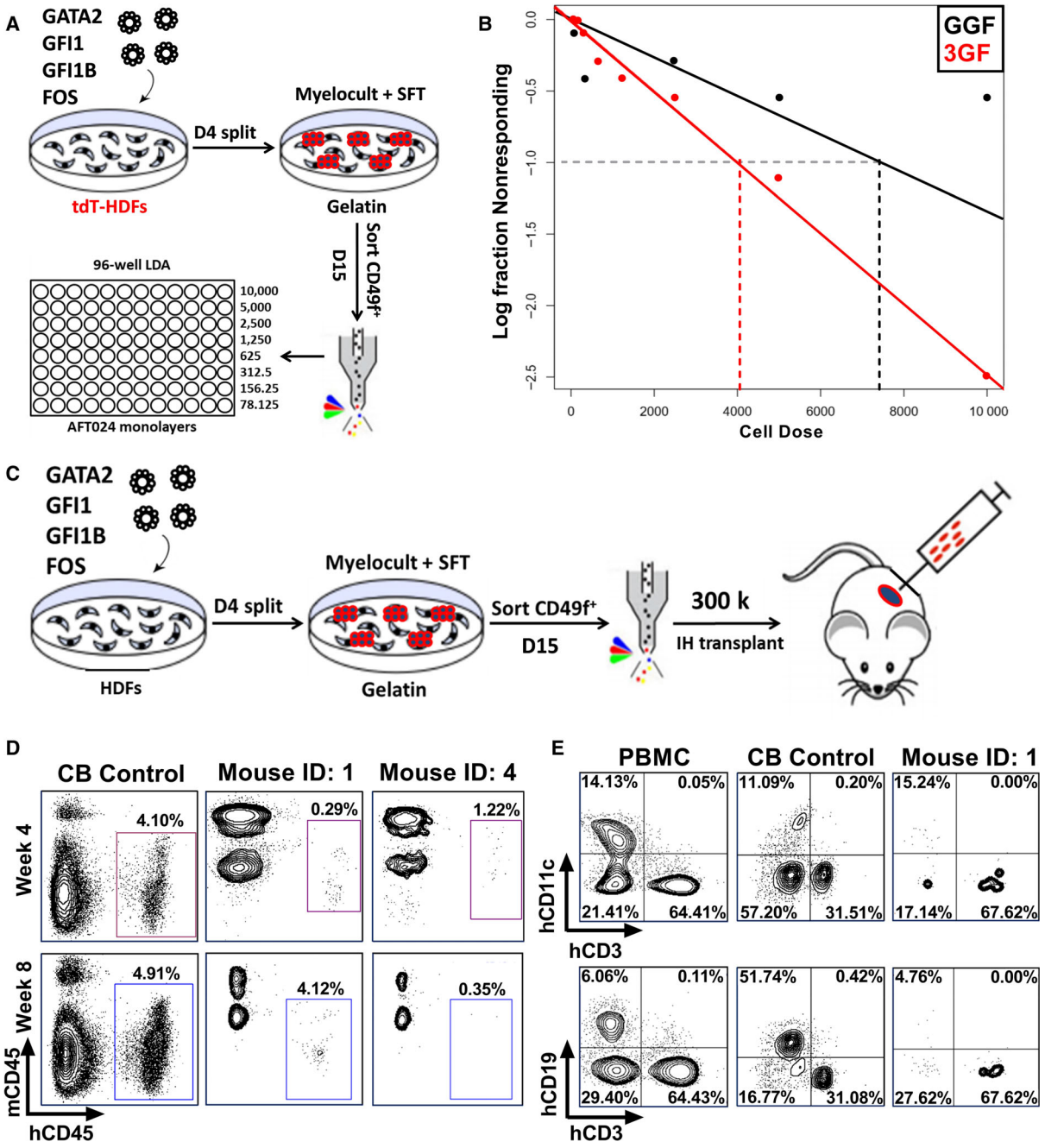


Fig. 7. 3GF reprogramming generates a greater frequency of HSC-like cells capable of short-term multilineage engraftment. (A) Experimental scheme for GGF and 3GF LDA using AFT024 monolayers. (B) ELDA of GGF and 3GF cells after 5 weeks of AFT024 LTC, with a 95% CI of 1/2740–1/5899 (frequency of 1/4020) for 3GF and 1/4834–1/11527 (frequency of 1/7465) for GGF. (C) Experimental scheme for intrahepatic transplants of D15 CD49f$^+$ sorted 3GF cells into newborn NSG mice. (D) Representative flow cytometry contour plots of hCD45 chimerism in CB controls and 2 mice that showed engraftment using D15 3GF cells. (E) Representative flow cytometry contour plots showing the multilineage distribution of Week 8 hCD45$^+$ cells from CB controls and D15 3GF IH transplants. CD3$^+$ T cells, CD19$^+$ B cells, and CD11c$^+$ myeloid cells. PB mononclear cells were used as a staining control (nontransplanted).

niche signals to support the further maturation of the reprogrammed cells similar to other published work [25]. The mothers were supplied with 1 mg·mL^{-1} of DOX in the drinking water for 2 weeks to maintain activation of the transgenes as the reprogramming cells interact with the *in vivo* niche.

