Experimental Hematology

Experimental Hematology 2019;73:25–37

KLF1 mutation E325K induces cell cycle arrest in erythroid cells differentiated from congenital dyserythropoietic anemia patient-specific induced pluripotent stem cells

Hiroshi Kohara[a], Taiju Utsugisawa[b], Chika Sakamoto[c], Lisa Hirose[a], Yoshie Ogawa[a], Hiromi Ogura[b], Ai Sugawara[a], Jiyuan Liao[a], Takako Aoki[b], Takuya Iwasaki[b], Takayoshi Asai[d], Sayoko Doisaki[e], Yusuke Okuno[e], Hideki Muramatsu[e], Takaaki Abe[f], Ryo Kurita[f], Shohei Miyamoto[a], Tetsushi Sakuma[g], Masayuki Shiba[f], Takashi Yamamoto[g], Shouichi Ohga[h], Kenichi Yoshida[i], Seishi Ogawa[i], Etsuro Ito[j], Seiji Kojima[e], Hitoshi Kanno[b], and Kenzaburo Tani[a,c,k]

[a]*Project Division of ALA Advanced Medical Research, The Institute of Medical Science, University of Tokyo, Tokyo, Japan;* [b]*Department of Transfusion Medicine and Cell Processing, Tokyo Women's Medical University, Tokyo, Japan;* [c]*Division of Molecular and Clinical Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan;* [d]*Japanese Red Cross Chiba Blood Center, Chiba, Japan;* [e]*Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan;* [f]*Department of Research and Development, Central Blood Institute, Japanese Red Cross Society, Tokyo, Japan;* [g]*Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Hiroshima, Japan;* [h]*Department of Pediatrics, Yamaguchi University Graduate School of Medicine, Ube, Japan;* [i]*Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan;* [j]*Department of Pediatrics, Hirosaki University Graduate School of Medicine, Hirosaki, Japan;* [k]*Department of Advanced Molecular and Cell Therapy, Kyushu University Hospital, Fukuoka, Japan*

(Received 31 August 2018; revised 5 March 2019; accepted 6 March 2019)

Krüppel-like factor 1 (KLF1), a transcription factor controlling definitive erythropoiesis, is involved in sequential control of terminal cell division and enucleation via fine regulation of key cell cycle regulator gene expression in erythroid lineage cells. Type IV congenital dyserythropoietic anemia (CDA) is caused by a monoallelic mutation at the second zinc finger of *KLF1* (c.973G>A, p.E325K). We recently diagnosed a female patient with type IV CDA with the identical missense mutation. To understand the mechanism underlying the dyserythropoiesis caused by the mutation, we generated induced pluripotent stem cells (iPSCs) from the CDA patient (CDA-iPSCs). The erythroid cells that differentiated from CDA-iPSCs (CDA-erythroid cells) displayed multinucleated morphology, absence of CD44, and dysregulation of the *KLF1* target gene expression. In addition, uptake of bromodeoxyuridine by CDA-erythroid cells was significantly decreased at the CD235a$^+$/CD71$^+$ stage, and microarray analysis revealed that cell cycle regulator genes were dysregulated, with increased expression of negative regulators such as *CDKN2C* and *CDKN2A*. Furthermore, inducible expression of the KLF1 E325K, but not the wild-type KLF1, caused a cell cycle arrest at the G1 phase in CDA-erythroid cells. Microarray analysis of CDA-erythroid cells and real-time polymerase chain reaction analysis of the KLF1 E325K inducible expression system also revealed altered expression of several *KLF1* target genes including erythrocyte

H Kohara designed the project, performed experimental work, analyzed data, and prepared the manuscript; TU and HO analyzed data and prepared the manuscript; Y Ogawa performed experimental work and analyzed data; AS, T Aoki, LH, SM and TI performed experimental work; CS established iPSCs; T Asai, SD, Y Okuno, HM, S Ohga, KY, S Ogawa, EI, and SK diagnosed the patient and performed targeted resequencing; T Abe, RK and MS performed mass spectrometry analysis; TS and TY constructed TALEN vectors and provided instructions on genome editing technology; H Kanno planned the project, analyzed data, and prepared the manuscript; KT planned the project and prepared the manuscript; and all authors read and approved the manuscript.

Offprint requests to: Hitoshi Kanno, Department of Transfusion Medicine and Cell Processing, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan; E-mail: kanno.hitoshi@twmu.ac.jp and Kenzaburo Tani, Project Division of ALA Advanced Medical Research, The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan; E-mail: k-tani@ims.u-tokyo.ac.jp

Supplementary data related to this article can be found online at https://doi.org/10.1016/j.exphem.2019.03.001.

https://doi.org/10.1016/j.exphem.2019.03.001

membrane protein band 4.1 (*EPB41*), *EPB42*, glutathione disulfide reductase (*GSR*), glucose phosphate isomerase (*GPI*), and ATPase phospholipid transporting 8A1 (*ATP8A1*). Our data indicate that the E325K mutation in KLF1 is associated with disruption of transcriptional control of cell cycle regulators in association with erythroid membrane or enzyme abnormalities, leading to dyserythropoiesis. © 2019 Published by Elsevier Inc. on behalf of ISEH – Society for Hematology and Stem Cells.

Krüppel-like factor 1 (*KLF1*), also known as erythroid Krüppel-like factor (EKLF), is a transcription factor controlling erythroid lineage commitment and differentiation. Analysis of *KLF1*-deficient human neonates revealed that *KLF1* regulated more than 800 erythroid genes directly or indirectly [1]. *KLF1* knockout mice were embryonic lethal because of critical roles of *KLF1* in the fate decision and maturation of erythroid cells [2,3]. Erythroid progenitor cells accumulate in the fetal liver of *KLF1* knockout mice, and their cell cycle defect is attributable to reduced expression of E2F transcription factor 2 (E2F2) [4,5]. Recent studies revealed that *KLF1* induces expression of the cyclin-dependent kinase (CDK) inhibitors Cdkn2c and Cdkn1b, also known as $p18^{INK4C}$ and $p27^{KIP1}$, respectively, at the late stage, leading to enucleation [6].

Congenital dyserythropoietic anemias (CDAs) are inherited red blood cell disorders consisting of ineffective erythropoiesis and dyserythropoietic changes in the bone marrow. Three classic types of CDAs—type I, type II, and type III—have been characterized with somewhat different but overlapping patterns of symptoms [7–9], and the gene responsible for these CDA subgroups has been discovered already: *CDAN1* or *C15orf41* is mutated in type I, *SEC23B* is mutated in type II, and *KIF23* is mutated in type III [10–12]. Although several mutations have been identified in human *KLF1*, the phenotypes are quite heterogeneous and include hereditary persistence of fetal hemoglobin (HPFH), mild thalassemia, and severe dyserythropoietic anemia. In the most severe cases, unclassified CDA is caused by a heterozygous missense mutation consisting of G>A transition (c.973G>A) in exon 3 of *KLF1* [13–17]. This mutation alters the amino acid sequence of *KLF1* (p. E325K) in the second of the three zinc-finger domains, which are key domains for interacting with DNA promoter regions, at the carboxy terminus. Currently the *KLF1* mutation is classified as type IV CDA. The heterozygous c.973G>A mutation results in high expression of fetal hemoglobin and lack of some erythroid cell surface markers, including CD44 and aquaporin 1 (AQP1), in circulating erythrocytes and erythroblasts [13,16]. However, the key mediators of dyserythropoiesis in bone marrow erythroblasts in type IV CDA patients remain unclear.

Cell reprogramming technology enables the generation of induced pluripotent stem cells (iPSCs) from a limited number of primary patient cells. iPSCs generated from congenital anemia patients have been used to assess the possibility of novel gene therapies [18–20], to screen drug candidates [21,22], and to obtain mechanistic insights into the pathophysiology of human diseases [23]. We recently diagnosed a female patient with undiagnosed congenital anemia as having type IV CDA, and generated iPSCs from peripheral blood mononuclear cells of this patient carrying the heterozygous mutation in *KLF1*. In addition, by applying genome editing technology, we generated CDA-iPSCs with inducible expression cassettes for wild-type or mutant *KLF1* and further investigated the pathologic significance of *KLF1* E325K during erythroid differentiation, especially focusing on cell cycle regulation. Here, we report that impaired erythropoiesis by type IV CDA patient-specific iPSCs (CDA-iPSCs) is associated with dysregulated cell cycle controls.

