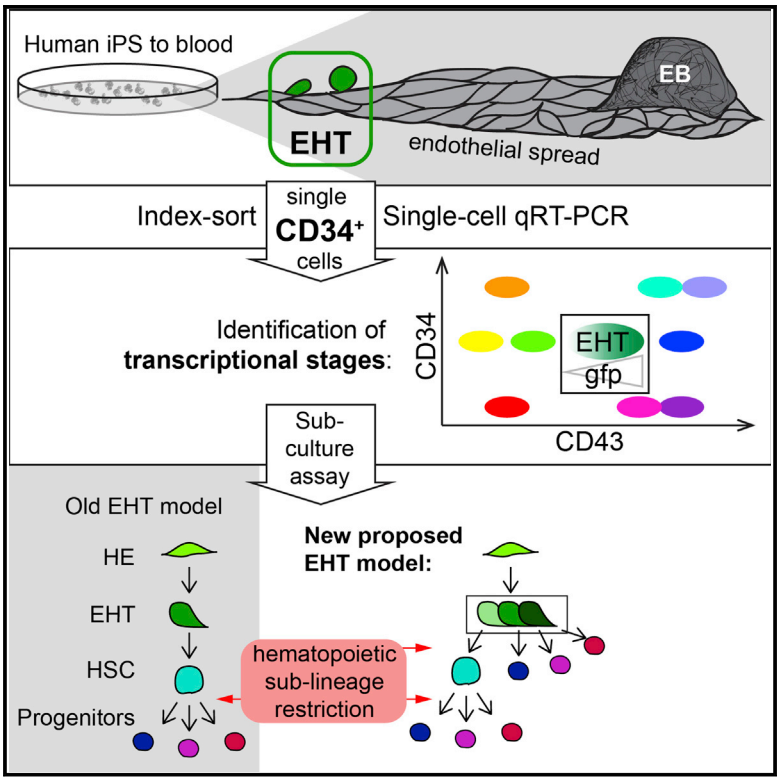
Report

Cell Reports

Single-Cell Analysis Identifies Distinct Stages of Human Endothelial-to-Hematopoietic Transition

Graphical Abstract



Authors

Carolina Guibentif, Roger Emanuel Rönn, Charlotta Böiers, ..., Tariq Enver, Göran Karlsson, Niels-Bjarne Woods

Correspondence

goran.karlsson@med.lu.se (G.K.), niels-bjarne.woods@med.lu.se (N.-B.W.)

In Brief

Molecular understanding of the developmental process of endothelial-to-hematopoietic transition (EHT) is critical for ultimately deriving transplantable hematopoietic stem cells from human pluripotent stem cells (hPSCs). Guibentif et al. provide a single-cell transcriptional characterization of this process using hPSC-derived cultures, revealing hematopoietic sublineage restriction before completion of EHT.

Highlights

- Single-cell transcriptional analysis identifies novel cell populations undergoing EHT
- EHT cluster expresses both endothelial and key hematopoietic genes
- Hematopoietic differentiation potential is restricted prior to completion of EHT
- Identification of sequential transcriptional signatures in human EHT

Accession Numbers

GSE87422

Guibentif et al., 2017, Cell Reports *19*, 10–19
April 4, 2017
http://dx.doi.org/10.1016/j.celrep.2017.03.023

OPEN ACCESS
CellPress

Cell Reports
Report

Single-Cell Analysis Identifies Distinct Stages of Human Endothelial-to-Hematopoietic Transition

Carolina Guibentif,[1] Roger Emanuel Rönn,[1] Charlotta Böiers,[1] Stefan Lang,[2] Shobhit Saxena,[1] Shamit Soneji,[2] Tariq Enver,[1,3] Göran Karlsson,[2,*] and Niels-Bjarne Woods[1,4,*]

[1]Section of Molecular Medicine and Gene Therapy, Lund Stem Cell Center, Lund University, BMC A12, 221 84 Lund, Sweden
[2]Division of Molecular Hematology, Lund Stem Cell Center, Lund University, BMC B12, 221 84 Lund, Sweden
[3]Stem Cell Laboratory, UCL Cancer Institute, University College London, London W1CE 6BT, UK
[4]Lead Contact
*Correspondence: goran.karlsson@med.lu.se (G.K.), niels-bjarne.woods@med.lu.se (N.-B.W.)
http://dx.doi.org/10.1016/j.celrep.2017.03.023

SUMMARY

During development, hematopoietic cells originate from endothelium in a process known as endothelial-to-hematopoietic transition (EHT). To study human EHT, we coupled flow cytometry and single-cell transcriptional analyses of human pluripotent stem cell-derived CD34$^+$ cells. The resulting transcriptional hierarchy showed a continuum of endothelial and hematopoietic signatures. At the interface of these two signatures, a unique group of cells displayed both an endothelial signature and high levels of key hematopoietic stem cell-associated genes. This interphase group was validated via sort and subculture as an immediate precursor to hematopoietic cells. Differential expression analyses further divided this population into subgroups, which, upon subculture, showed distinct hematopoietic lineage differentiation potentials. We therefore propose that immediate precursors to hematopoietic cells already have their hematopoietic lineage restrictions defined prior to complete downregulation of the endothelial signature. These findings increase our understanding of the processes of de novo hematopoietic cell generation in the human developmental context.