Cells sorted from Day 12 or Day 18 reprogramming cultures for CD49f did not result in detectable hCD45 chimerism. D15 CD49f$^+$ 3GF cells engrafted two out of eight mice. These two mice, labeled as Mouse ID:1 and ID:4, had detectable levels of hCD45 chimerism in their PB at Week 4 of analysis post-transplant. At Week 8, however, Mouse ID:4 began to lose hCD45 engraftment, while mouse ID:1 hCD45$^+$ cells appeared to expand. Lin-CD34+ CB transplants are shown as a positive control (Fig. 7D). After these time points, however, the hCD45$^+$ chimerism became undetectable (at Weeks 12 and 16 of analysis, as well as in subsequent organ harvesting experiments, data not shown), suggesting short-term engraftment. Multilineage contribution of the hCD45$^+$ cells from Mouse ID:1 at Week 8, however, reveals multilinaege engraftment with a predominance of CD3$^+$ T cells, with smaller levels of CD11c$^+$ myeloid cells and CD19$^+$ B cells (Fig. 7E). This indicates that D15 3GF CD49f$^+$ cells grown on gelatin appear to possess limited hematopoietic functional potential. As such, it will be of great interest to expose 3GF reprogrammed cells to the AFT024 *in vitro* maturation system and subsequently transplant purified reprogrammed cells based on cell surface immunophenotype as previously discussed.

Discussion

A replenishable source of engraftable, autologous human blood cells can provide a potential foundation to study and ultimately cure a multitude of hematologic disorders. To address this issue, several studies sought to generate various blood products *de novo*. Previous reprogramming strategies, in both human iPSCs and somatic cells, remain limited in the identity of their final derived cells or have practical issues with either their starting cell populations or TF cocktails [17–28,32]. Thus far, the low efficiency and poor engraftment capabilities of cells derived through iPSC differentiation restrict the utility of this method, although recent breakthroughs reveal possible derivation of engraftable cells from human iPSCs utilizing both TF overexpression and directed differentiation [31]. In this study, we optimized our hemogenic induction process without going through pluripotency to yield HSC-like cells that parallel endogenous HSCs in their cell surface phenotype, gene expression profile, and functional potential. The new 3GF cocktail, as well as coculture on AFT024 stroma, improves the yield and functional output of the derived cells.

Previous work [30], as well as our current data, demonstrates that we can induce the same developmental program in both mouse and human fibroblasts to derive hematopoietic cells. Addition of GFI1 to the reprogramming cocktail (Fig. 1) highlights the importance of the axis formed by GFI1 and GFI1B in regulating human hematopoiesis and EHT, likely *via* RUNX and other pathways [73]. Both factors act as transcriptional repressors that recruit histone-modifying genes to the promoter and enhancer regions of their target genes [74]. HSCs, MPP1, and MPP2 cell populations all express GFI1 and GFI1B, with GFI1B expression being the highest in the earliest HSC compartment while GFI1 expression was found to increase as cells differentiate [75,76]. GFI1 targets several previously reported TFs used in HSC reprogramming, such as Hoxa9, Pbx1, Meis1, and PU.1 [77]. Knockout studies show that GFI1 prevents HSC proliferation and plays a critical role in maintaining HSC self-renewal capacity [78]. GFI1B knockout results in massive HSC proliferation and maintenance of self-renewal, but concomitant removal of GFI1 and GFI1B impairs HSC survival [76]. GATA2 directly activates GFI1B *via* promoter and distal enhancer element binding in ChIP-seq studies of GATA2 in primary mast cells. Data in this ChIP-seq set also suggest that GATA2, GFI1, and GFI1B exist in a regulatory triad, with GFI1 and GFI1B acting in a mutually inhibitory manner, GFI1 inhibiting GATA2, and GATA2 activating GFI1B [79]. Other studies show the absolute requirement of GFI1 and GFI1B together in EHT, where GFI1 in particular specifically defines the subset of HE that gives rise to emergent HSCs. A set of EHT genes bound and repressed by GFI1 and GFI1B also show cobinding with RUNX1, as well as a prevalence of RBPJ binding motifs, signifying a potential role for NOTCH signaling [80]. This suggests that these factors, along with NOTCH and RUNX1 signaling [81], orchestrate the derivation of hematopoietic cells from HE.

Our analyses focused, in part, on the markers CD49f and ACE, together with the known human HSC marker CD34 [53,82], all shown to enrich for the functional population of HSCs (Fig. 2). ACE marks the para-aortic splanchnopleura and the AGM regions that possess all the hematopoietic potential of the developing human embryo [52], as well as derived hemangioblasts from human pluripotent stem cells [51,83]. CD49f expression in human cells enriches for the engraftable HSC population [54]. Recent studies show that EPCR marks the functional subset of CB CD34$^+$ progenitors exposed to UM171 [84], a small molecule shown to expand HSCs *ex vivo* [85]. Interestingly, our derived cells stain positive for this newly defined CB HSC marker, yet stain negative for GPI-80, a glycophosphatidylinositol-anchored surface

protein that marks human FL HSPCs [86] (Fig. 3 and data not shown). This suggests possible subpopulations of phenotypic HSPCs marked by either EPCR or GPI-80, and that our cells more closely resemble those derived from CB.