Methods

Generation of iPSCs

This study was approved by the research ethics committees of The Institute of Medical Science, University of Tokyo; Tokyo Women's Medical University Graduate School of Medicine; and Kyushu University. The blood sample of the type IV CDA patient was obtained at Tokyo Women's Medical University, and blood samples from healthy donors were obtained at Kyushu University. Mononuclear cells were isolated from blood samples by density gradient centrifugation using Ficoll–Paque PLUS (GE Healthcare, Little Chalfont, UK), and T lymphocytes were expanded by culture in interleukin-2 (IL-2)-containing KBM-502 medium (Kohjin Bio, Saitama, Japan) in the presence of Dynabeads Human T-Expander CD3/CD28 (Thermo Fisher Scientific, Waltham, MA). After 3 to 5 days, expanded T lymphocytes were reprogrammed using a Sendai virus (SeV) vector (Medical & Biological Laboratories, Nagoya, Japan) encoding four pluripotency-associated transcription factors (OCT4, KLF4, SOX2, and c-MYC). The cells were transferred onto irradiated mouse embryonic fibroblasts (GlobalStem, Gaithersburg, MD) and cultured in human embryonic stem cell (ESC) medium (DMEM/F12, Sigma-Aldrich, St Louis, MO) supplemented with 20% knockout serum replacement (Thermo Fisher Scientific), 1 mmol/L glutamine, 1% nonessential amino acids, penicillin–streptomycin (all Nacalai Tesque, Kyoto, Japan), 0.1 mmol/L β-mercaptoethanol (Sigma-Aldrich), and 4 ng/mL FGF2 (Peprotech, Rocky Hill, NJ). After an additional 2 to 4 weeks of culture, a single primary colony was obtained and picked for the following experiments. One iPSC line was established from a single patient. CDA-iPSCs will be available from RIKEN BioResource Center (https://cell.brc.riken.jp/en/).

Generation of targeted iPSCs

Genome editing of CDA-iPSCs was performed as follows. Transcription activator-like effector nuclease (TALEN) pairs targeting the adeno-associated virus site 1 (AAVS1) locus, that is, AAVS1-TALEN-L and AAVS1-TALEN-R, were

reported previously [24]. Plasmid vectors harboring TALENs were co-transfected with a donor vector for homology-directed insertion of tet-on expression cassettes into an AAVS1 safe harbor locus using the Neon electroporation transfection system (Thermo Fisher Scientific). The cells were maintained on iMatrix-511 in StemMACS, and the cells with successful gene insertion were selected with 0.5 μg/ml puromycin and 500 μg/mL G418 (Thermo Fisher Scientific).

An enhanced green fluorescent protein (EGFP) reporter gene was used to test these expression cassettes (Supplementary Figure E1A, online only, available at www.exphem.org). Fluorescence imaging and flow cytometric analysis of reporter iPSCs revealed that EGFP fluorescence increased over time and was strong enough for detection at 48 hours after doxycycline treatment (Supplementary Figures E1B,D, and E3E, online only, available at www.exphem.org). At 48 hours in the presence of doxycycline, the EGFP fluorescence intensity was controllable by the concentration of doxycycline, at least in the range 40 to 1,000 ng/mL, in the reporter iPSCs (Supplementary Figures E1C–E, online only, available at www.exphem.org). Successful targeting of these expression cassettes to the AAVS1 locus of CDA-iPSCs was verified by polymerase chain reaction (PCR) and genomic DNA sequencing analysis (Supplementary Figure E2A,C,D, online only, available at www.exphem.org).

Erythroid induction

In vitro differentiation of CDA-iPSCs into CD34+ hematopoietic progenitor cells was induced via embryoid body (EB) formation in the presence of a cocktail of cytokines and growth factors as previously described, with some modification [25]. Colonies of CDA-iPSCs were detached with 0.5 × TrypLE Select (Thermo Fisher Scientific) diluted 1:1 with 0.5 mmol/L ethylenediamine tetraacetic acid (EDTA) in phosphate-buffered saline (PBS), resuspended in Dulbecco's modified Eagle's medium (DMEM; Nacalai Tesque) supplemented with 10 ng/mL BMP-4 (Peprotech, Rocky Hill, NJ), 10% fetal bovine serum (Thermo Fisher Scientific), 1 × nonessential amino acids, 2 mmol/L L-glutamine (both Nacalai Tesque), 100 μmol/L 2-mercaptoethanol, 50 mg/mL ascorbic acid (both Sigma-Aldrich), 100 U/mL penicillin and 100 μg/mL streptomycin, and cultured on ultralow-attachment six-well plates (Corning, Tewksbury, MA). On day 1, the culture medium was replaced with the complete medium additionally containing bone morphogenetic protein 4 (BMP-4), 15 ng/mL activin A, and 5 ng/mL basic fibroblast growth factor (bFGF). On day 5, the culture medium was replaced with the same medium additionally containing bFGF, 10 ng/mL vascular endothelial growth factor (VEGF), and 5 ng/mL interleukin (IL)-6. On day 10, the culture medium was replaced with the same medium containing VEGF, IL-6, and 50 ng/mL SCF and then cultured until day 14.

On day 14, EBs were dissociated by 2 hours of incubation in collagenase B (Roche, Basel, Switzerland), followed by 15 min of incubation in Cell Dissociation Buffer (Thermo Fisher Scientific). The cells were magnetically labeled using a combination of anti-CD34-phycoerythrin (PE) antibody and anti-PE microbeads, and CD34+ cells were isolated using the MACS system (Miltenyi, Bergisch Gladbach, Germany).

CD34+ cells were then cultured in Stem Span SFEM II containing 1 × Erythroid Expansion Supplement (both Stem Cell Technologies, Vancouver, BC, Canada). On day 21, the culture medium was replaced with Stem Span containing 10 ng/mL erythropoietin (EPO) and 100 ng/mL stem cell factor (SCF), and on day 28 the culture medium was replaced with Stem Span containing EPO.

Flow cytometry

The following antibodies were used for this study: TRA-1-60-PE, TRA-1-81, BCAM-PE (Miltenyi, Bergisch Gladbach, Germany), bromodeoxyuridine (BrdU)–fluorescein isothiocyanate (FITC; Biolegend, San Diego, CA), SSEA-1, CD34-PE, CD45-PECy7, CD235a-PE, CD235a-FITC, CD71-APC, and CD44-PE (all BD Biosciences, Franklin Lakes, NJ).

Cells were labeled with antibodies and marker expression was analyzed using a FACSVerse flow cytometer (BD Biosciences). To detect BrdU incorporation, a BrdU Flow Kit was used according to the manufacturer's instructions (BD Biosciences).

Immunofluorescence analysis and microscopy

iPSCs cultured in 12-well cell culture plates were fixed in 4% paraformaldehyde (Nacalai Tesque, Kyoto, Japan) for 30 min. Following incubation in 0.3% Triton X-100 (Nacalai Tesque) in PBS, cells were blocked with 3% bovine serum albumin (BSA; Sigma-Aldrich, St Louis, MO) in PBS, and stained with primary antibodies against NANOG and OCT3/4 diluted in 0.03% Triton X-100 in PBS. The cells were further stained with secondary antibodies conjugated with Alexa Fluor 488 (Thermo Fisher Scientific). For morphologic analysis of CDA-iPSC-derived erythroid cells, the cells were cytospun onto slide glasses and stained with Hemacolor (Merck Millipore, Darmstadt, Germany).

Immunofluorescence and live fluorescence images were acquired using a BZ-9000 fluorescence microscope with 20 × objective lens (Keyence, Osaka, Japan), and the images were processed using ImageJ software (National Institutes of Health, Bethesda, MD). Bright-field images were taken using a CKX41 microscope equipped with an E330 camera (all Olympus, Tokyo, Japan).

Microarray analysis

Two independent RNA samples were combined in a single analysis. Total RNA was extracted using Sepasol-RNA I Super G (Nacalai Tesque) and purified by using the miRNeasy Micro Kit (Qiagen). Transcriptional profiling was performed with the SurePrint 3G Human GE system based on the One-Color Microarray-Based Gene Expression Analysis protocol provided by Agilent, using the One Color Spike Mix Kit, Low Input Quick Amp Labeling Kit, and Expression Hybridization Kit (Agilent, Santa Clara, CA). Data were analyzed by using GeneSpring GX 14.5 software (Agilent). The microarray data analysis was performed by the Chemicals Evaluation and Research Institute (Tokyo, Japan) and Cell Innovator (Fukuoka, Japan). We used the following criteria for differential regulation of genes: upregulated genes, Z score ≥ 2.0 and ratio ≥ 1.5-fold; downregulated genes, Z score ≤ -2.0 and ratio ≤ 0.66. Our data have been uploaded to the Gene Expression Omnibus database (Accession No. GSE102985).

