

RED CELLS, IRON, AND ERYTHROPOIESIS

Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E

Jie Li,¹ John Hale,² Pooja Bhagia,^{2,3} Fumin Xue,¹ Lixiang Chen,¹ Julie Jaffray,^{2,3} Hongxia Yan,² Joseph Lane,⁴ Patrick G. Gallagher,⁵ Narla Mohandas,² Jing Liu,⁶ and Xiuli An^{1,7}

¹Laboratory of Membrane Biology, ²Red Cell Physiology Laboratory, New York Blood Center, New York, NY; ³Pediatric Oncology, Memorial Sloan-Kettering Cancer Center, New York, NY; ⁴Department of Orthopedic Surgery, The New York Presbyterian Cornell Hospital, New York, NY; ⁵Departments of Pediatrics, Pathology and Genetics, Yale University School of Medicine, New Haven, CT; ⁶State Key Laboratory of Medical Genetics and School of Life Science, Central South University, Changsha, Hunan, China; and ⁷College of Life Science, Zhengzhou University, Zhengzhou, Henan, China

Key Points

- Purification and quantification of human erythroid progenitors provides a powerful means for studying normal and disordered erythropoiesis.
- Transcriptome data provides a resource for the mechanistic understanding of the generation of erythroid progenitors from hematopoietic stem cells (HSCs).

Burst-forming unit-erythroid (BFU-E) and colony-forming unit-erythroid (CFU-E) cells are erythroid progenitors traditionally defined by colony assays. We developed a flow cytometry-based strategy for isolating human BFU-E and CFU-E cells based on the changes in expression of cell surface markers during in vitro erythroid cell culture. BFU-E and CFU-E are characterized by CD45⁺GPA⁻IL-3R⁻CD34⁺CD36⁻CD71^{low} and CD45⁺GPA⁻IL-3R⁻CD34⁻CD36⁺CD71^{high} phenotypes, respectively. Colony assays validated phenotypic assignment giving rise to BFU-E and CFU-E colonies, both at a purity of ~90%. The BFU-E colony forming ability of CD45⁺GPA⁻IL-3R⁻CD34⁺CD36⁻CD71^{low} cells required stem cell factor and erythropoietin, while the CFU-E colony forming ability of CD45⁺GPA⁻IL-3R⁻CD34⁻CD36⁺CD71^{high} cells required only erythropoietin. Bioinformatic analysis of the RNA-sequencing data revealed unique transcriptomes at each differentiation stage. The sorting strategy was validated in uncultured primary cells isolated from bone marrow, cord blood, and peripheral blood, indicating that marker expression is not an artifact of in vitro cell culture, but represents an in vivo characteristic of erythroid progenitor populations. The ability to isolate highly pure human BFU-E and CFU-E progenitors will enable detailed cellular and molecular characterization of these distinct

progenitor populations and define their contribution to disordered erythropoiesis in inherited and acquired hematologic disease. Our data provides an important resource for future studies of human erythropoiesis. (Blood. 2014;124(24):3636-3645)

Introduction

Erythropoiesis is the process by which hematopoietic stem cells (HSCs) proliferate and differentiate to produce mature red blood cells. It is a tightly regulated process that can be divided into 2 stages, early and late. During the early stage of erythropoiesis, HSCs sequentially give rise to common myeloid progenitor, megakaryocyte-erythrocyte progenitor, burst-forming unit-erythroid (BFU-E), and colony-forming unit-erythroid (CFU-E) cells. BFU-E and CFU-E cells have been traditionally defined by colony assays.¹⁻⁶ During the late stage (also referred to as terminal erythroid differentiation), morphologically recognizable proerythroblasts undergo mitosis to produce basophilic, polychromatic, and orthochromatic erythroblasts. Orthochromatic erythroblasts expel their nuclei to generate reticulocytes. Finally, reticulocytes mature into red blood cells, initially in bone marrow (BM) and then in the circulation. Reticulocyte maturation includes the loss of intracellular organelles, such as mitochondria⁷⁻⁹ and ribosomes, and extensive membrane remodeling.¹⁰⁻¹²

To study the process of erythropoiesis, it is important to be able to isolate erythroid progenitors and erythroblasts at distinct stages of development. In this regard, considerable progress has been made in the murine system. Initial separation of BFU-E and CFU-E in mouse BM was achieved by unit gravity sedimentation.¹³ Isolation of mouse BFU-E and/or CFU-E by cell surface expression phenotype has also been described. Terszowski et al reported that lin⁻c-Kit⁺Sca-1⁻IL-7Rα⁻IL3Ra⁻CD41⁻CD71⁺ cells account for most of the CFU-E activity in mouse BM.¹⁴ In day 10.5 embryonic blood, aorta-gonad-mesonephros, or yolk sac, c-Kit⁺CD45⁺Ter119⁻CD71^{low} cells gave rise to BFU-Es and c-Kit⁺CD45⁻Ter119⁻CD71^{high} cells gave rise to CFU-Es.¹⁵ More recently, from embryonic day 14.5 to 15.5 fetal liver cells, Flygare et al isolated BFU-E and CFU-E cells by negative selection for Ter119, B220, Mac-1, CD3, Gr1, Sca-1, CD16/CD32, CD41, and CD34 cells, followed by separation based on the expression levels of CD71.¹⁶ Methods to isolate late stages of murine erythroid cells have also

Submitted July 14, 2014; accepted October 13, 2014. Prepublished online as *Blood* First Edition paper, October 22, 2014; DOI 10.1182/blood-2014-07-588806.

J.L., J.H., and P.B. contributed equally to the study.

The data reported in this article have been deposited in the Gene Expression Omnibus database (accession number GSE61566).

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2014 by The American Society of Hematology

been reported.^{17,18} By systemically examining changes in the expression pattern of more than 30 red-cell membrane proteins during murine terminal erythroid differentiation, we noted that the adhesion molecule CD44 exhibited a progressive and dramatic decrease from proerythroblasts to reticulocytes. This observation, in conjunction with cell size and the erythroid-specific marker Ter119, enabled us to devise a strategy for unambiguously distinguishing erythroblasts at all developmental stages during murine terminal erythroid differentiation,^{19,20} in a much more homogenous state than achieved in earlier work, based on expression levels of the transferrin receptor, CD71.¹⁸

In contrast to the extensive work on mouse erythropoiesis, our knowledge of the molecular markers for isolating distinct stages of human erythroid progenitors and erythroblasts is less well studied. We recently identified surface markers for isolating terminally differentiating erythroblasts at distinct developmental stages.²¹ Despite previous efforts,²²⁻²⁷ currently there is no established method to obtain highly pure human BFU-E and CFU-E cells. It has been reported that CD34 and CD45 are expressed in human hematopoietic progenitor cells²⁸ and that the expression of CD34 is lost at the CFU-E stage.^{26,29} It has also been noted that CD36 and CD71 are earlier erythroid markers than glycophorin A (GPA).^{30,31} In addition, different levels of interleukin (IL)-3R α expression on CD34⁺ cells are associated with preferential lineage readout, as IL-3R α ⁻ cells are enriched for erythroid, IL-3R α ^{lo} cells for multipotent, and IL-3R α ⁺ cells for granulocyte/macrophage (GM) colony-forming cells.³²⁻³⁴ These findings suggest that the above molecules could be potentially used as markers for isolating human BFU-E and CFU-E cells. However, the dynamic changes in the surface expression of these molecules during early human erythropoiesis have not been systematically examined.

In the present study, we examined the dynamic changes in expression levels of CD34, IL-3R, CD36, CD71, CD45, and GPA during the early stages of human erythropoiesis. The use of these surface markers enabled us to develop a fluorescence-activated cell sorter (FACS)-based strategy to isolate highly enriched populations of BFU-E and CFU-E cells. We further performed RNA sequencing (RNA-seq) on the sorted cell populations. Principal component analysis revealed that transcriptomes from differing stages, even between biologically different replicates were tightly clustered, validating the utility of the FACS-based sorting strategy. Bioinformatic analysis of RNA-seq data revealed significant differences between differentiation stages. The ability to isolate highly homogeneous populations of human BFU-E and CFU-E progenitors will facilitate detailed molecular and functional studies on erythroid progenitor populations, and help define the contribution of alterations in these progenitor populations in disorders with perturbed erythropoiesis, such as BM failure syndromes and myelodysplasia. The transcriptome data are likely to be a significant resource for studying the early stages of human erythropoiesis.