RNA sequencing of four different 3GF populations (D15 $CD49f^+CD34^-$, D15 $CD49f^+CD34^+$, D25 $CD49f^+CD34^-$, and D25 $CD49f^+CD34^+$) reveals that inclusion of GFI1 to the reprogramming process results in acquisition of an endothelial fate that leads to emergence of a hematopoietic profile as cells mature or activate CD34 (Figs 4 and 5 and Fig. S3). While these cells follow the same developmental trajectory as GGF reprogrammed cells, they remain distinct from their GGF counterparts based on the day of analysis or cell surface marker sorting (Fig. 4B). The upregulated TF profiles, however, appear to differ as 3GF reprogramming demonstrates upregulation of multiple TFs known to play important roles in HSC biology (Fig. 4C). Interestingly, 3GF cells begin to downregulate machinery needed for cellular replication, suggesting that more mature and/or $CD34^+$ cells become more quiescent compared to the D15 $CD49f^+CD34^-$ population (Fig. S3A). Notably, normalized read count comparisons reveal a significant reduction of CDK6 in $CD49f^+CD34^+$ vs. $CD49f^+CD34^-$ populations within both D15 and D25 subsets (Fig. S3B). CDK6 is a known regulator of cell cycle progression absent in human long-term HSCs but found in abundance in short-term HSCs [87]. Adult HSCs are known to remain quiescent in the BM to maintain their function [63,88]. Through our reprogramming, it is possible that the early cells we derive resemble those that emerge from the AGM and/or the placenta. The reprogrammed cells upregulate the machinery required for processing various glycosaminoglycans (GAGs) and proteoglycans, suggesting that they are primed to take in these signals to further mature as they become hematopoietic (Fig. S3A). It is widely known that a variety of HSC *in vitro* niche systems, such as AFT024, provide key GAGs and proteoglycans that help maintain biological function *in vitro* [9,70,71,89–92]. All these findings align with our hypothesis that the reprogrammed cells are ready to process signals provided by tissues such as the FL, and thus need to expand and further mature to become functional after exposure to these signals.

To date, the full HSC *in vivo* niche remains incompletely understood, complicating attempts to reconstruct this complex signaling system *in vitro*. Several known signals play a large role in inducing and sustaining hematopoiesis, such as the NOTCH [81], mTOR [93], and Wnt/β-catenin pathways [94–96]. Using a vascular niche known to express physiological levels of key angiocrine signals such as NOTCH, BMP, and c-KIT, engraftable HSC-like cells emerge from reprogrammed endothelial cells [32,33]. Other work demonstrates improved expansion of iPSC-derived hematopoietic progenitors on endothelial monolayers overexpressing the NOTCH ligands JAG1 and DL4 [69]. Previous work from our lab demonstrated the importance of NOTCH activation in hemogenic precursor $Prominin1^+Sca1^+CD34^+CD45^-$ ($PS34CD45^-$) cells isolated from mouse placentas. In 0.1% gelatin or OP9 coculture, these $PS34CD45^-$ cells cannot engraft immunodeficient mice. Only after 4 days of stromal coculture on OP9-DL1, which provides a canonical NOTCH signal, can these $PS34CD45^-$ cells give rise to all blood lineages and engraft primary and secondary immunodeficient mice [34].

The stromal cell line AFT024, derived from murine FL, can support both mouse [36] and human [90] hematopoiesis *in vitro*. Shown to express key signals for sustaining hematopoiesis such as delta-like [97]—which constitutes a noncanonical ligand for NOTCH [98]—and dermatopontin [9], this cell line represents a component of the *in vivo* stem cell niche. AFT024 supports the *ex vivo* maintenance of human $CD34^+CD38^-$ HSPCs significantly more efficiently than other human-derived cell lines in a contact-dependent manner, as supported by our data as well (Fig. 6 and Fig. S4), highlighting the plethora of signals these cells specifically express to support hematopoiesis *in vitro* [70,71,91]. Conversion of Pro-B cells to HSC-like cells required maturation with the mouse *in vivo* niche prior to full functionality [25], demonstrating the importance of an instructive niche to assist and/or complete the reprogramming process, regardless of the starting cell population or the TF panel used in the reprogramming. It will be of great interest to further analyze our reprogrammed cells after AFT024 LTC. More comprehensive purification of the reprogrammed cells will permit a stronger determination of stem cell frequency in AFT024 LDA assays. Additionally, key analyses after AFT024 LTC will include RNAseq and transplantation of isolated reprogrammed cells at various levels of phenotypic HSC purity.

