REGULAR ARTICLE

blood advances

Modeling primitive and definitive erythropoiesis with induced pluripotent stem cells

Giulia Pavani,[1,8] Joshua G. Klein,[1] Catriana C. Nations,[1,2] Jonathan H. Sussman,[3] Kai Tan,[4] Hyun Hyung An,[2] Osheiza Abdulmalik,[4] Christopher S. Thom,[5,6] Peter A. Gearhart,[7] Camryn M. Willett,[1] Jean Ann Maguire,[1] Stella T. Chou,[6] Deborah L. French,[1,8] and Paul Gadue[1,8]

[1]Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Philadelphia, PA; [2]Department of Cell and Molecular Biology and [3]Department of Genomics and Computational Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; [4]Division of Hematology and [5]Division of Neonatology, Children's Hospital of Philadelphia, Philadelphia, PA; [6]Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA; [7]Department of Obstetrics and Gynecology, Pennsylvania Hospital, University of Pennsylvania Health System, Philadelphia, PA; and [8]Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine and Children's Hospital of Philadelphia, Philadelphia, PA

Key Points

- iPSC-derived primitive and definitive RBCs display different characteristics, red cell antigen expression, and disease-modeling capabilities.
- Isogenic iPSC-derived definitive erythroblasts closely mimic primary fetal liver–derived erythroblasts.

During development, erythroid cells are produced through at least 2 distinct hematopoietic waves (primitive and definitive), generating erythroblasts with different functional characteristics. Human induced pluripotent stem cells (iPSCs) can be used as a model platform to study the development of red blood cells (RBCs) with many of the differentiation protocols after the primitive wave of hematopoiesis. Recent advances have established that definitive hematopoietic progenitors can be generated from iPSCs, creating a unique situation for comparing primitive and definitive erythrocytes derived from cell sources of identical genetic background. We generated iPSCs from healthy fetal liver (FL) cells and produced isogenic primitive or definitive RBCs which were compared directly to the FL-derived RBCs. Functional assays confirmed differences between the 2 programs, with primitive RBCs showing a reduced proliferation potential, larger cell size, lack of Duffy RBC antigen expression, and higher expression of embryonic globins. Transcriptome profiling by scRNA-seq demonstrated high similarity between FL- and iPSC-derived definitive RBCs along with very different gene expression and regulatory network patterns for primitive RBCs. In addition, iPSC lines harboring a known pathogenic mutation in the erythroid master regulator *KLF1* demonstrated phenotypic changes specific to definitive RBCs. Our studies provide new insights into differences between primitive and definitive erythropoiesis and highlight the importance of ontology when using iPSCs to model genetic hematologic diseases. Beyond disease modeling, the similarity between FL- and iPSC-derived definitive RBCs expands potential applications of definitive RBCs for diagnostic and transfusion products.

Submitted 18 September 2023; accepted 11 January 2024; prepublished online on *Blood Advances* First Edition 30 January 2024; final version published online 15 March 2024. https://doi.org/10.1182/bloodadvances.2023011708.

Raw sequencing data and processed reads associated with this work have been deposited in the Gene Expression Omnibus under accession number GSE242101.

Any other data relevant to the study are available upon reasonable request from the corresponding authors, Paul Gadue (gaduep@chop.edu) and Giulia Pavani (pavanig@chop.edu).

The full-text version of this article contains a data supplement.

Introduction

Red blood cells (RBCs) are currently used as an important biomedical product, both for transfusion medicine and for diagnostic purposes. COVID-19–related blood shortages,[1] increasing demand,[2] safety challenges, and requirements for specific RBC phenotypes[3] have created a global need for alternatives to volunteer blood donation, including ex vivo and in vitro manufacturing of RBCs.[4-6] Human induced pluripotent stem cells (iPSCs) are an attractive source for the generation of blood products because they can effectively differentiate into erythroid cells, potentially providing an unlimited source of RBCs.[7,8] Moreover, as they are amenable to gene editing, iPSCs can be engineered to express specific blood types[9] (including "universal" donor type O-[10]) or receptors for drug delivery,[11,12] in addition to providing a platform for disease-modeling and basic research.

During development, erythroid cells are produced through at least 2 hematopoietic waves, primitive and definitive, which generate RBCs with different functional characteristics. The first wave generates primitive erythroblasts that are detectable in the yolk-sac blood islands of the human conceptus as early as day 18 of gestation.[13] Primitive RBCs are relatively large, nucleated, transient, and express embryonic globins (ε- and ζ-globin).[14] Definitive RBCs arise from different subsets of hematopoietic progenitor cells (HPCs) that are formed later in the arterial vasculature at various sites in the embryo[15-17] and subsequently colonize the fetal liver at 4 postconception weeks (PCWs). Definitive RBCs are smaller, can enucleate, and predominantly express fetal globin subunits, such as γ-globin to fulfill the oxygen requirements of the growing fetus.[18]

Many current methodologies that generate iPSC-derived RBCs (iRBCs) mimic the first wave of embryonic hematopoiesis with the production of yolk sac-like primitive HPCs or generate a mixed population of primitive and definitive progenitors[19-22] that can be further differentiated into erythroblasts. These models have been extremely valuable in characterizing pathological phenotypes, initiating scale-up manufacturing processes and developing diagnostic applications.[9,23-26] However, primitive iRBCs may not recapitulate some fetal-specific phenotypes[27] and their limited proliferation potential, antigen expression, embryonic globin content, and larger size can be a hinderance to clinical translation. Although the switch between embryonic and fetal erythropoiesis has been extensively studied in mice,[16,28,29] many aspects of this process during human development remain elusive. Understanding how RBCs are generated through development in humans is a critical step in the production of pure functional iRBCs for therapeutic, diagnostic, and research use.

Recent advances in the directed differentiation of pluripotent stem cells have identified methodology that better recapitulates definitive, fetal-like hematopoiesis with protocols available to generate predominantly primitive or definitive HPCs.[22,30-33] A direct comparison of the erythroid potential of these iPSC-derived HPCs with primary fetal erythropoiesis has not been done, limiting our knowledge on how closely this cellular model recapitulates human development.

To gain insight into this process, we reprogrammed healthy fetal liver (FL) cells into iPSC lines and generated isogenic primitive and definitive HPCs. Using the same primary FL sample, we differentiated the FL- and iPSC-derived HPCs into red cells and compared them with functional assays, immunophenotyping, and transcriptomics, including scRNA-seq. This experimental design eliminates phenotypic variability caused by differences in genetic background and allowed the first direct comparison of iPSC-derived erythroid cells with genetically identical erythroid cells generated from primary FL. To highlight the importance of ontology in disease modeling, we also characterized a loss-of-function mutation in the erythroid master regulator *KLF1* and found that certain phenotypes are only seen in the definitive RBC program. By using an isogenic FL sample, this study validates iPSC-derived models of definitive erythropoiesis and deciphers differences between embryonic (primitive) and fetal (definitive) erythropoiesis in the human system.

Methods

Cell sources

Fetal liver. A healthy 18 PCW FL sample was obtained after informed consent under a protocol approved by the University of Pennsylvania Institutional Review Board, protocol #824810.

The FL iPSC line was generated by infection of isolated FL mononuclear cells with integration-free Sendai virus containing 4 reprogramming genes (*OCT4/SOX2/KLF4/MYC*).[34,35] Two independent clones were isolated, expanded, and checked for normal karyotype. Pluripotency was confirmed by flow cytometry and trilineage differentiation. Sendai clearance was checked at different passages using qPCR,[36] see supplemental Figure 1.

To predict antigen expression, blood group antigen genotyping was performed on the PreciseType Human Erythrocyte Antigen BeadChip (Immucor).

Other iPSC lines. Bone marrow derived iPSC line CHOPWT6 (WTBM1-8) and embryonic stem cell line H1 were previously described[37,38]

Hematopoietic and erythroid differentiation of iPSCs

For primitive hematopoietic differentiation, we used a previously established protocol.[32,33] Definitive hematopoietic progenitors were generated as described,[30,31,39] with minor modifications. Erythroid differentiation of hematopoietic progenitors was performed as previously described.[9] Detailed information about culture conditions, flow cytometry analysis, cytospin preparation, colony assays, globin analysis, and quantitative RT-PCR are provided in supplemental Methods.

Single-cell RNA sequencing

Experimental details and bioinformatic analyses are described in supplemental Methods.

Results

Generation of primitive and definitive HPCs from a FL–derived iPSC line

A major roadblock to the use of iRBCs for clinical and research applications is our limited understanding of the fidelity of in vitro cell models to recapitulate in vivo hematopoiesis. Specifically, a

thorough validation between primary FL- and iPSC-derived erythroblasts has been hindered by intrinsic genetic variability among iPSC lines[40] and technical challenges in obtaining a pure population of definitive HPCs to generate erythrocytes.[21] To compare primary FL- and iPSC-derived RBCs, iPSC lines were generated from an 18 PCW karyotypically normal FL sample (Figure 1A; supplemental Figure 1A-D) and compared to RBCs cultured from the same FL.