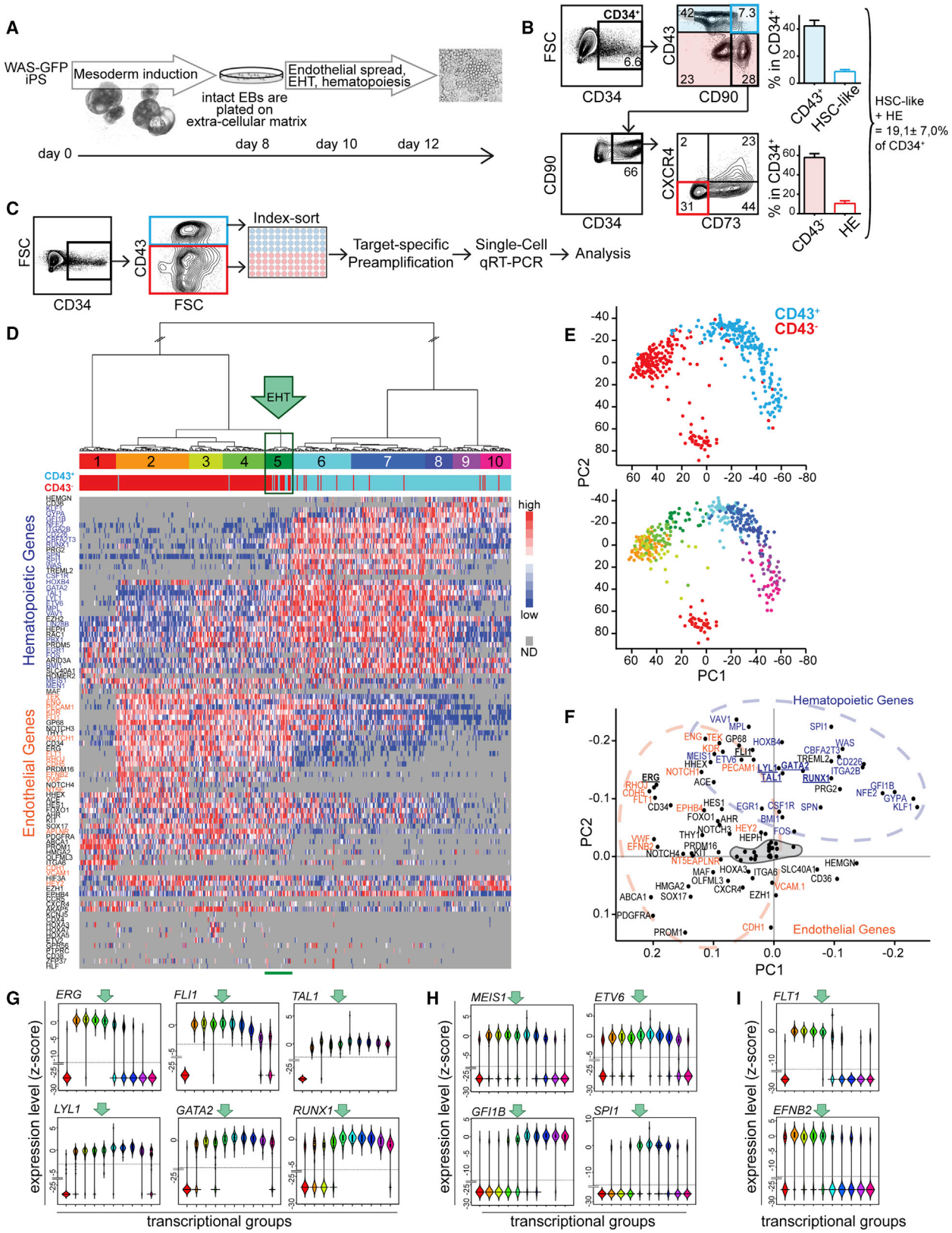
INTRODUCTION

During embryonic development, hematopoietic cells emerge from a precursor in the endothelial layer lining the ventral side of the dorsal aorta in the aorta-gonad-mesonephros (AGM) region. This conversion of endothelium into hematopoietic cells, termed endothelial-to-hematopoietic transition (EHT), has been extensively verified in various animal models (Bertrand et al., 2010; Boisset et al., 2010; Chen et al., 2009; Jaffredo et al., 1998; Kissa and Herbomel, 2010; Zovein et al., 2008) and has been observed in live-cell imaging studies of human pluripotent stem cell (hPSC) differentiation in vitro (Choi et al., 2012; Ditadi et al., 2015; Rafii et al., 2013). Although EHT was reported in the human embryo (Oberlin et al., 2002), no detailed molecular analysis of human cells undergoing EHT has been performed. Limitations to the study of EHT in human embryos make in vitro differentiation of hPSCs an attractive model for dissecting this developmental process. Two independent studies using hPSC differentiation reported a population enriched for the endothelial precursor of definitive hematopoiesis, termed hemogenic endothelium (HE), and highlight its transient nature and low frequency (Choi et al., 2012; Kennedy et al., 2012).

However, the molecular mechanisms regulating de novo emergence of blood cannot be fully understood until the transitional states in EHT are identified and characterized. Moreover, murine AGM imaging and transplantation studies show that while numerous hematopoietic clusters emerge on the walls of the dorsal aorta, very few of these cells are hematopoietic stem cells (HSCs) capable of adult multi-lineage reconstitution (Boisset et al., 2015; Kumaravelu et al., 2002; Taoudi et al., 2008; Yokomizo and Dzierzak, 2010), indicating heterogeneity among in vivo emerging blood.

Single-cell transcriptional analysis (reviewed in Hoppe et al., 2014) is an ideal tool to investigate lineage hierarchies in the heterogeneous populations undergoing EHT. Recent single-cell studies in the mouse have identified the onset of a hematopoietic regulatory program in the early stages of embryonic development (Moignard et al., 2015) and characterized a putative HE population from murine AGM (Swiers et al., 2013). In this latter study, distinct stages in the EHT process can be visualized for different populations at successive developmental time-points. Evaluation of the roles of specific transcription factors (TFs) in EHT at the single-cell level identified novel players in this process (Thambyrajah et al., 2016; Wilkinson et al., 2014). Recently, single-cell RNA sequencing was used to study the heterogeneity of prospectively isolated murine populations, including endothelial cells, pre-HSCs, and HSCs (Zhou et al., 2016).

To characterize EHT in the human developmental context, we performed single-cell transcriptional analysis of hPSC-derived CD34$^+$ cells. We identified a rare subset of cells sharing high numbers of transcripts of both endothelial and hematopoietic signatures, and determined this EHT subpopulation to be developmentally between HE and newly emerged hematopoietic cells. Moreover, we demonstrate that hematopoietic lineage differentiation potentials are already restricted in this narrow EHT window,



(legend on next page)

concurrent with the upregulation of the hematopoietic and downregulation of the endothelial programs. Using single-cell analysis, we provide further insight into endothelial and hematopoietic lineage commitment in the human developmental context.

RESULTS

hPSC-Derived CD34$^+$ Fraction Is a Suitable Population for Single-Cell Transcriptional Analysis of EHT

To dissect human EHT, we used in vitro differentiation of induced pluripotent stem cells (iPSCs) to the hematopoietic lineage (Rönn et al., 2015) (Figures 1A and S1A–S1C). Time-course fluorescence-activated cell sorting (FACS) analysis for the expression of HE markers and CD43$^+$ developing hematopoietic cells (Vodyanik et al., 2006) suggested that day 10 was an ideal time point to capture EHT cells (Figure S1D). At this time point, the iPSC-derived CD34$^+$ fraction contained the CD43$^-$CD34^{hi}CD90^{hi}CD73$^-$CXCR4$^-$ HE phenotype (10% of CD34$^+$) as well as our previously described HSC-like phenotype, CD43$^+$CD34$^+$CD90$^+$ (also 10% of CD34$^+$) (Majeti et al., 2007; Rönn et al., 2017) (Figure 1B). The remaining 80% of cells in the CD34$^+$ fraction include a broad repertoire of other known human endothelial and hematopoietic cells (Ivanovs et al., 2014; Kaufman et al., 2001; Kennedy et al., 2012; Tavian et al., 1996) and presumably as-yet-unidentified EHT-related cells.

The hPSC-Derived CD34$^+$ Cell Fraction Contains a Continuum Spanning Endothelial and Hematopoietic Transcriptional States, Including a Cell Population with a Dual Endothelial and HSC Signature

Day 10 iPSC-derived CD34$^+$ cells containing both endothelial and hematopoietic fractions (CD43-negative and positive cells, respectively) were index sorted for subsequent single-cell transcriptional analyses (Figure 1C). The iPSC line used had previously been transduced with a lentiviral vector expressing GFP under the control of the pan hematopoietic *WAS* gene promoter sequences. This WAS-GFP reporter was previously shown to track early hematopoietic commitment during hPSC differentiation (Muñoz et al., 2012) (Figures S1E–S1G). As index sorting allows us to determine the FACS phenotype of each transcriptionally analyzed single cell, the cells were simultaneously analyzed for WAS-GFP and for CD90, CD73, and CXCR4, the additional markers relevant in HE and HSC-like phenotypes. As expected, a fraction of the 437 analyzed CD34$^+$ cells consisted of either HE or HSC-like cells (Figures S1H and S1I). Expression levels of 91 genes related to endothelium, hematopoietic stem and progenitor cells (HSPCs), as well as EHT were assessed on each cell (Table S1).