Conclusions

Altogether, our results demonstrate that inclusion of GFI1 to the original GGF TF cocktail, as well as coculture on AFT024 hematopoiesis-supporting stromal layers, is sufficient to generate HSC-like cells from human fibroblasts capable of multilineage functionality. This process remains dynamic and travels

through a HE-like intermediate characterized by previously identified markers that identify endothelium with hematopoietic potential. Our results further support the notions that hematopoietic specification is a step-wise process [99], that it traverses through endothelial intermediates [100], and that it requires variety of signals that can be provided by stromal cells for successful EHT and maturation [101]. To summarize, we show that manipulations to the previously established reprogramming cocktail and strategy in mouse and human fibroblasts can induce hemogenesis that leads to the production of HSC-like cells capable of multilineage function. The optimization of this process can also provide an *in vitro* platform for drug testing and hematopoietic disease modeling with the goal of identifying putative treatments for hematopoietic disorders and eventual avenues for autologous HSC transplants.

Acknowledgements

The authors would like to thank the Sunita L. D'Souza and the Pluripotent Stem Cell Core at the Icahn School of Medicine at Mount Sinai for helpful discussions and reagents. We also thank the flow cytometry shared resource Core as well as the Genomics Core Facility at ISMMS. KM was supported by NIH 5RO1HL119404 and NYSTEM C32597GG. MGD was supported by the grants 5RO1HL11904, 4R33AI116191-031 (through BC), and F31 HL136148-01.

Author contributions

MGD designed and performed experiments, acquired and analyzed resultant data, and wrote the manuscript. DS, JMB, YF, NS, KL, FP, AG, HSK, and KR performed experiments and edited the manuscript. CFP, BC, and IRL performed experiments, edited the manuscript, and guided the project. KAM designed and performed experiments, edited the manuscript, and managed the project.

References

1 Babovic S and Eaves CJ (2014) Hierarchical organization of fetal and adult hematopoietic stem cells. *Exp Cell Res* **329**, 185–191.
2 Notta F, Zandi S, Takayama N, Dobson S, Gan OI, Wilson G, Kaufmann KB, McLeod J, Laurenti E, Dunant CF *et al.* (2016) Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, 218–221.
3 Haas S, Trumpp A and Milsom MD (2018) Causes and consequences of hematopoietic stem cell heterogeneity. *Cell Stem Cell* **22**, 627–638.
4 Hirschi KK (2012) Hemogenic endothelium during development and beyond. *Blood* **119**, 4823–4827.
5 Zhen F, Lan Y, Yan B, Zhang W and Wen Z (2013) Hemogenic endothelium specification and hematopoietic stem cell maintenance employ distinct Scl isoforms. *Development* **140**, 3977–3985.
6 Guibentif C, Rönn RE, Böiers C, Lang S, Saxena S, Soneji S, Enver T, Karlsson G and Woods N-B (2017) Single-cell analysis identifies distinct stages of human endothelial-to-hematopoietic transition. *Cell Rep* **19**, 10–19.
7 Zovein AC, Hofmann JJ, Lynch M, French WJ, Turlo KA, Yang Y, Becker MS, Zanetta L, Dejana E, Gasson JC *et al.* (2008) Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* **3**, 625–636.
8 Easterbrook J, Rybtsov S, Gordon-Keylock S, Ivanovs A, Taoudi S, Anderson RA and Medvinsky A (2019) Analysis of the spatiotemporal development of hematopoietic stem and progenitor cells in the early human embryo. *Stem Cell Rep* **12**, 1056–1068
9 Kokkaliaris KD, Drew E, Endele M, Loeffler D, Hoppe PS, Hilsenbeck O, Schauberger B, Hinzen C, Skylaki S, Theodorou M *et al.* (2016) Identification of factors promoting *ex vivo* maintenance of mouse hematopoietic stem cells by long-term single-cell quantification. *Blood* **128**, 1181–1192.
10 Park B, Yoo KH and Kim C (2015) Hematopoietic stem cell expansion and generation: the ways to make a breakthrough. *Blood Res* **50**, 194–203.
11 Daniel MG, Pereira C-F, Lemischka IR and Moore KA (2016) Making a hematopoietic stem cell. *Trends Cell Biol* **26**, 202–214.
12 Daniel MG, Lemischka IR and Moore K (2016) Converting cell fates: generating hematopoietic stem cells *de novo* via transcription factor reprogramming. *Ann N Y Acad Sci* **1370**, 24–35.
13 Wilkinson AC, Ishida R, Kikuchi M, Sudo K, Morita M, Crisostomo RV, Yamamoto R, Loh KM, Nakamura Y, Watanabe M *et al.* (2019) Long-term *ex vivo* haematopoietic-stem-cell expansion allows nonconditioned transplantation. *Nature* **571**, 117–121.
14 Kanda J (2013) Effect of HLA mismatch on acute graft-versus-host disease. *Int J Hematol* **98**, 300–308.
15 Takahashi K and Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676.
16 Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K and Yamanaka S (2007)

Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872.
17 Wang Y, Yates F, Naveiras O, Ernst P and Daley GQ (2005) Embryonic stem cell-derived hematopoietic stem cells. *Proc Natl Acad Sci USA* **102**, 19081–19086.
18 McKinney-Freeman SL, Naveiras O and Daley GQ. Isolation of hematopoietic stem cells from mouse embryonic stem cells. *Curr Protoc Stem Cell Biol* 2008, **Chapter 1**: Unit 1F 3.
19 Vereide DT, Vickerman V, Swanson SA, Chu L-F, McIntosh BE and Thomson JA (2014) An expandable, inducible hemangioblast state regulated by fibroblast growth factor. *Stem Cell Rep* **3**, 1043–1057.
20 Bowles KM, Vallier L, Smith JR, Alexander MRJ and Pedersen RA (2006) HOXB4 overexpression promotes hematopoietic development by human embryonic stem cells. *Stem Cells* **24**, 1359–136.
21 Amabile G, Welner RS, Nombela-Arrieta C, D'Alise AM, Di Ruscio A, Ebralidze AK, Kraytsberg Y, Ye M, Kocher O, Neuberg DS *et al.* (2013) *In vivo* generation of transplantable human hematopoietic cells from induced pluripotent stem cells. *Blood* **121**, 1255–1264.
22 Suzuki N, Yamazaki S, Yamaguchi T, Okabe M, Masaki H, Takaki S, Otsu M and Nakauchi H (2013) Generation of engraftable hematopoietic stem cells from induced pluripotent stem cells by way of teratoma formation. *Mol Ther* **21**, 1424–1431.
23 Doulatov S, Vo LT, Chou SS, Kim PG, Arora N, Li H, Hadland BK, Bernstein ID, Collins JJ, Zon LI *et al.* (2013) Induction of multipotential hematopoietic progenitors from human pluripotent stem cells via respecification of lineage-restricted precursors. *Cell Stem Cell* **13**, 459–470.
24 Elcheva I, Brok-Volchanskaya V, Kumar A, Liu P, Lee JH, Tong L, Vodyanik M, Swanson S, Stewart R, Kyba M *et al.* (2014) Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. *Nat Commun* **5**, 4372.
25 Riddell J, Gazit R, Garrison BS, Guo G, Saadatpour A, Mandal PK, Ebina W, Volchkov P, Yuan G-C, Orkin SH *et al.* (2014) Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors. *Cell* **157**, 549–564.
26 Batta K, Florkowska M, Kouskoff V and Lacaud G (2014) Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *Cell Rep* **9**, 1871–1884.
27 Szabo E, Rampalli S, Risueño RM, Schnerch A, Mitchell R, Fiebig-Comyn A, Levadoux-Martin M and Bhatia M (2010) Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* **468**, 521–526.
28 Pulecio J, Nivet E, Sancho-Martinez I, Vitaloni M, Guenechea G, Xia Y, Kurian L, Dubova I, Bueren J, Laricchia-Robbio L *et al.* (2014) Conversion of human fibroblasts into monocyte-like progenitor cells. *Stem Cells* **32**, 2923–2938.
29 Pereira CF, Chang B, Qiu J, Niu X, Papatsenko D, Hendry CE, Clark NR, Nomura-Kitabayashi A, Kovacic JC, Ma'ayan A *et al.* (2013) Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* **13**, 205–218.
30 Gomes AM, Kurochkin I, Chang B, Daniel M, Law K, Satija N, Lachmann A, Wang Z, Ferreira L, Ma'ayan A *et al.* (2018) Cooperative transcription factor induction mediates hemogenic reprogramming. *Cell Rep* **25**, 2821–2835.e7.
31 Sugimura R, Jha DK, Han A, Soria-Valles C, da Rocha EL, Lu Y-F, Goettel JA, Serrao E, Rowe RG, Malleshaiah M *et al.* (2017) Haematopoietic stem and progenitor cells from human pluripotent stem cells. *Nature* **545**, 432–438.
32 Sandler VM, Lis R, Liu Y, Kedem A, James D, Elemento O, Butler JM, Scandura JM and Rafii S (2014) Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature* **511**, 312–318.
33 Lis R, Karrasch CC, Poulos MG, Kunar B, Redmond D, Duran JGB, Badwe CR, Schachterle W, Ginsberg M, Xiang J *et al.* (2017) Conversion of adult endothelium to immunocompetent haematopoietic stem cells. *Nature* **545**, 439–445.
34 Pereira CF, Chang B, Gomes A, Bernitz J, Papatsenko D, Niu X, Swiers G, Azzoni E, de Bruijn MF and Schaniel C *et al.* (2016) Hematopoietic reprogramming *in vitro* informs *in vivo* identification of hemogenic precursors to definitive hematopoietic stem cells. *Dev Cell* **36**, 525–539.
35 Tsukada M, Ota Y, Wilkinson AC, Becker HJ, Osato M, Nakauchi H and Yamazaki S (2017) *In vivo* generation of engraftable murine hematopoietic stem cells by Gfi1b, c-Fos, and Gata2 overexpression within teratoma. *Stem Cell Rep* **9**, 1024–1033.
36 Moore KA, Ema H and Lemischka IR (1997) *In vitro* maintenance of highly purified, transplantable hematopoietic stem cells. *Blood* **89**, 4337–4347.
37 Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H and Jaenisch R (2008) Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151–159.
38 Magnusson M, Sierra MI, Sasidharan R, Prashad SL, Romero M, Saarikoski P, Van Handel B, Huang A, Li X and Mikkola HKA (2013) Expansion on stromal cells preserves the undifferentiated state of human hematopoietic stem cells despite compromised reconstitution ability. *PLoS ONE* **8**, e53912.
39 Hu Y and Smyth GK (2009) ELDA: extreme limiting dilution analysis for comparing depleted and enriched