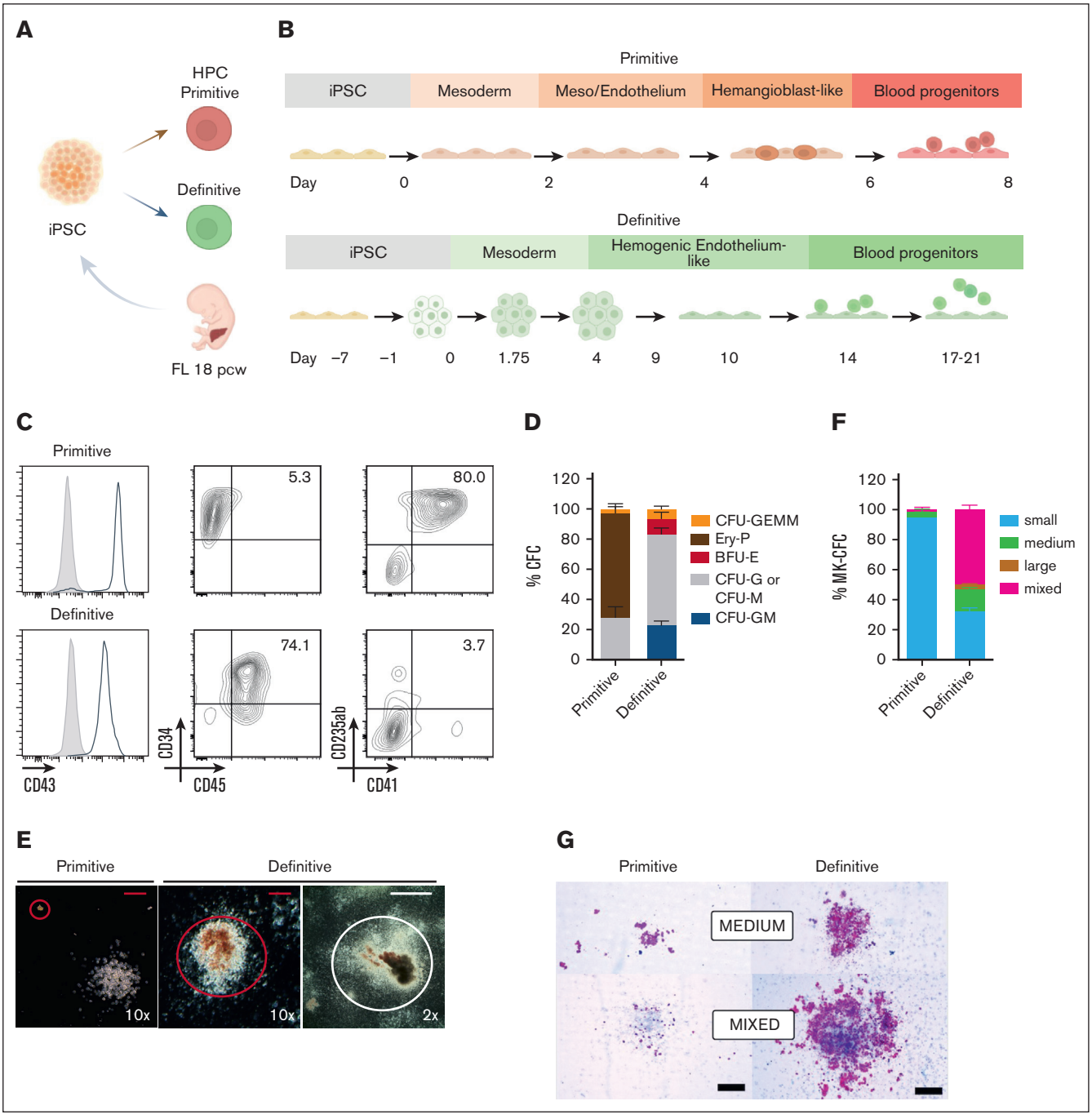

Figure 1. Generation of primitive and definitive HPCs from a FL–derived iPSC line. (A) Study design overview; iPSCs were produced from an 18 PCW healthy FL sample, and primitive and definitive HPCs were generated. (B) Schematic representation of the differentiation protocols and timeline used to generate primitive or definitive HPCs. (C) Representative flow cytometry analysis of primitive (top) and definitive (bottom) iPSC-derived HPCs. Percentage of positive cells is indicated; negative control is shown in gray. (D) Colony-forming cell frequency in primitive and definitive HPC (n = 6-9). (E) Representative images (bright field) of different progenitor-derived colonies at different magnification. Red colonies are circled. Red scale bar is 250 μm and white scale bar is 2 mm. (F) Megakaryocytic colony-forming cell frequency in primitive and definitive HPC (n = 2-4). G) Representative images (bright field) of megakaryocite colonies. Black scale bar is 200 μm.

To generate primitive and definitive HPCs, iPSCs were differentiated according to previously described protocols (Figure 1B; supplemental Methods) that specifically pattern primitive-posterior or definitive-lateral plate mesoderm.[30-32] As expected, both primitive and definitive iPSC-derived HPCs expressed the pan-hematopoietic marker CD43 and the progenitor marker CD34 (Figure 1C). Primitive HPCs also expressed CD41 and CD235ab markers typical of either megakaryocytes or erythrocytes that have been previously described to mark primitive HPCs.[41] In contrast, definitive HPCs mainly coexpressed CD34 and CD45 with no additional lineage-commitment markers, without transitioning through a CD41+CD235ab+ stage as they differentiate (data not shown).

To test the multilineage potential of the primitive and definitive HPCs, colony-forming unit (CFU) assays were performed (Figure 1D). Primitive HPCs generated small erythroid colonies (Ery-P; Figure 1E) and myeloid colonies (granulocytes, CFU-G or macrophages, CFU-M). Definitive progenitors generated larger colonies, including erythroid BFU-E, mixed granulomyelocytic colonies (CFU-GM), and very large multilineage colonies (CFU-GEMM, granulocyte, erythroid, macrophage, and megakaryocyte). In addition, megakaryocytic colony potential was assayed, as megakaryocytes from primitive and definitive origins exhibit substantial diversity in size and proliferation.[42,43] Significant differences in colony composition and morphology (Figure 1E-G) were observed, with primitive HPCs generating smaller, uni-lineage colonies compared with those generated by definitive HPCs. Taken together, these data show that iPSC-derived primitive and definitive HPCs have different characteristics, display distinct cellular markers, and give rise to different progeny.

Characterization of primitive and definitive RBC

To assess the functional differences between primitive and definitive iRBCs, HPCs were expanded in erythroid liquid cultures. Definitive iRBCs were more proliferative, expanding more than 60-fold by day 15 of differentiation (Figure 2A), whereas primitive iRBCs stopped dividing at approximately day 10 of differentiation and progressively underwent cell death. The kinetics of erythroid marker expression were markedly different between primitive and definitive erythroblasts. The primitive iRBCs highly expressed the erythroid markers CD36 (fatty acid translocase), CD71 (transferrin receptor), and CD235a (Glycophorin A) by day 3 of differentiation. The definitive iRBCs expressed these markers, but their expression peaked after day 9 of differentiation (Figure 2B). Terminally differentiated FL iRBCs, both definitive and primitive, were examined for maturation markers Band3 and CD49d to assure that samples were at similar stages of RBC maturation (supplemental Figure 2A). Cytospin and flow assays of erythroid cultures demonstrated low rates of enucleation (5%-8%) in the FL RBCs and definitive iRBCs, but no enucleation was detected in the primitive iRBCs (supplemental Figure 2B-C). Lastly, cell morphology and size were different between the developmental programs, with primitive iRBCs being larger as would be expected for this population (Figure 2C; supplemental Figure 3A).

A major feature that distinguishes primitive and definitive erythroblasts is globin gene expression. In humans, primitive erythroid cells express predominantly ε-globin and ζ-globin, whereas definitive erythroid cells switch to γ-globin and α-globin production.[18] To assess globin content at the transcriptional and protein levels, FL cells and iPSCs were terminally differentiated to primitive and/or definitive RBCs. In the primitive iRBCs, the ε-globin transcripts represented 42.2% ± 2.3 of total β-like transcripts compared with 7.9% ± 3.8 in the definitive iRBCs and 2.8% ± 1.1 in FL-derived RBCs, respectively (Figure 2D). At the protein level, a strong reduction of ε-globin and ζ-globin was accompanied by an increase of γ-globin and α-globin in both FL- and iPSC–derived definitive RBCs, indicating that these cells underwent the embryonic-to-fetal globin switch (Figure 2E; supplemental Figure 2D). This was also confirmed by hemoglobin tetramer analysis (Figure 2F; supplemental Figure 2E), in which definitive RBCs predominantly produced fetal hemoglobin (HbF, $\alpha_2\ \gamma_2$) compared with primitive RBCs that primarily produced embryonic forms (Gower 1, $\zeta_2\varepsilon_2$ and Gower 2, $\alpha_2\varepsilon_2$). Similar results were also obtained in other PSC lines (supplemental Figure 2F), validating the robustness of the definitive differentiation protocol in generating iRBCs that express fetal globins comparable with FL-derived RBCs.

Red cell antigen expression and typing

In addition to globins, expression of blood group antigens is developmentally regulated and differs in embryonic, fetal, and adult RBCs.[44,45] In particular, many blood antigens that are determined by immunodominant sugars are not fully glycosylated at birth. Blood group antigen expression was assessed on FL-derived RBCs and iRBCs by flow cytometry and correlated to RBC antigen phenotypes determined by genotyping the iPSC line (supplemental Table 3). As expected, the FL-derived RBCs and both primitive and definitive iRBCs expressed Kell (early embryonic) but not BCAM/Lu (after birth) (Figure 3A). Duffy was expressed exclusively on the FL- and iPSC–derived definitive RBCs and not the primitive iRBCs.