Spanning endothelial and hematopoietic programs, the cells were divided into ten groups by unsupervised hierarchical clustering (Figure 1D). Principal-component analysis (PCA) distributed the cells from nine of the ten groups in a continuum (Figure 1E). Retrospective coloring based on CD43 sorting phenotype revealed that hematopoietic and non-hematopoietic cells grouped separately, with groups 1–5 being almost exclusively CD43$^-$ and groups 6–10 being CD43$^+$ (Figures 1D and 1E). No plate-specific bias of independently run qRT-PCR analyses was seen (Figure S1J), and index-sort data confirmed consistency between protein expression levels of surface markers and their transcriptional state. Notably, a similar correlation was observed between the WAS-GFP reporter fluorescence and *WAS* transcript levels (Figure S1K).

We proceeded to analyze the identified transcriptional signatures. Groups 2–5 had a clear endothelial program, where group 2 expressed all endothelial genes, and as the hierarchically ordered groups proceeded, there was a progressive downregulation of endothelial gene subsets (Figures 1D and S2Ai). Groups 6–8 had a hematopoietic signature, including many genes relevant for HSC function (Figures 1D and S2Aii). Most of these HSC-affiliated genes were downregulated in groups 9 and 10, which in turn expressed higher levels of genes associated with erythro-myeloid maturation (Figures 1D and S2Aiii). Interestingly, group 5 represented a population at the interface of endothelial and hematopoietic groups that, in addition to an endothelial signature, also displayed increased levels of key hematopoietic genes (green arrow in Figure 1D; see also Figure S2Aiv). Group 1 clustered separately from the transcriptional continuum (Figures 1D and 1E), lacked expression of most endothelial and

Figure 1. The hPSC-Derived CD34$^+$ Cell Fraction Contains a Continuum of Transcriptional States Spanning Endothelium to Blood

(A) Overview of iPSC-to-blood differentiation protocol.

(B) FACS analysis of iPSC-to-blood cultures at day 10 of the differentiation protocol with gating strategy for the immunophenotypes corresponding to HSC-like (CD34$^+$CD43$^+$CD90$^+$, blue outline) and HE (CD43$^-$CD34^{hi}CD90^{hi}CD73$^-$CXCR4$^-$, red outline). Bar graphs show the percentage of CD43$^+$ and CD43$^-$ fractions (in blue fill and pink fill, respectively, corresponding to the shades in the FACS plot), as well as the HSC-like and HE immunophenotypes (in empty columns, blue and red outlines, respectively) within the CD34$^+$ fraction of day 10 iPSC-to-blood cultures (n = 3 independent experiments; data represent mean ± SEM).

(C) Workflow for single-cell surface marker and transcriptional analysis of CD34$^+$CD43$^{+/-}$ fractions of day 10 iPSC-to-blood cultures. Using index sort, the cells were also analyzed for WAS-GFP, CD90, CD73, and CXCR4.

(D) Unsupervised hierarchical clustering, where each group is colored. Second row: retrospective coloring according to CD43-based sorting. On the left, the list of genes is color-coded; hematopoietic genes are in blue, endothelial genes are in orange, and genes that were not expected to help the distinction between endothelium and blood (i.e., implicated in both or none of these two cell types) are in black.

(E) PCA of single cells. Top: cells retrospectively colored by CD43 sorting phenotype. Bottom: cells colored by groups as identified by unsupervised hierarchical clustering in (D).

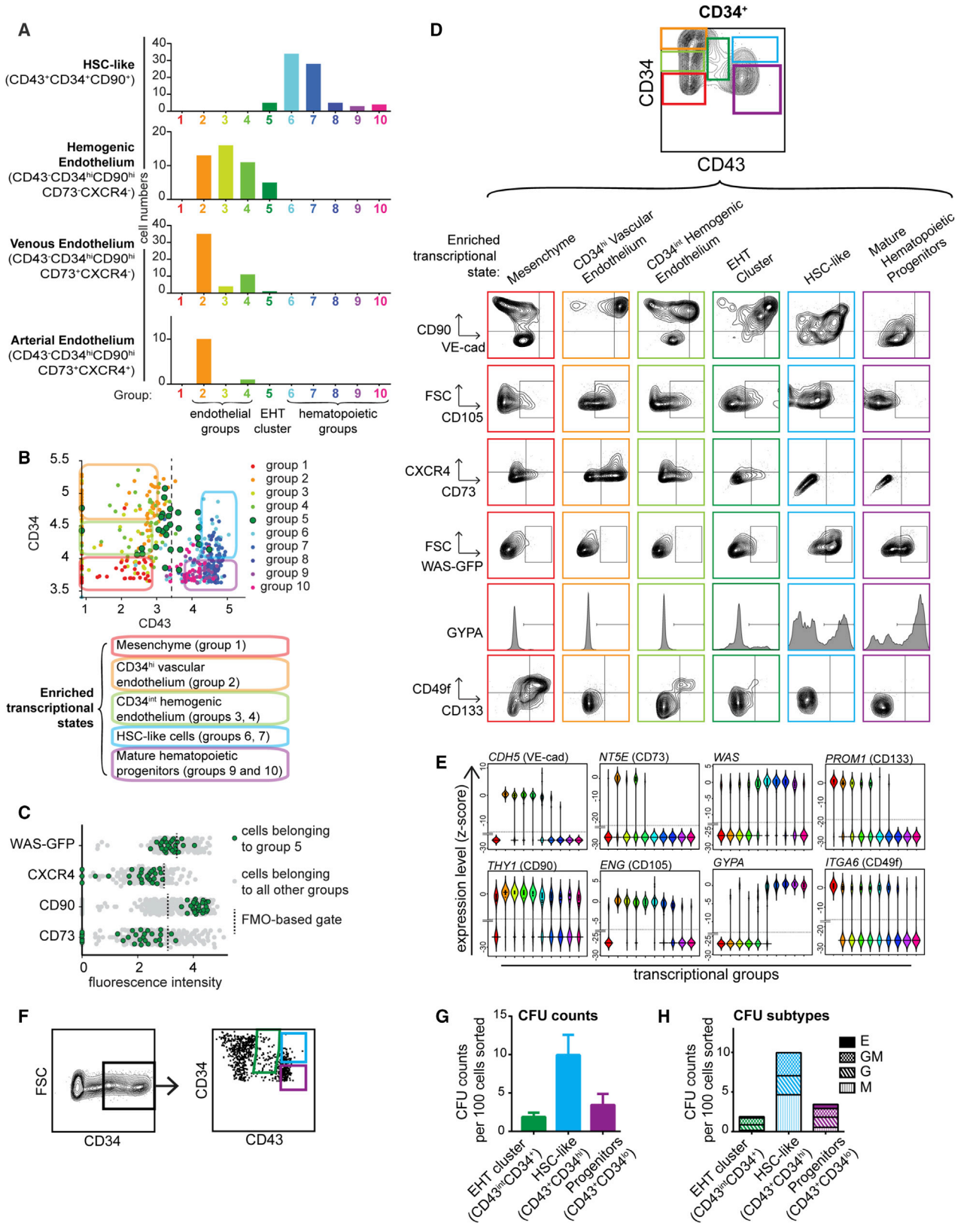
(F) Gene loadings of the PCA analysis. Members of the heptad are underlined. Genes in the center of the plot (with loading values close to 0) are *ZFP37*, *RAC1*, *PTPRC*, *PRDM5*, *PBX1*, *MEN1*, *LIN28B*, *KCNJ5*, *HOXA7*, *HOXA5*, *HOMER2*, *HLF*, *HIF3A*, *GPR56*, *EZH2*, *ETV2*, *CDX4*, *CD38*, *CCR5*, *ARID3A*, and *AKAP5*.