populations in stem cell and other assays. *J Immunol Methods* **347**, 70–78.
40 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M and Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
41 Liao Y, Smyth GK and Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930.
42 Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.
43 Luo W, Friedman MS, Shedden K, Hankenson KD and Woolf PJ (2009) GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161.
44 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25–29.
45 The Gene Ontology Consortium (2017) Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Res* **45**(D1), D331–D338.
46 Torre D, Lachmann A and Ma'ayan A (2018) BioJupies: automated generation of interactive notebooks for RNA-Seq data analysis in the cloud. *Cell Syst* **7**, 556–561.e3.
47 Kyba M, Perlingeiro RC and Daley GQ (2002) HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29–37.
48 Tavian M, Coulombel L, Luton D, Clemente HS, Dieterlen-Lièvre F and Péault B (1996) Aorta-associated CD34+ hematopoietic cells in the early human embryo. *Blood* **87**, 67–72.
49 Sommer CA, Stadtfeld M, Murphy GJ, Hochedlinger K, Kotton DN and Mostoslavsky G (2009) Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* **27**, 543–549.
50 Rasmussen MA, Holst B, Tümer Z, Johnsen MG, Zhou S, Stummann TC, Hyttel P and Clausen C (2014) Transient p53 suppression increases reprogramming of human fibroblasts without affecting apoptosis and DNA damage. *Stem Cell Rep* **3**, 404–413.
51 Zambidis ET, Sinka L, Tavian M, Jokubaitis V, Park T s, Simmons P and Peault B (2007) Emergence of human angiohematopoietic cells in normal development and from cultured embryonic stem cells. *Ann N Y Acad Sci* **1106**, 223–232.
52 Jokubaitis VJ, Sinka L, Driessen R, Whitty G, Haylock DN, Bertoncello I, Smith I, Peault B, Tavian M and Simmons PJ (2008) Angiotensin-converting enzyme (CD143) marks hematopoietic stem cells in human embryonic, fetal, and adult hematopoietic tissues. *Blood* **111**, 4055–4063.
53 Julien E, El Omar R and Tavian M (2016) Origin of the hematopoietic system in the human embryo. *FEBS Lett* **590**, 3987–4001.
54 Notta F, Doulatov S, Laurenti E, Poeppl A, Jurisica I and Dick JE (2011) Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science* **333**, 218–221.
55 Fares I,Chagraoui J, Lehnertz B, MacRae T, Mayotte N, Tomellini E, Aubert L, Roux PP and Sauvageau G (2017) EPCR expression marks UM171-expanded CD34(+) cord blood stem cells. *Blood* **129**, 3344–3351.
56 Balazs AB, Fabian AJ, Esmon CT and Mulligan RC (2006) Endothelial protein C receptor (CD201) explicitly identifies hematopoietic stem cells in murine bone marrow. *Blood* **107**, 2317–2321.
57 Gomes I, Sharma TT, Edassery S, Fulton N, Mar BG and Westbrook CA (2002) Novel transcription factors in human CD34 antigen-positive hematopoietic cells. *Blood* **100**, 107–119.
58 Gao L, Tober J, Gao P, Chen C, Tan K and Speck NA (2018) RUNX1 and the endothelial origin of blood. *Exp Hematol* **68**, 2–9.
59 Bockamp EO, McLaughlin F, Göttgens B, Murrell AM, Elefanty AG and Green AR (1997) Distinct mechanisms direct SCL/tal-1 expression in erythroid cells and CD34 positive primitive myeloid cells. *J Biol Chem* **272**, 8781–8790.
60 Bielczyk-Maczynska E, Serbanovic-Canic J, Ferreira L, Soranzo N, Stemple DL, Ouwehand WH and Cvejic A (2014) A loss of function screen of identified genome-wide association study Loci reveals new genes controlling hematopoiesis. *PLoS Genet* **10**, e1004450.
61 Gazit R, Garrison BS, Rao TN, Shay T, Costello J, Ericson J, Kim F, Collins JJ, Regev A, Wagers AJ *et al.* (2013) Transcriptome analysis identifies regulators of hematopoietic stem and progenitor cells. *Stem Cell Rep* **1**, 266–280.
62 Radtke S, Adair JE, Giese MA, Chan Y-Y, Norgaard ZK, Enstrom M, Haworth KG, Schefter LE and Kiem H-P (2017) A distinct hematopoietic stem cell population for rapid multilineage engraftment in nonhuman primates. *Sci Transl Med* **9**, eaan1145.
63 Seita J and Weissman IL (2010) Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med* **2**, 640–653.
64 Poulos MG, Crowley MJP, Gutkin MC, Ramalingam P, Schachterle W, Thomas J-L, Elemento O and Butler JM (2015) Vascular platform to define hematopoietic stem cell factors and enhance regenerative hematopoiesis. *Stem Cell Rep* **5**, 881–894.
65 King KY and Goodell MA (2011) Inflammatory modulation of HSCs: viewing the HSC as a