Droplet digital PCR analysis

To evaluate the effect of inducible expression of wild-type or E325K variant *KLF1*, we established a PCR-based system using Droplet Digital PCR system (Bio-Rad, Hercules, CA) to differentially quantitate the copy number of wild-type and mutant *KLF1* transcripts. At steady state, genome-edited CDA-iPSCs expressed wild-type and mutant *KLF1* at comparable levels, whereas control iPSCs expressed only wild-type *KLF1*, as expected (Supplementary Figure E2B). In our system, doxycycline was added at concentrations of 0, 100, and 1,000 ng/mL 7 days after $CD34^+$ cell isolation, after the emergence of $CD235a^+CD71^+$ erythroblastic cells; after 7 days of additional culture, the cells were collected for analyses (Supplementary Figure E2E). In both genome-edited CDA-iPCS lines expressing wild-type (tet-on *KLF1* Wt) and mutant (tet-on *KLF1* E325K) *KLF1*, the transgene was expressed after doxycycline treatment in a concentration-dependent manner (Supplementary Figure E2F). When doxycycline was added at concentration of 3,000 ng/mL, we detected clear upregulation of the intrinsic *KLF1* allele, which might represent positive feedback expression of intrinsic wild-type or mutant *KLF1* (data not shown).

Statistical analysis

Statistical analyses were performed with the GraphPad Prism 5.0d software package (GraphPad Software, La Jolla, CA). Statistical analysis among groups was performed using the two-tailed unpaired Student's t test or one-way analysis of variance followed by Tukey's multiple comparison test. P values < 0.05 were considered to indicate statistical significance.

Results

Clinical features

A 35-year-old woman with hemolytic anemia of unknown etiology was referred to the Tokyo Woman's Medical University Hospital. She was diagnosed with severe congenital hemolytic anemia in the infantile period, and received multiple red cell transfusions until the age of 6 years. Examination of the blood smear revealed abnormal red cell morphology with anisopoikilocytosis (Figure 1A), and splenomegaly and gallstones were observed on abdominal computed tomography, suggesting that she was affected by congenital hemolytic anemia, presumably because of a red cell membrane defect.

Screening tests for hemoglobinopathies, red cell membrane defects, and red cell enzymopathies were performed, and a markedly elevated fetal hemoglobin (HbF) level (21.9%) was detected. The eosin 5′-maleimide binding test revealed a slightly lower than normal level, and the activity of several red cell enzymes such as 6-phosphogluconate dehydrogenase (6PGD), glutathione peroxidase (GSH-Px), and adenylate kinase (AK) was simultaneously decreased (Table 1). These results implied that the anemia of this patient could not be explained by a single defect in hemoglobin, the red cell membrane, or enzymes.

Bone marrow aspirate examination revealed marked erythroid hyperplasia with hypercellular marrow and morphological abnormality with binucleation or multinucleation of erythroblasts (Figure 1B), suggesting that she had CDA. To diagnose CDA, we analyzed 184 genes associated with inherited bone marrow failure syndromes, as previously described [26], and found that she had a heterozygous missense mutation of *KLF1*, c.973G>A.

Generation of iPSCs from type IV CDA patient

We generated iPSCs from the peripheral blood mononuclear cells of the type IV CDA patient. CDA-iPSCs were able to be maintained on mouse embryonic fibroblasts for more than 30 passages without any change in morphology (Supplementary Figure E3A). Flow cytometric analysis and immunofluorescence analysis revealed that the pluripotency markers TRA-1-60, TRA-1-81, stage-specific embryonic antigen 4 (SSEA-4), NANOG, and OCT4 were expressed in the CDA-iPSCs (Supplementary Figure E3B). The c.973G>A heterozygous mutation at exon 3 of *KLF1* in CDA-iPSCs was confirmed by DNA sequencing analysis

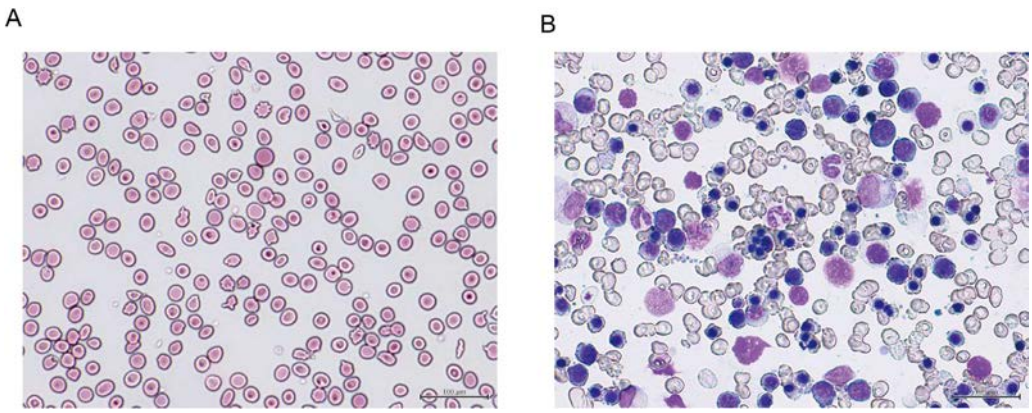

Figure 1. Analysis of the peripheral blood and bone marrow of the CDA patient. (**A**) Giemsa-stained blood sample of the patient at 35 years of age. (**B**) Giemsa-stained bone marrow sample of the patient at 31 years of age.

Table 1. Hematologic and serologic data and screening tests for congenital hemolytic anemia

RBC (10^6/lL)	Hb (g/dL)	Ht (%)	Retic (%)	MCV (fL)	MCH (pg)	MCHC (g/dL)	PLT (10^3/L)	Total bilirubin (mg/dL)	LDH (IU/L)	Iron (μg/dL)	Ferritin (ng/mL)	Haptoglobin (mg/dL)	HbF (%)	EMA binding (46.6–57.5 MCF)	AGLT (<30 min)	Isopropanol test (negative)
2.98*	8.8*	26.3*	2.7*	88	29.5	33.5	282	1.4*	251*	26*	15.7	21.9	21.9*	44.9	<30	negative
HK	GPI	PFK	ALD	TPI	PGK	ENOL	PK	G6PD	6PGD	GSH-Px	AK	ADA	AchE	P5N (CMP)	P5N (UMP)	GSH
(1.08	(57.2	(14.1	(2.62	(1,052	(264	(3.89	(13.0	(7.61	(9.0	(37.2	(165	(0.87	(28.6	(6.90	(9.75–15.5)	(65.9–88.5)
–1.46)	–70.3)	–20.0)	–6.30)	–1,567)	–326)	–6.30)	–19.8)	–9.81)	–10.7)	–51.4)	–307)	–1.59)	–42.7)	–10.8)		
3.17*	101*	14.9	8.89*	2,079*	387*	18.8*	63.7*	11.2*	7.08	31.1	105*	6.28*	37.7	13.6*	18.8*	81.7

RBC=red blood cells; Hb=hemoglobin; Ht=hematocrit; Retic=reticulocytes; MCV=mean corpuscular volume; MCH=mean corpuscular hemoglobin; MCHC=mean corpuscular hemoglobin concentration; PLT=platelets; LDH=lactate dehydrogenase; HbF = fetal hemoglobin; EMA=eosin 5′-maleimide; MCF=mean channel fluorescence=IU/g Hb for all red cell enzymes except for P5Ns (μmol P_i liberated/hr/g Hb) and GSH (mg/dL RBC); AGLT=acidified glycerol lysis time; HK=hexokinase; GPI=glucose phosphate isomerase; PFK=phosphofructokinase; ALD=aldolase; TPI=triosephosphate isomerase; PGK=phosphoglycerate kinase; ENOL=enolase; PK=pyruvate kinase; G6PD=glucose-6-phosphate dehydrogenase; 6PGD=6-phosphogluconate dehydrogenase; GSH-Px=glutathione peroxidase; AK=adenylate kinase; ADA=adenosine deaminase; AchE=acetylcholinesterase; P5N=pyrimidine 5′–nucleotidase; CMP=cytidine monophosphate; UMP=uridine monophosphate; GSH=reduced glutathione.
*Abnormal values.

(Supplementary Figure E3C). CDA-iPSCs retained a normal karyotype (Supplementary Figure E3D) and showed induction of teratoma derived from all three germ layers when injected into immunodeficient mice (Supplementary Figure E3E).

Defective erythropoiesis and dysregulated cell cycle in CDA-iPSC-derived erythroblasts

To generate hematopoietic progenitor cells from CDA-iPSCs, we induced differentiation via EB formation in the presence of a cocktail of cytokines and growth factors (Figure 2A). Among the EB cells derived from CDA-iPSCs, the $CD34^+$ cell fraction, composed of hematopoietic progenitor cells and endothelial cells, was comparable to that of healthy donor control iPSCs (Figure 2B,C). There was also no significant difference in the proportion of $CD45^+$ cells in the $CD34^+$ cell fraction, indicating that CDA-iPSCs and control iPSCs have comparable ability to generate primitive hematopoietic cells (Figure 2B,C).