(G) Bean plots for the heptad genes included in this analysis.

(H) Bean plots for additional HSC-affiliated genes.

(I) Bean plots for endothelial genes.

Green arrows above bean plots highlight EHT cluster (group 5). See also Figures S1 and S2 and Table S1.



(legend on next page)

hematopoietic genes, and expressed genes suggestive of a mesenchymal cell (Figures 1D and S2Av) (Bakondi et al., 2009; Lee et al., 2009; Mabuchi et al., 2013; Moignard et al., 2015).

To gain further insight into the gene expression dynamics, we examined the gene loadings of the PCA and observed a separation between endothelial and hematopoietic genes. Heptad complex members (Beck et al., 2013; Wilson et al., 2010) were distributed in the upper portion of the gene loadings plot (Figure 1F, genes in underlined font). *TAL1*, *LYL1*, *GATA2*, and *RUNX1* appeared to cluster in the gene loadings plot, suggesting a cooperation in human EHT, and were notably expressed in group 5 (Figure 1G). *ERG* and *FLI1* were also expressed in group 5 and were clearly defining the endothelial groups, while *FLI1* was also expressed in the first hematopoietic groups (Figure 1G). Other key hematopoietic TFs (e.g., *MEIS1*, *ETV6*, *GFI1B*, and *SPI1*) displayed distinct expression patterns in group 5 and surrounding groups (Figure 1H). Finally, two endothelial genes (*FLT1* and *EFNB2*) were both highly expressed in groups 2–5 while being less expressed or not detected in the hematopoietic groups 6–10 (Figure 1I).

Taken together, we hypothesize that the group 5 cells, at the interface of the endothelial and hematopoietic programs, are in a narrow window of the EHT process, just prior to becoming bona fide hematopoietic cells, and we henceforth designate this transcriptional cluster as the EHT cluster.

The EHT Cluster and Surrounding Transcriptional States Identified by Single-Cell qRT-PCR Correspond to Distinct Endothelial and Hematopoietic Populations as Assessed by FACS

To elucidate how previously identified FACS-based populations related to our identified transcriptional groups, we used the index-sort data for backtracking the surface phenotype of each analyzed cell. We observed that cells with the HSC-like immunophenotype were predominantly found downstream of the EHT cluster, with groups 6 and 7 (Figure 2A) showing highest expression levels of HSC-affiliated genes (Figures 1D, 1G, and 1H). We also observed that HE cells were found scattered throughout groups 2 to 5 (highest in group3), indicating heterogeneity of this immunophenotype. Cells corresponding to arterial and venous endothelium as described in Ditadi et al. (2015) were found almost exclusively in group 2 (Figure 2A).

We then used the index-sort data to define sortable populations of EHT cluster cells and of other identified transcriptional states and observed that the different transcriptional groups already clustered in distinct populations based on CD34 and CD43 surface expression (Figure 2B). Specifically, the $CD43^{-}CD34^{hi}$ population enriched for the vascular endothelium (group 2), $CD43^{-}CD34^{int}$ enriched for HE cells (groups 3 and 4), $CD43^{+}CD34^{hi}$ enriched for HSC-like transcriptional signature (groups 6 and 7), and $CD43^{+}CD34^{low}$ enriched for the erythro-myeloid-committed signature (groups 9 and 10) (Figure 2B). Interestingly, we observed that the EHT cluster localized mostly at the interface between the $CD43^{-}$ and $CD43^{+}$ populations (Figure 2B, large green dots). These EHT cluster cells were almost exclusively $CD90^{hi}$, $CXCR4^{-}$, and $CD73^{-}$ and contained both WAS-GFP^{-} and WAS-GFP^{+} cells (Figure 2C). This enrichment strategy and the endothelial and hematopoietic cell identities of the transcriptional groups were confirmed by FACS analyses in additional independent experiments (Figure 2D). Notably, the transcript and protein levels as assessed by FACS for new surface markers were highly correlated and similarly up- or downregulated in the corresponding groups, including the WAS-GFP reporter and endogenous *WAS* transcript, whose expression was initiated in some cells within the EHT cluster and increased in the hematopoietic groups (Figures 2D and 2E). Interestingly, glycophorin A (GYPA) expression was also initiated in the EHT cluster. Moreover, sorting the identified transcriptional groups (EHT cluster, HSC-like, and erythro-myeloid committed cells) and plating into a colony-forming unit (CFU) assay revealed differences in both colony numbers and subtypes, demonstrating functional differences between the populations (Figures 2F–2H).

The EHT Transcriptional Cluster Comprises Subgroups with Distinct Levels of Hematopoietic Commitment Both Transcriptionally and in Differentiation Potential

To further dissect EHT, we performed a new single-cell qRT-PCR analysis on EHT cluster-enriched cells ($CD43^{int}CD34^{+}CD90^{hi}$) (Figure 3A). Combining this expression data with that of the initially sorted material increased the complexity of the EHT cluster, which could then be divided into three subgroups by hierarchical clustering (Figure 3B). These subgroups had progressively increased expression of a subset of hematopoietic genes, including endogenous *WAS*, *GFI1B*, and *ETV6* (Figure 3C). These differences correlated with a progressive increase in CD43

Figure 2. The Different Transcriptional Groups Identified by Single-Cell Analysis Correspond to Distinct Populations as Assessed by FACS

(A) Distribution, throughout the ten transcriptional groups identified in Figure 1D, of the single cells displaying the previously characterized surface phenotypes (using index-sort data) stated on the left.

(B) Scatterplot of the single cells according to surface expression of CD34 and CD43 using index-sort data. The single cells are colored according to the transcriptional group they belong to. Putative FACS gates were drawn to enrich for the transcriptional groups stated in the legend below. Fluorescence intensity in $\log_{10}$.

(C) FACS phenotype characterization of transcriptional group 5. Expression levels for cells belonging to group 5 EHT cluster cells are in dark green, and the remaining cells are in gray. Fluorescence intensity in $\log_{10}$.

(D) Validation of putative gates drawn in (B). Below, the color of the frame of the plots displays their correspondence to the gate of same color in the above CD34/CD43 FACS plot. Plots are representative of two to four independent experiments.