Materials and methods

Cell culture system and antibodies

Please see supplemental Methods, available on the *Blood* Web site, for detailed description of the antibodies and materials for culture system.

Flow cytometry and FACS of primary human samples

To perform flow cytometry analysis and to sort BFU-E and CFU-E cells from primary human samples including BM, cord blood, and peripheral blood, we isolated CD45⁺ cells from mononuclear cells (MNC) derived by Ficoll

density gradient separation, followed by positive selection using CD45 magnetic beads. The CD45⁺ cells were stained with the following mouse anti-human antibodies: PE-conjugated anti-CD34, fluorescein isothiocyanate-conjugated anti-CD36, PE-Cy7-conjugated anti-IL-3R (CD123), V450-conjugated anti-CD45, APC-conjugated anti-CD41 and anti-GPA, PerCp-conjugated anti-CD3, anti-CD4, anti-CD14, and anti-CD19, APC-H7-conjugated anti-CD71, or appropriate isotype controls (all from BD Biosciences) for 30 minutes in the dark. Cells were washed twice with 40 mL phosphate-buffered saline/0.5% bovine serum albumin, resuspended in 5 mL phosphate-buffered saline/0.5% bovine serum albumin, and stained with the viability marker 7-AAD on ice for 10 minutes in the dark. The analysis was performed using BD FACSDiva. Sorting was performed using a MoFlo high-speed cell sorter (Beckman Coulter). The number of BFU-E or CFU-E cells in one million MNCs was quantitated.

Colony assay

Cultured cells or sorted cells were plated in triplicate at a density of 200 cells in 1 ml of MethoCult H4434 classic medium (completed medium containing stem cell factor [SCF], IL-3, erythropoietin (EPO), and GM-colony stimulating factor), or in 1 mL of MethoCult H4330 medium with EPO only. The CFU-E and BFU-E colonies were defined according to the criteria described by An et al.³⁵

Other methods

Western blot analysis, cytospin preparation and staining, RNA-seq, RNA-seq analysis, and real-time quantitative reverse-transcription polymerase chain reaction were performed as described.^{21,35}

Results

Expression of surface proteins and colony forming ability of the cultured CD34⁺ cells during early erythropoiesis

To identify surface markers that distinguish human BFU-E and CFU-E cells, we examined expression levels of CD34, IL-3R, CD36, CD71, CD45, and GPA in cultured cells as a function of time by both western blotting (Figure 1A) and flow cytometric analysis (Figure 1B). In western blotting, we noted: (1) progressively decreased expression of CD34 and IL-3R; (2) progressively increased expression of CD36 and CD71; (3) unchanged expression levels of CD45; and (4) expression of GPA beginning on day 7. Flow cytometry demonstrated: (1) progressive decrease in the fraction of cells expressing CD34⁺ and IL-3R⁺; (2) progressive increase in the population of cells expressing CD36⁺ and CD71⁺; (3) no changes in the surface expression of CD45; and (4) no surface expression of GPA until day 7, when a small fraction of cells expressing a low level of GPA appeared.

To correlate the changes in cell surface expression with the ability to form erythroid colonies, we studied the time course of generation of these progenitor cells during the culture of the purified human CD34⁺ cells. Cells were taken every day for 7 days during the first phase of the 2-phase culture system and 200 cells were plated on semisolid MethoCult medium. Under our experimental conditions, 3 types of colonies were seen: BFU-E (which can be further classified as large BFU-E, BFU-E, and small BFU-E), CFU-E, and CFU-GM. Representative images of various colonies are shown in supplemental Figure 1. Quantitative enumeration of the number of colonies for 200 plated cells as a function of time in culture from 7 independent experiments is shown in Figure 1C. Both BFU-E and CFU-E colonies increased, with the number of BFU-E and CFU-E colonies peaking on day 4 and day 5, respectively. In contrast, CFU-GM colonies decreased gradually and almost disappeared on day 6.

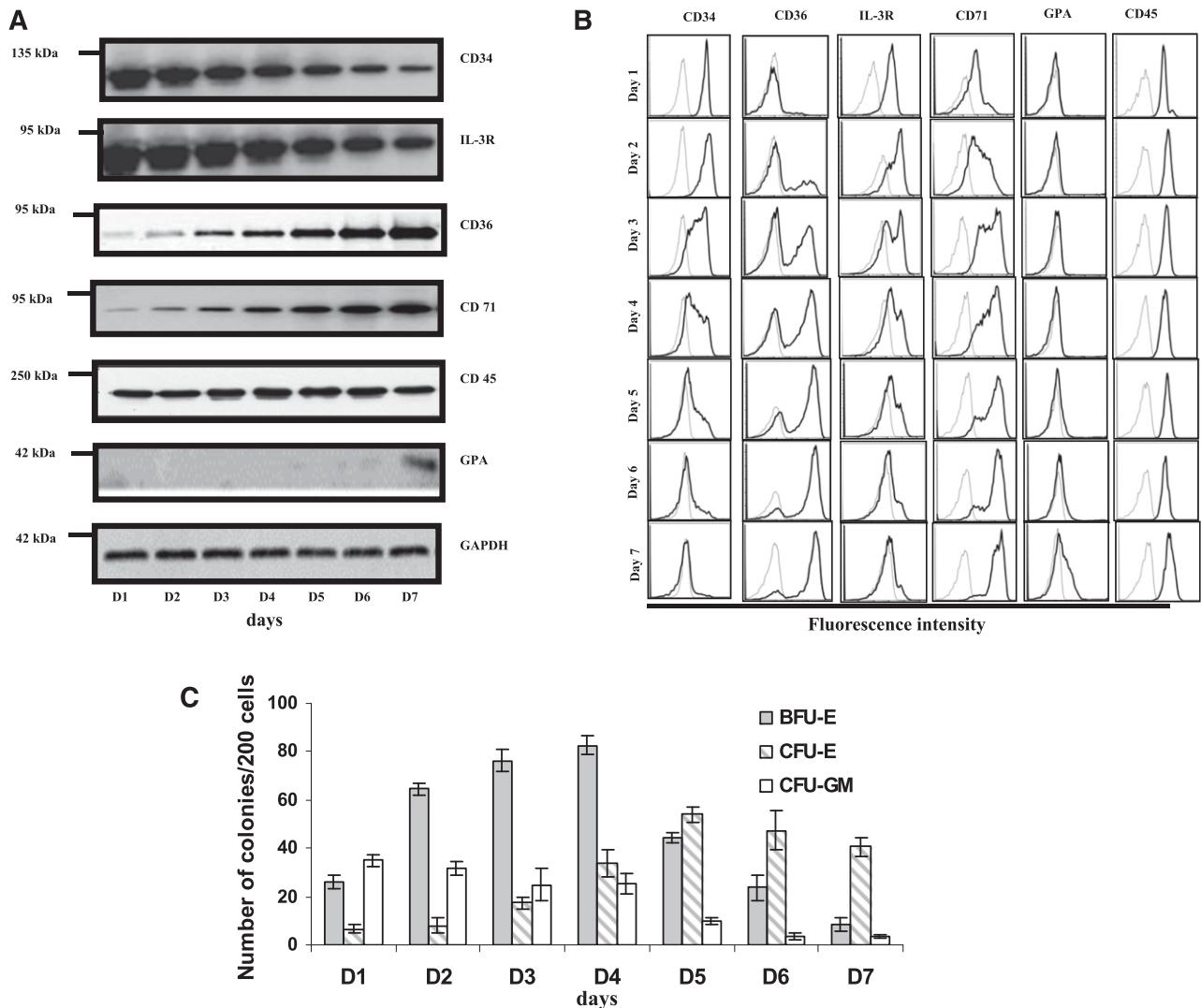


Figure 1. Expression of surface proteins and colony forming ability of the cultured CD34⁺ cells during early erythropoiesis. (A) Immunoblot analysis. Blots of sodium dodecyl sulfate-polyacrylamide gel electrophoresis of total cellular protein prepared from cells cultured from day 1 to day 7 were probed with antibodies against the indicated proteins. Note the decreased expression of CD34 and IL-3R, and the increased expression of CD36 and CD71. CD45 remained constant throughout and GPA was negative until day 6. Glyceraldehyde-3-phosphate dehydrogenase was used as loading control. (B) Flow cytometric analysis. The surface expression of indicated proteins was measured by flow cytometry. The representative profiles are shown. Gray line: autofluorescence control from unstained cells; black line: fluorescence from cells stained with indicated antibody. (C) Quantitation of colony forming ability. BFU-E colonies peak at day 4 then gradually decrease, whereas CFU-E colonies peak at day 5 and then decrease. The data are from 7 independent experiments.