foundation for the immune response. *Nat Rev Immunol* **11**, 685–692.
66 Ramshaw HS, Haylock D, Swart B, Gronthos S, Horsfall MJ, Niutta S and Simmons PJ (2001) Monoclonal antibody BB9 raised against bone marrow stromal cells identifies a cell-surface glycoprotein expressed by primitive human hemopoietic progenitors. *Exp Hematol* **29**, 981–992.
67 Anzai N, Lee Y, Youn BS, Fukuda S, Kim YJ, Mantel C, Akashi M and Broxmeyer HE (2002) C-kit associated with the transmembrane 4 superfamily proteins constitutes a functionally distinct subunit in human hematopoietic progenitors. *Blood* **99**, 4413–4421.
68 Ishikawa T, Factor VM, Marquardt JU, Raggi C, Seo D, Kitade M, Conner EA and Thorgeirsson SS (2012) Hepatocyte growth factor/c-met signaling is required for stem-cell-mediated liver regeneration in mice. *Hepatology* **55**, 1215–1226.
69 Gori JL, Butler JM, Chan Y-Y, Chandrasekaran D, Poulos MG, Ginsberg M, Nolan DJ, Elemento O, Wood BL, Adair JE *et al.* (2015) Vascular niche promotes hematopoietic multipotent progenitor formation from pluripotent stem cells. *J Clin Invest* **125**, 1243–1254.
70 Nolta JA, Thiemann FT, Arakawa-Hoyt J, Dao M, Barsky LW, Moore KA, Lemischka IR and Crooks GM (2002) The AFT024 stromal cell line supports long-term *ex vivo* maintenance of engrafting multipotent human hematopoietic progenitors. *Leukemia* **16**, 352–361.
71 Thiemann FT, Moore KA, Smogorzewska EM, Lemischka IR and Crooks GM (1998) The murine stromal cell line AFT024 acts specifically on human CD34 + CD38- progenitors to maintain primitive function and immunophenotype *in vitro*. *Exp Hematol* **26**, 612–619.
72 van Os RP, Dethmers-Ausema B and de Haan G (2008) *In vitro* assays for cobblestone area-forming cells, LTC-IC, and CFU-C. *Methods Mol Biol* **430**, 143–157.
73 Lancrin C, Mazan M, Stefanska M, Patel R, Lichtinger M, Costa G, Vargel O, Wilson NK, Moroy T, Bonifer C *et al.* (2012) GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood* **120**, 314–322.
74 Moroy T, Vassen L, Wilkes B and Khandanpour C (2015) From cytopenia to leukemia: the role of Gfi1 and Gfi1b in blood formation. *Blood* **126**, 2561–2569.
75 Hock H, Hamblen MJ, Rooke HM, Schindler JW, Saleque S, Fujiwara Y and Orkin SH (2004) Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* **431**, 1002–1007.
76 Khandanpour C, Sharif-Askari E, Vassen L, Gaudreau M-C, Zhu J, Paul WE, Okayama T, Kosan C and Moroy T (2010) Evidence that growth factor independence 1b regulates dormancy and peripheral blood mobilization of hematopoietic stem cells. *Blood* **116**, 5149–5161.
77 Phelan JD, Shroyer NF, Cook T, Gebelein B and Grimes HL (2010) Gfi1-cells and circuits: unraveling transcriptional networks of development and disease. *Curr Opin Hematol* **17**, 300–307.
78 Zeng H, Yücel R, Kosan C, Klein-Hitpass L and Möröy T (2004) Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *EMBO J* **23**, 4116–4125.
79 Moignard V, Macaulay IC, Swiers G, Buettner F, Schütte J, Calero-Nieto FJ, Kinston S, Joshi A, Hannah R, Theis FJ *et al.* (2013) Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat Cell Biol* **15**, 363–372.
80 Thambyrajah R, Mazan M, Patel R, Moignard V, Stefanska M, Marinopoulou E, Li Y, Lancrin C, Clapes T, Möröy T *et al.* (2016) GFI1 proteins orchestrate the emergence of haematopoietic stem cells through recruitment of LSD1. *Nat Cell Biol* **18**, 21–32.
81 Burns CE, Traver D, Mayhall E, Shepard JL and Zon LI (2005) Hematopoietic stem cell fate is established by the Notch-Runx pathway. *Genes Dev* **19**, 2331–232.
82 Civin CI, Strauss LC, Brovall C, Fackler MJ, Schwartz JF and Shaper JH (1984) Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells. *J Immunol* **133**, 157–165.
83 Zambidis ET, Soon Park T, Yu W, Tam A, Levine M, Yuan X, Pryzhkova M and Peault B (2008) Expression of angiotensin-converting enzyme (CD143) identifies and regulates primitive hemangioblasts derived from human pluripotent stem cells. *Blood* **112**, 3601–3614.
84 Fares I, Chagraoui J, Lehnertz B, MacRae T, Mayotte N, Tomellini E, Aubert L, Roux PP and Sauvageau G (2017) EPCR expression marks UM171-expanded CD34 + cord blood stem cells. *Blood* **129**, 3344–3351.
85 Fares I, Chagraoui J, Gareau Y, Gingras S, Ruel R, Mayotte N, Csaszar E, Knapp DJ, Miller P and Ngom M *et al.* (2014) Cor blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science* **345**, 1509–1512.
86 Prashad SL, Calvanese V, Yao CY, Kaiser J, Wang Y, Sasidharan R, Crooks G, Magnusson M and Mikkola HKA (2015) GPI-80 defines self-renewal ability in hematopoietic stem cells during human development. *Cell Stem Cell* **16**, 80–87.
87 Laurenti E, Frelin C, Xie S, Ferrari R, Dunant CF, Zandi S, Neumann A, Plumb I, Doulatov S, Chen J *et al.* (2015) CDK6 levels regulate quiescence exit in human hematopoietic stem cells. *Cell Stem Cell* **16**, 302–313.