$CD34^+$ cells were then enriched by magnetic bead sorting to a purity of at least 95% (data not shown). Erythroid differentiation was induced under serum-free liquid culture conditions in the presence of IL-3, SCF, and EPO (Figure 2A). At days 21, 28, and 35 of differentiation, there was no significant difference in the total *KLF1* mRNA expression level between control and CDA-erythroid cells when the assay probe used for gene expression analysis did not recognize the mutation site (Supplementary Figure E4A, online only, available at www.exphem.org). Microscopic images revealed that the control iPSC-derived cells in erythroid differentiation culture at day 35 contained hemoglobin-expressing red cells, whereas the CDA-iPSC-derived cells had a paler color, indicating a low efficiency of erythroid cell induction from CDA-iPSCs or lower globin expression (Figure 3A). Erythroid lineage cells approaching terminal differentiation can be divided into subfractions based on expression of CD71 and CD235a, also known as glycophorin A (GPA); $CD71^+/CD235a^-$ cells obtain CD235a expression, and then $CD71^+/CD235a^+$ cells decrease their CD71 expression [27]. Erythroblastic cells expressing CD71 and CD235a were detected by flow cytometric analysis during erythroid differentiation culture (Figure 3B). In accordance with the microscopic observations, the proportion of $CD235a^+$ cells was significantly lower in CDA-iPSC-derived cells on days 28 and 35 (Figure 3B,D). CD71 expression on $CD235a^+$ cells continuously decreased over time in cells from control iPSCs (Figure 3B,C). In contrast, CD71 expression was maintained in CDA-erythroid cells throughout the culture period (Figure 3B,C), indicating inappropriate maturation.

Cell cycle status is reported to be well organized during late erythropoiesis, and *KLF1* is critically involved in this process [6,28–31]. A BrdU incorporation assay revealed a

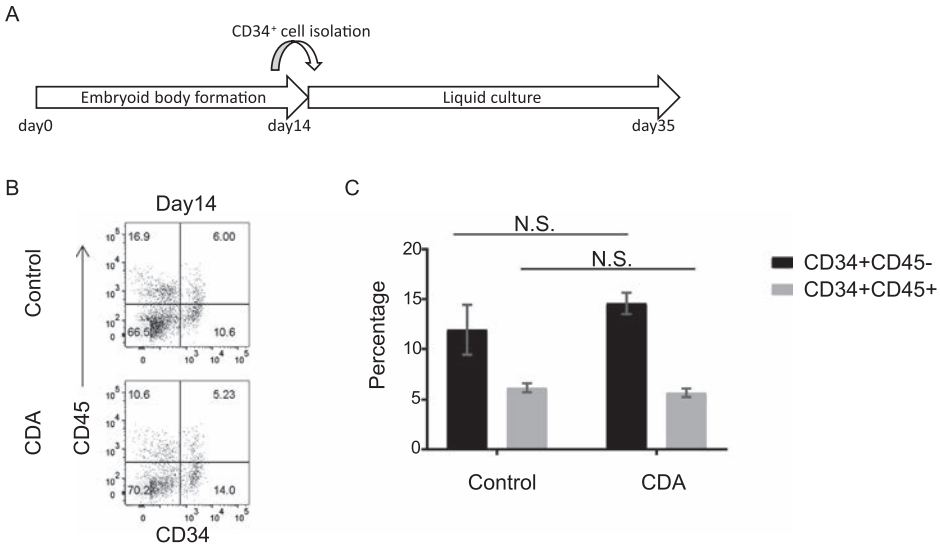

Figure 2. Induction of $CD34^+$ primitive hematopoietic cells from CDA-iPSCs. (**A**) Schematic representation of the differentiation procedure. EBs were differentiated in the presence of cytokine cocktails. $CD34^+$ cells isolated from day 14 embryoid bodies were cultured for an additional 21 days in the presence of EPO, IL-3, and SCF. (**B**) Representative flow cytometry dotplots of CD34 and CD45 on day 14 for EBs derived from CDA-iPSCs and control iPSCs. (**C**) Percentage of subsets of $CD34^+CD45^-$ cells and $CD34^+CD45^+$ cells among live EB cells derived from CDA- and control iPSCs. Data are the means ± SD of four experimental replicates. *N.S.*=no significance between samples.

decrease in cycling $CD235a^+$ cells during erythroid differentiation culture, and a significantly lower proportion of cells with BrdU uptake were detected in CDA-erythroid cells on day 28 (Figure 3E and Supplementary Figure E4B). These data indicated abnormal G0/G1 arrest at the $CD71^+/CD235a^+$ stage of CDA-erythroid cells.

We tested for the expression of Lutheran blood group glycoprotein (LU), also known as basal cell adhesion molecule (BCAM) or CD239, because its absence results in the In(Lu) phenotype and is linked to *KLF1* mutations [32]. However, LU expression was undetectable by flow cytometry during erythroid differentiation culture of both control iPSCs and CDA-iPSCs, indicating the limitation of our in vitro system as a model of erythropoiesis (Supplementary Figure E4C).

Disease-relevant phenotypes in CDA-iPSC-derived erythroblasts

The dysregulated cell cycle status in our in vitro model prompted us to further evaluate the characteristics of the $CD235a^+$ erythroid cell fraction derived from CDA-iPSCs. Morphologic analysis of the $CD235^+$ cell fraction revealed that many erythroblastic cells derived from CDA-iPSCs were multinucleated similarly to bone marrow erythroblasts from the CDA patient (Figures 1B and 4A). CD44, a receptor for hyaluronic acid, has been reported to be absent or specifically decreased in erythroid lineage cells in the peripheral blood of type IV CDA patients [13,16]. Flow cytometric analysis revealed that the $CD235a^+/CD71^+$ cell fraction derived from CDA-iPSCs lacked CD44 expression, whereas $CD45^+$ myeloid lineage cells that emerged simultaneously during erythroid differentiation expressed CD44 at levels comparable to those of the control group (Figure 4B). In addition, quantitative real-time (qRT)-PCR analysis indicated significantly decreased *AQP1* expression (Figure 4C), which is compatible with the clinical observations of other groups [13,16]. Gene expression analysis also revealed the trend for upregulation of gamma globin genes (*HBG1* and *HBG1*) and downregulation of beta globin gene (*HBB*), although we did not detect any clear change in the proportions of the different globin chains in each sample by mass spectrometry (Supplementary Figure E4D). In total, these data indicate that our in vitro model of CDA-erythroid cells mimics the disease process and is helpful in analyzing the pathophysiology of type IV CDA.

Disruption of transcriptional control in CDA-iPSC-derived erythroblasts

To understand how the heterozygous mutation in *KLF1* leads to dyserythropoietic changes and loss of cell cycle control in erythroblastic cells, we performed a comprehensive microarray analysis of $CD235a^+$ cells at day 35 of erythroid differentiation culture. In total, 293 genes were differentially expressed between CDA-iPSC-derived cells and control cells. First, we confirmed that among the genes with the *KLF1* binding motif NCNCNCCCN in their promoter regions [33–37], *BCL11A, AQP1,* and *CD44* were downregulated in the CDA sample, although there were no clear changes in the expression of intercellular adhesion molecule 4 (*ICAM4*) and *CDKN1A* ($p21^{CIP1}$; Supplementary Table E1, online only, available at www.exphem.org).