To assess functional expression of red cell antigens by primitive and definitive iRBCs, a standard blood bank assay of gel column agglutination was used. Previous work from our group demonstrated that primitive iRBCs could be utilized in this assay to identify uncommon antibodies against specific Rh antigens in alloimmunized patients.[9] To test if Duffy expression could be functionally used in an agglutination assay, primitive and definitive iRBCs were incubated with monoclonal typing antibodies and tested on gel columns (Figure 3B). Both iRBCs showed the predicted agglutination pattern for the Kell system (K-k+), but only the definitive iRBCs showed agglutination to Duffy (Fya+ Fyb+). As S and s antigens are highly glycosylated, agglutination was much higher on the definitive than primitive iRBCs. These results identify blood group antigens as specific markers of definitive erythropoiesis and expand the possibilities of using iRBCs for diagnostic applications.

scRNA-seq of primitive, definitive, and FL-derived RBCs

To determine if our iRBC model of in vitro erythroid development truly recapitulates fetal definitive erythropoiesis, FL- and iPSC-derived RBCs were analyzed using single-cell RNA sequencing (scRNA-seq). The FL- and iPSC-derived HPCs were used to generate RBCs using the same differentiation media and analyzed on day 6 to prevent excess hemoglobinization (Figure 4A). Erythroid surface marker expression was comparable (Figure 4B), due to stage matching the erythroblasts from the different sources. FL- and iPSC-derived definitive RBCs were smaller, displayed

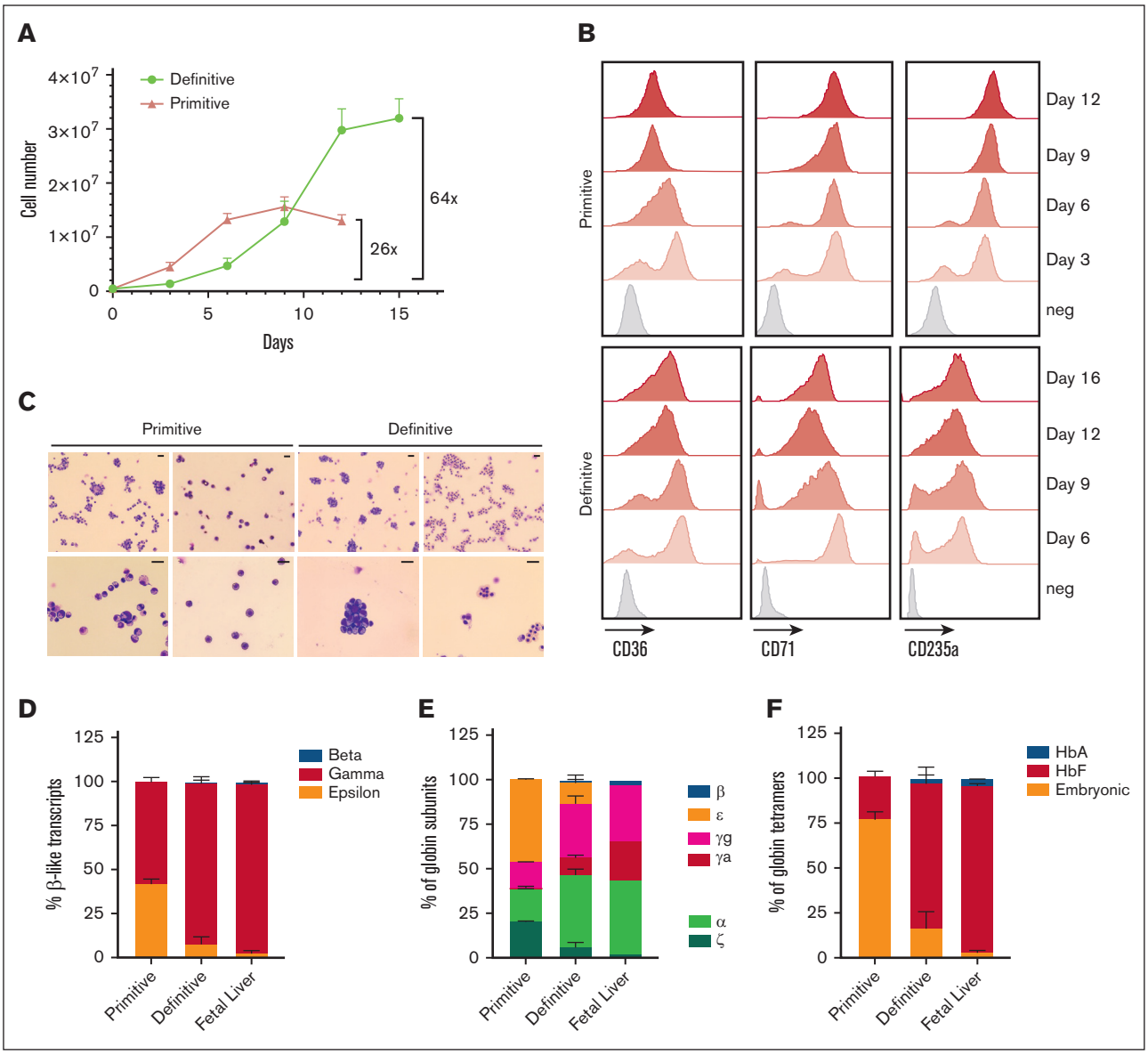

Figure 2. Functional characterization of primitive and definitive iRBCs. (A) Erythroid expansion of primitive and definitive HPCs in erythroid liquid cultures. (B) Representative histograms of CD36, CD71, and CD235a expression of primitive (top) or definitive erythroblasts during erythroid differentiation. Differentiation days are indicated on the right and negative control is shown in gray. (C) Cytospins showing May-Grünwald Giemsa staining of erythroid cells from primitive and definitive cultures at day 6 and day 12 of differentiation. Scale bars represent 20 μm. (D) Relative abundance of β-like transcripts ($\beta+\varepsilon+\gamma_g+\gamma_a$) at day 12 of erythroid liquid culture (mean ± SD, n = 3-4). (E) HPLC analysis of globin monomers in differentiated erythroblasts, relative quantification of α-like and β-like subunits are shown (primitive and definitive, mean ± SD, n = 3; FL n = 1). (F) HPLC analysis of globin tetramers in differentiated erythroblasts, relative quantification of adult hemoglobin (adult, $\alpha_2\beta_2$), HbF (fetal, $\alpha_2\ \gamma_2$), and embryonic (Gower 1 $\zeta_2\varepsilon_2$ and Gower 2, $\alpha_2\varepsilon_2$) globins (mean ± SD). HPLC, high-performance liquid chromatography.

slightly higher CD36 levels during differentiation, and showed a similar morphology compared with primitive iRBCs (Figure 4B; supplemental Figure 3A-B). Following data processing and quality control (supplemental Figure 4A), a total of 11 598, 9547, and 8808 cells were captured for FL-derived RBCs, primitive iRBC and definitive iRBCs, respectively.

The FL- and iPSC-derived definitive samples overlapped considerably on the uniform manifold approximation and projection, whereas the primitive samples remained distinct (Figure 4C). After quality control (supplemental Figure 4A) and clustering, 14 cell clusters were identified (Figure 4D) and differentially expressed genes were calculated (supplemental Figure 4B). All samples were comprised predominantly of erythroid cells, with small populations of megakaryocytes, myeloid, and mast cells also identified based on marker gene expression (Figure 4D-E; supplemental Figure 4B-C) and morphological examination (data not shown). Notably, only FL- and iPSC–derived definitive samples contained a distinct

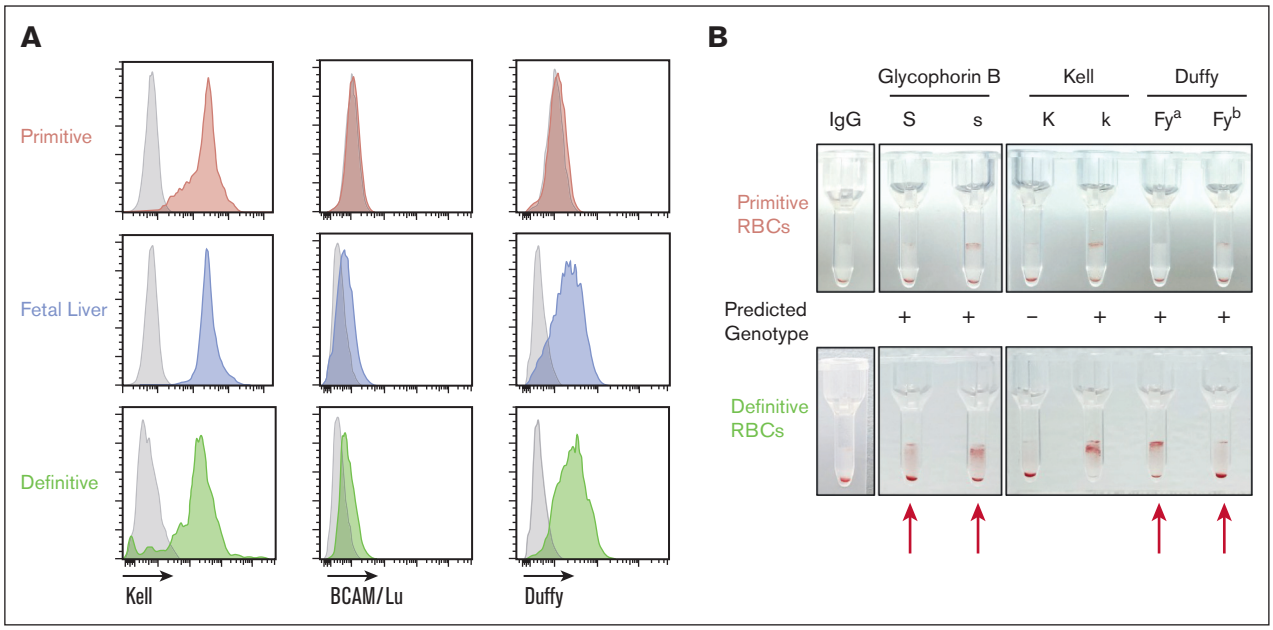

Figure 3. Red cell antigen expression in different developmental programs. (A) Expression of different red cell antigens by flow cytometry in primitive, definitive, and FL-derived RBC. Negative control is shown in gray. (B) Representative pictures of gel card assay using primitive and definitive isogenic iRBCs. RBCs were incubated with IgG against specific red cell antigens or control IgG. A positive reaction (agglutination) is indicated by the retention of RBCs at the top of the gel, whereas a negative reaction results in RBCs pelleted at the bottom of the tube. Expected outcome based on iPSC genotype is indicated.

