

ORIGINAL ARTICLE

Molecular pathways of early CD105-positive erythroid cells as compared with CD34-positive common precursor cells by flow cytometric cell-sorting and gene expression profiling

S Machherndl-Spandl¹, S Suessner², M Danzer², J Proell², C Gabriel², J Lauf³, R Sylie⁴, H-U Klein⁵, MC Béné⁶, A Weltermann¹ and P Bettelheim¹

Special attention has recently been drawn to the molecular network of different genes that are responsible for the development of erythroid cells. The aim of the present study was to establish in detail the immunophenotype of early erythroid cells and to compare the gene expression profile of freshly isolated early erythroid precursors with that of the CD34-positive (CD34⁺) compartment. Multiparameter flow cytometric analyses of human bone marrow mononuclear cell fractions ($n = 20$) defined three distinct early erythroid stages. The gene expression profile of sorted early erythroid cells was analyzed by Affymetrix array technology. For 4524 genes, a differential regulation was found in CD105-positive erythroid cells as compared with the CD34⁺ progenitor compartment (2362 upregulated genes). A highly significant difference was observed in the expression level of genes involved in transcription, heme synthesis, iron and mitochondrial metabolism and transforming growth factor- β signaling. A comparison with recently published data showed over 1000 genes that as yet have not been reported to be upregulated in the early erythroid lineage. The gene expression level within distinct pathways could be illustrated directly by applying the Ingenuity software program. The results of gene expression analyses can be seen at the Gene Expression Omnibus repository.

Blood Cancer Journal (2013) 3, e100; doi:10.1038/bcj.2012.45; published online 11 January 2013

INTRODUCTION

The molecular network of genes involved in the development of erythroid cells has begun to be better understood for the past few years.^{1–9} The underlying molecular mechanism of erythroid cellular differentiation is a complex process, which is subject to a number of physiological conditions.^{10–17} Most importantly, erythroid differentiation is regulated by the erythroid transcription factor globin transcription factor (GATA)-1, which represses GATA-2 and PU.1 and impacts on early and late differentiation.^{18,19} At the terminal stage of differentiation, the erythroid program is defined by the genes that are still expressed. Their high level of expression depends on specific epigenetic modifications, recruitment of transcription factors and posttranscriptional effects.²⁰

Extensive studies have demonstrated the influence of various cytokines and growth factors, particularly erythropoietin, on the different stages of erythroid maturation. Erythropoietin transcription depends on such physiological processes as the regulation of transcriptional responses to hypoxia and involves signaling by transforming growth factor- β (TGF- β), which regulates the expression of the erythropoietin gene.^{12,13}

Maturation in the erythroid series can be examined through immunophenotyping using multiparameter flow cytometry (FCM).²¹ Emerging from the early compartment of CD34-positive (CD34⁺) progenitors, maturing erythroid cells express various antigens, such as the transferrin receptor (CD71), the thrombospondin receptor glycoprotein IIIb (gpIIIb; CD36),

adenosine diphosphate-ribosyl cyclase (CD38) and glycophorin A (CD235a). Most of these markers are not lineage-restricted (CD71, CD36, CD38), or can be found almost throughout erythroid differentiation (CD235a), and their use has proven to be limited to clearly delineate the different stages of early erythroid differentiation, especially for the very immature compartment. The additional use of CD45 and CD105 (endoglin), demonstrating the loss of CD45 and increased expression of CD105 together with CD235a acquisition, has however, allowed to better delineate early erythropoiesis from myeloid commitment.^{22,23} At later stages, more mature erythroid cells are further easily defined by blood group antigens.²⁴

Better knowledge of normal erythroid cell differentiation is mandatory to examine pathological conditions. In this respect, the immunophenotypic profile of pathological erythroid cells deviates from the cells of healthy controls, although the specificity and sensitivity of these findings remain to be clarified.^{25–29}

The aim of the present study was to investigate the immunophenotypic profile of distinct stages of early erythroid differentiation. For this purpose, the cellular immunophenotype of freshly obtained and uncultured normal bone marrow (BM) cells was examined beyond the stage of CD34⁺ precursor cells and along the erythropoietic lineage by FCM analysis. The CD105-positive (CD105⁺) erythroid compartment was then separated by cell-sorting, and the gene expression profile of these cells was compared with that of CD34⁺ common progenitor cells.

¹Division of Hematology/Oncology, Elisabethinen Hospital Linz, Linz, Austria; ²Department of Molecular Biology, Transfusion Service of Upper Austria, Linz, Austria; ³Labor Europaplatz, Linz, Austria; ⁴Institute for Pathology, Wagner Jauregg Hospital, Linz, Austria; ⁵Institute of Medical Informatics, University of Muenster, Muenster, Germany and ⁶Hematology Laboratory, CHU de Nantes, Nantes, France. Correspondence: Dr S Machherndl-Spandl, Division of Hematology/Oncology, Elisabethinen Hospital, Fadingerstrasse 1, Linz A-4020, Austria.

E-mail: sigrid.machherndl-spandl@elisabethinen.or.at

Received 3 December 2012; accepted 7 December 2012

MATERIALS AND METHODS

Patients

Fourteen patients who underwent BM aspiration and biopsy in the context of routine clinical practice and six healthy volunteers were included in this study after giving informed consent. The patients were suffering from non-Hodgkin lymphoma ($n=8$), Hodgkin lymphoma ($n=4$), idiopathic thrombocytopenic purpura (ITP; $n=1$) or cutaneous mastocytosis ($n=1$). Their age range was 37–79 years old (median 53). The samples were obtained at the Division of Hematology and Oncology of the Elisabethinen Hospital, Linz, Austria. The study was approved by the local ethics committee.

All BM samples showed normal cellularity and cell distribution. Pathological alteration was ruled out by cytology, histopathology, cytogenetic and molecular examination. All patients' and volunteers' peripheral blood cell counts were within reference values.

Sample collection and cell separation

The analysis of erythropoietic cells by FCM is associated with a variety of methodological problems. Lysis of erythrocytes leads to the loss of mature erythroid cells, and fragments of erythrocytes can adhere to other cells giving false signals. We therefore decided to use BM mononuclear cell (MNC) fractions to investigate early erythroid cells. This fraction contains all progenitor cells, and mature erythrocytes remain to a lesser extent.

FCM studies on BM erythroid cells

Ethylendiaminetetraacetic acid-anticoagulated BM MNC fractions were obtained by Ficoll (density 1.077) gradient centrifugation. For FCM analyses and cell-sorting experiments, washed cells from MNC fractions were stained for 30 min with appropriate amounts of directly conjugated antibodies at 4 °C. The cells were then washed twice with buffered saline solution.

The reaction of precursor cells with monoclonal antibodies (MoAbs) was evaluated by 8- to 10-color (FACS Aria III; Becton Dickinson, Franklin Lakes, NJ, USA) FCM. For this purpose, a backbone of four MoAbs (APC-CD34, Horizon V500-CD45, Horizon V500-CD105 and PerCP-Cy5.5-CD117; BD Biosciences, San Jose, CA, USA) was established. This backbone was used in combination with MoAbs labeled with Fluorescein isothiocyanate (FITC) (CD11b, CD13, CD33, CD36, CD71 and CD235a), Phycoerythrin (PE) (CD13, CD33 and CD150), eFluor 450 (CD14 and CD123), eFluor 605NC (CD19 and CD45) or APC (CD105 and CD117) from eBioscience (San Diego, CA, USA). PE-Cy7 conjugated anti-CD13, CD38 and human leukocyte antigen-DR (HLA-DR) and Allophycocyanin (APC) conjugated anti-CD34 were purchased from BD Biosciences Pharmingen (San Diego, CA, USA). MoAbs were also obtained from Becton Dickinson (FITC-CD41, PE-CD36 and CD235a, Horizon V500-CD45, Horizon V450-CD105, PerCP-Cy5.5-CD117), IBGRL Research (Bristol, UK: FITC-conjugated CD173, CD233, CD238 and CD239), Biologend (San Diego, CA, USA: PE-CD105 and CD318) and DAKO (Glostrup, Denmark: FITC-CD61).

FACSDiva software Version 6.1.3 (BD Biosciences) was applied to evaluate multiparameter analyses. A minimum of 300 000 MNC were analyzed in each case studied. The gating strategy and stepwise analysis of antibody reactions are presented in the Results section (Figures 1a–g).

The immunophenotypic profile of more mature erythroid cells was established by 3-color FCM analysis. For this purpose, a combination was applied consisting of CD45 and CD105 with a third MoAb directed against erythroid-associated structures, and stained cells were analyzed with a FACSCalibur (Becton Dickinson) flow cytometer.

Cell-sorting experiments

Double cell-sorting was performed using the Aria III cell sorter (Becton Dickinson) and FACSDiva software Version 6.1.3 (BD Biosciences). Pure (>95%) subsets were obtained by FCM sorting with specific marker combinations. Cytospin preparations were stained with Wright-Giemsa in order to demonstrate the morphological properties of the different sorted compartments.

A sufficient amount of CD34⁺/CD45^{weak} cells was obtained in seven cases. Pure CD105⁺/CD45⁻ cells were sorted in 11 cases. As controls, mature CD14⁺ monocytes were sorted in six cases.