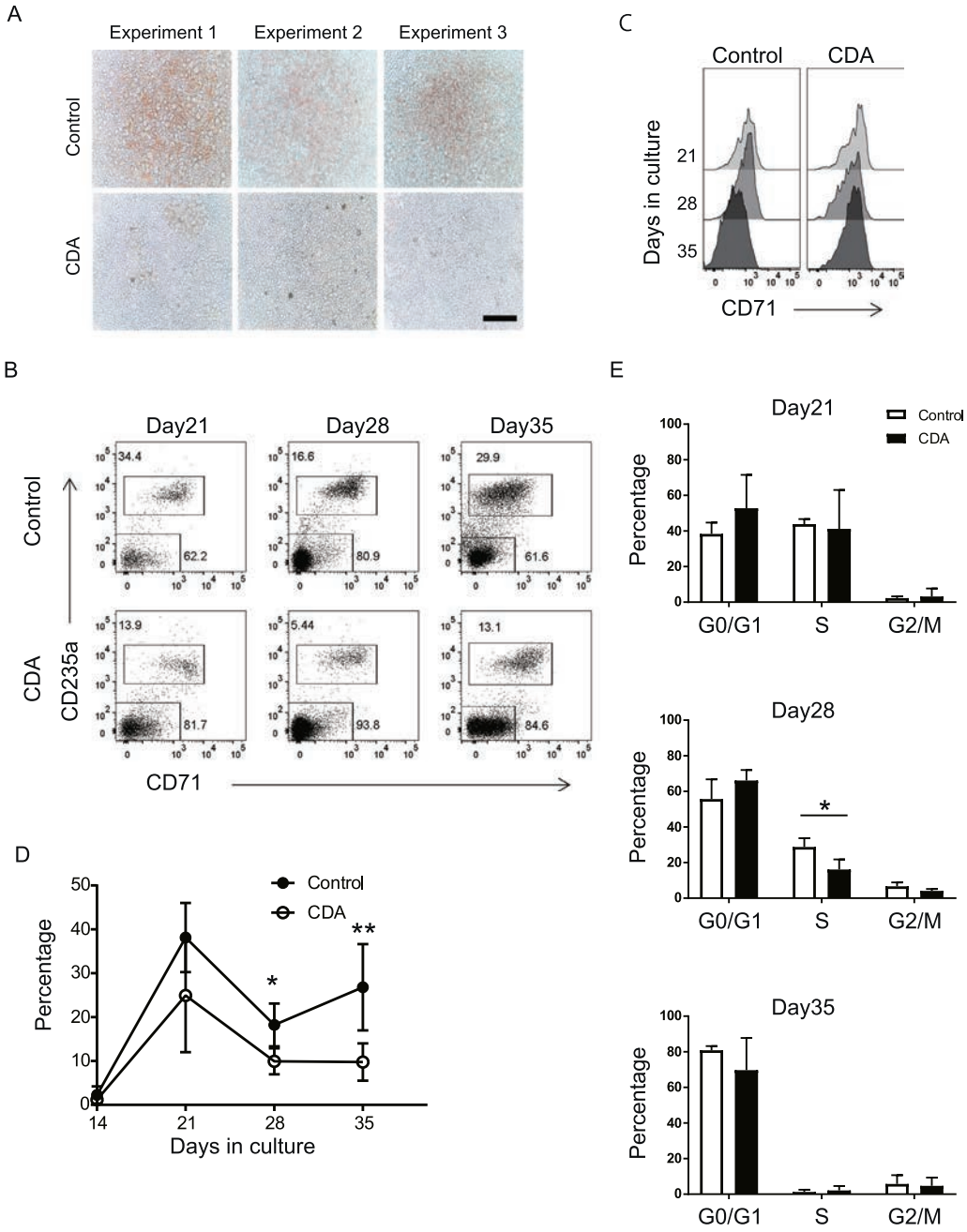

Figure 3. Defective erythropoiesis and dysregulated cell cycle status in erythroblastic cells derived from CDA-iPSCs. (**A**) Bright-field images of cells in erythropoietic liquid culture at day 35. Bar = 200 μm. (**B**) Representative flow cytometry dotplots of CD71 and CD235a during erythropoietic liquid culture started from $CD34^+$ cells derived from CDA-iPSCs and control iPSCs. The numbers indicate the percentages of the gated cells. (**C**) Representative flow cytometric histograms of CD71 expression on $CD235a^+$ cells during erythropoietic liquid culture started from $CD34^+$ cells derived from CDA-iPSCs and control iPSCs. (**D**) Percentage of $CD235a^+$ cells among live cells during erythropoietic liquid culture started from $CD34^+$ cells derived from CDA-iPSCs and control iPSCs. Data are the means ± SD of four experimental replicates. $*P < 0.05$; $**P < 0.02$. (**E**) Percentage of $BrdU^+$ cells among iPSC-derived erythroblasts after a pulse of BrdU was applied for 3 hours. Data are the means ± SD of three experimental replicates. $*P < 0.05$.

Signals for *HBB* and *BCAM* were not reliably detected because of their low expression in the samples.

In CDA-erythroid cells, many positive and negative regulators of the cell cycle were differentially expressed, indicating disruption of cell cycle control during the late stage of erythropoiesis from CDA-iPSCs (Supplementary Table E2, online only, available at www.exphem.org). Among these regulators, *E2F2* and *CDKN2C* are known

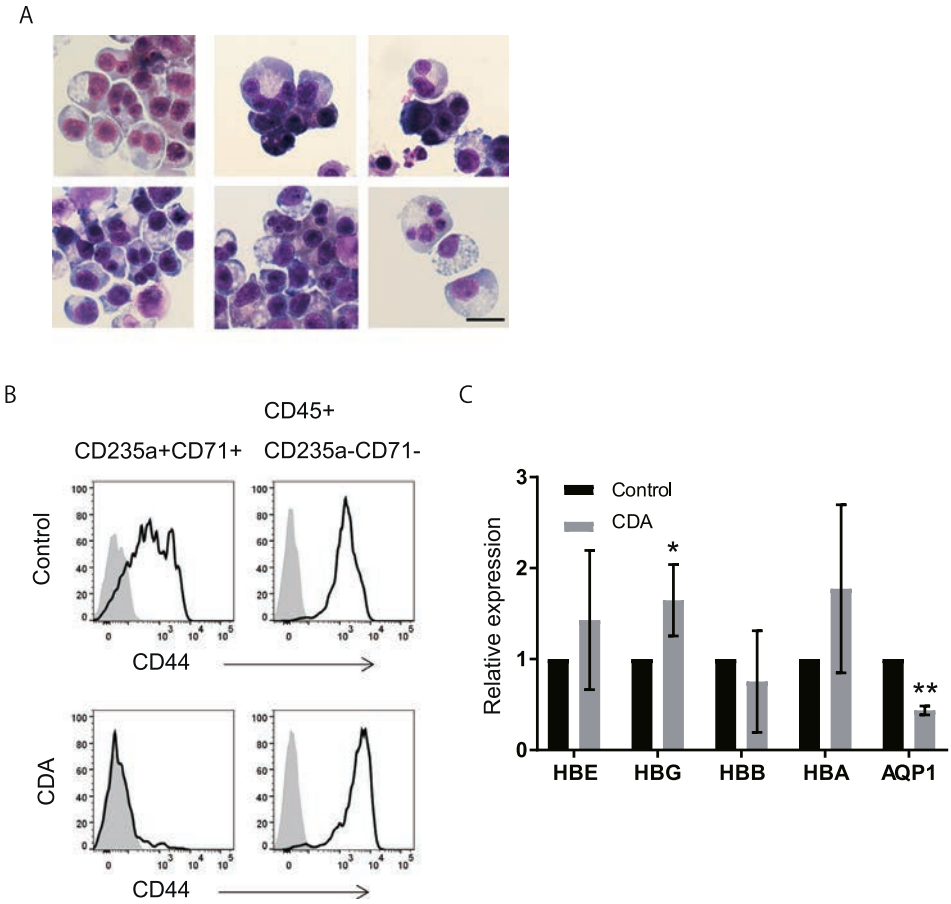

Figure 4. Disease-relevant phenotypes in CDA-iPSC-derived erythroblasts. (**A**) Cytospin specimen of erythroblastic cells on day 28 of differentiation culture of CDA-iPSCs with rapid Giemsa staining. Bar =10 μm. (**B**) Representative flow cytometric histograms of CD44 expression on $CD235a^{+}CD71^{+}$ cells and $CD45^{+}CD235a^{-}CD71^{-}$ cells derived from CDA-iPSCs and control iPSCs on day 35 in culture. (**C**) qRT-PCR analysis of the expression of *HBB, HBG1/HBG2, HBE, HBA*, and *AQP1*. Data are the means ± SD of three experimental replicates. $*P < 0.05$; $**P < 0.02$.

target genes of human *KLF1*. *E2F1* and cyclin genes, including *CCNB1, CCNB2, CCDD2*, and *CCND3,* were also upregulated. In the subset of negative cell cycle regulators, the CDK inhibitors *CDKN2C* and *CDKN2A* were upregulated, whereas *CDKN1C* ($p57^{KIP2}$) was downregulated. Although many positive and negative cell cycle regulators were upregulated, the upregulation of the CDK inhibitors *CDKN2C* and *CDKN2A* might explain the dyserythropoiesis in CDA-erythroid cells.

We then compared the differentially regulated genes with the set of *KLF1*-dependent genes previously identified as erythroid genes poorly expressed in circulating erythroblasts from patients with *KLF1*-null mutations [1], and unexpectedly found that approximately half of these *KLF1* target genes were upregulated in the CDA sample (Figure 5A). The subsets of *KLF1*-dependent genes upregulated (fold increase > 1.50, *Z* score > 2.00) and downregulated (fold increase < 0.66; Z score < −2.00) in CDA-erythroid cells are shown in Supplementary Tables E3 and E4 (online only, available at www.exphem.org), respectively.

In our panel of gene sets for genetic testing of hemolytic and dyserythropoietic anemia, erythrocyte membrane protein bands 4.1 (*EPB41*) and 4.2 (*EPB42*) were strikingly downregulated, indicating the instability of the membrane skeleton in erythrocytes (Figure 5B). Glutathione disulfide reductase (*GSR*), glucose phosphate isomerase (*GPI*), and ATPase phospholipid-transporting 8A1 (*ATP8A1*) were also downregulated in CDA-erythroid cells. In contrast, there was no clear change in the expression of CDA-related genes including *CDAN1, SEC23B*, and *KIF23* (Figure 5B). We also did not find any critical change in apoptosis pathway genes (Supplementary Figure E5, online only, available at www.exphem.org).