Isolation and characterization of human BFU-E and CFU-E cells

It has been previously shown that high levels of IL-3R are expressed on B-lymphoid and myeloid progenitors and that erythroid progenitors express little or no surface IL-3R.^{32,34} However, it is not clear whether IL-3R⁻ cells can give rise to BFU-E, CFU-E, or both. To test this, we sorted IL-3R⁺ and IL-3R⁻ cells on day 4 from the cultured CD34⁺ cells and examined their colony forming ability. Figure 2A shows that although ~85% of colonies formed by IL3R⁺ cells were CFU-GM colonies, more than 90% of colonies formed by IL-3R⁻ cells were either BFU-E or CFU-E colonies.

Because both BFU-E and CFU-E cells reside in the IL-3R⁻ population, additional markers are needed to distinguish BFU-E from CFU-E. The observation that surface expression of CD34 progressively decreased while expression of CD36 progressively increased during early erythropoiesis, suggested that changes in surface expression of these proteins could be potential markers for distinguishing between BFU-E and CFU-E cells. To test this, we performed

multicolor labeling of cells cultured for 5 days with antibodies against IL-3R, CD34, and CD36, from collected IL-3R⁻CD34⁺CD36⁻ and IL-3R⁻CD34⁻CD36⁺ populations.

We characterized the colony forming ability of the sorted cells under different culture conditions. Figure 2Bi shows that in the presence of complete medium, ~90% of the colonies formed by IL-3R⁻CD34⁺CD36⁻ cells were BFU-Es, with 5% CFU-Es and 5% CFU-GM. However, in the EPO-only medium, IL-3R⁻CD34⁺CD36⁻ cells failed to form BFU-E colonies and only a few CFU-E colonies were observed (Figure 2Bii). Interestingly, the BFU-E colony forming ability of IL-3R⁻CD34⁺CD36⁻ cells was rescued by SCF but was not affected by IL-3 (Figure 2Biii). The lack of an IL-3 effect is expected because IL-3R is not expressed on their surface. In contrast to IL-3R⁻CD34⁺CD36⁻ cells, in complete medium, ~90% of the colonies formed by IL-3R⁻CD34⁻CD36⁺ cells were CFU-Es and very few CFU-GM were observed (Figure 2Ci). Importantly, in the EPO only medium, IL-3R⁻CD34⁻CD36⁺ cells continued to give rise to CFU-E colonies at a similar efficiency and purity as the complete medium (Figure 2Cii).

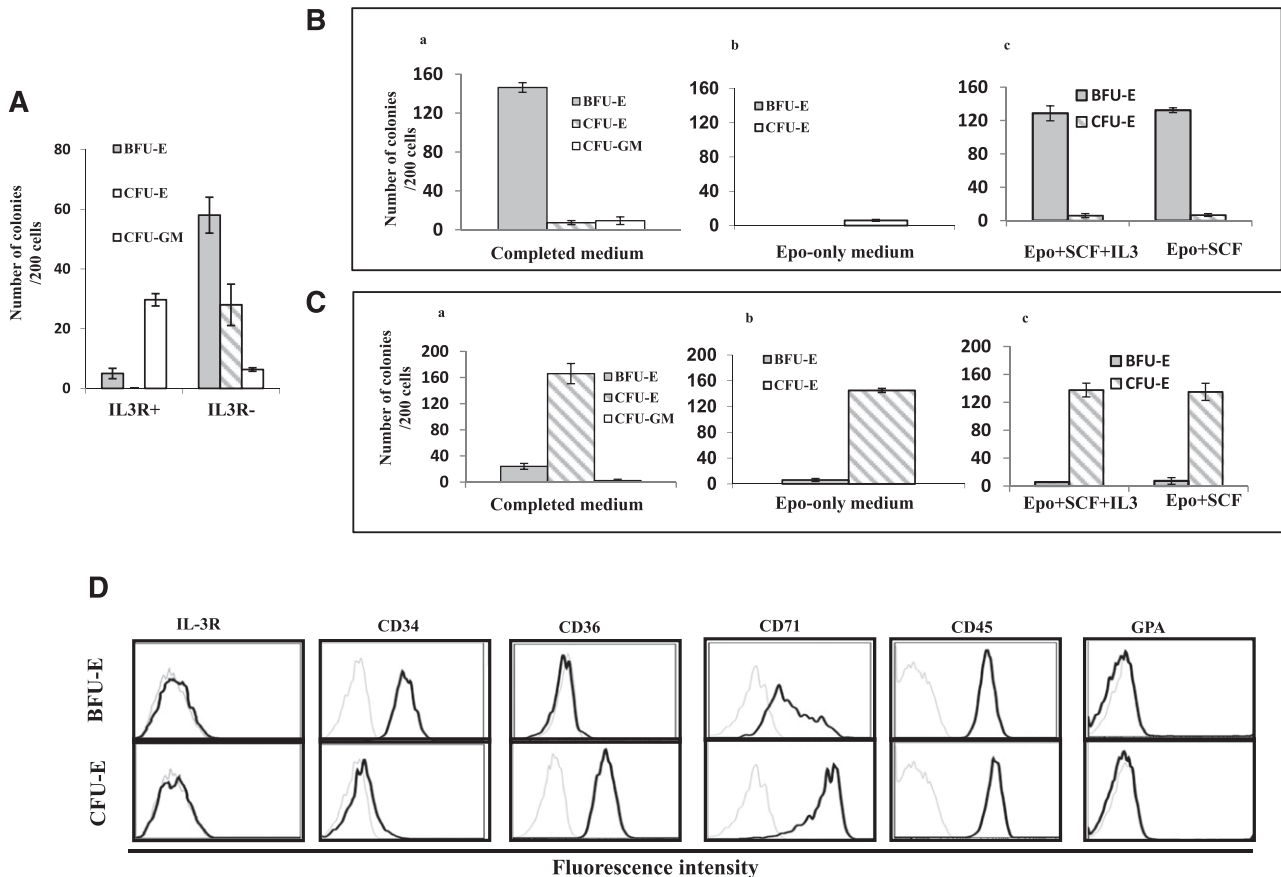


Figure 2. Isolation of BFU-E and CFU-E cells by cell-sorting using CD34, CD36, and IL-3R as markers. (A) Colony forming ability of the sorted IL-3R⁺ and IL-3R⁻ cells. (B) Colony forming ability of the sorted IL-3R⁻ CD34⁺ CD36⁻ in complete medium, EPO-only medium, and EPO+SCF+IL3 or EPO+SCF medium. (C) Colony forming ability of the sorted IL-3R⁻ CD34⁻ CD36⁺ cells in complete medium, EPO-only medium, and EPO+SCF+IL3 or EPO+SCF medium. (D) Expression of surface markers of sorted IL-3R⁻ CD34⁺ CD36⁻ and IL-3R⁻ CD34⁻ CD36⁺ cells. The surface expression of indicated proteins was measured by flow cytometry. The representative profiles are shown. Gray line: autofluorescence control from unstained cells; black line: fluorescence from cells stained with indicated antibody. Results from 3 independent experiments are shown.

Moreover, neither SCF nor IL-3 had an effect on the numbers or purity of colonies formed by IL-3R⁻ CD34⁻ CD36⁺ cells (Figure 2Ciii). Our findings are consistent with previously well-documented findings that BFU-E colony formation requires a combination of several cytokines including EPO and SCF, while EPO alone is sufficient for CFU-E colony formation.^{5,6,36} Together, our findings demonstrate that it is feasible to obtain highly pure human BFU-E and CFU-E cells using IL-3R, CD34, and CD36 as cell surface markers. Overall, our culture and sorting conditions allowed us to simultaneously obtain ~1 × 10⁶ BFU-E cells and 5 × 10⁶ CFU-E cells from a starting population of 1 × 10⁶ CD34⁺ cells.

After establishing that IL-3R⁻ CD34⁺ CD36⁻ cells are BFU-E cells and IL-3R⁻ CD34⁻ CD36⁺ cells are CFU-E cells by colony assays and by their distinct response to EPO, we examined the expression of various surface markers on sorted BFU-E and CFU-E cells. As shown in Figure 2D, both BFU-E and CFU-E cells are CD45-positive and GPA and IL-3R are negative, although they differ in the expression of CD34, CD36, and CD71. Whereas BFU-E cells are CD34-positive, CD36-negative, and CD71 low, CFU-E cells are CD34-negative and both CD36 and CD71 are positive. Thus, human BFU-E cells are CD45⁺ GPA⁻ IL-3R⁻ CD34⁺ CD36⁻ CD71^{low}, whereas CFU-E cells are CD45⁺ GPA⁻ IL-3R⁻ CD34⁻ CD36⁺ CD71^{high}.