Gene expression profile

Trizol (Invitrogen, Life Technologies, Carlsbad, NM, USA) was applied for total RNA extraction of pure sorted subsets according to the

manufacturer's instructions, with a phenol chloroform step (RNeasy Plus Mini Kit, Qiagen, Hilden, Germany). RNA integrity was confirmed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

Isolated RNA (100–500 ng) was used to produce biotinylated complementary DNA (cDNA). The two-cycle cDNA synthesis and two-cycle target labeling and control reagent kits (Affymetrix, Santa Clara, CA, USA) were applied. Consequently, the cDNA was hybridized to GeneChip Human Genome U133 Plus 2.0 arrays (Affymetrix), following the manufacturer's recommendations. All procedures of isolation, labeling, hybridization, staining and scanning were performed by the microarray analysis core facility of the Transfusion Service of Upper Austria, Linz.

The Affymetrix GeneChip Fluidics Station 450 was used to wash and stain the arrays with streptavidin-phycoerythrin, according to the standard protocol for eukaryotic targets (IHC kit, Affymetrix). Arrays were scanned with an Affymetrix GeneChip scanner 3000 at 570 nm. Intensity values were determined using GeneChip operating software (Affymetrix). The raw intensity values were then normalized by the Robust Multiarray Average algorithm and analyzed using Partek Genomics Suite software (Partek Incorporated, St Louis, MO, USA).³⁰ Analysis of variance models were used to detect differences in gene expression between cell populations. Genes with a significantly different expression level ($P \leq 0.01$; fold change ≥ 2.0) were selected. Hierarchical clustering was performed based on normalized expression values of the differentially expressed genes. Networks were generated by applying Ingenuity pathways analysis (Ingenuity Systems, Redwood City, CA, USA). A stringent cutoff ($P \leq 0.01$; fold change ≥ 3.0) was used to detect differentially expressed genes between CD34⁺ and CD105⁺ cells for network analyses. These genes were overlaid onto a global molecular network from Ingenuity's knowledge base. Subnetworks enriched by differentially regulated genes were then detected (Figures 5a–c).

Raw and normalized expression data are available at the Gene Expression Omnibus repository.

Real-time quantitative PCR

Gene expressions observed in microarray were verified for five selected genes from five samples by quantitative real-time PCR. Total RNA was purified from CD34⁺, CD105⁺ and CD14⁺ subsets using the RNeasy kit (Qiagen) protocol. The RNA concentration and integrity of the samples were assessed by an Agilent 2100 Bioanalyzer using the RNA 6000 Nano Chips Kit (Agilent Technologies). Reverse transcription from 100 ng total RNA was performed using the high capacity cDNA archive kit protocol (Life Technologies). Quantitative PCR was performed using a LightCycler 480 (Roche Diagnostics, Basel, Switzerland) and Taqman gene expression assays (Life Technologies) for KEL (metallo-endopeptidase, Hs00166270_m1), GATA-1 (Hs01085823_m1), CD36 (thrombospondin receptor, Hs00169627_m1), erythropoietin receptor (Hs00959427_m1), ANPEP (alanyl (membrane) aminopeptidase, CD13, Hs00174265_m1), ENG (endoglin, CD105, Hs00923996_m1), and RUNX2 (runt-related transcription factor 2, Hs00231692_m1). The PCR program used initial denaturation at 95 °C 10 min followed by 95 °C 10 s, 60 °C 45 s, for 50 cycles. Standard curves were prepared for each assay to determine amplification efficiency. Expression values of the transcripts were normalized to the housekeeping gene 18S ribosomal RNA (Life Technologies). A calibrator was additionally introduced to normalize all samples within one run. PCR efficiency correction and normalization to the calibrator and housekeeping gene was performed using LightCycler 480 relative quantification software (Roche Diagnostics).

RESULTS

FCM analyses

Twenty BM samples were investigated for FCM. The backbone of CD34, CD117, CD105 and CD45, as illustrated in Figure 1, associated with other markers, allowed to clearly define the various stages of erythroid differentiation. BM MNCs were first analyzed for CD45, CD34 and CD105 expression respectively (Figures 1a–d). About half of the CD105 compartment expresses CD117 (Figure 1e). The inclusion of CD71 (Figures 1f–j) disclosed a 'transition stage' between the CD34⁺ compartment and the erythroid cells strongly stained by CD105. What was dubbed as stage A of erythroid differentiation is further characterized by a faint expression of CD36 (Figure 1m), CD173 (Figures 1o–s), CD105 (Figures 1f–x) and CD238 (not shown). At this stage of differentiation the myeloid marker CD33 is necessary to separate

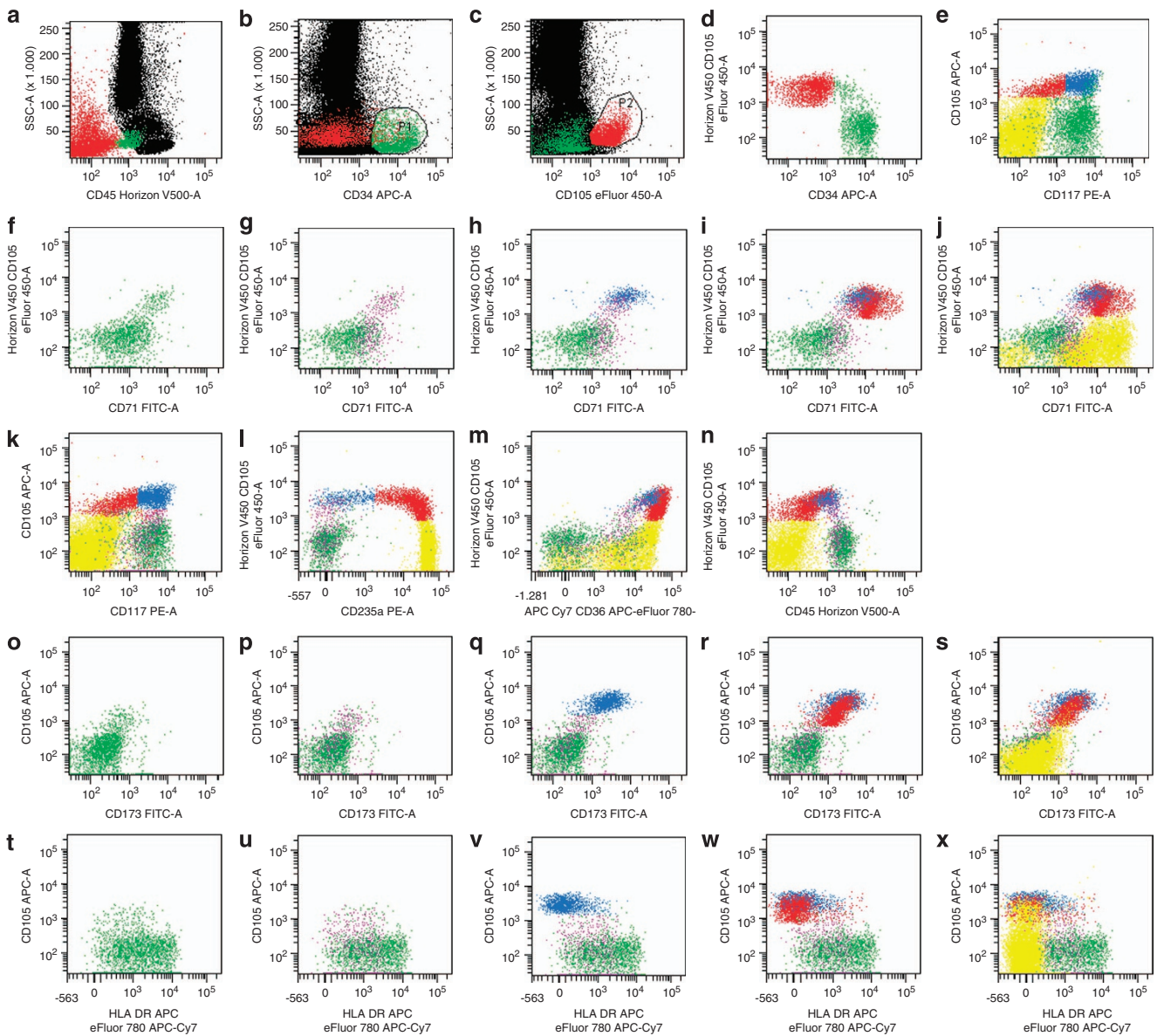


Figure 1. Immunophenotypic definition of erythroid maturation in flow cytometry. (a–c) CD45/CD34 (green)/CD105 (red) combination versus SSC. (d) CD105 versus CD34 (exclusion of mature granulocytes, monocytes and lymphoid cells). (e) CD117 expression on common progenitor cells (green) and on part of the CD105⁺ cells (blue). Mature erythroid cells are marked yellow. (f–j) a small proportion of the CD71⁺ cells are part of the CD34 compartment. Violet marked cells defines a separate compartment representing erythroid stage A. (k) CD117 expression within the erythroid differentiation (blue = erythroid stage B, red = erythroid stage C). (l–n) CD235a, CD36 and CD45 expression. CD173 (o–s) and HLA-DR (t–x) during the erythroid differentiation.

CD71⁺ erythroid cells from CD71⁺ myeloid cells (not shown). This was crucial for subsequent experiments, including cell-sorting of pure erythroid populations. A similar observation was made for CD36 present both in erythroid and monocytic differentiation.