mast-cell population, in line with previous observations that identified an erythroid-megakaryocyte-mast progenitor in the FL.[46-48] Cell-cycle analysis revealed that primitive cells were less proliferative, with most cells in G1, whereas FL- and iPSC–derived definitive RBCs were predominantly cycling (G2M and S; supplemental Figure 4D). Erythroid maturation staging was confirmed using previously published gene lists[49,50] and by CytoTRACE analysis[51] (supplemental Figure 4E-F). Ten erythroid clusters were identified according to differentiation markers and cell-cycle stage, 3 of which were comprised mainly of primitive iRBCs (P_Ery_Mid_G2M, P_Ery_Late and P_Ery_Ter) and 7 that contained both FL- and iPSC–derived definitive RBCs (Ery_Early, Ery_Mid_S, Ery_Mid_G2M, Ery_Late1, Ery_Late2, Ery_Ter1, Ery_Ter2) (Figure 4D-E; supplemental Figure 4E).

An examination of individual erythroid genes identified core erythroid markers and transcription factors, such as Glycophorin B (*GYPB*), GATA-binding protein 1 (*GATA1*), and nuclear factor-erythroid 2 (*NFE2*) that were expressed at similar levels in all samples (Figure 4F, top panel; supplemental Figure 4G, *KLF1*), confirming our stage-matching strategy. As expected, the expression of globin genes remained significantly different, with embryonic globins such as ζ (*HBZ*) highly expressed in primitive iRBCs, and fetal or adult globins (γ, *HBG1* and β, *HBB*, respectively) expressed more abundantly in FL- and iPSC-derived RBCs (Figure 4F, bottom panel). To quantify similarity between the different samples, a Pearson correlation analysis was performed. Importantly, a higher correlation was observed between the FL- and iPSC-derived definitive RBCs (0.96) compared with the primitive iRBCs (0.63) (Figure 4G). Notably, no FL cells and only a small portion of definitive iRBCs (1.5%) were observed in the primitive erythroid clusters (Figure 4H), indicating the purity of samples.

Differentially expressed genes s within primitive and FL-definitive erythroid clusters were identified (Figure 4I and supplemental Data). In primitive iRBCs, the top 25 upregulated genes included *HBZ*, as expected, and genes involved in purine metabolism (*ADA*), response to stress (*HSF1*, *ATF5*) and redox regulation of the cell (*PRXL2A*). In the FL and definitive erythroid clusters, there was higher expression of cell-cycle genes (*megakaryocite I67*, *TK1*, *CCNA2*, *MYBL*, and *RRM2*), confirming their proliferative state, in addition to β-globin (*HBB*), *HES6,* and *CD36* as previously observed by flow cytometry analysis. Importantly, the expression of the known erythroid fetal gene *SOX6,* was only detected in definitive and FL RBCs (supplemental Figure 4H), as was increased expression of genes involved in globin switching such as *ZBTB7A*[52] and NFI-factors.[53,54]

We used a recently published human yolk-sac data set[55] to further explore similarities between our iPSC-derived RBCs and human primitive erythropoiesis. To compare the 2 datasets, we projected our scRNA-seq data onto the yolk-sac atlas and assigned each cell of our dataset to the most similar cluster in the reference. We then calculated prediction scores within our erythroid clusters and observed a higher value for the primitive sample compared to the definitive and FL erythroblasts (supplemental Figure 4I-J), indicating a higher level of similarity between primitive and human yolk-sac red cells.

Our results suggest that human FL- and iPSC–derived definitive RBCs share similar transcriptomic and cellular composition, whereas primitive iRBCs remain distinct.

Regulon and gene network analysis

To better understand the transcriptional differences across erythroid developmental programs, we used the single-cell

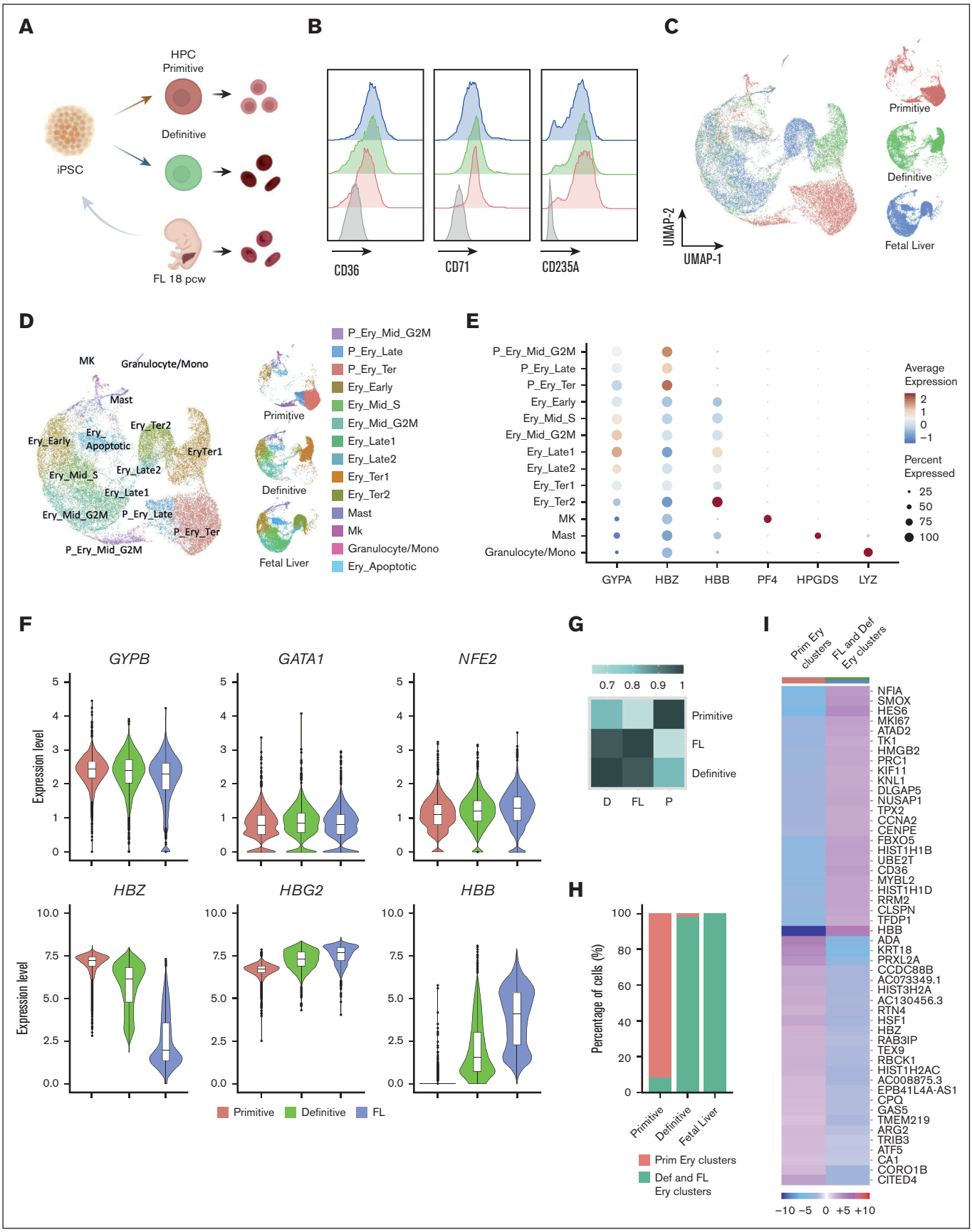

Figure 4.

regulatory network inference and clustering algorithm[56,57] to infer transcription factor activity and gene-regulatory networks (GRNs) in our samples (Figure 5A). A total of 361 active transcriptional regulons were identified. This analysis further validated previous cluster identities, with nonerythroid groups enriched for known transcriptional regulons specific to megakaryocytes[58] (ETV6, FLI1, and MEIS1), granulocytes[59] (CEBPE and IRF8), and mast cells[60] (GATA2 and BATF) (supplemental Figure 5A).

Within the erythroid clusters, we observed high activity of many core erythroid transcription factors (TFs) (GATA1, LMO2, NFE2, KLF1, and TAL1; supplemental Figure 5B). In the 3 primitive erythroid clusters, 148 GRNs with a regulon specificity score above 0.1 were observed, mostly active in the more differentiated cluster (P_Ery_Ter), of which 14 were specifically active in primitive iRBCs (Figure 5B and extended data). Among these regulons we found genes implicated in endoplasmic reticulum (ER) stress response (CREB3L1, ATF4, and HSF1) and stress–induced globin regulation (ATF4,[61] HSF1,[62] and HNF4A[63]).