Cell cycle arrest at G1 phase by induction of the KLF1 E325K variant

To further study the impact of the E325K mutation on cell cycle status and *KLF1* activity in iPSC-derived erythroid cells, we constructed a tet-on controlled gene expression system expressing wild-type or mutant *KLF1* in CDA-iPSCs using genome editing technology

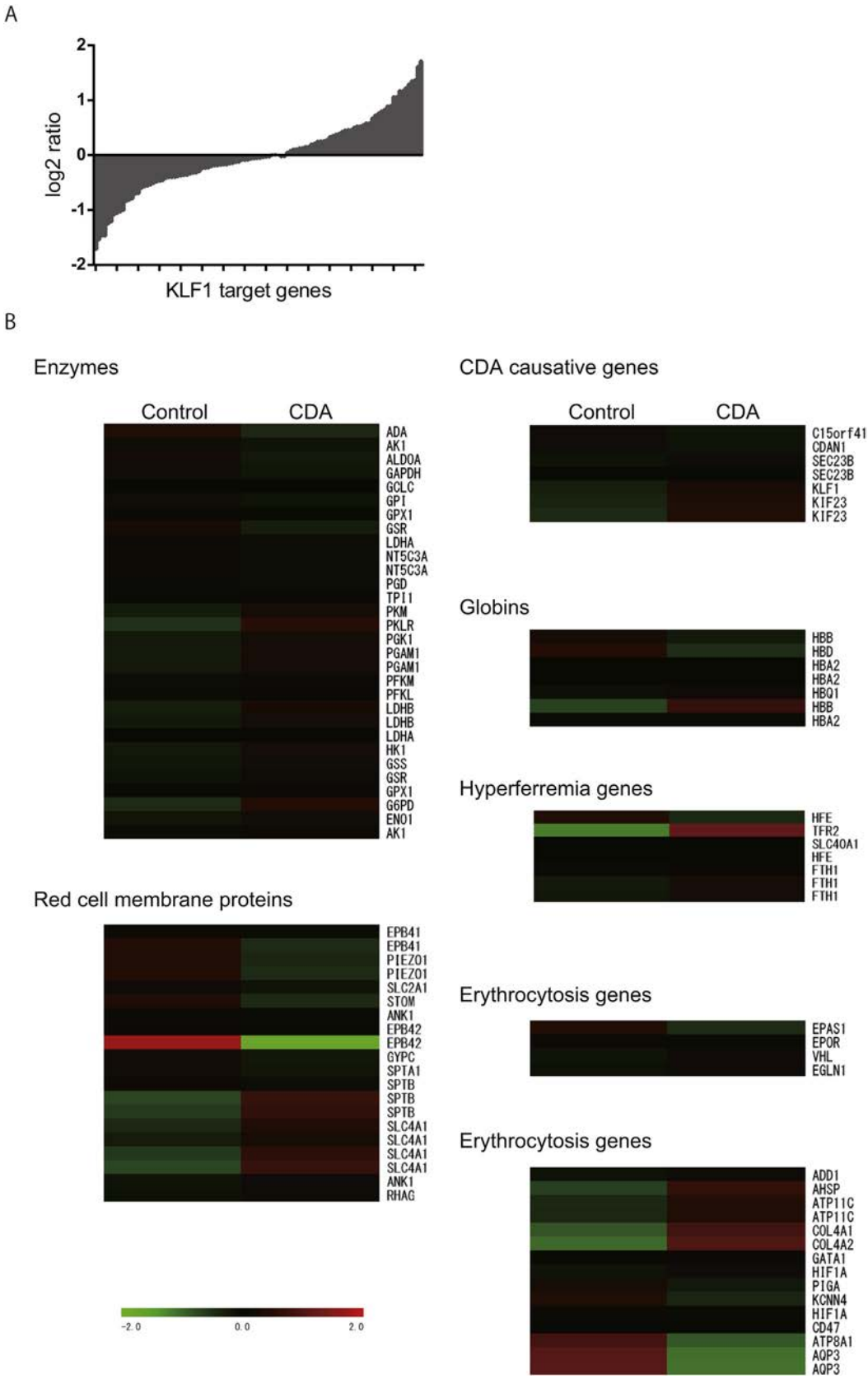

Figure 5. Disruption of transcriptional control in CDA-iPSC-derived erythroblasts. (**A**) Log2-fold change in the subset of KLF1 target genes expressed on the cells in erythroid differentiation culture. (**B**) Microarray heatmaps of congenital anemia-related gene expression in erythroblasts derived from CDA-iPSCs.

(Supplementary Figure E2A). Inducible expression of wild-type *KLF1* partially but clearly restored CD44 expression in $CD235a^+/CD71^+$ cells derived from CDA-iPSCs at day 28 of erythroid differentiation culture (Figure 6A).

Importantly, inducible expression of wild-type *KLF1* significantly increased, whereas *KLF1* E325K decreased, the proportion of $BrdU^+$ cycling cells in erythroid cells at day 28 when doxycycline was added at the concentration of 100 ng/mL (Figure 6B). The proportion of cells in each phase of the cell cycle indicated that *KLF1* E325K induced cell cycle arrest at G1 phase in erythroid cells during erythropoietic liquid culture (Figure 6B). These results were in accordance with the observation of cell cycle status in CDA-erythroid cells illustrated in Figure 3E. To assess the involvement of cell cycle-related genes in the cell cycle arrest shown in Figure 6B, we performed quantitative real-time PCR analysis of the panel of cell cycle–related genes. Most cell cycle–negative regulators tested were upregulated by expression of *KLF1* E325K (Supplementary Figure E6A, online only, available at www.exphem.org). We also evaluated the expression of some *KLF1* target genes other than cell cycle–related genes. The expression level of erythroid membrane protein *EPB41* was increased in response to induced wild-type *KLF1*, whereas it was decreased by *KLF1* E325K (Supplementary Figure E6B). Decreased transcriptional activity by *KLF1* E325K were also observed for mechanosensitive ion channel protein *PIEZO1*, membrane lipid flippase *ATP8A1*, and water channel aquaporin 3 (*AQP3*), although there were no statistical significance. Transcriptional activity of *KIF23* and *HBB* were not changed by *KLF1* E325K (Supplementary Figure E6B, online only, available at www.exphem.org).

Discussion

Clinical features of abnormal red cell morphology with anisopoikilocytosis, elevated HbF, and lower eosin 5′-maleimide binding in the CDA patient strongly led us to expect the dysregulation of *KLF1* target genes. From 184 genes related to inherited bone marrow failure, we identified a heterozygous missense mutation of *KLF1* [26], c.973G>A, which has been reported in CDA patients by other groups [13–17]. To identify the molecular basis of this type IV CDA, we took advantage of iPSC technology in combination with targeted genome editing technology. The patient's peripheral blood cells were reprogrammed into iPSCs successfully but with low efficiency; we obtained a single iPSC clone with normal morphology and growth characteristics. Because results variability of iPSCs even from the same patient is well known, analyzing multiple lines would have been more informative. This result motivated us to carefully consider the generalizability of our data on CDA-iPSCs and prompted us to generate