At stage B of erythroid differentiation, a strong staining of CD117 (Figure 1k), CD105 (Figures 1k–x), CD36 (Figure 1m), CD71 (Figures 1h–j), CD173 (Figures 1q–s) and CD238 (not shown) can be observed, whereas CD45 expression (Figure 1n) is clearly diminished and HLA-DR is no more detectable (Figure 1v).

Stage C is defined by the appearance of CD235a (Figure 1l) and disappearance of CD117 (Figure 1k). CD105 remains highly expressed at stage C but is lost at the terminal stage of erythroid differentiation (Figures 1e–x).

Relationships between the CD34⁻/CD105⁺ compartment and mature erythroid cells expressing blood group antigens are shown in Figures 2a–i. Of all tested markers, the expression of the blood

group antigen Lu (85 and 78 kDa Lutheran glycoproteins) detected by MoAb BRIC221³¹ was found to be the latest antigen to appear during terminal erythroid differentiation. CD36, CD38, CD71 and CD173 showed decreased staining within the mature erythroid compartment, whereas the expression of CD238, CD233, CD235a and CD239 remained stable.

A schematic presentation summarizing the immunophenotypic features of erythroid differentiation is proposed in Figure 3.

Cellular distribution of the different erythroid compartments

Overall, the CD34⁺ and CD34⁻/CD105⁺ compartments represented 0.5% (range: 0.2–0.7%) and 0.9% (range: 0.5–1.6%) of all nucleated MNC, respectively. Within these precursor cells, stage A is the smallest compartment with a mean of 5% (range: 2–8%), whereas stage B has a mean of 14% (range: 8–24%) and stage C is

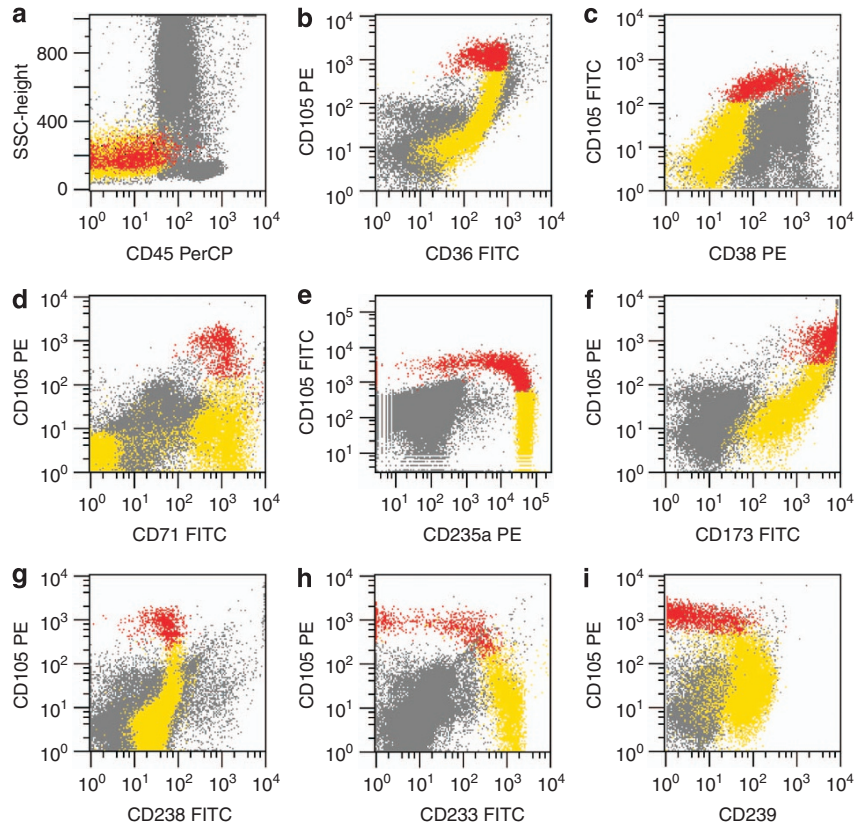


Figure 2. Surface marker profile of mature erythroid cells. The relationships between the CD34⁻/CD105⁺ compartment and mature erythroid cells expressing blood group antigens are shown (gating strategy according to the CD45 and CD105 high staining): Mature erythroid cells (yellow) lack CD45 (a), decrease progressively CD36, CD38 and CD173 expression (b, c, f), loose CD105 (b–i) and finally CD71 (d) and acquire CD235, CD233 and CD239 (Lu-Antigen) expression (e, h, i).

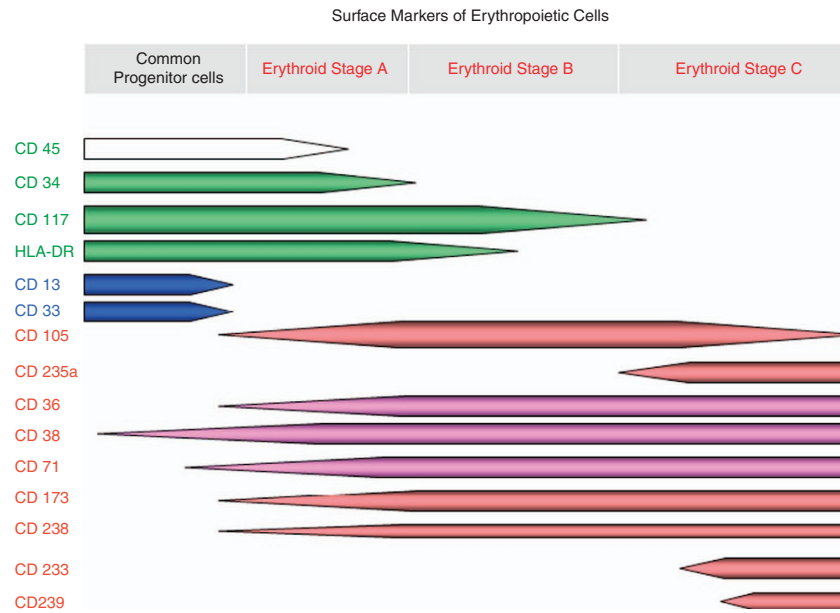


Figure 3. Schematic presentation of the marker profile of erythropoiesis. Green arrows represent precursor-associated markers, red arrows represent erythroid-associated markers, blue arrows represent myeloid markers, white arrow represent common leukocyte antigen CD45 and pink arrows represent non-specific markers.

the largest (mean 56%, range: 51–63%). CD34⁺ common progenitor cells are also quite abundant with a mean of 25% (range: 18–33).

Cytology of the different cell compartments

The cytology of sorted cell preparations (Supplementary Figure 1) is consistent with typical images of erythroid progenitors.

Gene expression profile of erythroid cells

Gene expression experiments concentrated on CD34⁺ progenitor cells and the CD34⁻/CD105⁺ erythroid compartment after sorting freshly obtained BM. Monocytes, separated according to side scatter characteristics and CD14 expression, were used as controls and for the recognition of gene equally regulated in both CD34 and CD105 compartments.

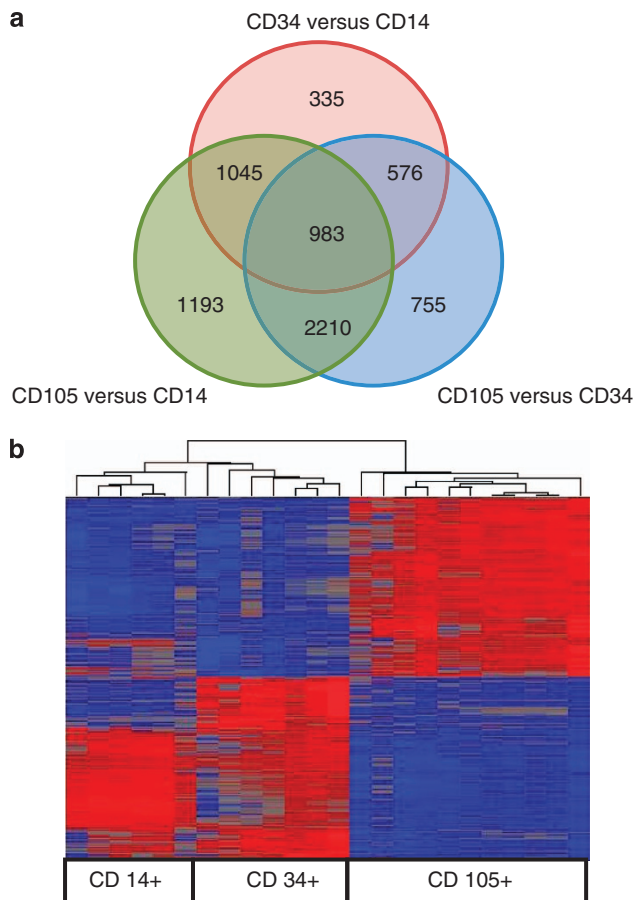


Figure 4. Comparative transcriptomic analysis of sorted precursor (CD34⁺), early erythroid (CD34⁻/CD105⁺) cells and CD14⁺ monocytes as control. **(a)** Three-way Venn diagram. Number of common or specific genes based on the calculated gene lists ($P \leq 0.01$; \geq two-fold change) comparing CD34⁺ progenitor cells, CD105⁺ erythroid cells and CD14⁺ monocytes, respectively. Cell populations were compared pairwise regardless of up- or down-regulation. **(b)** Heat map visualization of expression patterns of genes in CD14⁺ monocytic, CD34⁺ common progenitor and CD105⁺ erythroid cell populations, respectively. Supervised hierarchical clustering is based on a total of 5394 genes, which are at least two-fold overexpressed (red) or two-fold underexpressed (blue) ($P \leq 0.01$), comparing the CD34⁻/CD105⁺ and the CD34⁺ cell populations. Lines represent gene probes; columns represent individual cell populations.