The FL- and iPSC–derived definitive RBC clusters showed more complexity, in line with a heterogeneous and multistep maturation process of these cells. We inferred TF regulon activity in specific phases of erythropoiesis (early, mid, late, and terminal) corresponding to previously identified erythroid clusters (Figures 4D and 5B). In early phases of differentiation, we observed specific activity of known GRNs including LIN28B,[64] MAZ (a direct binder of α-globin promoter involved in human erythroid traits),[65] and BCLAF1 (implicated in fetal hematopoiesis[66,67]). Mid and late erythroid clusters showed increased activity of regulons associated with cell cycle and genomic stability (BRCA1), chromatin 3D structure (CTCF), and oxidative stress regulation (GABPB1), also in addition to erythroid regulators BCL11A and SOX6 (Figure 5C), known to be important in fetal and adult erythropoiesis.[68] However, the definitive iRBCs showed reduced activity of these regulons compared to FL RBCs from the same cluster, whereas other erythroid core factors like *KLF1* were similar (Figure 5D), possibly explaining the noncomplete silencing of ζ- and ε-globin genes observed in the definitive iRBCs. During terminal erythroid differentiation, we observed higher activity for many GRNs including NFIX,[69] EPO-responsive chromatin remodeler YY1 and known regulators of terminal erythropoiesis and enucleation such as FOXO3.[70]

To identify changes in biological processes across different samples, we also performed gene set enrichment analysis (GSEA) in erythroid subsets (Figure 5E). Primitive RBCs showed an enrichment in TGF-β signaling and platelet genes (hallmark coagulation, extended data), suggesting a more promiscuous transcriptional program indicative of the erythro-megakaryocytic nature of embryonic HPCs. In contrast, definitive iPSC-derived and FL RBCs were enriched in cell-cycle pathways (E2F targets, G2M checkpoints, and MYC targets), mTorc1 signaling and the unfolded protein response. Moreover, definitive and FL RBCs upregulated fatty acid metabolism and glycolysis, a feature commonly found in proliferating cells.[71] Taken together, our data suggest core erythroid similarities between primitive and definitive iRBCs but also highlight a differential activation of transcriptional units and biological processes between the 2 developmental programs.

Disease modeling with different developmental programs

To highlight the importance of cell ontology in disease modeling, studies were undertaken to characterize a loss-of-function mutation in the erythroid master regulator Krüppel-like factor 1 (*KLF1*). This transcription factor regulates erythroid cell specification [72] and may have differential involvement in primitive and definitive erythropoiesis as *KLF1*-null mice die in utero by E15.5 due to severe anemia.[73,74] In addition, heterozygous carriers of *KLF1* mutations produce well studied erythroid-specific phenotypes, such as low CD44 expression and the In(Lu) blood type.[75,76]

To investigate the role of *KLF1* in these developmental programs, iPSC lines expressing a known pathogenic missense mutation (L300P[77]) in the DNA-binding domain of *KLF1* was generated. Both heterozygous and homozygous mutations were created using base editing technology[78] (supplemental Figure 6A). Primitive and definitive HPCs followed by expansion to iRBCs were generated from the isogenic control and mutant lines (Figure 6A). No differences were observed in HPC yield between any of the lines (supplemental Figure 6B; data not shown) and only primitive KLF1$^{L300P/L300P}$ HPCs had a reduction in CD235ab expression with a slight increase in megakaryocytic markers (supplemental Figure 6C).

As expected, the KLF1$^{L300P/L300P}$ homozygous mutation led to limited expansion, lack of hemoglobinization, and reduced expression of red cell markers in both primitive and definitive iRBCs (Figure 6B; supplemental Figure 6D). A strong reduction in *KLF1* expression was also observed during erythroid differentiation of mutant lines, suggesting abnormal differentiation (supplemental Figure 6E). This impairment in erythropoiesis was confirmed in colony-forming cell -assays, where an almost complete loss of erythroid colonies was observed (Figure 6C). In comparison, the KLF1$^{\text{wild type(WT)}/L300P}$ heterozygous mutation led to an intermediate phenotype, with significantly reduced numbers of erythroid colonies (Figure 6B-C). With the limited generation of iRBCs from the KLF1$^{L300P/L300P}$ line, further studies were performed comparing WT and KLF1$^{WT/L300P}$ iRBCs only.

During development, the expression of many erythroid-specific antigens is regulated by *KLF1*, including CD44 and Duffy.[76] To

Figure 4. Transcriptional analysis by scRNA-seq of primitive, definitive, and FL-derived RBCs. (A) Schematic outline of scRNA-seq experiment. (B) Flow cytometry histograms of erythroid markers at the time of library preparation in primitive (pink), definitive (green), and FL-derived (blue) RBCs. Negative control is shown in gray. (C) Uniform manifold approximation and projection of primitive (n = 9547), definitive (n = 8808), and FL-derived (n = 11 598) RBCs. Subsets of individual samples are shown on the left (primitive iRBCs, pink; definitive iRBCs, green; and FL-derived RBCs, blue). (D) Distinct clusters of primitive, definitive, and FL-derived RBCs as defined by Seurat. (E) Dot plot showing expression of selected marker genes in each cluster. The dot size represents the percentage of cells within a cell cluster in which the gene was detected, and the dot color intensity represents the relative normalized average expression level of that marker. (F) Differential expression of a selection of erythroid genes in primitive, definitive, and FL-derived RBCs. (G) Correlation heat map of FL and definitive and primitive-derived RBCs in erythroid clusters. (H) Histogram showing the percentage of cells in each erythroid clusters in different samples. (I) Heat map of the top 25 differentially expressed genes between primitive and FL-definitive clusters.

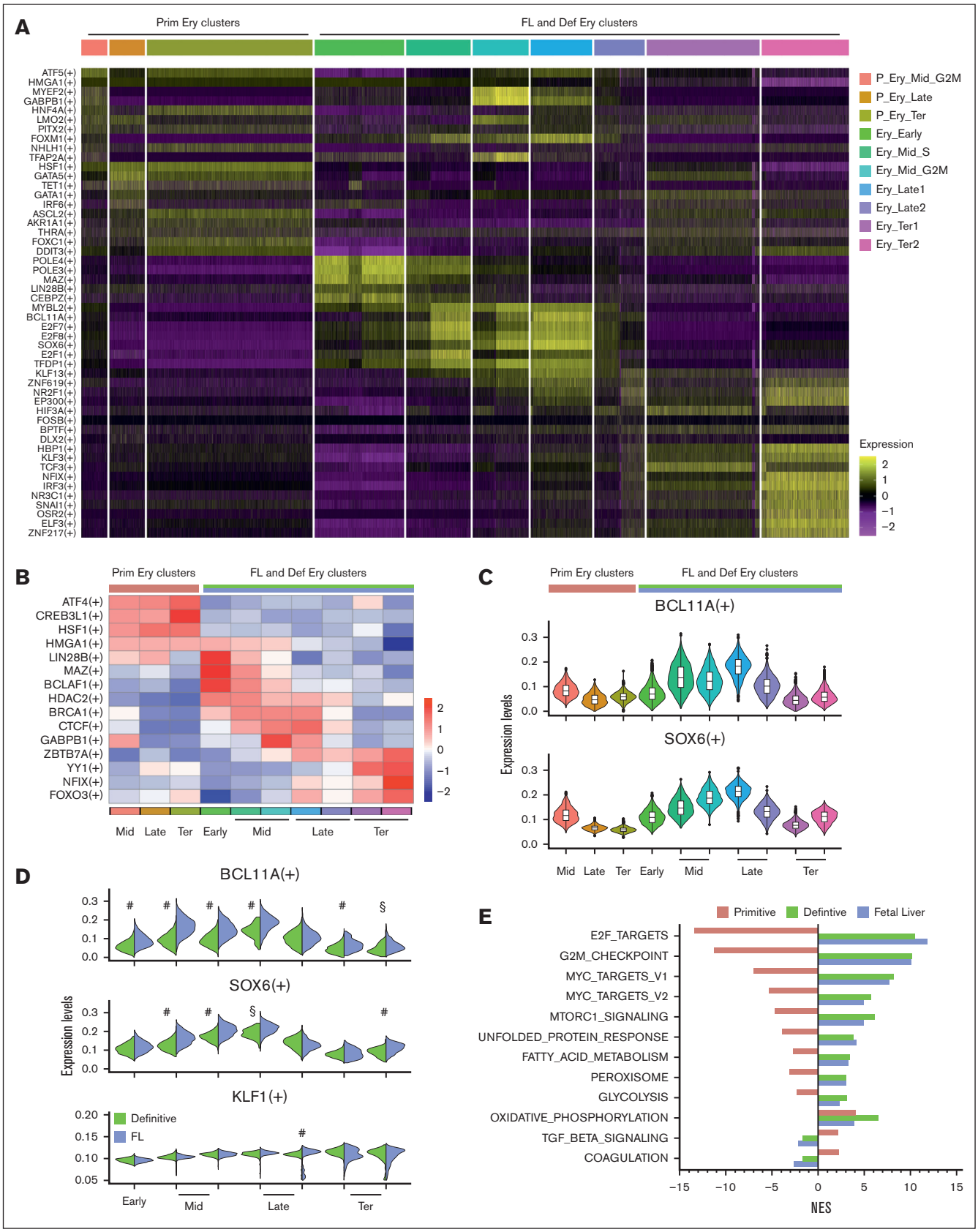

Figure 5.

assess expression of *KLF1*-independent (CD47) and *KLF1*-dependent (CD44 and Duffy) red cell markers, primitive, and definitive WT and mutant iRBCs were analyzed by flow cytometry. As expected, no difference in CD47 expression was observed (Figure 6D), whereas CD44 expression was reduced in both primitive and definitive mutant iRBCs, Duffy expression was reduced in definitive mutant iRBCs only.