RNA-seq

To examine the changes in gene expression across the entire transcriptome during early stage erythropoiesis, we performed deep

sequencing on purified CD34⁺, BFU-E, and CFU-E cells. Data were analyzed as described previously.³⁵ For comparison, we also included deep sequence data of proerythroblasts from our previous study.³⁵ Initially, we analyzed how many annotated genes were expressed at each differentiation stage, with counts >0. We found that on average, 11 147, 10 380, 10 067, and 9822 RefSeq genes were expressed in CD34⁺, BFU-E, CFU-E, and proerythroblasts, respectively. Thus, 41% to 47% of known human genes were expressed during early erythroid differentiation. Quantitative real-time polymerase chain reaction was performed to validate expression levels of representative messenger RNA transcripts detected by RNA-seq (supplemental Figure 2). It is interesting to note that although the expression of GATA2 was dramatically decreased from BFU-E to CFU-E stage, that of α-hemoglobin stabilizing protein was significantly increased.

Transcriptome profiles support FACS-based sorting methodology

We then performed principle component analysis (PCA) using all genes with counts >0 shown using the edgeR package.³⁷ Figure 3A shows that 3 biological replicates at the same stage of development are tightly clustered, demonstrating that the samples were reproducible and exhibit similar transcription profiles. Figure 3A also shows the clear separation of each staged population, indicating that the transcription profiles of each population are distinct from each other. Correlation plots in Figure 3B indicate that adjacent samples show higher similarity

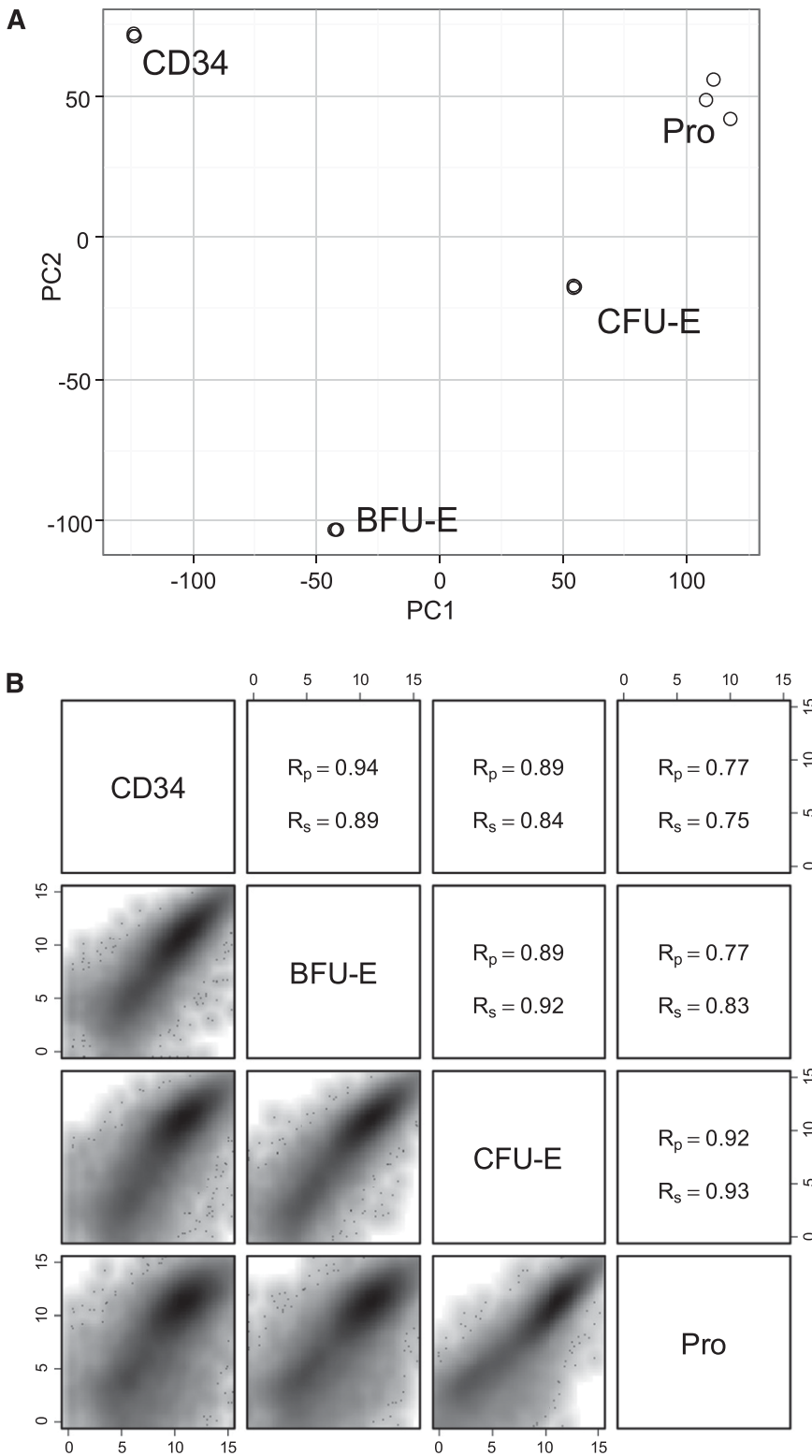


Figure 3. PCA of expressed genes and pairwise comparisons of different stages of early human erythroid progenitors. (A) PCA plot of 3 biological replicates at 4 different stages of early erythroid differentiation. (B) Pairwise comparison of the gene expression at different stages of early erythroid progenitors. (Upper right) Pearson and Spearman coefficient values. (Lower left) Scatter density plot. PC, principle component; R_p , Pearson product-moment correlation coefficient; R_s , Spearman rank correlation coefficient.

with corresponding Pearson coefficients between 0.89 and 0.94 and Spearman coefficients between 0.89 and 0.93, whereas distant samples have lower coefficients. The above findings demonstrate that transcriptome analyses support the FACS-based methodology used to purify distinct differentiation stages. It also indicates that differentiation stages are distinct, with both shared and dissimilar gene expression profiles defining each stage within the transcriptional space.

Temporal patterns of gene expression during early stage erythropoiesis

To further analyze the differences in transcriptome profiles during early stage erythropoiesis, we performed pairwise comparison of differentially expressed genes of adjacent stages. Differentially expressed genes were defined as those with a log twofold change >2 and fpkm >10 on

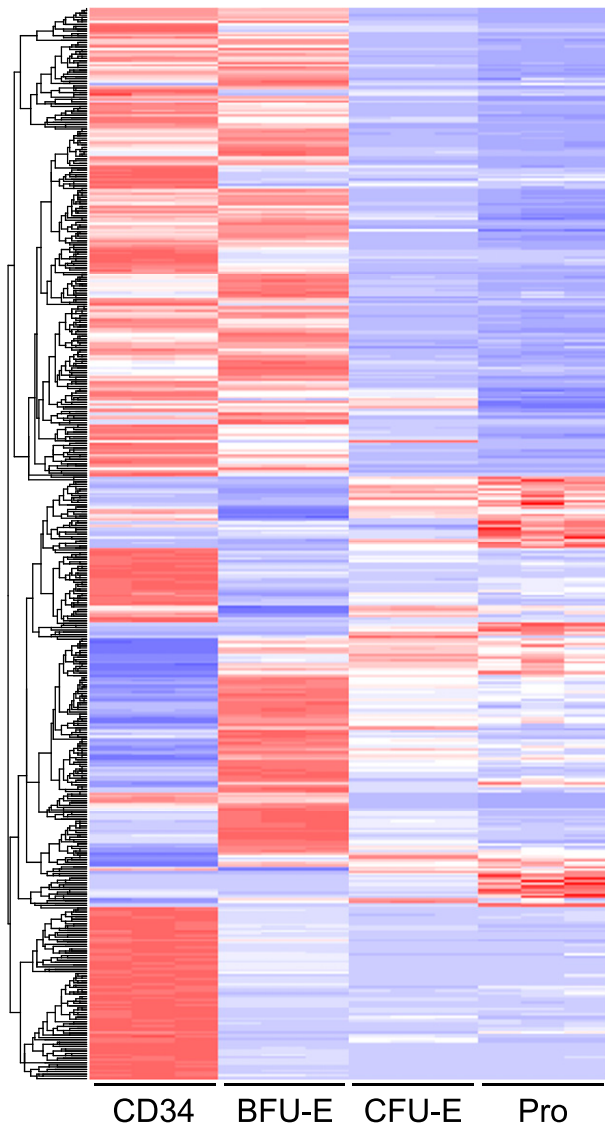


Figure 4. Global gene expression at different stages of human early erythroid progenitor. Expression values of differentially expressed genes between the different stages are shown as a heat map. In each row, the red, white, and blue colors represent the expression level (from high to low) of a particular gene. The rows are organized by hierarchical clustering with complete linkage and Euclidean distance metric.