88 Hao S, Chen C and Cheng T (2016) Cell cycle regulation of hematopoietic stem or progenitor cells. *Int J Hematol* **103**, 487–497.
89 Punzel M, Gupta P, Roodell M, Mortari F and Verfaillie CM (1999) Factor(s) secreted by AFT024 fetal liver cells following stimulation with human cytokines are important for human LTC-IC growth. *Leukemia* **13**, 1079–1084.
90 Punzel M, Gupta P and Verfaillie CM (2002) The microenvironment of AFT024 cells maintains primitive human hematopoiesis by counteracting contact mediated inhibition of proliferation. *Cell Commun Adhes* **9**, 149–159.
91 Hackney JA, Charbord P, Brunk BP, Stoeckert CJ, Lemischka IR and Moore KA (2002) A molecular profile of a hematopoietic stem cell niche. *Proc Natl Acad Sci USA* **99**, 13061–13066.
92 Papy-Garcia D and Albanese P (2017) Heparan sulfate proteoglycans as key regulators of the mesenchymal niche of hematopoietic stem cells. *Glycoconj J* **34**, 377–391.
93 Kentsis A and Look AT (2012) Distinct and dynamic requirements for mTOR signaling in hematopoiesis and leukemogenesis. *Cell Stem Cell* **11**, 281–282.
94 Lopez-Ruano G, Prieto-Bermejo R, Ramos TL, San-Segundo L, Sánchez-Abarca LI, Sánchez-Guijo F, Pérez-Simón JA, Sánchez-Yagüe J, Llanillo M and Hernández-Hernández Á (2015) PTPN13 and beta-catenin regulate the quiescence of hematopoietic stem cells and their interaction with the bone marrow niche. *Stem Cell Rep* **5**, 516–531.
95 Khan NI and Bendall LJ (2006) Role of WNT signaling in normal and malignant hematopoiesis. *Histol Histopathol* **21**, 761–774.
96 Lento W, Congdon K, Voermans C, Kritzik M and Reya T (2013) Wnt signaling in normal and malignant hematopoiesis. *Cold Spring Harb Perspect Biol* **5**, a008011. https://doi.org/10.1101/cshperspect.a008011
97 Moore KA, Pytowski B, Witte L, Hicklin D and Lemischka IR (1997) Hematopoietic activity of a stromal cell transmembrane protein containing epidermal growth factor-like repeat motifs. *Proc Natl Acad Sci USA* **94**, 4011–4016.
98 Traustadottir GA, Jensen CH, Thomassen M, Beck HC, Mortensen SB, Laborda J, Baladrón V, Sheikh SP and Andersen DC (2016) Evidence of non-canonical NOTCH signaling: Delta-like 1 homolog (DLK1) directly interacts with the NOTCH1 receptor in mammals. *Cell Signal* **28**, 246–254.
99 Pearson S, Sroczynska P, Lacaud G and Kouskoff V (2008) The stepwise specification of embryonic stem cells to hematopoietic fate is driven by sequential exposure to Bmp4, activin A, bFGF and VEGF. *Development* **135**, 1525–1535.
100 Choi KD, Vodyanik MA, Togarrati PPriya, Suknuntha K, Kumar A, Samarjeet F, Probasco MD, Tian S, Stewart R, Thomson JA *et al.* (2012) Identification of the hemogenic endothelial progenitor and its direct precursor in human pluripotent stem cell differentiation cultures. *Cell Rep* **2**, 553–567.
101 Weiss L and Sakai H (1984) The hematopoietic stroma. *Am J Anat* **170**, 447–463.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. TF expression analyses in human cells and tissues. Information related to Figures 1 and 2.

Fig. S2. CD49f + ACE + or CD49f + EPCR + cells within the positive and negative CD34 populations.

Fig. S3. GO term analysis and CDK6 quantification. Information related to Figures 4 and 5.

Fig. S4. *In vitro* analysis of Lin-CD34 + CB cells. Information related to Figures 6 and 7.

Fig. S5. Observed differences of GGF LDA and comparative flow analysis of 3GF LDA from gelatin and AFT024 LTC. Information related to Figures 6 and 7.

Table S1. TF cloning and selection.

Table S2. Endothelial and hematopoietic gene lists for RNAseq analysis.

Table S3. Cell counts and ELDA for CB *in vitro* functional analysis.