In addition to expression of red cell antigens, *KLF1* is also involved in globin expression and switching during development. In primitive mutant iRBCs, the embryonic (ε-globin) and fetal (γ-globin) globin composition was unaffected (Figure 6E). In definitive mutant compared to WT iRBCs, a relative increase in the expression of embryonic globins was observed at the expense of fetal globins, indicating a different ontogenetic involvement of *KLF1* at the globin loci. This observation was confirmed by high-performance liquid chromatography, with a significant increase of ε-globin in single-chain analysis and concomitant reduction of HbF (Figure 6F-G). Our findings show that iPSCs can model erythroid defects caused by pathogenic *KLF1* mutations, such as reduced proliferation, hemoglobinization, and CD44 expression; however, only definitive iRBCs could capture differences in globin switching and Duffy expression in *KLF1* mutants.

These data highlight the importance of cell ontology in disease modeling and demonstrate the strength of an appropriate platform to identify specific developmental features of a disease phenotype.

Discussion

Due to their unlimited proliferation and differentiation capabilities, iPSCs are a promising source to generate blood products ex vivo.[79] The first methods to generate erythroblasts from iPSCs mainly recapitulated primitive erythropoiesis, producing iRBCs with embryonic characteristics.[19,20,32,33] Utilizing manipulation of signaling inputs during mesoderm patterning, more recent methods have emerged to generate pure primitive or definitive HPCs[22,30,31] but the iRBCs generated from these protocols have not been studied in depth.

In this study, we sought to define the purity and ability of primitive and definitive iRBC populations to recapitulate human developmental erythropoiesis and compare them to isogenic FL-derived RBCs. Our data show significant differences between the 2 iRBC programs, with definitive iRBCs being smaller and more proliferative then primitive iRBCs, expressing definitive red cell antigens, and undergoing embryonic-to-fetal globin switching. The definitive iRBCs were very similar to genetically-matched FL-derived RBCs in regard to gene expression as well as globin composition. Although very similar, the definitive iRBCs are not identical to FL RBCs, as they show lower expression of some key genes and regulatory networks involved in globin switching such as BCL11A and SOX6, which may inhibit these cells from complete silencing of embryonic globin expression. Future optimization efforts focused on increasing expression of these factors during erythroid differentiation will further improve the quality of iPSC-derived RBCs.

Transcriptional regulatory analysis of erythroid populations uncovered a number of factors which are differentially active throughout red cell maturation and across developmental origins. Although core erythroid factor GRNs (supplemental Figure 5B) are similarly active in all samples, definitive and FL RBCs show a more complex, dynamic, and stage-specific utilization of transcriptional units. Many regulatory networks involved in this maturation process include histone-modifying enzymes (HDAC2, KDM5A), chromatin insulators (CTCF), and chromatin remodelers (EP300, HMGA1). Investigation of chromatin accessibility and architecture in primitive and definitive iRBCs at specific maturation stages can be the topic of further studies and unravel additional differences between these developmental programs.

A major outcome of assessing blood antigen expression during our erythroid differentiations is the identification of the RBC antigen Duffy (*ACKR1*) that can distinguish definitive human erythroblasts. Interestingly, *ACKR1* was also one of the top 100 upregulated genes in definitive vs primitive mouse erythroblasts identified by a previous study.[80] We previously generated iRBCs with specialized Rh phenotypes including high and low prevalence Rh antigens (hr^S, Go^a, V/VS), which are difficult but important to identify because they may lead to delayed hemolytic transfusion reactions.[3,81] Although the primitive iRBCs generated express Rh and Kell antigens at sufficient levels for diagnostic testing, the ability to express additional blood group antigens including Duffy and MNS, enhances the diagnostic scope of these in vitro-derived iRBCs to identify clinically significant antibodies and improve transfusion safety.[82]

The generation of definitive fetal liver–like erythroblasts from iPSCs further enhances the use of these cells for disease modeling. We demonstrated that iRBCs harboring pathologic mutations in *KLF1* reveal specific phenotypes only when a definitive differentiation protocol is used. This finding has important consequences both for basic research and drug discovery approaches, because pathological phenotypes or drug targets might be evident only when a specific developmental stage is modeled. This could be the case for many hematologic disorders that start in fetal life including transient myeloproliferative disorder,[83] thrombocytopenia-absent radius[84] (OMIM #274000), Diamond-Blackfan anemia,[85] or Hb Bart's hydrops fetalis.[86]

Overall, our work demonstrates that definitive in vitro differentiation protocols faithfully mimic human FL erythropoiesis and uncover a larger set of phenotypes when using iPSCs for hematopoietic disease modeling. These findings pave the way for a better understanding of erythropoiesis in human development, and they will help to generate a more uniform population of the most mature RBCs to be used as a clinical diagnostic reagent and potentially, as future transfusion products.

Figure 5. Identification of active regulons and gene signatures in primitive and definitive iRBCs. (A) Heat map of top 10 differentially expressed regulons across Seurat erythroid clusters. (B) Heat map of a selection of differentially active regulons in erythroid clusters. The color intensity represents the average regulon activity, z-score normalized by row. (C) Violin plot illustrating regulon activity of BCL11A and SOX6 in erythroid clusters. (D) Split violin plot of BCL11A, SOX6, and KLF1 regulon activity in FL-definitive erythroid clusters. Definitive iRBCs are shown in green and FL sample is shown in blue. Adjusted *P*-values are calculated using Wilcoxon rank sum test with Bonferroni correction ($^{\S}P_{adj} <$ 1e-4, $^{\#}P_{adj} <$ 1e-8). (E) GSEA analysis of FL and definitive and primitive-derived RBC. Normalized enrichment scores of significant pathways ($P_{adj} < .01$) are shown.

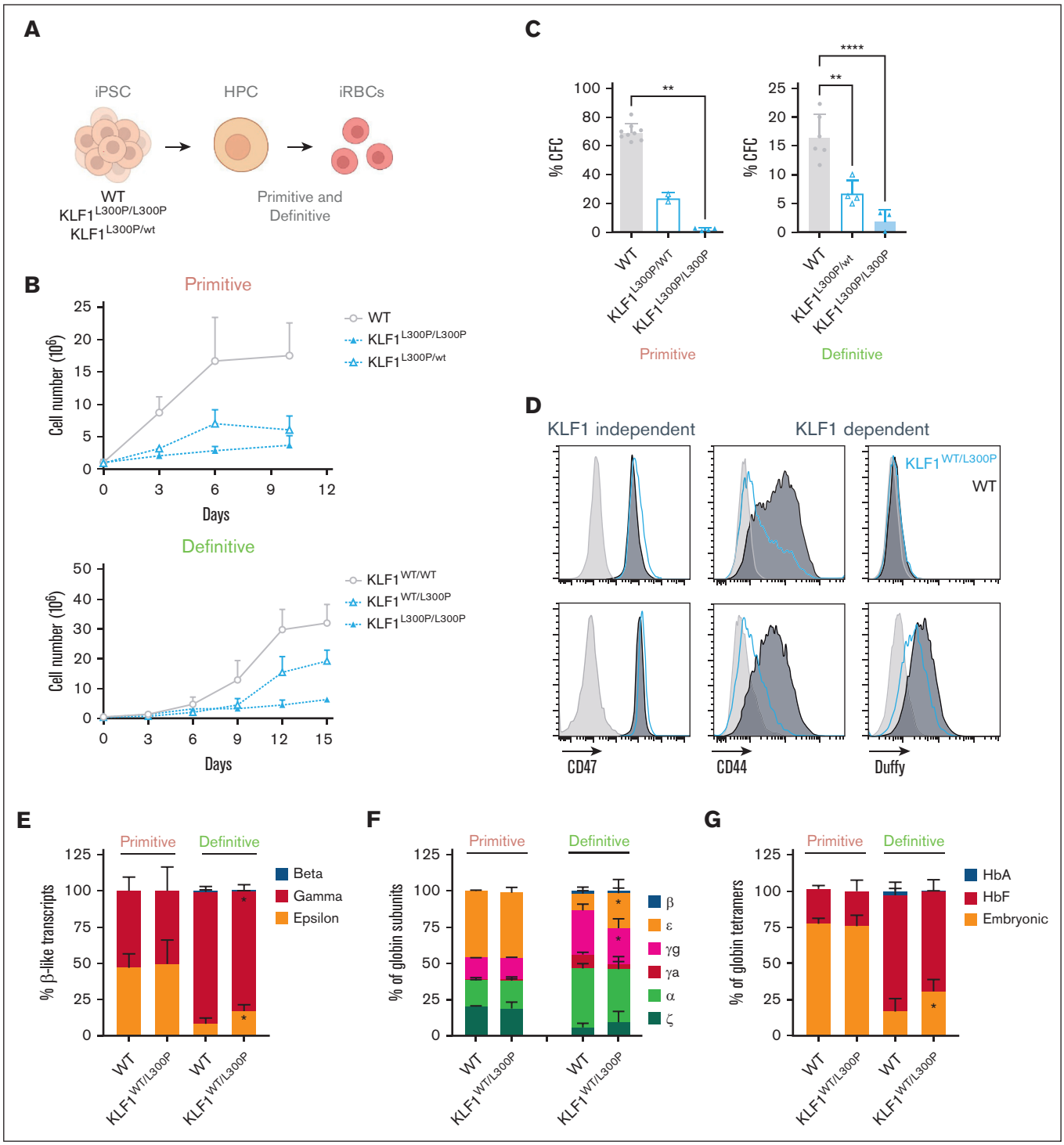

Figure 6. Functional characterization of primitive and definitive RBCs from KLF1 edited iPSC lines. (A) Schematic representation of the experimental layout. (B) Erythroid expansion of primitive or definitive erythroid liquid cultures (mean ± SD, n = 2-5). (C) Percentage of erythroid colonies derived from primitive and definitive HPCs from wild type (WT) and KLF1 edited lines (mean ± SD, n = 2-9, 1-way analysis of variance, Tukey test; $^{**}P < .01$ and $^{***}P < .001$). (D) Expression of KLF1-dependent or independent red cell antigens by flow cytometry in primitive or definitive day 12 RBCs from WT or KLF1$^{WT/L300P}$ lines. (E) Relative abundance of β-like transcripts ($\beta+\varepsilon+\gamma_g+\gamma_a$) at day 12 of erythroid liquid culture (mean ± SD, n = 3-5, $^{*}P < .05$, paired 2-tailed *t* test WT vs KLF1$^{WT/L300P}$). (F-G) HPLC quantification of globin subunits (F) and tetramers (G) in primitive and definitive RBCs from WT or KLF1$^{WT/L300P}$ iPSC lines (mean ±SD, n = 2-4, * = $P < .05$, paired 2-tailed *t* test WT vs KLF1$^{WT/L300P}$). HPLC, high-performance liquid chromatography.