The Venn diagram and overall results of hierarchical clustering are presented in Figures 4a and b.

A total of 4524 genes was identified as differentially regulated between the CD34⁺ common progenitor compartment and CD34⁻/CD105⁺ erythroid cells. Significant upregulation was obvious in 2362 genes and significant downregulation in 2172. Quantitative PCR was used to confirm the expression profiles of selected, differentially regulated transcripts (Supplementary Table 1).

As shown in Table 1, the increased (CD105, CD36, CD71, CD173, CD233, CD235a, CD238 and CD239) or decreased (CD13, CD33, CD34, CD38, CD44, CD45 and HLA-DR) expression of specific markers detected by FCM was corroborated by microarray analyses. CD117 is expressed at high levels equally in both CD34⁺ and CD34⁻/CD105⁺ populations, but is upregulated in comparison with CD14⁺ monocytes. A direct comparison of the mean fluorescence intensity detected by FCM with the gene expression (fold change) is presented in Supplementary Table 2.

Special attention was then drawn to significant up- or down-regulated genes of functional relevance. Table 2 lists the most important transcription factors and their expression levels comparing CD34⁻/CD105⁺ erythroid cells and the CD34⁺ common progenitor compartment. One of the best-defined transcription factors for the regulation of hematopoiesis is GATA-2. As expected, a stage-specific decrease was seen for this gene, which is necessarily downregulated to initiate erythroid differentiation. GATA-2 downregulation is accompanied by that of Pu.1 (SPI1), known to trigger myeloid differentiation as evidenced by the comparison of both CD34⁺ cells and CD34⁻/CD105⁺ cells with mature monocytes. At the same stage of differentiation, a remarkable increase in GATA-1 gene expression can be observed. FOG-1 (*Friend of GATA-1*), another key regulator for erythroid differentiation, is also upregulated. Eleven genes encoding transcription factors were found to be upregulated in both compartments as compared with the CD14⁺ monocytes used as controls.

The search was then extended to significantly regulated genes of different pathways and functionally relevant molecules, including the TGF- β signal transduction pathway (Table 3). Gene expression levels of a number of other pathways (heme synthesis, iron and mitochondrial metabolism, erythropoietin receptor) are presented in Supplementary Tables S3,S4. To differentiate erythropoiesis from megakaryopoiesis, this investigation included relevant genes for the latter lineage (Supplementary Table S5).

Considerable results were obtained for cell cycle-relevant genes and for genes initiating proliferation or apoptosis (Supplementary Table S6). Distinct genes were found to be orchestral-related as confirmed (Figure 5a) by using the special program Ingenuity software. This demonstrated the influence of transcription factor Dp-1 (TFDP1) on the upregulation of a large number of genes regulating the cell cycle, including E2F-dependent transcription, cyclin E1, retinoblastoma-like 1, cell division cycle 6 homolog and origin recognition complex subunits 1 and 2. This program was also applied to demonstrate a variety of other linked proteins and genes by uploading all data on gene expression levels. Another two pathways with a significant up- or downregulation of gene expression levels are presented in Figures 5b and c. Figure 5b illustrates the relationship between endoglin and the TGF- β receptor/Smad/Smurf pathway. The Smad7/Smurf2 cytoplasmic complex has been shown to inhibit TGF- β and activin signaling by associating with their receptors, thus preventing SMAD2 access. The network of distinct transcription factors linked to heme synthesis is provided in Figure 5c.

Beside the above-mentioned transcription factors, this data analysis revealed upregulation in 723 gene probes of both the CD34⁺ and CD34⁻/CD105⁺ compartments as compared with CD14⁺ monocytes. Twenty of these genes with the highest fold change are listed in Supplementary Table 7.

Table 1. Surface markers/blood group antigens

Gene designation	Entrez gene ID	P-value vs 34 ⁺ CP vs 105 ⁺ EP	FC 105 ⁺ EP vs 34 ⁺ CP	P-value 105 ⁺ EP vs 14 ⁺ mono	FC 105 ⁺ EP vs 14 ⁺ mono	P-value CD34 ⁺ CP vs 14 ⁺ mono	FC 34 ⁺ CP vs 14 ⁺ mono	Functional properties (provided by Gene Cards)
<i>RHD</i> Rh blood group, D antigen	6007	3.03E-10	39.3	3.64E-08	18.6	NS		Part of an oligomeric complex that is likely to have a transport or channel function in the >erythrocyte membrane
<i>AQP1</i> Aquaporin 1 (Colton blood group)	358	6.06E-14	32.1	1.47E-12	22.0	NS		Forms a water-specific channel that provides the plasma membranes of red cells with high permeability to water
<i>CD36</i> Thrombospondin receptor	948	8.35E-09	27.5	NS		8.85E-06	-11.2	Glycoprotein of the platelet surface and serves as a receptor for thrombospondin in platelets and various cell lines; may have important functions as a cell adhesion molecule
<i>EPB42</i> Erythrocyte membrane protein band 4.2	2038	8.41E-10	24.5	2.74E-08	15.5	NS		ATP-binding protein that may regulate the association of protein 3 with ankyrin. It probably has a role in erythrocyte shape and mechanical property regulation
<i>CTSE</i> Cathepsin E, erythrocyte membrane aspartic proteinase	1510	3.45E-10	24.3	1.13E-09	22.9	NS		May have a role in immune function. Probably involved in the processing of antigenic peptides during MHC class II-mediated antigen presentation
<i>KEL/CD238</i> Kell blood group, metallo-endopeptidase	3792	5.12E-12	20.7	4.90E-12	24.3	NS		This gene encodes a type II transmembrane glycoprotein that is the highly polymorphic Kell blood group antigen
<i>TFRC/CD71</i> Transferrin receptor	7037	5.06E-06	19.8	1.49E-04	10.8	NS		Cellular uptake of iron occurs via receptor-mediated endocytosis of ligand-occupied transferrin receptor into specialized endosomes
<i>SLC4A1/CD233</i> Solute carrier family 4, anion exchanger, member 1	6521	2.50E-10	16.9	1.71E-09	14.4	NS		Major integral glycoprotein of the erythrocyte membrane. Its integral domain mediates an exchange of inorganic anions; its cytoplasmic domain provides binding sites for cytoskeletal proteins, glycolytic enzymes and hemoglobin
<i>FUT11/CD173</i> Fucosyltransferase 1 (galactoside 2-alpha-L-fucosyltransferase, H blood group)	2523	3.04E-13	11.3	1.71E-12	10.3	NS		Golgi stack membrane protein that is involved in the creation of a precursor of the H antigen, which is required for the final step in the soluble A and B antigen synthesis pathway
<i>EPB41L4A</i> Erythrocyte membrane protein band 4.1 like 4A	64097	3.25E-06	6.1	3.74E-05	4.8	NS		Members of the band 4.1 protein superfamily are thought to regulate the interaction between the cytoskeleton and plasma membrane
<i>ENG/CD105</i> Endoglin	2022	7.64E-06	3.9	8.11E-06	4.1	NS		Transmembrane protein that is a major glycoprotein of the vascular endothelium. This protein is a component of the TGF-β receptor complex and binds TGFβ1 and TGFβ3
<i>SIAMF1/CD150</i> Signaling lymphocytic activation molecule family member 1	6504	9.16E-03	-2.1	NS		NS		High-affinity self-ligand important in bidirectional T-cell to B-cell stimulation
<i>PTPRCAP/CD45</i> Protein tyrosine phosphatase, receptor type	5790	2.56E-04	-2.7	NS		3.33E-03	2.3	Transmembrane phosphoprotein specifically associated with tyrosine phosphatase PTPRC/CD45, a key regulator of T- and B-lymphocyte activation
<i>CD44</i> Indian blood group	960	1.20E-03	-3.7	3.57E-08	-23.2	2.18E-04	-6.2	Receptor for HA. Mediates cell-cell and cell-matrix interactions; role in cell migration, tumor growth and progression
<i>CD38</i> CD38 molecule	952	1.50E-04	-4.6	NS		NS		Multifunctional ectoenzyme widely expressed in cells and tissues especially in leukocytes, functions in cell adhesion, signal transduction and calcium signaling
<i>CD33</i> CD33 molecule	945	2.49E-07	-5.6	4.41E-12	-29.3	3.51E-06	-5.2	Adhesion molecule of myelomonocytic-derived cells that mediates sialic acid-dependent binding to cells
<i>CD13</i> Alanyl (membrane) aminopeptidase	290	1.12E-07	-6.9	2.78E-13	-62.9	1.10E-07	-9.2	Broad-specificity aminopeptidase. Has a role in the final digestion of peptides generated from hydrolysis of proteins by gastric and pancreatic proteases
<i>CD34</i> CD34 molecule	947	4.08E-11	-7.2	NS		5.21E-09	5.6	Possible adhesion molecule with a role in early hematopoiesis by mediating the attachment of stem cells to the bone marrow extracellular matrix or directly to stromal cells
<i>FUT4/CD15</i> Fucosyltransferase 4 (α-(1,3) fucosyltransferase, myeloid-specific)	2526	7.69E-06	-7.4	5.26E-06	-8.7	NS		Transfers fucose to N-acetyllactosamine polysaccharides to generate fucosylated carbohydrate structures. It catalyzes the synthesis of the non-sialylated antigen, Lewis x (CD15).