at least one side of the comparison. Genes meeting these criteria were extracted from the output of CuffDiff using the cummeRbund package.³⁸ The heat map in Figure 4 shows that both upregulation and downregulation of gene expression occur during the transition from one stage to the next. From CD34⁺ cell to BFU-E, 783 genes were differentially expressed, out of which 399 genes were upregulated and 444 were downregulated. The transition from BFU-E to CFU-E was characterized predominantly by downregulation of gene expression, with only 164 genes being upregulated and 479 genes being downregulated. Relatively fewer genes were changed from CFU-E stage to proerythroblasts, with 126 genes upregulated and 187 genes downregulated. The differentially expressed genes are listed in supplemental Table 1.

Differentiation stages are transcriptionally enriched for genes of differing function

To gain insights into the functional consequences of the differentially expressed genes, we analyzed the expression patterns of all differentially

expressed, coregulated genes using k-means clustering by means of self-organized maps.³⁹ As shown in Figure 5, the differentially expressed genes clustered into 4 major groups. Group 1 contained 430 genes and had dramatic downregulation at the CD34 to BFU-E transition. In contrast, the expression of group 2 genes (225 genes) was significantly increased from CD34 to BFU-E. Group 3 contained 174 genes and showed the greatest downregulation at the transition from BFU-E to CFU-E. Finally, group 4 genes (268 genes) had significant upregulation from BFU-E onwards. We then performed Gene Ontology (GO) analysis of differentially expressed genes within these 4 groups and identified associated enriched GO terms to gain insights into the biological processes regulated during early erythroid differentiation.^{40,41} In group 1, genes involved in immune response and cell activation were significantly downregulated during the commitment of HSCs to erythroid lineage. Interestingly, during the same transition period, the genes related to cell division and cell cycle were significantly upregulated (group 2). In group 3, with high levels of expression in CD34 and BFU-E that decreased during differentiation, GO terms significantly enriched for differentially expressed genes related to protein synthesis, including translation and ribosome biogenesis, were identified. In group 4, genes enriched for erythroid-specific terms associated with hemoglobin complex and spectrin membrane skeleton were identified.

In addition to the 4 major groups, other minor clusters were also identified. The profiles of these clusters are shown in supplemental Figure 3 and the corresponding heat maps of the genes in each cluster are shown in supplemental Figure 4. The list of genes in each cluster is shown in supplemental Table 2. GO terms enrichment analysis results for each cluster are shown in supplemental Table 3.

Changes of transcription factor gene expression during early erythropoiesis

Transcription factors are proteins that detect and bind to DNA regulatory sequences and regulate gene expression. They play important roles in cell fate decision and cell differentiation. To examine the changes during transition from the previous stage to the next stage, we performed pairwise comparison. Figure 6A shows that from CD34 to BFU-E, about equal numbers of transcription factors were upregulated or downregulated. Figure 6B shows that during the transition from BFU-E to CFU-E, most of the transcription factors were downregulated. In contrast, during the transition from CFU-E to Pro, more transcription factors were upregulated (Figure 6C). The list of up or downregulated transcription factors is shown in supplemental Table 4. Analysis of transcription factors by changes in absolute expression revealed that the top and the most differentially expressed transcription factors in each stage are different (Table 1).

Isolation and quantification of BFU-E and CFU-E cells from primary human samples

Having developed a flow cytometry-based strategy that distinguishes BFU-E and CFU-E cells using cells obtained from an *in vitro* CD34⁺ culture system, we applied this strategy to the isolation and quantification of BFU-E and CFU-E populations from primary human samples, including BM, cord blood, and peripheral blood. Since both BFU-E and CFU-E cells are CD45⁺, we first obtained CD45⁺ cells from MNCs of BM, cord blood, or peripheral blood by positive selection using CD45⁺ beads. We then performed multicolor labeling of the CD45⁺ cells with antibodies against CD45, CD3, CD4, CD14, CD19, CD41, IL-3R, CD36, CD34, GPA, and CD71. The entire gating strategy to identify and quantify BFU-E and CFU-E following labeling with various antibodies is shown in Figure 7A,

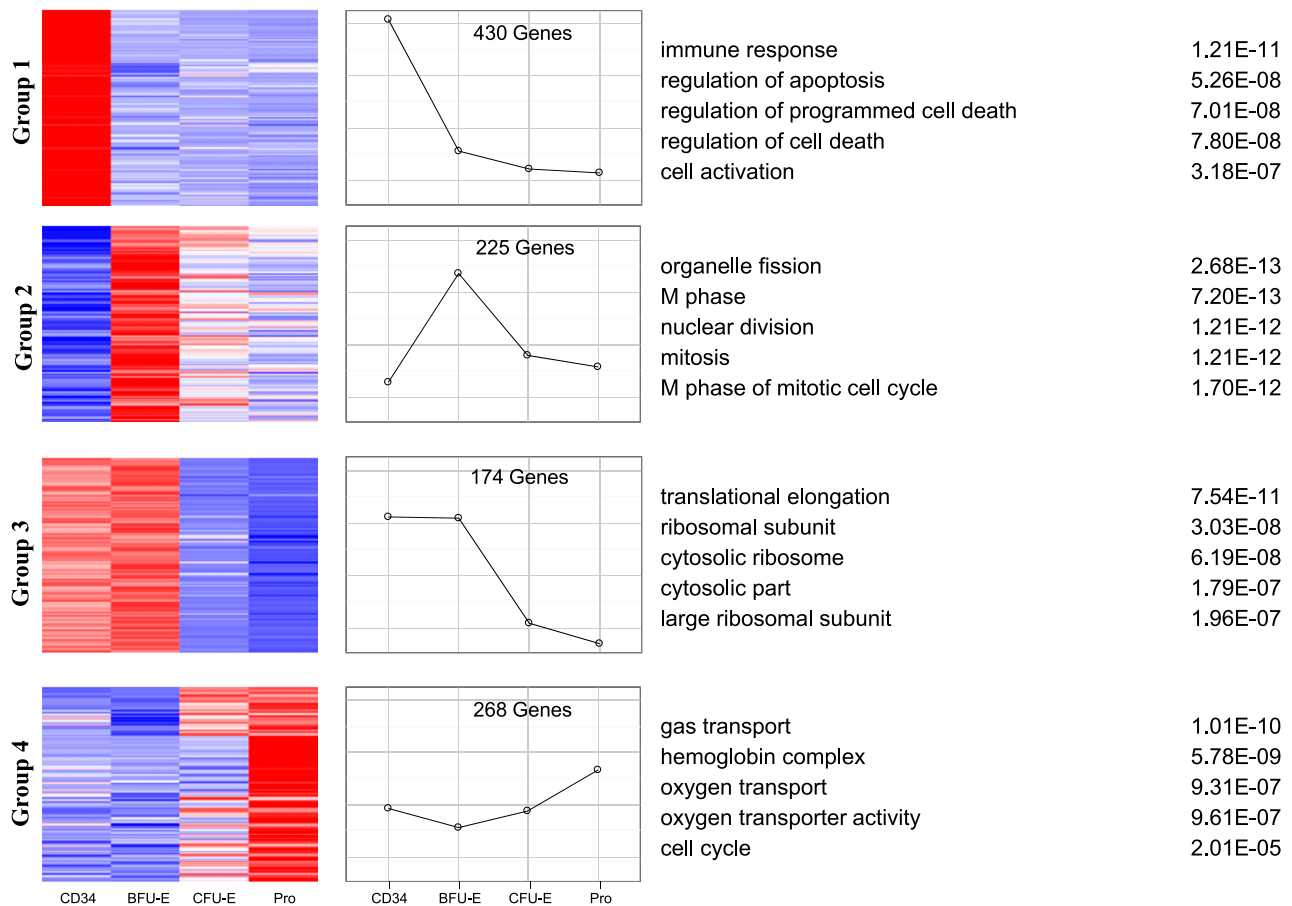


Figure 5. Clusters of gene expression across the different stages of early erythroid differentiation. Differentially expressed genes between adjacent stages with log fold >2 and fpkm >10 were clustered into 16 groups by the process of self-organized maps. The 4 major groups are illustrated with heat maps of the gene members of the cluster (left) and the expression profile is shown graphically (middle). GO analysis of the genes within the cluster identified the top 5 GO terms and their corresponding *P*-values (right).