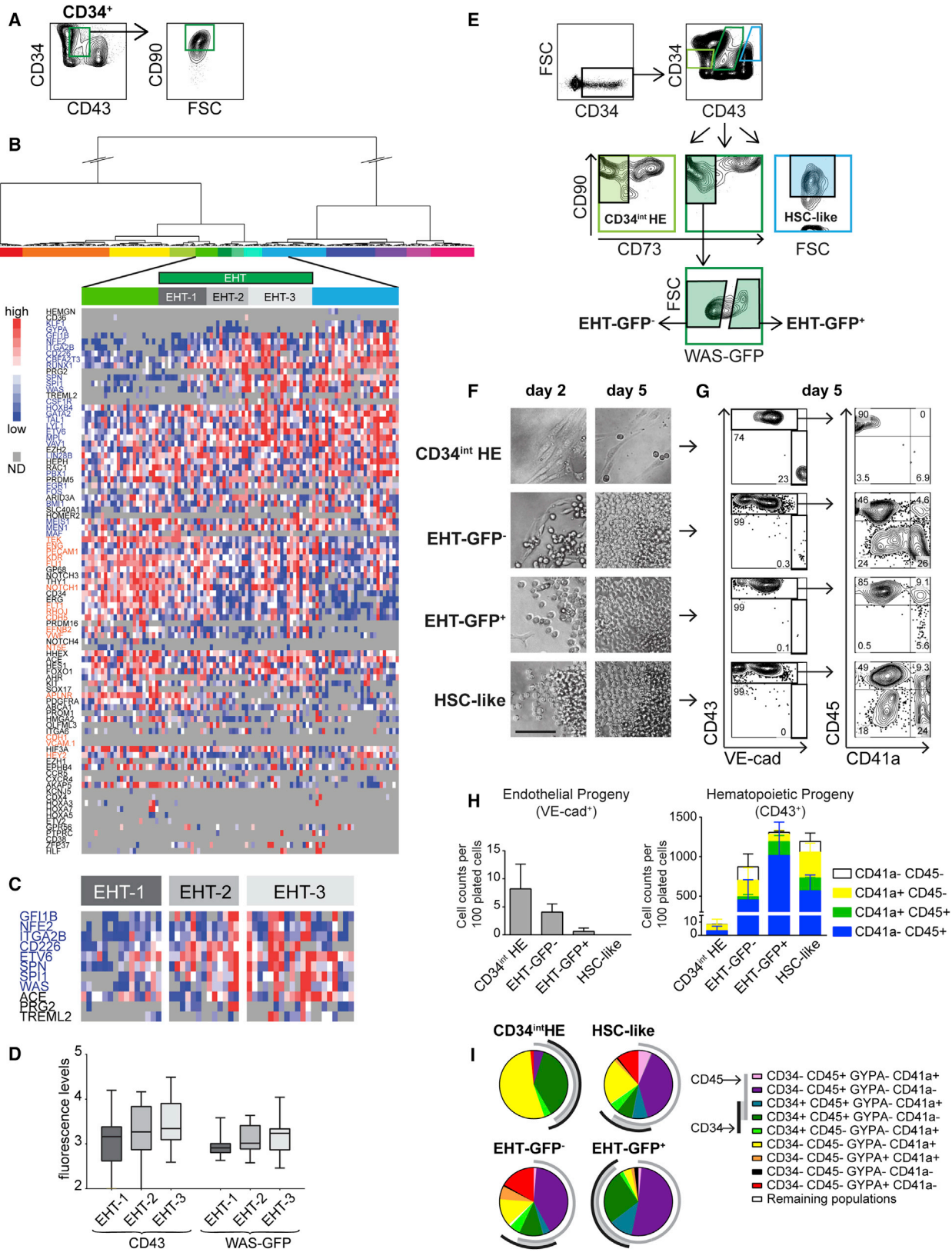
(E) Bean plots from the transcriptional analysis of Figure 1, the transcripts of the surface proteins analyzed in (D).

(F) Gating strategy for sorting the $CD43^{+}CD34^{+}$ identified subpopulations corresponding to EHT cluster (outlined in green), HSC-like (blue), and erythro-myeloid-committed cells (purple).

(G) CFU counts per 100 plated cells sorted using the gating strategy outlined in (F).

(H) CFU counts per 100 plated cells sorted using the gating strategy outlined in (F), with average representation of colony subtypes.

Data represent mean ± SEM; n = 2 independent experiments.



(legend on next page)

surface expression and in the WAS-GFP reporter expression (Figure 3D), suggesting that the originally identified EHT cluster contained cells with different levels of hematopoietic commitment. We therefore sought to elucidate if sorted WAS-GFP$^-$ and WAS-GFP$^+$ EHT cluster cells would exhibit different hematopoietic potentials.

We sorted EHT cluster cells (CD34$^+$CD43^{int}CD90^{hi}CD73$^-$) separating WAS-GFP$^-$ (EHT-GFP$^-$) from WAS-GFP$^+$ (EHT-GFP$^+$) fractions for subculture assays (Figures 3E and S3A–S3C). These EHT cluster sorted populations displayed cells of both spindle-shaped endothelial and round hematopoietic morphologies 2 days after plating (Figure 3F). Tracking WAS-GFP expression in the cultured EHT fractions, we observed instances of WAS-GFP$^+$ cells with endothelial morphology, including from EHT-GFP$^-$ sorted cells (Figure S3D, arrowheads). Surface marker analysis revealed high frequency of hematopoietic cells from both these sorted fractions by day 5 (Figure 3G). In comparison, the CD34^{int} HE sorted cells had exclusively endothelial morphology shortly after plating (day 2) and developed round hematopoietic cells by day 5 of subculture. As expected, the HSC-like population subculture resulted almost exclusively in blood cells and no endothelial progeny (Figures 3F–3H). Interestingly, there was increased blood/endothelium ratio from the analyzed populations the more transcriptionally committed they were toward the hematopoietic fate (Figures 3G and 3H, left).

These results confirm enrichment for cells undergoing EHT in the EHT-cluster-sorted populations and suggest a developmental progression within the EHT cluster.

The EHT Cluster Subpopulations Possess Distinct and Overlapping Hematopoietic Development Markers and Lineage Potentials

To evaluate the hematopoietic differentiation potentials of the above EHT-related populations, we analyzed the hematopoietic (CD43$^+$) fraction resulting from the 5-day-subcultured cells for the expression of CD45 (as an indicator of developmental progression within hematopoietic ontogeny), CD41a and GYPA (as indicators of erythro-myeloid lineage commitment), and CD34 as a marker of progenitors. The HSC-like population gave rise to blood cells with all combinations of CD45/CD41a immunophenotypes, suggesting a broad repertoire of developmentally distinct and lineage committing hematopoietic cells (Figures 3G and 3H, right). However, the progeny of EHT-GFP$^+$, the closest EHT cluster population according to gene expression, was almost exclusively CD45$^+$, suggesting a developmentally more mature hematopoietic progression (Figures 3G and 3H, right). The more distant EHT-GFP$^-$ cells had a progeny very similar to that of HSC-like cells, suggesting EHT-GFP$^-$ cells to be full contributors to all developmental stages and committed lineages from the HSC-like fraction. Analyzing GYPA further confirmed the similarities between the HSC-like and EHT-GFP$^-$ cells; both are capable of generating an early-committing erythroid lineage cell, but the EHT-GFP$^+$ sorted fraction did not have this potential (Figure 3I, red dataset in pie charts). This suggests that the EHT-GFP$^-$ cells possess a broader hematopoietic differentiation potential, generating more developmentally immature CD45$^-$ as well as GYPA$^+$ cells, as compared to the EHT-GFP$^+$ cells. These results also indicate that hematopoietic lineage differentiation potential is already defined prior to complete downregulation of the endothelial program during EHT.