Table 1. (Continued)

Gene designation	Entrez gene ID	P-value 105 ⁺ EP vs 34 ⁺ CP	FC 105 ⁺ EP vs 34 ⁺ CP	P-value 105 ⁺ EP vs 14 ⁺ mono	FC 105 ⁺ EP vs 14 ⁺ mono	P-value CD34 ⁺ CP vs 14 ⁺ mono	FC 34 ⁺ CP vs 14 ⁺ mono	Functional properties (provided by Gene Cards)
<i>HLA-DRB1</i> Major histocompatibility complex, class II, DR β-1	3123	1.06E-08	-9.7	5.37E-11	-24.7	3.79E-03	-2.6	Binds peptides derived from antigens that access the endocytic route of APC and presents them on the cell surface for recognition by the CD4 T-cells
<i>ITGAM/CD11b</i> Integrin, alpha M (complement component 3 receptor 3 subunit)	3684	1.09E-04	-14.5	3.19E-07	-76.4	NS		Implicated in various adhesive interactions of monocytes, macrophages and granulocytes as well as in mediating the uptake of complement-coated particles
<i>MPO</i> Myeloperoxidase	4353	4.03E-05	-30.5	4.65E-03	-9.0	NS		Heme protein synthesized during myeloid differentiation that constitutes the major component of neutrophil azurophilic granules
<i>HLA-DRA</i> Major histocompatibility complex, class II, DR-α	3122	1.36E-08	-25.7	7.74E-10	-55.8	NS		Binds peptides derived from antigens that access the endocytic route of APC and presents them on the cell surface for recognition by the CD4 T-cells
<i>KIT/CD117</i> KIT ligand	3815	NS		3.82E-06	10.1	5.19E-07	18.4	Tyrosine-protein kinase that acts as cell-surface receptor for the cytokine KITLG/SCF and has an essential role in the regulation of cell survival and proliferation, hematopoiesis, stem cell maintenance, gametogenesis mast cell development
<i>IL3RA/CD123</i> Interleukin 3 receptor-α (low affinity)	3563	NS		NS		NS		Interleukin 3 specific subunit of a heterodimeric cytokine receptor

Abbreviations: APC, antigen-presenting cell; ATP, adenosine triphosphate; CP, common progenitor cells; EP, erythroid progenitor cells; FC, erythroid progenitors; FC, flow cytometry; HA, hyaluronic acid; MHC, major histocompatibility complex; NS, not significant; TGF, transforming growth factor.

The gene expression level of cultured erythroid cells was recently evaluated according to their stage of differentiation within the erythroid lineage.⁴ A comparison of the present study with these results (Supplementary Table S8) suggest that 1259 genes are upregulated not only within distinct stages of erythroid differentiation, but also in comparison to the CD34⁺ compartment. As expected, however, a large number of genes was differentially regulated. For example, comparison between the CD34⁺ compartment and the CD34⁻/CD105⁺ erythroid cells includes B-cell precursors in the former and excludes late stage of erythropoiesis in the latter. Therefore, a high number of genes that are responsible for engagement in the B- or myeloid-lineage were included in this study and found to be significantly downregulated. By contrast, a large amount of genes reported by Merryweather-Clarke *et al.*⁴ as having a relevant function at the late stage of erythropoiesis was not found to be differentially regulated by comparing the CD34⁺ compartment with CD34⁻/CD105⁺ early erythroid cells.

DISCUSSION

This study investigated stage-specific changes in the expression level of genes, which are important regulators of early erythropoiesis.

A detailed surface immunophenotypic profile of early erythroid progenitors was first established, using highly sensitive FCM analyses. This approach allowed to delineate accurately the erythroid lineage from the myeloid branch and to define immunophenotypically different erythroid compartments.

The discrimination between early erythroid cells and myeloid precursor cells as well as their exact characterization are of clinical significance, especially in such pathological conditions as myelodysplastic syndromes, in which FCM analyses are of growing diagnostic and prognostic relevance.²⁵⁻²⁹ Here, we provide evidence that the immunophenotype of early erythroid cells can be analyzed in more detail by FCM, which identified three different stages of early erythropoiesis. Emerging from the common progenitor level, a part of the cells differentiate into erythroid cells, which can be first recognized in FCM by the expression of CD105, CD71, CD36 and CD238. With ongoing erythroid differentiation, precursor-associated markers undergo stepwise downregulation, first of CD34 and HLA-DR, followed by CD117. Simultaneously, erythroid-associated structures such as blood group antigens emerge on the surface of the cells. In the present work, special attention was given to endoglin (CD105). The expression of endoglin as a marker for erythropoiesis was first observed by Bühring *et al.*³² Endoglin is a TGF-β coreceptor expressed specifically on endothelial cells, which has been indicated to be similarly important for the development of erythropoiesis.³³⁻³⁷ Pimanda *et al.*³⁷ have reported in a mouse model that endoglin expression in blood and endothelial cells is differentially regulated via the ETs/GATA hemangioblast code. In this model, Friend leukemia integration 1 (Flil1), GATA-2, Scl (STIL) and to a lesser effect Pu.1 (Spi1) have been shown to be upregulated with ongoing hematopoietic differentiation and to bind to the endoglin hematopoietic enhancers. With the exception of Scl (STIL) gene upregulation, all these genes (Flil1, GATA-2 and Pu.1), which are important initiating regulators for myeloid and megakaryocytic differentiation,³⁸ were indeed found here to be significantly downregulated in CD105⁺ human erythroid cells. It could be that upregulation of these genes in the mouse model takes place at an earlier stage of differentiation, at which endoglin is not yet expressed on the surface of the precursor cells. This is underlined by the finding that absence of the endoglin gene reduces the levels of key transcription factors GATA-2, Scl and Lmo2.³⁹ In the study by Zhang *et al.*,³⁹ endoglin was suggested to regulate primitive hematopoiesis by modulating

Table 2. Transcription Factors

Gene designation	Entrez gene ID	P-value 105 ⁺ EP vs 34 ⁺ CP	FC 105 ⁺ EP vs 34 ⁺ CP	P-value 105 ⁺ EP vs 14 ⁺ mono	FC 105 ⁺ EP vs 14 ⁺ mono	P-value CD34 ⁺ CP vs 14 ⁺ mono	FC 34 ⁺ CP vs 14 ⁺ mono	Functional properties (provided by Gene Cards)
<i>SOX6</i>								Transcriptional activator
SRY (sex determining region Y)-box 6	55553	1.26E-12	33.6	1.62E-11	25.2	NS		
<i>KLF1</i>								Transcription regulator of erythrocyte development, fetal-to-adult globin switching
Kruppel-like factor 1 (erythroid)	10661	1.52E-08	16.6	8.93E-10	32.4	NS		
<i>GFI1b</i>								transcriptional regulator (differentiation of erythroid and megakaryocytic lineages)
Growth factor independent 1B transcription repressor	8328	1.46E-09	12.1	1.34E-11	29.5	4.78E-03	2.4	
<i>GATA-1</i>								Erythroid development, regulating the switch of fetal hemoglobin to adult hemoglobin
GATA binding protein 1 (globin transcription factor 1)	2623	4.21E-09	9.4	1.00E-09	12.8	NS		
<i>E2F4</i>								E2F transcription factors, such as E2F8, are essential for orchestrating expression of genes required for cell cycle progression and proliferation (Christensen et al. ⁴⁰ (PubMed 16179649)).(supplied by OMIM)
E2F transcription factor 4, p107/p130-binding	1874	1.15E-08	5.4	1.30E-05	3.0	NS		
<i>STIL</i>								Embryonic development as well as in cellular growth and proliferation;long-term silencing affects cell survival
SCL/TAL1 interrupting locus	6491	9.20E-06	4.9	2.33E-08	12.2	9.80E-03	2.5	
<i>TFDP1</i>								DP2/E2F complex functions in the control of cell-cycle progression from G1 to S phase, the E2F-1/DP complex appears to mediate both cell proliferation and apoptosis
Transcription factor Dp-1, E2F dimerization partner	7027	1.91E-06	4.5	5.02E-06	4.4	NS		
<i>LDB1</i>								Regulation of red blood cell development, maintaining in an immature state
LIM domain binding 1	8861	6.00E-07	4.3	1.27E-05	3.4	NS		
<i>MBD1</i>								Repression of transcription from methylated gene promoters
Methyl-CpG binding domain protein 1	4152	3.12E-05	3.1	4.20E-03	2.0	NS		
<i>ZBTB7A</i>								Essential downstream target of GATA1 in erythroid development
Zinc-finger and BTB domain-containing 7A	51341	5.92E-04	2.3	NS		3.06E-03	-2.2	
<i>EED</i>								Maintenance of transcriptional repressive state of genes over successive cell generations
Embryonic ectoderm development	8726	2.66E-04	2.3	3.06E-07	4.3	NS		
<i>LMO2</i>								Required for yolk sac erythropoiesis, role in hematopoietic development
LIM domain only 2 (rhombotin-like 1)	4005	5.01E-06	2.0	2.56E-08	2.8	NS		
<i>RUNX1</i>								α -subunit of CBF, involved in the development of normal hematopoiesis
Runt-related transcription factor 1	861	-2.28E-04	-2.7	NS		NS		essential role in regulating transcription of genes involved in the development and proliferation of hematopoietic and endocrine cell lineages
<i>GATA-2</i>								Negative regulation of the JAK/STAT signalling pathway
GATA-binding protein 2	2624	2.38E-03	-3.4	NS		3.24E-04	5.7	
<i>SOC3</i>								Transcriptional activator
Suppressor of cytokine signaling 3	9021	1.49E-04	-4.2	9.16E-11	-48.7	1.01E-06	-11.5	
<i>MAFF</i>								Coordinating transcription activation and repression by MAFK
V-maf musculoaponeurotic fibrosarcoma oncogene homolog F (avian)	23764	1.42E-05	-11.7	5.73E-07	-25.8	NS		
<i>BACH2</i>								Coordinating transcription activation and repression by MAFK
BTB and CNC homology 1, basic leucine zipper transcription factor 2	60468	8.01E-10	-16.1	NS		8.04E-09	16.4	