using BM cells as an example. The same gating strategies were used for cord blood and peripheral blood samples (supplemental Figure 6). The colony forming ability of the BFU-E and CFU-E cells sorted from BM, cord blood, and peripheral blood is shown in Figure 7B, which demonstrated that they gave rise to the corresponding colonies at a purity of more than 90% with the plating efficiency of 50% to 60%. Quantitative analysis performed using 4 BM samples showed that there were more CFU-E cells (ranging from 1500 to 5000/ 10^6 MNCs) than BFU-E cells (ranging from 300 to 1700/ 10^6 MNCs) in human BM. The quantitative analysis of BFU-E and CFU-E numbers in peripheral blood and cord blood is summarized in Figure 7C. It shows that there are more BFU-E cells than CFU-E cells in both peripheral blood and cord blood. The BFU-E numbers ranged from 130 to 1500/ 10^6 MNCs in peripheral blood and CFU-E cells were almost undetectable in peripheral blood (only 4 to 40/ 10^6 MNCs). In cord blood, the numbers of both BFU-E and CFU-E cells were significantly higher than in adult peripheral blood, with 900 to 3500 BFU-E/ 10^6 MNCs and 40 to 420 CFU-E/ 10^6 MNCs.

Discussion

BFU-E and CFU-E cells, the two earliest committed erythroid progenitor populations, have been primarily characterized by colony

assays.¹⁻⁶ Detailed cellular and molecular studies of human BFU-E and CFU-E cells have been hindered due to the inability to obtain sufficient cells of high purity. The present study aimed to address this issue. We demonstrated that BFU-E and CFU-E cells were phenotypically identified as CD45⁺GPA⁻IL-3R⁻CD34⁺CD36⁻CD71^{low} and CD45⁺GPA⁻IL-3R⁻CD34⁻CD36⁺CD71^{high}, respectively. Using combinations of these markers, we were able to simultaneously obtain about 1×10^6 BFU-E cells and 5×10^6 CFU-E cells starting from 1×10^6 CD34⁺ cells. The identities of isolated BFU-E and CFU-E cells were demonstrated by their ability to form distinct colonies, their distinct requirement for EPO, and their distinct gene expression profiles.

Examination of the surface expression of CD45, GPA, and CD71 on the sorted erythroid progenitor cell populations revealed that BFU-E and CFU-E cells were CD45-positive and GPA-negative. In the case of CD71, whereas BFU-E cells were CD71 low, CFU-E cells were CD71 high. The findings regarding GPA expression during human erythropoiesis have been controversial. While Loken et al reported that GPA was present on CFU-E,⁴² Gahmberg et al reported that it is present on the basophilic erythroblasts.⁴³ Several other studies have reported that GPA begins to appear on proerythroblasts.^{30,44,45} Our data demonstrate that GPA is not present on CFU-E cells. The sorted GPA-positive cells from our culture system were unable to give rise to CFU-E colonies (our unpublished data). Thus, it is important to note that although GPA is considered as a specific marker for erythroid lineage, it is not expressed on BFU-E or CFU-E cells. Similarly,

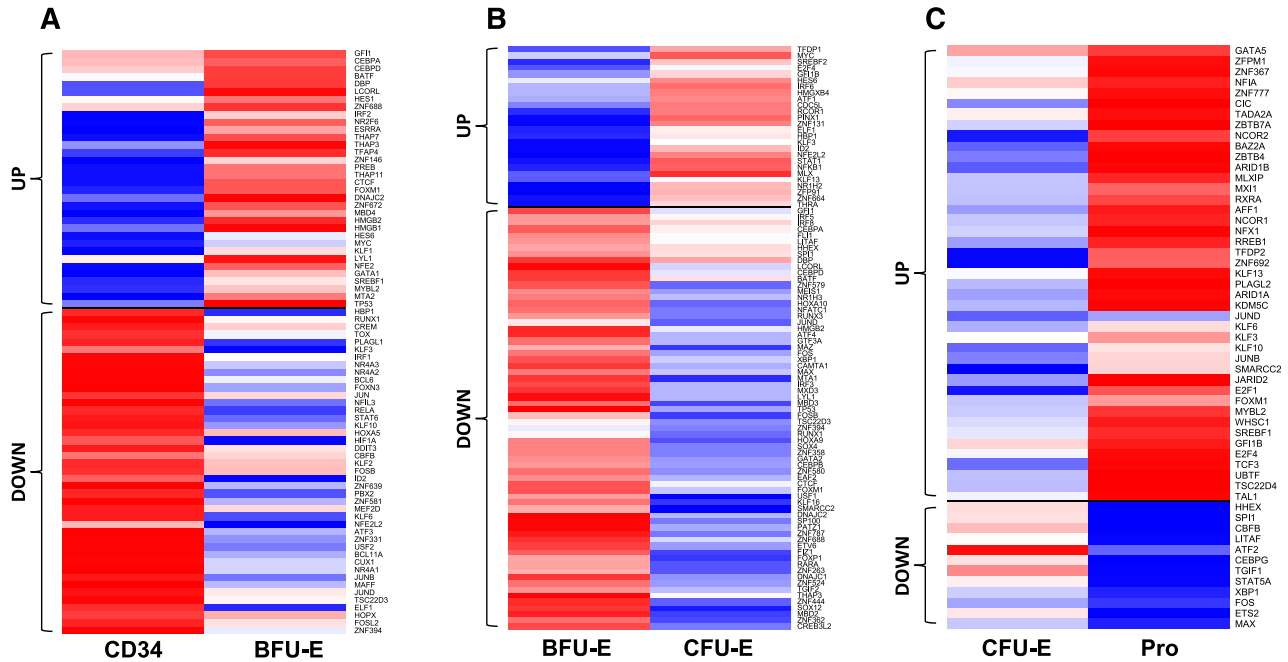


Figure 6. Pairwise comparison of transcription factor genes in human early stage erythroid progenitors. Transcription factor genes with log twofold >1 and fpkm >10 in at least 1 stage for CD34 to BFU-E (A), BFU-E to CFU-E (B), and the CFU-E to Pro (C) transitions are shown. In each row, the red, white, and blue colors represent the expression level (from high to low) for each gene.

although CD45 is considered to be a pan-leukocyte marker, both BFU-E and CFU-E cells also express this marker.

Several global gene expression studies have been performed on differentiating erythroid cells, representing cells at different developmental stages.⁴⁶⁻⁴⁸ In this study, we present the transcriptome data of highly enriched human erythroid progenitor cells. These data sets not only strongly support the FACS-based strategy for purification

of human BFU-E and CFU-E cells, but also provide significant resources for detailed cellular and molecular characterization of these cell populations. Our comprehensive bioinformatics analysis of RNA-seq data revealed potential novel pathways for erythroid lineage commitment, as well as erythroid progenitor differentiation. For example, the upregulation of cell-cycle-related genes from CD34 to BFU-E suggested the transition from a quiescent status of HSC to a more proliferative status of the erythroid progenitor cell. Furthermore, transcription factor analysis revealed that distinct sets of transcription factors were differentially upregulated or downregulated during early stage erythropoiesis. These findings suggest that proliferation and differentiation of these cell populations are probably mediated by a different transcriptional network.