We then set out to combine this differentiation potential data with gene expression correlation analysis in order to infer lineage relationships between the identified transcriptional groups (Figure 4A). We observed high correlation coefficients within the endothelial groups 2–5, hematopoietic groups 6–8, and mature hematopoietic groups 9 and 10. We then combined this with the index-sort data and differentiation potential assays to propose a model with the different transcriptional states identified in the hPSC-derived CD34$^+$ population as well as lineage relationships using gene expression correlation as well as validation through subculture (Figure 4B).

DISCUSSION

Using qRT-PCR analysis of index-sorted hPSC-derived CD34$^+$ single-cells enabled molecular characterization of transcriptional states throughout human EHT as well as the identification and sorting of populations of cells undergoing this process that upon culture possess differing hematopoietic lineage potentials. Our results show heterogeneity of the cells undergoing EHT, as they downregulate their endothelial programs and upregulate their hematopoietic programs. The EHT cluster could be subdivided based on increasing expression of eight genes, including *GFI1B*, shown in the murine system to be a key player

Figure 3. Differentiation Potential of the Identified EHT Transcriptional States

(A) Gating strategy enriching for the EHT cluster.

(B) Top: hierarchical clustering integrating EHT-cluster-enriched cells together with overall analysis of CD34$^+$ cells. Bottom: heatmap of gene expression for the three subgroups identified within the EHT cluster.

(C) Expression levels of genes upregulated along the three EHT cluster subgroups identified in (B).

(D) CD43 surface expression and WAS-GFP levels detected by FACS (using the index-sort data) in the three EHT cluster subgroups identified in (B).

(E) Sorting strategy used for subculturing populations enriched for CD34^{int} HE, EHT cluster further divided in WAS-GFP$^-$ and WAS-GFP$^+$, and HSC-like transcriptional groups.

(F) Phase-contrast pictures of sorted cells during subculture. Day 2 corresponds to the day following aggregate plating; day 5 is the day of harvest for FACS analysis. Scale bar, 100 μm.

(G) FACS analysis of day 5 sub-cultures, corresponding to the micrographs in the same rows in (F). Plots are representative of three independent experiments.

(H) Left: endothelial progeny per 100 plated cells. Right: stacked bar graph of the CD41a/CD45 hematopoietic progeny resulting from each subpopulation. Data represent mean ± SEM; n = 3 independent experiments.

(I) Pie charts representing the average composition, in three independent experiments, of the CD43$^+$ blood fraction at the end of subculture for each sorted population. Each color corresponds to each population resulting from all CD45, CD41a, CD34, and GYPA surface marker combinations.

See also Figure S3.

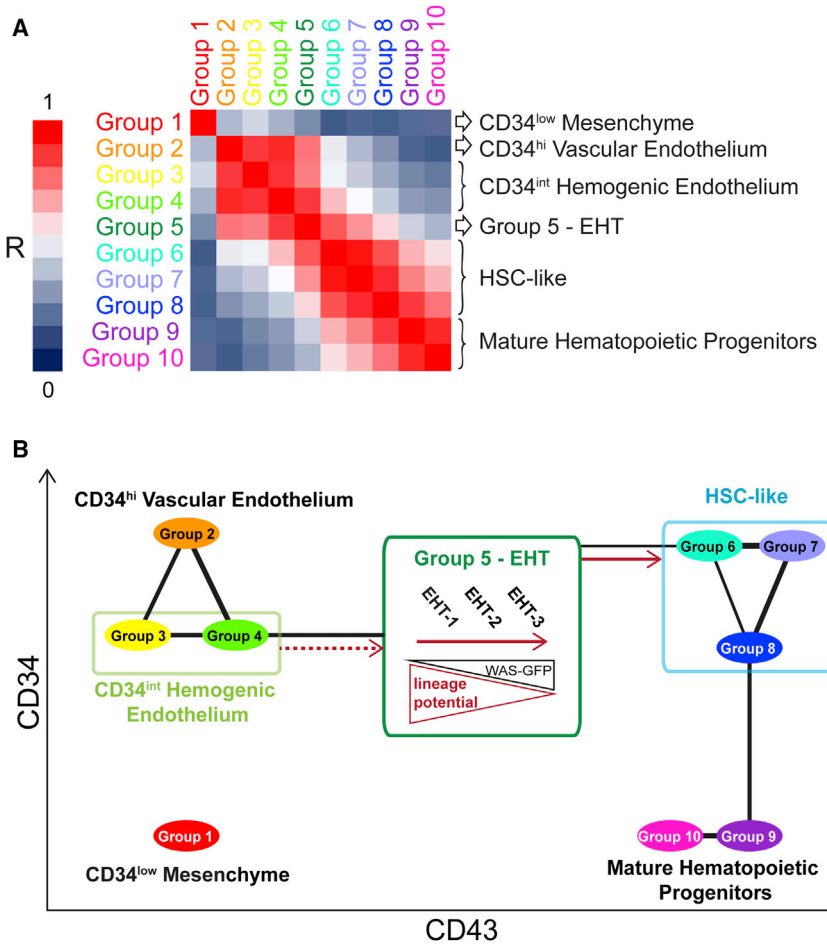


Figure 4. Inferring Population Hierarchy Using Single-Cell Analysis

(A) Pearson's correlation analysis calculated with the average expression values for each gene within each group. FACS-sorted populations, listed on the right, correspond to visibly correlating clusters. (B) Combining differentiation potential assay and Pearson's correlation data to infer lineage relationships between the identified groups. Groups are placed according to their CD34/CD43 phenotype. Black connections correspond to Pearson's correlation between groups with R > 0.8 (thickness proportional to R value). Red arrows summarize lineage relationships confirmed (solid arrows) or inferred (dashed arrows) from the differentiation potential assay.

for the loss of endothelial identity during EHT (Lancrin et al., 2012) and for hematopoietic maturation within AGM cells undergoing EHT (Thambyrajah et al., 2016). The *WAS*-negative (EHT-GFP$^-$) cells within this EHT population displayed the broadest hematopoietic lineage potential similar to that of the HSC-like population, while the *WAS*-expressing (EHT-GFP$^+$) fraction gave rise to a more restricted hematopoietic progeny, almost exclusively CD45$^+$, with no GYPA$^+$ cells. This led us to propose that the hematopoietic differentiation potential in the developmental context is already determined within unique cell populations at the EHT stage. This also implies that separate outcrops of hematopoietic cells emerge from EHT subsets with different lineage propensities depending on the expression levels of specific hematopoietic genes. In support of our findings, a recent study of murine AGM pre-HSCs also demonstrates heterogeneity during emergence at the single-cell level (Zhou et al., 2016). The heterogeneity within EHT could allow for yet another level of control of hematopoietic lineage contributions during development.

The CD34^{int}HE population is enriched in transcriptional group 3, the most abundant transcriptional group in the cells displaying the immunophenotype corresponding to that of previously characterized HE (Ditadi et al., 2015). Interestingly, among the endothelial transcriptional groups, group 3 had specifically high expression of genes, including a subset of hematopoietic TFs and epigenetic modifiers (see Figure S2Avi). These genes may be part of the regulatory network conferring hemogenic capacity to these cells. The hematopoietic progeny of this CD34^{int}HE subset contained similar populations as that of EHT cluster and HSC-like fractions (albeit with different frequencies), with nearly every CD45$^+$ cell co-expressing CD34 (Figure 3I). This suggests that CD34^{int}HE population may give rise to the EHT cluster and HSC-like populations (from which only a small subset of resulting CD45$^+$ cells co-expressed CD34).