Acknowledgments

The authors extend special thanks to the members of the U01 NHLBI Progenitor Cell Translational Consortium, especially Gordon Keller for sharing the definitive hematopoietic progenitor cell protocol. Sequencing was conducted at the Center for Applied Genomics at the Children's Hospital of Philadelphia.

This work was supported by National Heart, Lung, and Blood Institute U01 HL134696 (S.T.C., P.G., and D.L.F.) and DK 100854 (S.T.C.). The fetal liver sample was reprogrammed by the CHOP Human Pluripotent Stem Cell Core through the support of a Distinguished Chair of Pediatrics (S.T.C.).

Authorship

Contribution: G.P. conceived the study, designed and performed experiments, analyzed data, and wrote the manuscript; J.G.K. performed experiments and analyzed data; C.C.N. performed megakaryocyte experiments; J.H.S. and C.S.T. performed analysis on scRNA-seq data; H.H.A. performed gel card experiments; O.A. performed high-performance liquid chromatography analyses; P.A.G. collected the fetal liver sample; J.A.M. reprogrammed fetal liver cells; C.W. performed experiments; D.L.F. and K.T. provided scientific advice and edited the manuscript; and S.T.C. and P.G. conceived the study, designed experiments, analyzed data, and wrote the manuscript.

Conflict-of-interest disclosure: S.T.C. has filed a US patent (application number 16/757,815), entitled "Engineered Red Blood Cells Having Rare Antigen Phenotypes" in the name of The Children's Hospital of Philadelphia and New York Blood Center, Inc (national stage entry of international application number PCT/US2018/057932); the manuscript describes using the erythroid differentiation protocol of the engineered red blood cells covered in patient application. The remaining authors declare no competing financial interests.

ORCID profiles: G.P., 0000-0002-2522-2684; J.H.S., 0000-0002-3057-3550; K.T., 0000-0002-9104-5567; O.A., 0000-0003-2497-4291; S.T.C., 0000-0003-4333-6965; D.L.F., 0000-0002-7535-1716.

Correspondence: Paul Gadue, Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA 19104; email: gaduep@chop.edu; and Giulia Pavani, Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA 19104; email: pavanig@chop.edu.

References

1. American Red Cross faces severe blood shortage as coronavirus outbreak threatens availability of nation's supply. 17 March 2020. Accessed 14 February 2024. https://www.redcross.org/about-us/news-and-events/press-release/2020/american-red-cross-faces-severe-blood-shortage-as-coronavirus-outbreak-threatens-availability-of-nations-supply.html
2. Roberts N, James S, Delaney M, Fitzmaurice C. The global need and availability of blood products: a modelling study. *Lancet Haematol*. 2019;6(12):e606-e615.
3. Chou ST, Evans P, Vege S, et al. RH genotype matching for transfusion support in sickle cell disease. *Blood*. 2018;132(11):1198-1207.
4. Giarratana MC, Kobari L, Lapillonne H, et al. Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. *Nat Biotechnol*. 2005;23(1):69-74.
5. Giarratana MC, Rouard H, Dumont A, et al. Proof of principle for transfusion of in vitro-generated red blood cells. *Blood*. 2011;118(19):5071-5079.
6. Timmins NE, Athanasas S, Gunther M, Buntine P, Nielsen LK. Ultra-high-yield manufacture of red blood cells from hematopoietic stem cells. *Tissue Eng Part C Methods*. 2011;17(11):1131-1137.
7. Lu SJ, Feng Q, Park JS, et al. Biologic properties and enucleation of red blood cells from human embryonic stem cells. *Blood*. 2008;112(12):4475-4484.
8. Lapillonne H, Kobari L, Mazurier C, et al. Red blood cell generation from human induced pluripotent stem cells: perspectives for transfusion medicine. *Haematologica*. 2010;95(10):1651-1659.
9. An HH, Gagne AL, Maguire JA, et al. The use of pluripotent stem cells to generate diagnostic tools for transfusion medicine. *Blood*. 2022;140(15):1723-1734.
10. Yu S, Vassilev S, Lim ZR, et al. Selection of O-negative induced pluripotent stem cell clones for high-density red blood cell production in a scalable perfusion bioreactor system. *Cell Prolif*. 2022;55(8):e13218.
11. Brenner JS, Pan DC, Myerson JW, et al. Red blood cell-hitchhiking boosts delivery of nanocarriers to chosen organs by orders of magnitude. *Nat Commun*. 2018;9(1):2684.
12. Villa CH, Pan DC, Johnston IH, et al. Biocompatible coupling of therapeutic fusion proteins to human erythrocytes. *Blood Adv*. 2018;2(3):165-176.
13. Takashina T. Haemopoiesis in the human yolk sac. *J Anat*. 1987;151:125-135.
14. Palis J, Malik J, McGrath KE, Kingsley PD. Primitive erythropoiesis in the mammalian embryo. *Int J Dev Biol*. 2010;54(6-7):1011-1018.
15. Medvinsky AL, Dzierzak EA. Development of the definitive hematopoietic hierarchy in the mouse. *Dev Comp Immunol*. 1998;22(3):289-301.
16. Palis J. Primitive and definitive erythropoiesis in mammals. *Front Physiol*. 2014;5:3.
17. Calvanese V, Capellera-Garcia S, Ma F, et al. Mapping human haematopoietic stem cells from haemogenic endothelium to birth. *Nature*. 2022;604(7906):534-540.

18. Peschle C, Mavilio F, Care A, et al. Haemoglobin switching in human embryos: asynchrony of zeta−−alpha and epsilon−−gamma-globin switches in primitive and definite erythropoietic lineage. *Nature*. 1985;313(5999):235-238.
19. Olivier EN, Qiu C, Velho M, Hirsch RE, Bouhassira EE. Large-scale production of embryonic red blood cells from human embryonic stem cells. *Exp Hematol*. 2006;34(12):1635-1642.
20. Zambidis ET, Peault B, Park TS, Bunz F, Civin CI. Hematopoietic differentiation of human embryonic stem cells progresses through sequential hematoendothelial, primitive, and definitive stages resembling human yolk sac development. *Blood*. 2005;106(3):860-870.
21. Chang CJ, Mitra K, Koya M, et al. Production of embryonic and fetal-like red blood cells from human induced pluripotent stem cells. *PLoS One*. 2011;6(10):e25761.
22. Ng ES, Azzola L, Bruveris FF, et al. Differentiation of human embryonic stem cells to HOXA(+) hemogenic vasculature that resembles the aorta-gonad-mesonephros. *Nat Biotechnol*. 2016;34(11):1168-1179.
23. An HH, Poncz M, Chou ST. Induced pluripotent stem cell-derived red blood cells, megakaryocytes, and platelets: progress and challenges. *Curr Stem Cell Rep*. 2018;4(4):310-317.
24. Byrska-Bishop M, VanDorn D, Campbell AE, et al. Pluripotent stem cells reveal erythroid-specific activities of the GATA1 N-terminus. *J Clin Invest*. 2015;125(3):993-1005.
25. Garcon L, Ge J, Manjunath SH, et al. Ribosomal and hematopoietic defects in induced pluripotent stem cells derived from Diamond Blackfan anemia patients. *Blood*. 2013;122(6):912-921.
26. Olivier EN, Marenah L, McCahill A, Condie A, Cowan S, Mountford JC. High-efficiency serum-free feeder-free erythroid differentiation of human pluripotent stem cells using small molecules. *Stem Cells Transl Med*. 2016;5(10):1394-1405.
27. Chou ST, Opalinska JB, Yao Y, et al. Trisomy 21 enhances human fetal erythro-megakaryocytic development. *Blood*. 2008;112(12):4503-4506.
28. Silver L, Palis J. Initiation of murine embryonic erythropoiesis: a spatial analysis. *Blood*. 1997;89(4):1154-1164.
29. Kingsley PD, Malik J, Emerson RL, et al. "Maturational" globin switching in primary primitive erythroid cells. *Blood*. 2006;107(4):1665-1672.
30. Ditadi A, Sturgeon CM, Tober J, et al. Human definitive haemogenic endothelium and arterial vascular endothelium represent distinct lineages. *Nat Cell Biol*. 2015;17(5):580-591.
31. Sturgeon CM, Ditadi A, Awong G, Kennedy M, Keller G. Wnt signaling controls the specification of definitive and primitive hematopoiesis from human pluripotent stem cells. *Nat Biotechnol*. 2014;32(6):554-561.
32. Paluru P, Hudock KM, Cheng X, et al. The negative impact of Wnt signaling on megakaryocyte and primitive erythroid progenitors derived from human embryonic stem cells. *Stem Cell Res*. 2014;12(2):441-451.
33. Mills JA, Paluru P, Weiss MJ, Gadue P, French DL. Hematopoietic differentiation of pluripotent stem cells in culture. *Methods Mol Biol*. 2014;1185:181-194.
34. Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc Jpn Acad Ser B Phys Biol Sci*. 2009;85(8):348-362.
35. Tokusumi T, Iida A, Hirata T, Kato A, Nagai Y, Hasegawa M. Recombinant Sendai viruses expressing different levels of a foreign reporter gene. *Virus Res*. 2002;86(1-2):33-38.
36. Takasaki K, Kumar SS, Gagne A, French DL, Chou ST. Generation of 2 isogenic clones from a patient with trisomy 21 and a GATA1 mutation. *Stem Cell Res*. 2023;69:103098.
37. Sullivan SK, Mills JA, Koukouritaki SB, et al. High-level transgene expression in induced pluripotent stem cell-derived megakaryocytes: correction of Glanzmann thrombasthenia. *Blood*. 2014;123(5):753-757.
38. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998;282(5391):1145-1147.
39. Dege C, Sturgeon CM. Directed differentiation of primitive and definitive hematopoietic progenitors from human pluripotent stem cells. *J Vis Exp*. 2017;129.
40. Mills JA, Wang K, Paluru P, et al. Clonal genetic and hematopoietic heterogeneity among human-induced pluripotent stem cell lines. *Blood*. 2013;122(12):2047-2051.
41. Klimchenko O, Mori M, Distefano A, et al. A common bipotent progenitor generates the erythroid and megakaryocyte lineages in embryonic stem cell-derived primitive hematopoiesis. *Blood*. 2009;114(8):1506-1517.
42. Notta F, Zandi S, Takayama N, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*. 2016;351(6269):aab2116.
43. Wang Y, Hayes V, Jarocha D, et al. Comparative analysis of human ex vivo-generated platelets vs megakaryocyte-generated platelets in mice: a cautionary tale. *Blood*. 2015;125(23):3627-3636.
44. Toivanen P, Hirvonen T. Antigens Duffy, Kell, Kidd, Lutheran and Xg a on fetal red cells. *Vox Sang*. 1973;24(4):372-376.
45. Habibi B, Bretagne M, Bretagne Y, Forestier F, Daffos F. Blood group antigens on fetal red cells obtained by umbilical vein puncture under ultrasound guidance: a rapid hemagglutination test to check for contamination with maternal blood. *Pediatr Res*. 1986;20(11):1082-1084.
46. Franco CB, Chen CC, Drukker M, Weissman IL, Galli SJ. Distinguishing mast cell and granulocyte differentiation at the single-cell level. *Cell Stem Cell*. 2010;6(4):361-368.
47. Popescu DM, Botting RA, Stephenson E, et al. Decoding human fetal liver haematopoiesis. *Nature*. 2019;574(7778):365-371.