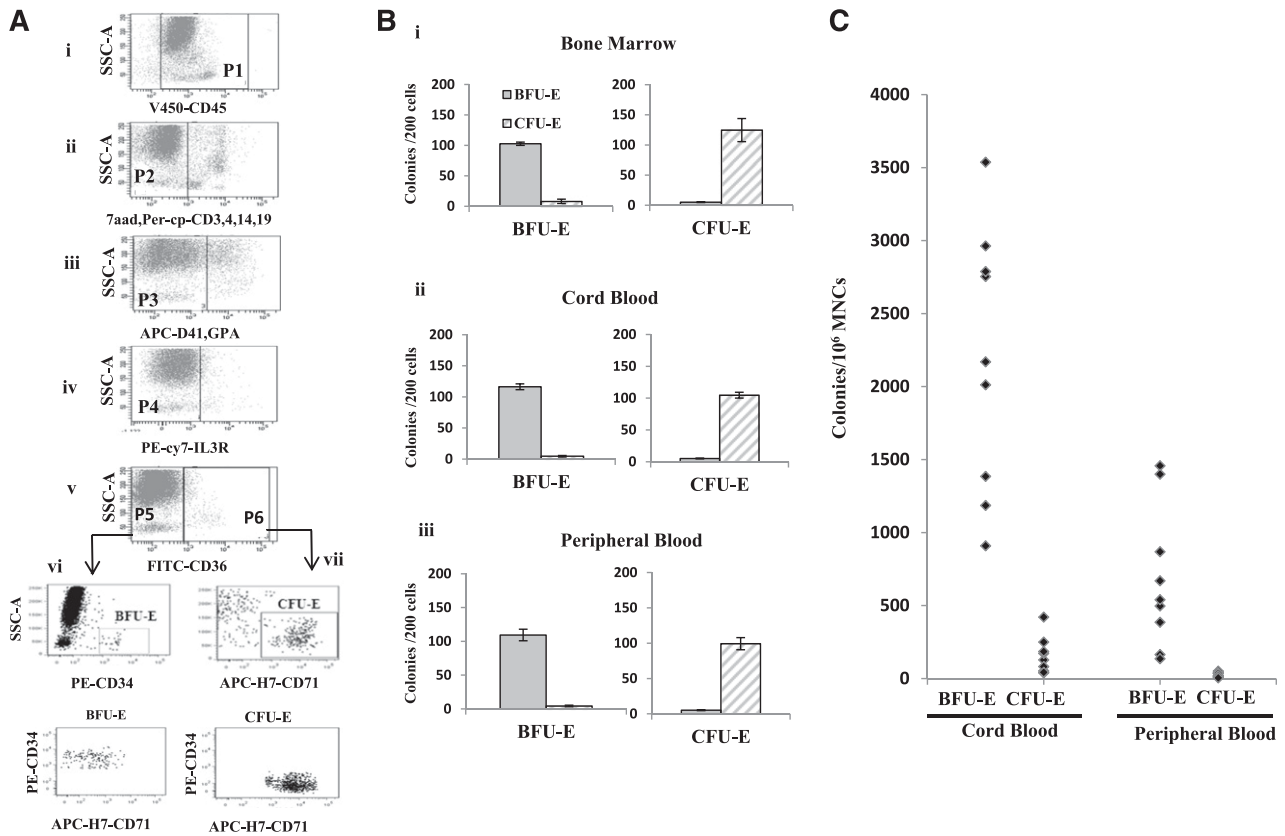
In the case of some erythroid-specific genes, it has been previously shown that in the mouse, early erythroid progenitors coexpressed both GATA-1 and GATA-2, whereas late erythroid progenitors did not express GATA-2 but highly expressed GATA-1.⁴⁹ Our findings that both GATA1 and GATA2 are expressed in BFU-E and that GATA1 is expressed but GATA2 is almost undetectable in CFU-E cells imply that the expression patterns of GATA1 and GATA2 in human erythroid progenitors are similar to that of the mouse. It has been demonstrated that a GATA2/GATA1 switch is important for normal erythropoiesis.⁵⁰ Our findings demonstrated that this decisive switching of GATA factors occurred between BFU-E and CFU-E. In contrast to GATA2, AHSP was not detectable in BFU-E but was highly expressed in CFU-E, suggesting that these 2 molecules can be used as markers for BFU-E and CFU-E, respectively.

Most importantly, the combination of surface markers we have identified using the in vitro erythroid culture system has enabled us, for the first time, to isolate and to quantify human BFU-E and CFU-E cells from primary human samples, including BM, cord blood, and peripheral blood. The ability to isolate BFU-E and CFU-E cells from primary human cell populations using our sorting strategy implies that the expression of the identified markers is not an artifact

Table 1. Top up or downregulated transcription factors between each stage

	CDM to BFU-E	BFU-E to CFU-E	CFU-E to Pro
1	JUND	ATF4	GF1B
2	TSC22D3	HMGB2	ATF4
3	HMGB2	HMGA1	VBX1
4	ATF4	HMGB1	E2F4
5	HMGB1	VBX1	MYC
6	HMGA1	GTF3A	TFDP1
7	NR4A1	FOS	MVBL2
8	MAFF	MYC	HMGA1
9	ELF1	XBPI	GTF3A
10	NFE2	MAZ	SREBF1
11	HOPX	JUND	HMQB2
12	KLF1	LYL1	MAZ
13	ZNF394	KLF1	KLF1
14	FOSL2	CAMTA1	TCF3
15	MAZ	NFE2	NFE2
16	LYL1	TFDP1	STAT5A
17	JUNB	TP53	MYB
18	XBPI	HE56	UBTT
19	TP53	MAX	XBPI
20	QATA1	MBD3	TALI

Significantly up or downregulated genes have fpkm >10 in at least one stage and log twofold change >1. Genes are ranked in order of decreasing absolute change in fpkm. Upregulated genes are denoted in bold and downregulated genes are shown in plain text.



of the *in vitro* cell culture, but represents an *in vivo* characteristic of native erythroid progenitor populations. BFU-E and CFU-E in normal, as well as diseased BM and blood, have been enumerated using colony assays.⁵¹⁻⁵³ However, the results have been inconsistent. Because colony assays are time-consuming and subjective, our flow cytometry based approach provides significant advantages over conventional colony assays. Furthermore, since CFU-E isolated directly from marrow and blood have experienced only normal concentrations of EPO (5 to 25 mU/mL) *in vivo*, they should be much better for studying EPO signaling, EPO-induced transcriptional events etc, than CFU-E populations obtained by culturing blood CD34⁺ cells for days in orders-of-magnitude higher EPO levels.

In conclusion, we have defined the surface markers for BFU-E and CFU-E cells that has enabled us to develop a method to simultaneously isolate BFU-E and CFU-E cells from cultured erythroid cells and from primary human samples. Our findings will be useful for detailed cellular and molecular characterization of these distinct erythroid progenitor populations and in delineating mechanisms of disordered erythropoiesis in inherited and acquired hematologic disease. The purified cells could also be used for screening drugs that would specifically act on BFU-E or CFU-E cells, which in turn could lead to better therapeutic approaches for patients with altered erythropoiesis.

Acknowledgments

This study was supported in part by the National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases (grants DK100810, DK26263, and HL106184).

Authorship

Contribution: J. Li, J.H., P.B., F.X., L.C., and H.Y. performed research and analyzed the data; J.J. and J. Lane provided normal human BM samples; and P.G.G., N.M., J. Liu, and X.A. designed experiments, analyzed the data, and wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Jing Liu, State Key Laboratory of Medical Genetics and School of Life Sciences, Central South University, Changsha, Hunan 410078, China; e-mail: jingliucs@hotmail.com; and Xiuli An, Laboratory of Membrane Biology, New York Blood Center, 310 E 67th St, New York, NY 10065; e-mail: xan@nybloodcenter.org.