Expression dynamics of the analyzed heptad members suggested a pattern with cooperation of *TAL1*, *LYL1*, *GATA2*, and *RUNX1* and separate activity of *FLI1* and *ERG*. This corroborates a recent study showing stronger correlation between *TAL1*, *LYL1*, *GATA2*, and *RUNX1* in terms of DNA-binding patterns, with *FLI1* and *ERG* displaying a separate binding partnership (Wilson et al., 2016). The first four TFs have been shown to synergize for HSC emergence in the murine setting (Chan et al., 2007; Chen et al., 2009; Nottingham et al., 2007), and it is therefore plausible that they also cooperate during human EHT.

Our approach of profiling a heterogeneous subset of cells at the single-cell level proved here to be highly useful to dissect sequential transcriptional states during human EHT. It can be applied to the study of primary material from developing human tissues to decipher the molecular mechanisms of human EHT in vivo and ultimately generate transplantable HSCs in vitro. Our findings set the stage for further exploration of heterogeneity during human hematopoietic cell emergence.

EXPERIMENTAL PROCEDURES

iPSC Expansion and Differentiation

The human iPSC-CB1RB9 cell line was expanded and differentiated toward the hematopoietic lineage as described previously (Rönn et al., 2015) (see Supplemental Experimental Procedures). Briefly, mesoderm-biased embryoid bodies were generated and plated on extracellular matrix to develop an

endothelial-like cell monolayer, from which presumptive hematopoietic cells emerge. Mesotodal medium (Primorigen Biosciences) was applied throughout the procedure.

Single-Cell qRT-PCR

At day 10 of the differentiation protocol, the cells were harvested and stained as described in Supplemental Experimental Procedures. Single $CD34^+$ cells were sorted directly into lysis buffer using the index-sorting function of the Diva Software. Target-specific pre-amplification mastermix was then added to each sample for reverse transcription (RT) and pre-amplification. Pre-amplified samples were diluted and analyzed in Fluidigm 96.96 arrays on a Biomark device (Fluidigm), with the TaqMan assays listed in Table S1. For further details and reagents, see Supplemental Experimental Procedures.

Single-Cell qRT-PCR Data Analysis

The data were first analyzed with Fluidigm Real-Time PCR Analysis Software for evaluation of the amplification curves, and Ct thresholds were set automatically using the "auto detector" function. Reactions with Ct > 30 were considered negative. In assays where amplification was detected in no-RT controls, all reactions displaying Ct > [min($Ct_{noRTcontrols}$) − 2] were considered negative. The data were then analyzed using the SCexV online tool (http://stemsysbio.bmc.lu.se/SCexV/; Lang et al., 2015), together with the index-sort data, for performing PCA and hierarchical clustering based on the gene expression data. Cells displaying a Ct value > 25 in the housekeeping gene GAPDH were excluded from the analysis, as well as the positive control samples. Subsequent analysis was then performed for the remaining cells (representing over 90% of the total sorted cells), using *Z* score transformed expression data of all assays excluding the housekeeping gene GAPDH. Unsupervised hierarchical clustering results were used for defining the groups of interest.

Subculture Assay

At day 10 of iPSC-to-blood differentiation, the indicated populations were sorted and cultured in an assay adapted from Ditadi and Sturgeon (2016). Specifically, the sorted cells aggregated overnight in ultra-low adherence U-shaped-bottom 96-well plates, and aggregates were then transferred onto Matrigel-coated flat-bottom 96-well plates. At 5 days post-sort, the cells were collected and analyzed by FACS. For details, see Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the single-cell qRT-PCR data for the hPSC-derived $CD34^+$ cells reported in this paper is GEO: GSE87422.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.03.023.

AUTHOR CONTRIBUTIONS

C.G. and N.-B.W. conceived and designed the project. The single-cell transcriptional analysis procedure was optimized by C.B. and G.K. in T.E.'s laboratory. R.E.R. and S. Saxena participated in the design of the experiments. C.G. performed experiments and analyses. C.G., S. Soneji, and S.L. performed in silico analyses. C.G. and N.-B.W. wrote the manuscript, which was reviewed and edited by all authors.

ACKNOWLEDGMENTS

The authors thank Alya Zriwil, Mikael Sommarin, and Eva Erlandsson for advice in experiment setup; Dr. Francisco Martín Molina from Granada University for providing the WAS-GFP vector; Roksana Moraghebi for advanced technical support and providing the iPSC-CB1RB9 line; Leal Oburoglu for critical reading of the paper; and Xiaojie Xian at the Flow Cytometry, Virus, and ES/iPSC core facilities at Lund Stem Cell Center in BMC for technical support. This study was funded by the Swedish Research Council (N.-B.W. and G.K.), the Swedish Cancer Society (N-B.W., G.K., and T.E.), the Swedish Childhood Cancer Foundation (N.-B.W., G.K., T.E., and C.B.), AFA Insurance (Sweden) (N.-B.W.), a HematoLinné Program grant (N.-B.W and T.E.), a Stem Therapy Program grant (N.-B.W.), and the Ragnar Söderberg Foundation (G.K.). N.-B.W. is a cofounder and board member of Longboat Explorer Inc.

Received: July 18, 2016
Revised: October 6, 2016
Accepted: January 17, 2017
Published: April 4, 2017

REFERENCES

Bakondi, B., Shimada, I.S., Perry, A., Munoz, J.R., Ylostalo, J., Howard, A.B., Gregory, C.A., and Spees, J.L. (2009). CD133 identifies a human bone marrow stem/progenitor cell sub-population with a repertoire of secreted factors that protect against stroke. Mol. Ther. *17*, 1938–1947.

Beck, D., Thoms, J.A., Perera, D., Schütte, J., Unnikrishnan, A., Knezevic, K., Kinston, S.J., Wilson, N.K., O'Brien, T.A., Göttgens, B., et al. (2013). Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. Blood *122*, e12–e22.

Bertrand, J.Y., Chi, N.C., Santoso, B., Teng, S., Stainier, D.Y., and Traver, D. (2010). Haematopoietic stem cells derive directly from aortic endothelium during development. Nature *464*, 108–111.

Boisset, J.C., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E., and Robin, C. (2010). In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature *464*, 116–120.

Boisset, J.C., Clapes, T., Klaus, A., Papazian, N., Onderwater, J., Mommaas-Kienhuis, M., Cupedo, T., and Robin, C. (2015). Progressive maturation toward hematopoietic stem cells in the mouse embryo aorta. Blood *125*, 465–469.

Chan, W.Y., Follows, G.A., Lacaud, G., Pimanda, J.E., Landry, J.R., Kinston, S., Knezevic, K., Piltz, S., Donaldson, I.J., Gambardella, L., et al. (2007). The paralogous hematopoietic regulators Lyl1 and Scl are coregulated by Ets and GATA factors, but Lyl1 cannot rescue the early Scl-/- phenotype. Blood *109*, 1908–1916.