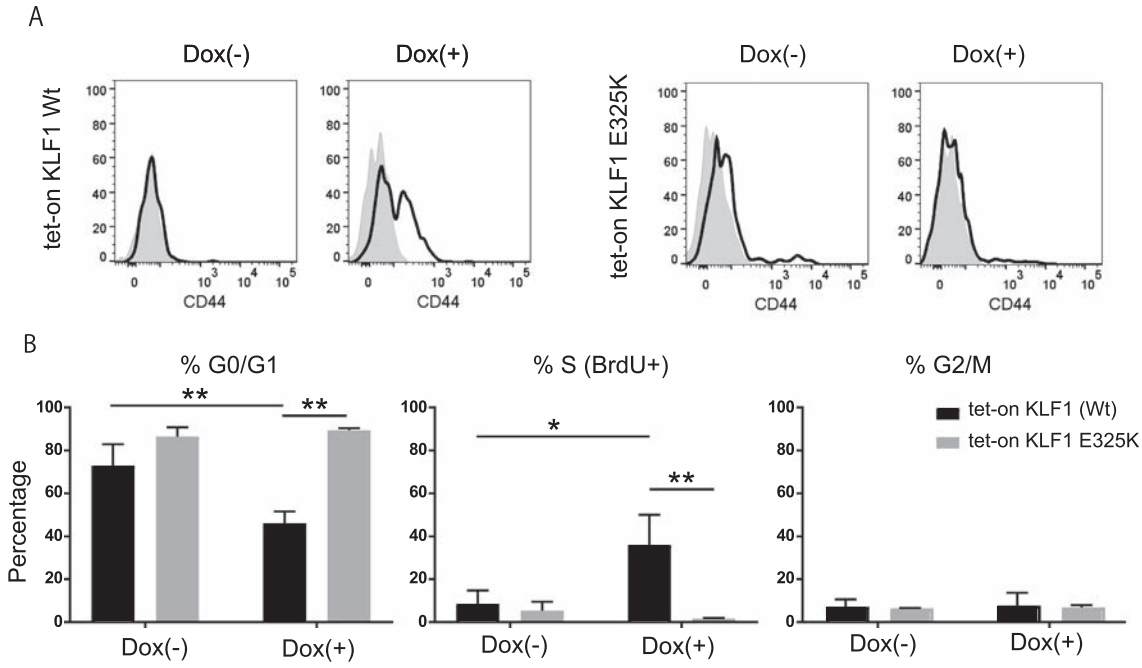

Figure 6. Dysregulation of the cell cycle by inducible expression of KLF1 E325K. (**A**) Flow cytometric analysis of erythroblastic cells cultured in the presence of 1,000 ng/mL doxycycline (Dox). (**B**) Percentage of erythroid cells in each phase of the cell cycle after a pulse of BrdU was applied for 3 hours. The cells were cultured in the presence of 100 ng/mL Dox for 1 week before experiments. Data are the means ± SD of three experimental replicates. $*P < 0.05$; $**P < 0.02$.

iPSCs with a tet-on inducible system to better understand the effect of *KLF1* E325K.

We first induced primitive hematopoietic cells from CDA-iPSCs. As expected because of the erythroid lineage-restricted expression of *KLF1*, comparable levels of $CD45^+$ and $CD45^-$ cells among $CD34^+$ cells were detected in control- and CDA-iPSC–derived cells. $CD34^+/CD45^+$ cells isolated from iPSCs during EB formation have been reported to retain their capability for hematopoiesis [22], and most $CD34^+/CD45^-$ cells were endothelial-lineage cells expressing vascular endothelial (VE)-cadherin (data not shown), which are thought to contain hemogenic endothelial cells [38,39]. The CDA-erythroid cells were well characterized specifically in known CDA properties. However, the lack of BCAM expression and the low efficiency of enucleation even in control erythroid cells indicate the limitations of our model.

KLF1 activity has been reported to be associated with genetic mutations and posttranscriptional modifications [40–42]. Most currently reported *KLF1* mutations with functional alterations are classified into hypomorphic variants or truncating loss-of-function variants. Some heterozygous frameshift or nonsense mutations resulting in truncated *KLF1* lacking zinc finger domains have also been reported, and these mutations are thought to cause haploinsufficiency for *KLF1*. Importantly, the clinical phenotypes of these mutations are milder than that of E325K [42]. *KLF1* E325K has been reported to actively interfere with transcription of *KLF1* target genes, including *CD44*, in a promoter–reporter assay using K562 chronic myelogenous leukemia cells, raising the possibility that *KLF1* E325K acts as a dominant–negative protein [13]. Alternatively, binding affinity of the mutant *KLF1* E325K to the consensus promoter sequence NCNCNCCCN [33–36] of *KLF1* target genes including *HBB, CD44*, and *AQP1* reportedly decreased [37]. In this study, we found that approximately half of the *KLF1* target genes that are downregulated in *KLF1*-null neonates [1] were upregulated in CDA-iPSC–derived cells. These observations suggest that the effects of the *KLF1* E325K interact with target DNA sequences in a multifactorial manner; that is, it acts through dominant-negative effects, dominant–positive effects, or haplo-insufficiency. It is important to note that expression of *KLF1* target genes could change dynamically during erythroid differentiation.

Detailed genetic studies are required to elucidate the mechanisms underlying this dysregulation of *KLF1* target genes. For example, the microarray analysis revealed no change in the expression of *ICAM4* and *CDKN1A*, which contain *KLF1* binding motifs in their promoter sequences. Although the recognition sequences in the promoter regions of these genes have been reported to show reduced binding affinity to *KLF1*, our data suggest that *ICAM4* and *CDKN1A* might be categorized in the subset of *KLF1* target genes without downregulation by *KLF1* E325K, as suggested in Figure 5A. Although RNA samples were prepared from erythroid cells at day 35, day 28 or slightly after that would have been optimum to assess *KLF1* effect, because the effect of the transcription might be confounded by other factors that help survival and maintenance of cells in culture at day 35.

Using microarray profiling and qRT-PCR analysis, we identified *EPB41* and *EPB42* as downregulated genes in CDA-erythroid cells. EPB41 and EPB42 are ATP-binding proteins that stabilize the membrane cytoskeleton of erythrocytes. Mutations in *EPB41* and *EPB42* are associated with hereditary elliptocytosis [43] and spherocytosis [44,45], respectively. Erythrocytes of the present case exhibited anisopoikilocytosis as well as decreased eosin 5′-maleimide binding. Therefore, downregulation of *EPB41* and *EPB42* might be associated with the abnormal morphology and the slight decrease in the surface area of mature erythrocytes in type IV CDA patients.

Microarray analysis also revealed that negative regulators of the cell cycle, *CDKN2C* and *CDKN2A*, were upregulated in CDA-erythroid cells. CDKN2C was recently reported to be one of the *KLF1* target cell cycle inhibitors regulating cell cycle exit and enucleation of murine erythroid cells [31]. To illustrate that cell cycle arrest is driven by *KLF1* E325K, we established an expression system for wild-type *KLF1* and *KLF1* E325K by inserting inducible expression cassettes into the AAVS1 locus of CDA-iPSCs. Gene complementation by induced expression of wild-type *KLF1* clearly restored the cell cycle defect in CDA-iPSCs. In contrast, induced expression of *KLF1* E325K exhibited no therapeutic effect on cell cycle status. *KLF1* has been reported to regulate terminal differentiation of erythroid cells by controlling the cell cycle regulators *E2F2, CDKN2C*, and *CDKN1B* [6]. It should be noted *CDKN2C* and *CDKN2A*, negative regulators of the cell cycle, were upregulated in CDA-erythroid cells (Supplementary Table E2). These observations were consistent with the upregulation of cell cycle regulators including *CDKN2A* and *CDKN2C* when *KLF1* E325K expression was induced. Based on these observations, we propose that *KLF1* E325K caused the cell cycle arrest of erythroblasts in type IV CDA patients. Further analysis is required to determine the molecular basis of cell cycle dysregulation in the CDA-erythroid cells.

Acknowledgments

We thank the CDA patient who participated in this study; the Flow Cytometry Core Facility and Center for Stem Cell Biology and Regenerative Medicine at The Institute of Medical Science, University of Tokyo,

for assistance with instrumentation; Dr. Satoshi Yamazaki for technical assistance with the teratoma assay; and Dr. Hiroaki Ono and Dr. Hidetoshi Takada for providing helpful comments for blood collection. This work was supported by the Japan Agency for Medical Research and Development (AMED), SBI Pharmaceutical, and Shinnihonseiyaku.

Conflict-of-interest disclosure

Hiroshi Kohara, Shohei Miyamoto and Kenzaburo Tani are supported by grants from Shinnihonseiyaku Company, SBI Pharmaceutical. Co, Ltd and neopharm Japan Co., Ltd. The terms of this arrangement have been reviewed and approved by the University of Tokyo in accordance with its conflict of interest policies. The remaining authors declare no competing financial interests.