Table 2. (Continued)

Gene designation	Entrez gene ID	P-value 105 ⁺ EP vs 34 ⁺ CP	FC 105 ⁺ EP vs 34 ⁺ CP	P-value 105 ⁺ EP vs 14 ⁺ mono	FC 105 ⁺ EP vs 14 ⁺ mono	P-value CD34 ⁺ CP vs 14 ⁺ mono	FC 34 ⁺ CP vs 14 ⁺ mono	Functional properties (provided by Gene Cards)
<i>FLI1</i> Friend leukemia virus integration 1	2313	1.11E-09	-23.3	1.27E-09	-26.5	NS		Transcriptional activator
<i>ERG</i> V-ets erythroblastosis virus E26 oncogene homolog (avian)	2078	4.83E-11	-38.0	NS		1.13E-09	33.7	Erythroblast transformation-specific transcription factor
<i>BIM</i> B-cell CLL/lymphoma 11A (zinc-finger protein)	53335	NS		NS		1.52E-05	8.4	Regulation of programmed cell death or apoptosis
<i>PU.1</i> [*] Spleen focus forming virus (SFFV) proviral integration oncogene spl	6688	NS		2.02E-04	-5.0	1.26E-04	-6.3	Transcriptional activator that may be specifically involved in the differentiation or activation of macrophages or B-cells; ¹¹
<i>EZH2</i> [*] Enhancer of zeste homolog 2 (Drosophila)	2146	NS		1.83E-07	4.7	1.05E-04	2.9	Maintenance of transcriptional repressive state of genes over successive cell generations
<i>KDM1a</i> [*] Lysin (K)-specific demethylase 1A	23028	NS		8.03E-07	3.1	1.22E-03	2.0	Component of a GFI/KDM1A complex that suppresses blood cell development
<i>E2F7</i> [*] E2F transcription factor 7	144455	NS		8.54E-08	9.1	1.67E-05	5.4	Inhibitor of E2F-dependent transcription that is important for the control of the E2F1-TP53 apoptotic pathway. Directly represses E2F1 transcription
<i>TCF3</i> [*] Transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	6929	NS		4.98E-08	7.0	1.97E-06	5.4	E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors
<i>TFDP2</i> [*] Transcription factor Dp-2 (E2F dimerization partner 2)	7029	NS		4.49E-10	6.8	3.32E-09	6.6	Can stimulate E2F-dependent transcription. The E2F-1/DP complex appears to mediate both cell proliferation and apoptosis
<i>SP4</i> [*] Sp4 transcription factor	6671	NS		1.21E-06	4.9	2.82E-04	3.1	Probable transcriptional activator
<i>TTF2</i> [*] Transcription termination factor, RNA polymerase II	8458	NS		5.60E-04	4.5	6.81E-02	2.2	Transcription termination factor; may contribute to mitotic transcription repression
<i>LZTF1</i> [*] Leucine zipper transcription factor-like 1	54585	NS		5.74E-03	2.8	5.42E-03	3.1	Novel gene with unknown biological functions; tumor suppressor
<i>GTF3A</i> [*] General transcription factor IIIA	2971	NS		2.83E-04	2.7	1.98E-03	2.4	May initiate transcription of the 5S ribosomal RNA gene and maintain the stability of transcription of other genes
<i>ATF7IP2</i> [*] Activating transcription factor 7 interacting protein 2	80063	NS		5.69E-04	2.6	1.54E-03	2.5	Recruiter that couples transcriptional factors to general transcription apparatus, and thereby modulates transcription regulation and chromatin formation. Both can act as an activator or a repressor depending on the context
<i>MTE2</i> [*] Metal response element binding transcription factor 2	22823	NS		1.24E-03	2.3	5.85E-05	3.5	Binds to the MRE of metallothionein-1A gene promoter; specific PRC2 regulator modulates embryonic stem cell pluripotency and somatic cell reprogramming

Abbreviations: CBF, core binding factor; DP, dimerization partner; GATA, globin transcription factor; JAK, Janus kinase; MAFK, musculoaponeurotic fibrosarcoma oncogene homolog K; MRE, metal-regulating element; NS, not significant; STAT, signaltransducers and activators of transcription. Eleven genes encoding transcription factors were found to be upregulated in both compartments as compared to the CD14⁺ monocytes used as controls. These genes were marked with ^{*}.

Table 3. TGF- β pathway

Gene designation	Entrez gene ID	P-value 105 ⁺ EP vs 34 ⁺ CP	FC 105 ⁺ EP vs 14 ⁺ mono	P-value 105 ⁺ EP vs 14 ⁺ mono	FC 105 ⁺ EP vs 14 ⁺ mono	P-value CD34 ⁺ CP vs 14 ⁺ mono	FC 34 ⁺ CP vs 14 ⁺ mono	Functional properties (provided by Gene Cards)
<i>TGFBR1</i> Transforming growth factor- β receptor 1 (ALK5)	7046	7.39E-05	-7.6	4.48E-06	-7.6	NS		Signaling promotes inhibition of hemangioblast and hematopoiesis
<i>ACVR1B</i> Activin A receptor, type IB (ALK2)	91	2.57E-10	-15.1	5.48E-14	-15.1	3.71E-06	-2.9	Signaling promotes hemangioblast development and hematopoiesis
<i>SMURF2</i> SMAD specific E3 ubiquitin protein ligase 2	64750	2.93E-08	-5.4	6.19E-08	-5.5	NS		Interacts with SMAD1 and SMAD7 in order to trigger their and proteasome-dependent degradation
<i>Smad7</i> SMAD family member 7	4092	1.46E-04	-18.1	1.68E-05	-18.1	NS		Antagonist of signaling by TGF- β type 1 receptor superfamily members
<i>TCF7L2</i> Transcription factor 7-like 2 (T-cell specific; HMG-box)	6934	1.22E-10	-20.7	4.38E-10	-20.7	NS		Participates in the Wnt signaling pathway
<i>Smad4</i> SMAD family member 4	4089	NS	3.0	3.33E-06	3.0	4.29E-05	2.7	Smad proteins are phosphorylated and activated by transmembrane serine-threonine receptor kinases in response to TGF- β signaling
<i>Smad1</i> SMAD family member 1	4086	NS	NS	NS	NS	3.73E-04	5.9	Transcriptional modulator activated by BMP

Abbreviations: BMP, bone morphogenetic protein; HMG-box, high mobility group box; NS, not significant; TGF, transforming growth factor.

the activity of the Smad1/5 signaling pathway at early stages of development.

CD105 surface expression on erythroid cells is clearly a stage-specific phenomenon. Here, endoglin appeared to be expressed at an early phase of the erythroid lineage, increase at the intermediate stage B, and then progressively decline as erythroid cells mature. Cell-sorting experiments showed that endoglin was expressed up to the stage of basophilic erythroblasts, but not at the terminal stage of erythroid maturation. The underlying molecular mechanisms of endoglin downregulation toward the end of erythroid differentiation are so far unknown. It may be speculated that the decreased transcriptional elongation of the gene could be responsible for this process. Another crucial factor could be TGF- β 1, known to be a strong inducer of erythroid differentiation and suppressor of proliferation,⁴¹ as well as a candidate for the initiation of this process via the TGF- β receptor/Smad axis. This hypothesis was tested using the Ingenuity software program, providing graphic information about protein networks and differentially expressed genes. Indeed, a number of genes were shown to interact significantly with the TGF- β receptor/Smad axis in CD34⁻/CD105⁺ cells.⁴² This pathway may also provide a novel therapeutic field, as a soluble receptor fusion protein consisting of the extracellular domain of a modified activin receptor Type IIB linked to human Fc has recently been shown to reduce anemia in patients suffering from multiple myeloma.⁴³

Beside this, the molecular study performed here on sorted BM progenitors confirmed a parallel expression of genes and surface antigens, comforting FCM investigations.

Moreover, among 4524 genes differentially expressed between the CD34⁻/CD105⁺ erythroid compartment and CD34⁺ common progenitors, functionally important transcription factors including GATA-1, LMO2, Zinc-finger and BTB domain-containing protein 7A (ZBTB7A) were found to be upregulated, whereas GATA-2, FLI1 and others were found to be significantly downregulated. Eleven genes showing functional activity as transcription factors were equally expressed by CD34⁺ and CD34⁻/CD105⁺ cells. Apart from Myb, none of these genes have a known function within the erythroid series. However, it is remarkable that some of them, such as transcription factors 3 and transcription factor Dp-2 (TFDP2) as well as the activating transcription factor 7 interacting protein 2,⁴⁴⁻⁴⁶ appear to mediate either cell proliferation or apoptosis, depending on the context. Likewise, they might control apoptotic pathways (E2F7),⁴⁷ modulate embryonic stem cell pluripotency via the metal response element binding transcription factor 2,⁴⁸ or act as tumor suppressors via the leucine zipper transcription factor-like 1.⁴⁹

Comparison of the expression levels of genes of functional relevance for the erythroid skeleton, heme synthesis and iron metabolism showed, as expected, significant upregulation for most of them in progenitors engaged in an erythroid program.