References

- Stephenson JR, Axelrad AA, McLeod DL, Shreeve MM. Induction of colonies of hemoglobin-synthesizing cells by erythropoietin in vitro. *Proc Natl Acad Sci USA*. 1971;68(7):1542-1546.
- McLeod DL, Shreeve MM, Axelrad AA. Improved plasma culture system for production of erythrocytic colonies in vitro: quantitative assay method for CFU-E. *Blood*. 1974;44(4):517-534.
- Iscove NN, Sieber F, Winterhalter KH. Erythroid colony formation in cultures of mouse and human bone marrow: analysis of the requirement for erythropoietin by gel filtration and affinity chromatography on agarose-concanavalin A. *J Cell Physiol*. 1974;83(2):309-320.
- Moriyama Y, Fisher JW. Effects of testosterone and erythropoietin on erythroid colony formation in human bone marrow cultures. *Blood*. 1975;45(5):665-670.
- Gregory CJ, Eaves AC. Human marrow cells capable of erythropoietic differentiation in vitro: definition of three erythroid colony responses. *Blood*. 1977;49(6):855-864.
- Gregory CJ, Eaves AC. Three stages of erythropoietic progenitor cell differentiation distinguished by a number of physical and biologic properties. *Blood*. 1978;51(3):527-537.
- Gronowicz G, Swift H, Steck TL. Maturation of the reticulocyte in vitro. *J Cell Sci*. 1984;71:177-197.
- Zhang J, Randall MS, Loyd MR, et al. Mitochondrial clearance is regulated by Atg7-dependent and -independent mechanisms during reticulocyte maturation. *Blood*. 2009;114(1):157-164.
- Kundu M, Lindsten T, Yang CY, et al. Ulk1 plays a critical role in the autophagic clearance of mitochondria and ribosomes during reticulocyte maturation. *Blood*. 2008;112(4):1493-1502.
- Chasis JA, Prenant M, Leung A, Mohandas N. Membrane assembly and remodeling during reticulocyte maturation. *Blood*. 1989;74(3):1112-1120.
- Waugh RE, Mantalaris A, Bauserman RG, Hwang WC, Wu JH. Membrane instability in late-stage erythropoiesis. *Blood*. 2001;97(6):1869-1875.
- Mel HC, Prenant M, Mohandas N. Reticulocyte motility and form: studies on maturation and classification. *Blood*. 1977;49(6):1001-1009.
- Heath DS, Axelrad AA, McLeod DL, Shreeve MM. Separation of the erythropoietin-responsive progenitors BFU-E and CFU-E in mouse bone marrow by unit gravity sedimentation. *Blood*. 1976;47(5):777-792.
- Terszowski G, Waskow C, Conradt P, et al. Prospective isolation and global gene expression analysis of the erythrocyte colony-forming unit (CFU-E). *Blood*. 2005;105(5):1937-1945.
- Stumpf M, Waskow C, Krötschel M, et al. The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proc Natl Acad Sci USA*. 2006;103(49):18504-18509.
- Flygare J, Rayon Estrada V, Shin C, Gupta S, Lodish HF. HIF1alpha synergizes with glucocorticoids to promote BFU-E progenitor self-renewal. *Blood*. 2011;117(12):3435-3444.
- McGrath KE, Bushnell TP, Palis J. Multispectral imaging of hematopoietic cells: where flow meets morphology. *J Immunol Methods*. 2008;336(2):91-97.
- Liu Y, Pop R, Sadegh C, Brugnara C, Haase VH, Socolovsky M. Suppression of Fas-FasL coexpression by erythropoietin mediates erythroblast expansion during the erythropoietic stress response in vivo. *Blood*. 2006;108(1):123-133.
- Chen K, Liu J, Heck S, Chasis JA, An X, Mohandas N. Resolving the distinct stages in erythroid differentiation based on dynamic changes in membrane protein expression during erythropoiesis. *Proc Natl Acad Sci USA*. 2009;106(41):17413-17418.
- Liu J, Zhang J, Ginzburg Y, et al. Quantitative analysis of murine terminal erythroid differentiation in vivo: novel method to study normal and disordered erythropoiesis. *Blood*. 2013;121(8):e43-e49.
- Hu J, Liu J, Xue F, et al. Isolation and functional characterization of human erythroblasts at distinct stages: implications for understanding of normal and disordered erythropoiesis in vivo. *Blood*. 2013;121(16):3246-3253.
- Lansdorp PM, Sutherland HJ, Eaves CJ. Selective expression of CD45 isoforms on functional subpopulations of CD34+ hemopoietic cells from human bone marrow. *J Exp Med*. 1990;172(1):363-366.
- Inada T, Iwama A, Sakano S, Ohno M, Sawada K, Suda T. Selective expression of the receptor tyrosine kinase, HTK, on human erythroid progenitor cells. *Blood*. 1997;89(8):2757-2765.
- Freyssinier JM, Lecoq-Lafon C, Amsellem S, et al. Purification, amplification and characterization of a population of human erythroid progenitors. *Br J Haematol*. 1999;106(4):912-922.
- Bonig H, Chang KH, Nakamoto B, Papayannopoulou T. The p67 laminin receptor identifies human erythroid progenitor and precursor cells and is functionally important for their bone marrow lodgment. *Blood*. 2006;108(4):1230-1233.
- Tirelli V, Ghinassi B, Migliaccio AR, et al. Phenotypic definition of the progenitor cells with erythroid differentiation potential present in human adult blood. *Stem Cells Int*. 2011;2011:602483.
- Sawada K, Krantz SB, Dai CH, et al. Purification of human blood burst-forming units-erythroid and demonstration of the evolution of erythropoietin receptors. *J Cell Physiol*. 1990;142(2):219-230.
- Iguchi T, Sogo S, Hisha H, et al. HGF activates signal transduction from EPO receptor on human cord blood CD34+/CD45+ cells. *Stem Cells*. 1999;17(2):82-91.
- Weissman IL, Shizuru JA. The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. *Blood*. 2008;112(9):3543-3553.
- Edelman P, Vinci G, Villeval JL, et al. A monoclonal antibody against an erythrocyte ontogenic antigen identifies fetal and adult erythroid progenitors. *Blood*. 1986;67(1):56-63.
- Okumura N, Tsuji K, Nakahata T. Changes in cell surface antigen expressions during proliferation and differentiation of human erythroid progenitors. *Blood*. 1992;80(3):642-650.
- Huang S, Chen Z, Yu JF, et al. Correlation between IL-3 receptor expression and growth potential of human CD34+ hematopoietic cells from different tissues. *Stem Cells*. 1999;17(5):265-272.
- Manz MG, Miyamoto T, Akashi K, Weissman IL. Prospective isolation of human clonogenic common myeloid progenitors. *Proc Natl Acad Sci USA*. 2002;99(18):11872-11877.
- Kohn LA, Hao QL, Sasidharan R, et al. Lymphoid priming in human bone marrow begins before expression of CD10 with upregulation of L-selectin. *Nat Immunol*. 2012;13(10):963-971.
- An X, Schulz VP, Li J, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*. 2014;123(22):3466-3477.
- Udupa KB, Reissmann KR. In vivo erythropoietin requirements of regenerating erythroid progenitors (BFU-e, CFU-e) in bone marrow of mice. *Blood*. 1979;53(6):1164-1171.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140.
- Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat Protoc*. 2012;7(3):562-578.
- Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc Natl Acad Sci USA*. 1999;96(6):2907-2912.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.
- Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1-13.
- Loken MR, Shah VO, Dattilio KL, Civin CI. Flow cytometric analysis of human bone marrow: I. Normal erythroid development. *Blood*. 1987;69(1):255-263.
- Gahmberg CG, Jokinen M, Andersson LC. Expression of the major sialoglycoprotein (glycophorin) on erythroid cells in human bone marrow. *Blood*. 1978;52(2):379-387.
- Fukuda M, Fukuda MN. Changes in cell surface glycoproteins and carbohydrate structures during the development and differentiation of human erythroid cells. *J Supramol Struct Cell Biochem*. 1981;17(4):313-324.
- Sieff C, Bicknell D, Caine G, Robinson J, Lam G, Greaves MF. Changes in cell surface antigen expression during hemopoietic differentiation. *Blood*. 1982;60(3):703-713.
- Gubin AN, Njoroge JM, Bouffard GG, Miller JL. Gene expression in proliferating human erythroid cells. *Genomics*. 1999;59(2):168-177.
- Merryweather-Clarke AT, Atzberger A, Soneji S, et al. Global gene expression analysis of human erythroid progenitors. *Blood*. 2011;117(13):e96-e108.
- Kingsley PD, Greenfest-Allen E, Frame JM, et al. Ontogeny of erythroid gene expression. *Blood*. 2013;121(6):e5-e13.
- Suzuki N, Suwabe N, Ohneda O, et al. Identification and characterization of 2 types of erythroid progenitors that express GATA-1 at distinct levels. *Blood*. 2003;102(10):3575-3583.
- Ikonomi P, Rivera CE, Riordan M, Washington G, Schechter AN, Noguchi CT. Overexpression of GATA-2 inhibits erythroid and promotes megakaryocyte differentiation. *Exp Hematol*. 2000;28(12):1423-1431.
- Nathan DG, Clarke BJ, Hillman DG, Alter BP, Housman DE. Erythroid precursors in congenital hypoplastic (Diamond-Blackfan) anemia. *J Clin Invest*. 1978;61(2):489-498.
- Eaves CJ, Eaves AC. Erythropoietin (Ep) dose-response curves for three classes of erythroid progenitors in normal human marrow and in patients with polycythemia vera. *Blood*. 1978;52(6):1196-1210.
- Urabe A, Murphy MJ Jr, Haghbin M, Gee TS. Erythroid progenitors (BFU-e and CFU-e) in acute leukaemia. *J Clin Pathol*. 1979;32(7):666-669.