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

Choi, K.D., Vodyanik, M.A., Togarrati, P.P., Suknuntha, K., Kumar, A., Samarjeet, F., Probasco, M.D., Tian, S., Stewart, R., Thomson, J.A., and Slukvin, I.I. (2012). Identification of the hemogenic endothelial progenitor and its direct precursor in human pluripotent stem cell differentiation cultures. Cell Rep. *2*, 553–567.

Ditadi, A., and Sturgeon, C.M. (2016). Directed differentiation of definitive hemogenic endothelium and hematopoietic progenitors from human pluripotent stem cells. Methods *101*, 65–72.

Ditadi, A., Sturgeon, C.M., Tober, J., Awong, G., Kennedy, M., Yzaguirre, A.D., Azzola, L., Ng, E.S., Stanley, E.G., French, D.L., et al. (2015). Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. Nat. Cell Biol. *17*, 580–591.

Hoppe, P.S., Coutu, D.L., and Schroeder, T. (2014). Single-cell technologies sharpen up mammalian stem cell research. Nat. Cell Biol. *16*, 919–927.

Ivanovs, A., Rybtsov, S., Anderson, R.A., Turner, M.L., and Medvinsky, A. (2014). Identification of the niche and phenotype of the first human hematopoietic stem cells. Stem Cell Reports *2*, 449–456.

Jaffredo, T., Gautier, R., Eichmann, A., and Dieterlen-Lièvre, F. (1998). Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. Development *125*, 4575–4583.

Kaufman, D.S., Hanson, E.T., Lewis, R.L., Auerbach, R., and Thomson, J.A. (2001). Hematopoietic colony-forming cells derived from human embryonic stem cells. Proc. Natl. Acad. Sci. USA *98*, 10716–10721.

Kennedy, M., Awong, G., Sturgeon, C.M., Ditadi, A., LaMotte-Mohs, R., Zúñiga-Pflücker, J.C., and Keller, G. (2012). T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures. Cell Rep. *2*, 1722–1735.

Kissa, K., and Herbomel, P. (2010). Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature *464*, 112–115.

Kumaravelu, P., Hook, L., Morrison, A.M., Ure, J., Zhao, S., Zuyev, S., Ansell, J., and Medvinsky, A. (2002). Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. Development *129*, 4891–4899.

Lancrin, C., Mazan, M., Stefanska, M., Patel, R., Lichtinger, M., Costa, G., Vargel, O., Wilson, N.K., Möröy, 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.

Lang, S., Ugale, A., Erlandsson, E., Karlsson, G., Bryder, D., and Soneji, S. (2015). SCExV: a webtool for the analysis and visualisation of single cell qRT-PCR data. BMC Bioinformatics *16*, 320.

Lee, R.H., Seo, M.J., Pulin, A.A., Gregory, C.A., Ylostalo, J., and Prockop, D.J. (2009). The CD34-like protein PODXL and alpha6-integrin (CD49f) identify early progenitor MSCs with increased clonogenicity and migration to infarcted heart in mice. Blood *113*, 816–826.

Mabuchi, Y., Morikawa, S., Harada, S., Niibe, K., Suzuki, S., Renault-Mihara, F., Houlihan, D.D., Akazawa, C., Okano, H., and Matsuzaki, Y. (2013). LNGFR(+)THY-1(+)VCAM-1(hi+) cells reveal functionally distinct subpopulations in mesenchymal stem cells. Stem Cell Reports *1*, 152–165.

Majeti, R., Park, C.Y., and Weissman, I.L. (2007). Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell *1*, 635–645.

Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A.J., Tanaka, Y., Wilkinson, A.C., Buettner, F., Macaulay, I.C., Jawaid, W., Diamanti, E., et al. (2015). Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nat. Biotechnol. *33*, 269–276.

Muñoz, P., Toscano, M.G., Real, P.J., Benabdellah, K., Cobo, M., Bueno, C., Ramos-Mejía, V., Menendez, P., Anderson, P., and Martín, F. (2012). Specific marking of hESCs-derived hematopoietic lineage by WAS-promoter driven lentiviral vectors. PLoS ONE *7*, e39091.

Nottingham, W.T., Jarratt, A., Burgess, M., Speck, C.L., Cheng, J.F., Prabhakar, S., Rubin, E.M., Li, P.S., Sloane-Stanley, J., Kong-A-San, J., and de Bruijn, M.F. (2007). Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. Blood *110*, 4188–4197.

Oberlin, E., Tavian, M., Blazsek, I., and Péault, B. (2002). Blood-forming potential of vascular endothelium in the human embryo. Development *129*, 4147–4157.

Rafii, S., Kloss, C.C., Butler, J.M., Ginsberg, M., Gars, E., Lis, R., Zhan, Q., Josipovic, P., Ding, B.S., Xiang, J., et al. (2013). Human ESC-derived hemogenic endothelial cells undergo distinct waves of endothelial to hematopoietic transition. Blood *121*, 770–780.

Rönn, R.E., Guibentif, C., Moraghebi, R., Chaves, P., Saxena, S., Garcia, B., and Woods, N.B. (2015). Retinoic acid regulates hematopoietic development from human pluripotent stem cells. Stem Cell Reports *4*, 269–281.

Rönn, R.E., Guibentif, C., Saxena, S., and Woods, N.B. (2017). Reactive oxygen species impair the function of CD90(+) hematopoietic progenitors generated from human pluripotent stem cells. Stem Cells *35*, 197–206.

Swiers, G., Baumann, C., O'Rourke, J., Giannoulatou, E., Taylor, S., Joshi, A., Moignard, V., Pina, C., Bee, T., Kokkaliaris, K.D., et al. (2013). Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. Nat. Commun. *4*, 2924.

Taoudi, S., Gonneau, C., Moore, K., Sheridan, J.M., Blackburn, C.C., Taylor, E., and Medvinsky, A. (2008). Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin+CD45+ pre-definitive HSCs. Cell Stem Cell *3*, 99–108.

Tavian, M., Coulombel, L., Luton, D., Clemente, H.S., Dieterlen-Lièvre, F., and Péault, B. (1996). Aorta-associated CD34+ hematopoietic cells in the early human embryo. Blood *87*, 67–72.

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.

Vodyanik, M.A., Thomson, J.A., and Slukvin, I.I. (2006). Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. Blood *108*, 2095–2105.

Wilkinson, A.C., Kawata, V.K., Schütte, J., Gao, X., Antoniou, S., Baumann, C., Woodhouse, S., Hannah, R., Tanaka, Y., Swiers, G., et al. (2014). Single-cell analyses of regulatory network perturbations using enhancer-targeting TALEs suggest novel roles for PU.1 during haematopoietic specification. Development *141*, 4018–4030.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell *7*, 532–544.

Wilson, N.K., Schoenfelder, S., Hannah, R., Sánchez Castillo, M., Schütte, J., Ladopoulos, V., Mitchelmore, J., Goode, D.K., Calero-Nieto, F.J., Moignard, V., et al. (2016). Integrated genome-scale analysis of the transcriptional regulatory landscape in a blood stem/progenitor cell model. Blood *127*, e12–e23.

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

Zhou, F., Li, X., Wang, W., Zhu, P., Zhou, J., He, W., Ding, M., Xiong, F., Zheng, X., Li, Z., et al. (2016). Tracing haematopoietic stem cell formation at single-cell resolution. Nature *533*, 487–492.

Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S., Zanetta, L., Dejana, E., Gasson, J.C., et al. (2008). Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell *3*, 625–636.