The data reported here are consistent for many genes with recent reports on gene regulation within erythropoiesis.^{2-5,50} For example, 1,465 of the genes described by Merryweather-Clarke *et al.*⁴ as differentially regulated were also picked up here. However, some differences emerged, which may be explained by the use of different methodologies. Indeed, Merryweather-Clarke *et al.*⁴ concentrated on erythroid cells at different stages of differentiation, while we compared CD34⁺ and CD34⁻/CD105⁺ cells, therefore concentrating on early erythroid commitment. Furthermore, Merryweather-Clarke *et al.*⁴ used erythropoietin-stimulated cells, whereas we used freshly sorted unmanipulated cells.

The Ingenuity program further allowed us to focus on selected networks with known functions, such as the GATA-2 and endoglin pathways. Surprisingly, a distinct part of the cell cycle regulation was the pathway with the highest ranking for upregulated genes when comparing CD34⁻/CD105⁺ cells to the CD34⁺ common

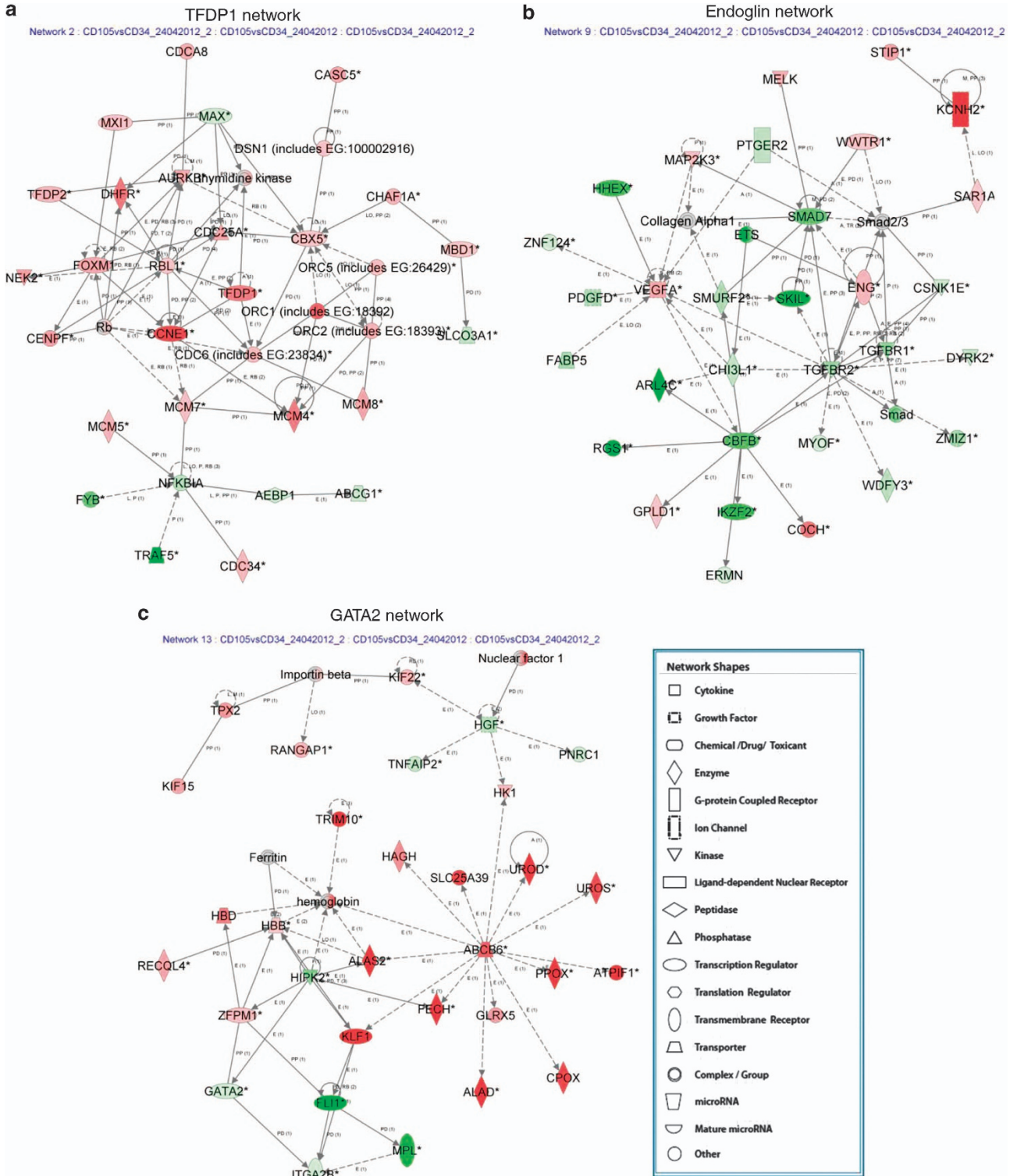


Figure 5. Pathways of relationships identified by feeding gene expression patterns identified in erythroid progenitors to the Ingenuity software. **(a)** TFDP1 network. The figure illustrates the relationship between TFDP1 and other key regulators of the cell cycle. **(b)** Endoglin network. Endoglin is closely related to the TGF- β receptor complex and to the Smad7/Smurf2 pathway. **(c)** GATA-2 network showing direct relationship with the ZFPM1 (FOG) gene known as a transcription regulator playing an essential role in erythroid (crucial role of ABCB6 for heme synthesis) and megakaryocytic cell differentiation (downregulation of FLI1, MPL, ITGA2B). Downregulated genes are shown in green, upregulated genes in red. Color intensity indicates fold changes (that is, deeper colors represent greater change). Solid arrows represent known physical interactions, dotted arrows represent indirect interactions.

progenitor compartment. This included transcription factor DP-1 (TFDP1), cyclin E (CCNE1), cell division cycle 6 homolog, origin recognition complex 1, retinoblastoma-like 1 and aurora kinase B.

In summary, this study establishes the immunophenotypic features of erythroid progenitors along three stages of maturation. Freshly obtained and separated CD34⁻/CD105⁺ early erythroid cells, evaluated for their gene expression profile in comparison with the CD34⁺ common progenitor compartment, allowed to observe a high number of differentially regulated genes. Although the function of some of these genes is as yet unknown, relevant functional pathways were nonetheless illustrated directly with the Ingenuity software program. These findings provide new information on the mechanisms and regulation of erythropoiesis and might find diagnostic/clinical applications in refractory anemia.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors are grateful to Karl Emile Thomanek and Julia Hess for preparing the manuscript. This work was supported in part by a grant from the Austrian Society of Hematology and Oncology.

AUTHOR CONTRIBUTIONS

SM-S, H-UK and PB designed and performed the research, analyzed and interpreted the data and wrote the manuscript; SS, MD, JP, CG, JL and RS performed research and contributed analytical tools; H-UK contributed analytical tools; MCB and AW wrote the paper.