48. Vanuytsel K, Murphy G. Induced pluripotent stem cell–based mapping of β-globin expression throughout human erythropoietic development. *Blood Adv*. 2018;2(15):1998-2011.

49. Ludwig LS, Lareau CA, Bao EL, et al. Transcriptional states and chromatin accessibility underlying human erythropoiesis. *Cell Rep*. 2019;27(11):3228-3240.e7.

50. Shi L, Lin YH, Sierant MC, et al. Developmental transcriptome analysis of human erythropoiesis. *Hum Mol Genet*. 2014;23(17):4528-4542.

51. Gulati GS, Sikandar SS, Wesche DJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*. 2020;367(6476):405-411.

52. Masuda T, Wang X, Maeda M, et al. Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science*. 2016;351(6270):285-289.

53. Chaand M, Fiore C, Johnston B, et al. Erythroid lineage chromatin accessibility maps facilitate identification and validation of NFIX as a fetal hemoglobin repressor. *Commun Biol*. 2023;6(1):640.

54. Qin K, Huang P, Feng R, et al. Dual function NFI factors control fetal hemoglobin silencing in adult erythroid cells. *Nat Genet*. 2022;54(6):874-884.

55. Goh I, Botting RA, Rose A, et al. Yolk sac cell atlas reveals multiorgan functions during human early development. *Science*. 2023;381(6659):eadd7564.

56. Aibar S, Gonzalez-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14(11):1083-1086.

57. Van de Sande B, Flerin C, Davie K, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc*. 2020;15(7):2247-2276.

58. Noris P, Pecci A. Hereditary thrombocytopenias: a growing list of disorders. *Hematology*. 2017;2017(1):385-399.

59. Kurotaki D, Nakabayashi J, Nishiyama A, et al. Transcription factor IRF8 governs enhancer landscape dynamics in mononuclear phagocyte progenitors. *Cell Rep*. 2018;22(10):2628-2641.

60. Ohmori Sy, Moriguchi T, Noguchi Y, et al. GATA2 is critical for the maintenance of cellular identity in differentiated mast cells derived from mouse bone marrow. *Blood*. 2015;125(21):3306-3315.

61. Boontanrart MY, Schroder MS, Stehli GM, et al. ATF4 regulates MYB to increase gamma-Globin in response to loss of beta-Globin. *Cell Rep*. 2020;32(5):107993.

62. Gao J, Liu J, Zhang L, et al. Heat shock transcription factor 1 regulates the fetal gamma-globin expression in a stress-dependent and independent manner during erythroid differentiation. *Exp Cell Res*. 2020;387(2):111780.

63. Chami N, Chen MH, Slater AJ, et al. Exome genotyping identifies pleiotropic variants associated with red blood cell traits. *Am J Hum Genet*. 2016;99(1):8-21.

64. Lee YT, de Vasconcellos JF, Yuan J, et al. LIN28B-mediated expression of fetal hemoglobin and production of fetal-like erythrocytes from adult human erythroblasts ex vivo. *Blood*. 2013;122(6):1034-1041.

65. Deen D, Butter F, Daniels DE, et al. Identification of the transcription factor MAZ as a regulator of erythropoiesis. *Blood Adv*. 2021;5(15):3002-3015.

66. Crowley SJ, Bednarski JJ, Magee JA, LI Y, White LS, Yang W. BCLAF1 regulates expression of AP-1 genes and fetal hematopoietic stem cell repopulation activity. *Blood*. 2022;140(Supplement 1):2852-2853.

67. White LS, Soodgupta D, Johnston RL, Magee JA, Bednarski JJ. Bclaf1 promotes maintenance and self-renewal of fetal hematopoietic stem cells. *Blood*. 2018;132(Supplement 1):1269-1269.

68. Xu J, Sankaran VG, Ni M, et al. Transcriptional silencing of gamma-globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev*. 2010;24(8):783-798.

69. Walker M, Li Y, Morales-Hernandez A, et al. An NFIX-mediated regulatory network governs the balance of hematopoietic stem and progenitor cells during hematopoiesis. *Blood Adv*. 2023;7(17):4677-4689.

70. Liang R, Camprecios G, Kou Y, et al. A systems approach identifies essential FOXO3 functions at key steps of terminal erythropoiesis. *PLoS Genet*. 2015;11(10):e1005526.

71. Ito K, Suda T. Metabolic requirements for the maintenance of self-renewing stem cells. *Nat Rev Mol Cell Biol*. 2014;15(4):243-256.

72. Siatecka M, Bieker JJ. The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood*. 2011;118(8):2044-2054.

73. Perkins AC, Sharpe AH, Orkin SH. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature*. 1995;375(6529):318-322.

74. Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature*. 1995;375(6529):316-318.

75. Helias V, Saison C, Peyrard T, et al. Molecular analysis of the rare in(Lu) blood type: toward decoding the phenotypic outcome of haploinsufficiency for the transcription factor KLF1. *Hum Mutat*. 2013;34(1):221-228.

76. Perkins A, Xu X, Higgs DR, et al. Kruppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. *Blood*. 2016;127(15):1856-1862.

77. Kawai M, Obara K, Onodera T, et al. Mutations of the KLF1 gene detected in Japanese with the In(Lu) phenotype. *Transfusion*. 2017;57(4):1072-1077.

78. Koblan LW, Doman JL, Wilson C, et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol*. 2018;36(9):843-846.

79. Hansen M, von Lindern M, van den Akker E, Varga E. Human-induced pluripotent stem cell-derived blood products: state of the art and future directions. *FEBS Lett*. 2019;593(23):3288-3303.

80. Kingsley PD, Greenfest-Allen E, Frame JM, et al. Ontogeny of erythroid gene expression. *Blood*. 2013;121(6):e5-e13.

81. Larson PJ, Lukas MB, Friedman DF, Manno CS. Delayed hemolytic transfusion reaction due to anti-Go(a), an antibody against the low-prevalence Gonzales antigen. *Am J Hematol*. 1996;53(4):248-250.

82. Boateng LA, Schonewille H, Ligthart PC, et al. One third of alloantibodies in patients with sickle cell disease transfused with African blood are missed by the standard red blood cell test panel. *Haematologica*. 2021;106(8):2274-2276.

83. Zipursky A, Brown EJ, Christensen H, Doyle J. Transient myeloproliferative disorder (transient leukemia) and hematologic manifestations of Down syndrome. *Clin Lab Med*. 1999;19(1):157-167, vii.

84. Petit F, Boussion S. Thrombocytopenia absent radius syndrome. In: Adam MP, Mirzaa GM, Pagon RA, et al; eds. *GeneReviews*. Seattle, WA: University of Washington; 2009:1993-2024.

85. Da Costa L, Leblanc T, Mohandas N. Diamond-Blackfan anemia. *Blood*. 2020;136(11):1262-1273.

86. Nicholls RD, Higgs DR, Clegg JB, Weatherall DJ. Alpha zero-thalassemia due to recombination between the alpha 1-globin gene and an AluI repeat. *Blood*. 1985;65(6):1434-1438.