REFERENCES

- Ney PA. Gene expression during terminal erythroid differentiation. *Curr Opin Hematol* 2006; **13**: 203–208.
- Doré LC, Crispino JD. Transcription factor networks in erythroid cell and megakaryocyte development. *Blood* 2011; **118**: 231–239.
- Hattangadi SM, Wong P, Zhang L, Flygare J, Lodish HF. From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood* 2011; **118**: 6258–6268.
- Merryweather-Clarke AT, Atzberger A, Soneji S, Gray N, Clark K, Waugh C et al. Global gene expression analysis of human erythroid progenitors. *Blood* 2011; **117**: e96–e108.
- Da Cunha AF, Brugnerotto AF, Duarte AS, Lanaro C, Costa GG, Saad ST et al. Global gene expression reveals a set of new genes involved in the modification of cells during erythroid differentiation. *Cell Prolif* 2010; **43**: 297–309.
- Cantù C, Lerardi R, Alborelli I, Fugazza C, Cassinelli L, Piconesi S et al. Sox6 enhances erythroid differentiation in human erythroid progenitors. *Blood* 2011; **117**: 3669–3679.
- Gregory GD, Miccio A, Bersenev A, Wang Y, Hong W, Zhang Z et al. FOG1 requires NuRD to promote hematopoiesis and maintain lineage fidelity within the megakaryocytic-erythroid compartment. *Blood* 2010; **115**: 2156–2166.
- Mancini E, Sanjuan-Pla A, Luciani L, Moore S, Grover A, Zay A et al. FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. *EMBO J* 2011; **31**: 351–365.
- Wontakal N Sandeep, Guo Xingyi, Will Britta, Shi M, Raha D, Mahajan MC et al. A large gene network in immature erythroid cells is controlled by the myeloid and B Cell transcriptional regulator PU.1. *PLoS Genet* 2011; **7**: e1001392.
- Fortunel NO, Hatzfeld A, Hatzfeld JA. Transforming growth factor- β : pleiotropic role in the regulation of hematopoiesis. *Blood* 2000; **96**: 2022–2036.
- Sánchez-Elsner T, Botella ML, Velasco B, Langa C, Bernabéu C. Endoglin expression is regulated by transcriptional cooperation between the hypoxia and transforming growth factor- β pathways. *J Biol Chem* 2002; **277**: 43799–43808.
- Sánchez-Elsner T, Ramirez JR, Rodriguez-Sanz F, Varela E, Bernabéu C, Botella LM. A cross-talk between hypoxia and TGF- β orchestrates erythropoietin gene regulation through sp1 and smads. *J Mol Biol* 2004; **336**: 9–24.
- Dong M, Blobe GC. Role of transforming growth factor- β in hematologic malignancies. *Blood* 2006; **107**: 4589–4596.
- Sivertsen EA, Hystad ME, Gutzkow KB, Døsen G, Smeland EB, Blomhoff HK et al. PI3K/Akt-dependent Epo-induced signalling and target genes in human early erythroid progenitor cells. *Br J Haematol* 2006; **135**: 117–128.
- Babitt JL, Huang FW, Xia Y, Sidis Y, Andrews NC, Lin HY. Modulation of bone morphogenetic protein signaling *in vivo* regulates systemic iron balance. *J Clin Invest* 2007; **117**: 1755–1758.
- Terskikh AV, Miyamoto T, Chang C, Diatchenko L, Weissman IL. Gene expression analysis of purified hematopoietic stem cells and committed progenitors. *Blood* 2003; **102**: 94–101.
- Perry JM, Harandi OF, Paulson RF. BMP4, SCF, and hypoxia cooperatively regulate the expansion of murine stress erythroid progenitors. *Blood* 2007; **109**: 4494–4502.
- Rylski M, Welch JJ, Chen YY, Letting DL, Diehl JA, Chodosh LA et al. GATA-1 mediated proliferation arrest during erythroid maturation. *Mol Cell Biol* 2003; **23**: 5031–5042.
- Munugalavada V, Dore LC, Tan BL, Hong L, Vishnu M, Weiss MJ et al. Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. *Mol Cell Biol* 2005; **25**: 6747–6759.
- Terszowski G, Waskow C, Conrad P, Lenze D, Koenigsmann J, Carstansen D et al. Prospective isolation and global gene expression analysis of the erythrocyte colony-forming unit (CFU-E). *Blood* 2005; **105**: 1937–1945.
- Loken MR, Shah VO, Dattilio KL, Civin CI. Flow cytometric analysis of human bone marrow: I. Normal erythroid development. *Blood* 1987; **69**: 255–263.
- Mirabelli P, Di Noto R, Lo Pardo C, Morabito P, Abate G, Gorrese M et al. Extended flow cytometry characterization of normal bone marrow progenitor cells by simultaneous detection of aldehyde dehydrogenase and early hematopoietic antigens: implication for erythroid differentiation studies. *BMC Physiol* 2008; **8**: 13.
- van der Velden VHJ, te Marvelde JG, Cullen M, Mejstrikova E, Flores-Mantero J, Sedek L et al. Section 7. Antibody panel for AML and MDS. In: van Dongen JJM, Lhermitte L, Böttcher S, Almeida J, van der Velden VHJ, Flores-Mantero J et al. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. *Leukemia*, 2012; **26**: 1908–1975.
- Southcott MJ, Tanner MJ, Anstee DJ. The expression of human blood group antigens during erythropoiesis in a cell culture system. *Blood* 1999; **93**: 4425–4435.
- Della Porta MG, Malcovati L, Invernizzi R, Travaglio E, Pascutto C, Maffioli M et al. Flow cytometry evaluation of erythroid dysplasia in patients with myelodysplastic syndrome. *Leukemia* 2006; **20**: 549–555.
- Xu F, Wu L, He Q, Zhang Z, Chang C, Li X. Immunophenotypic analysis of erythroid dysplasia and its diagnostic application in myelodysplastic syndromes. *Intern Med J* 2012; **42**: 401–411.
- Van de Loosdrecht AA, Alhan C, Béné MC, Della Porta MG, Dräger AM, Feuillard J et al. Standardization of flow cytometry in myelodysplastic syndromes: report from the first European LeukemiaNet working conference on flow cytometry in myelodysplastic syndromes. *Haematologica* 2009; **94**: 1124–1134.
- Malcovati L, Della Porta MG, Lunghi M, Pascutto C, Vanelli L, Travaglio E et al. Flow cytometry evaluation of erythroid and myeloid dysplasia in patients with myelodysplastic syndrome. *Leukemia* 2005; **19**: 776–783.
- Matarraz S, López A, Barrena S, Fernandez C, Jensen E, Flores-Mantero J et al. Bone marrow cells from myelodysplastic syndromes show altered immunophenotypic profiles that may contribute to the diagnosis and prognostic stratification of the disease: a pilot study on a series of 56 patients. *Cytometry B Clin Cytom* 2010; **78**: 154–168.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; **4**: 249–264.
- Anstee DJ. The functional importance of blood group-active molecules in human red blood cells. *Vox Sang* 2011; **100**: 140–149.
- Bühning HJ, Müller CA, Letarte M, Gougos A, Saalmüller A, van Agthoven AJ et al. Endoglin is expressed on a subpopulation of immature erythroid cells of normal human bone marrow. *Leukemia* 1991; **5**: 841–847.
- Rokhlin OW, Cohen MB, Kubagawa H, Letarte M, Cooper MD. Differential expression of endoglin on fetal and adult hematopoietic cells in human bone marrow. *J Immunol* 1995; **154**: 4456–4465.
- Cho SK, Bourdeau A, Letarte M, Zuniga-Pflücker JC. Expression and function of CD105 during the onset of hematopoiesis from Flk1⁺ precursors. *Blood* 2001; **98**: 3635–3642.
- Pierelli L, Bonanno G, Rutella S, Marone M, Scambia G, Leone G. CD105(endoglin) expression on hematopoietic stem/progenitor cells. *Leuk Lymphoma* 2001; **42**: 1195–1206.
- Moody JL, Singbrant S, Karlsson G, Blank U, Aspling M, Flygare J et al. Endoglin is not critical for hematopoietic stem cell engraftment and reconstitution but regulates adult erythroid development. *Stem Cells* 2007; **25**: 2809–2819.
- Pimanda JE, Chan WY, Wilson NK, Smith AM, Kinston S, Knezevic K et al. Endoglin expression in blood and endothelium is differentially regulated by modular assembly of the Ets/Gata hemangioblast code. *Blood* 2008; **112**: 4512–4522.

- 38 Tijssen MR, Cvejic A, Joshi A, Hannah RL, Ferreira R, Forrai A *et al*. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in Megakaryocytes identifies hematopoietic regulators. *Dev Cell* 2011; **20**: 597–609.
- 39 Zhang L, Magli A, Catanese J, Xu Z, Kyba M, Perlingeiro RC. Modulation of TGF- β signaling by endoglin in murine hemangioblast development and primitive hematopoiesis. *Blood* 2011; **118**: 88–97.
- 40 Christensen J, Cloos P, Toftegaard U, Klinkenberg D, Bracken AP, Trinh E *et al*. Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. *Nucleic Acids Res* 2005; **33**(17): 5458–5470.
- 41 Krystal G, Lam V, Dragowska W, Takahashi C, Appel J, Gontier A *et al*. Transforming growth factor beta 1 is an inducer of erythroid differentiation. *J Exp Med* 1994; **180**: 851–860.
- 42 Borges L, Iacovino M, Mayerhofer T, Koyano-Nakagawa N, Baik J, Garry DJ *et al*. A critical role for endoglin in the emergence of blood during embryonic development. *Blood* 2012; **119**: 5417–5428.
- 43 Raje N, Sotatercept VaileTS. a soluble activin receptor type ZA IgG-Fc fusion protein for the treatment of anemia and bone loss. *Curr Opin Mol Ther* 2010; **12**: 586–597.
- 44 Semerad CL, Mercer EM, Inlay MA, Weissman IL, Murre C. E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphoid and myeloerythroid progenitors. *Proc Natl Acad Sci USA*. 2009; **106**: 1930–1935.
- 45 Zhang Y, Venkatraj VS, Fischer SG, Warburton D, Chellappan SP. Genomic cloning and chromosomal assignment of the E2F dimerization partner TFDP gene family. *Genomics* 1997; **39**: 95–98.
- 46 Ichimura T, Watanabe S, Sakamoto Y, Aoto T, Fujita N, Nakao M. Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins. *J Biol Chem* 2005; **280**: 13928–13935.
- 47 Hu T, Ghazaryan S, Sy C, Wiedmeyer C, Chang V, Wu L. Concomitant inactivation of Rb and E2f8 in hematopoietic stem cells synergizes to induce severe anemia. *Blood* 2012; **119**: 4532–4542.
- 48 Jones A, Wang H. Polycomb repressive complex 2 in embryonic stem cells: an overview. *Protein Cell* 2010; **1**: 1056–1062.
- 49 Wei Q, Zhou W, Wang W, Gao B, Wang L, Cao J *et al*. Tumor-suppressive functions of leucine zipper transcription factor-like 1. *Cancer Res* 2010; **70**: 2942–2950.
- 50 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**: 17413–17418.



This work is licensed under a Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>

Supplementary Information accompanies this paper on Blood Cancer Journal website (<http://www.nature.com/bcj>)