References

1. Magor GW, Tallack MR, Gillinder KR, et al. KLF1-null neonates display hydrops fetalis and a deranged erythroid transcriptome. Blood. 2015;125:2405–2417.
2. Perkins AC, Sharpe AH, Orkin SH. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature. 1995;375:318–322.
3. Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature. 1995;375:316–318.
4. Pilon AM, Arcasoy MO, Dressman HK, et al. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. Mol Cell Biol. 2008;28:7394–7401.
5. Tallack MR, Keys JR, Humbert PO, Perkins AC. EKLF/KLF1 controls cell cycle entry via direct regulation of E2f2. J Biol Chem. 2009;284:20966–20974.
6. Gnanapragasam MN, McGrath KE, Catherman S, Xue L, Palis J, Bieker JJ. EKLF/KLF1-regulated cell cycle exit is essential for erythroblast enucleation. Blood. 2016;128:1631–1641.
7. Wickramasinghe SN. Congenital dyserythropoietic anaemias: clinical features, haematological morphology and new biochemical data. Blood Rev. 1998;12:178–200.
8. Heimpel H. Congenital dyserythropoietic anemias: epidemiology, clinical significance, and progress in understanding their pathogenesis. Ann Hematol. 2004;83:613–621.
9. Renella R, Wood WG. The congenital dyserythropoietic anemias. Hematol Oncol Clin North Am. 2009;23:283–306.
10. Bianchi P, Fermo E, Vercellati C, et al. Congenital dyserythropoietic anemia type II (CDAII) is caused by mutations in the SEC23B gene. Hum Mutat. 2009;30:1292–1298.
11. Renella R, Roberts NA, Brown JM, et al. Codanin-1 mutations in congenital dyserythropoietic anemia type 1 affect HP1{alpha} localization in erythroblasts. Blood. 2011;117:6928–6938.
12. Liljeholm M, Irvine AF, Vikberg AL, et al. Congenital dyserythropoietic anemia type III (CDA III) is caused by a mutation in kinesin family member, KIF23. Blood. 2013;121:4791–4799.
13. Arnaud L, Saison C, Helias V, et al. A dominant mutation in the gene encoding the erythroid transcription factor KLF1 causes a congenital dyserythropoietic anemia. Am J Hum Genet. 2010;87:721–727.
14. Ravindranath Y, Goyette G, Buck S, Gadgeel M, Dombkowski A, Boxer LA, Gallagher PG, Johnson RM. A new case of KLF1 G973A mutation and congenital dyserythropoietic anemia (CDA)—further definition of emerging new syndrome and possible association with gonadal dysgenesis. In: Abstracts, 2011 American Society of Hematology (ASH) Meeting. Blood. 2011;118:2101.
15. Singleton BK, Fairweather VS, Lau W, Parsons SF, Burton NM, Frayne J, Brady RL, Anstee DJ. A novel EKLF mutation in a patient with dyserythropoietic anemia: the first association of EKLF with disease in man. In: Abstracts, 2009 American Society of Hematology (ASH) Meeting. Blood. 2009;114:162.
16. Jaffray JA, Mitchell WB, Gnanapragasam MN, et al. Erythroid transcription factor EKLF/KLF1 mutation causing congenital dyserythropoietic anemia type IV in a patient of Taiwanese origin: review of all reported cases and development of a clinical diagnostic paradigm. Blood Cells Mol Dis. 2013;51:71–75.
17. de-la-Iglesia-Inigo S, Moreno-Carralero MI, Lemes-Castellano A, Molero-Labarta T, Mendez M, Moran-Jimenez MJ. A case of congenital dyserythropoietic anemia type IV. Clin Case Rep. 2017;5:248–252.
18. Bedel A, Taillepierre M, Guyonnet-Duperat V, et al. Metabolic correction of congenital erythropoietic porphyria with iPSCs free of reprogramming factors. Am J Hum Genet. 2012;91:109–121.
19. Garcon L, Ge J, Manjunath SH, et al. Ribosomal and hematopoietic defects in induced pluripotent stem cells derived from Diamond Blackfan anemia patients. Blood. 2013;122:912–921.
20. Garate Z, Quintana-Bustamante O, Crane AM, et al. Generation of a high number of healthy erythroid cells from gene-edited pyruvate kinase deficiency patient-specific induced pluripotent stem cells. Stem Cell Rep. 2015;5:1053–1066.
21. Liu GH, Suzuki K, Li M, et al. Modelling Fanconi anemia pathogenesis and therapeutics using integration-free patient-derived iPSCs. Nat Commun. 2014;5:4330.
22. Doulatov S, Vo LT, Macari ER, et al. Drug discovery for Diamond-Blackfan anemia using reprogrammed hematopoietic progenitors. Sci Transl Med. 2017;9:eaah5645.
23. Byrska-Bishop M, VanDorn D, Campbell AE, et al. Pluripotent stem cells reveal erythroid-specific activities of the GATA1 N-terminus. J Clin Invest. 2015;125:993–1005.
24. Takayama K, Igai K, Hagihara Y, et al. Highly efficient biallelic genome editing of human ES/iPS cells using a CRISPR/Cas9 or TALEN system. Nucleic Acids Res. 2017;45:5198–5207.
25. Kennedy M, Awong G, Sturgeon CM, et al. T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures. Cell Rep. 2012;2:1722–1735.
26. Muramatsu H, Okuno Y, Yoshida K, et al. Clinical utility of next-generation sequencing for inherited bone marrow failure syndromes. Genet Med. 2017;19:796–802.
27. Pop R, Shearstone JR, Shen Q, et al. A key commitment step in erythropoiesis is synchronized with the cell cycle clock through mutual inhibition between PU.1 and S-phase progression. PLoS Biol. 2010;8:e1000484.
28. Hsieh FF, Barnett LA, Green WF, et al. Cell cycle exit during terminal erythroid differentiation is associated with accumulation of p27(Kip1) and inactivation of cdk2 kinase. Blood. 2000;96:2746–2754.
29. Rylski M, Welch JJ, Chen YY, et al. GATA-1-mediated proliferation arrest during erythroid maturation. Mol Cell Biol. 2003;23:5031–5042.
30. Sankaran VG, Orkin SH, Walkley CR. Rb intrinsically promotes erythropoiesis by coupling cell cycle exit with mitochondrial biogenesis. Genes Dev. 2008;22:463–475.
31. Gnanapragasam MN, Bieker JJ. Orchestration of late events in erythropoiesis by KLF1/EKLF. Curr Opin Hematol. 2017;24:183–190.
32. Singleton BK, Burton NM, Green C, Brady RL, Anstee DJ. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. Blood. 2008;112:2081–2088.

33. Feng WC, Southwood CM, Bieker JJ. Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. J Biol Chem. 1994;269:1493–1500.
34. Crossley M, Whitelaw E, Perkins A, Williams G, Fujiwara Y, Orkin SH. Isolation and characterization of the cDNA encoding BKLF/TEF-2, a major CACCC-box-binding protein in erythroid cells and selected other cells. Mol Cell Biol. 1996;16:1695–1705.
35. Tallack MR, Whitington T, Yuen WS, et al. A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. Genome Res. 2010;20:1052–1063.
36. Siatecka M, Sahr KE, Andersen SG, Mezei M, Bieker JJ, Peters LL. Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. Proc Natl Acad Sci USA. 2010;107:15151–15156.
37. Singleton BK, Lau W, Fairweather VS, et al. Mutations in the second zinc finger of human EKLF reduce promoter affinity but give rise to benign and disease phenotypes. Blood. 2011;118:3137–3145.
38. Hirai H, Ogawa M, Suzuki N, et al. Hemogenic and nonhemogenic endothelium can be distinguished by the activity of fetal liver kinase (Flk)-1 promoter/enhancer during mouse embryogenesis. Blood. 2003;101:886–893.
39. Ogawa M, Kizumoto M, Nishikawa S, Fujimoto T, Kodama H, Nishikawa SI. Expression of alpha4-integrin defines the earliest precursor of hematopoietic cell lineage diverged from endothelial cells. Blood. 1999;93:1168–1177.
40. Siatecka M, Bieker JJ. The multifunctional role of EKLF/KLF1 during erythropoiesis. Blood. 2011;118:2044–2054.
41. Borg J, Patrinos GP, Felice AE, Philipsen S. Erythroid phenotypes associated with KLF1 mutations. Haematologica. 2011;96:635–638.
42. Perkins A, Xu X, Higgs DR, et al. Kruppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. Blood. 2016;127:1856–1862.
43. Delaunay J. Molecular basis of red cell membrane disorders. Acta Haematol. 2002;108:210–218.
44. Bouhassira EE, Schwartz RS, Yawata Y, et al. An alanine-to-threonine substitution in protein 4.2 cDNA is associated with a Japanese form of hereditary hemolytic anemia (protein 4.2NIPPON). Blood. 1992;79:1846–1854.
45. Rybicki AC, Heath R, Wolf JL, Lubin B, Schwartz RS. Deficiency of protein 4.2 in erythrocytes from a patient with a Coombs negative hemolytic anemia: evidence for a role of protein 4.2 in stabilizing ankyrin on the membrane. J Clin Invest. 1988;81:893–901.

Supplementary Methods

Construction of the targeting vector

Two targeting vectors for homologous recombination were generated. For PCR amplification of fragments from plasmid vectors for cloning, PrimeSTAR Max DNA Polymerase (Takara Bio, Otsu, Japan) was used, and for PCR amplification of a fragment form genomic DNA for cloning, KOD FX Neo Polymerase (Toyobo, Osaka, Japan) was used. The tetracycline-responsive element (TRE) was amplified from the vector provided in the Tet-On 3G Inducible Expression System EF1 Alpha Version (Takara). KLF1 was amplified from genomic DNA of healthy donor-derived iPSCs, and these fragments were cloned using the In-Fusion HD Cloning Kit (Takara) into an AAVS1 SA-PGK-neo-pA vector that was obtained from Addgene. The neomycin resistance gene in the AAVS1

SA-PGK-neo-pA vector was replaced with a puromycin resistance gene by using In-Fusion (AAVS1 SA-PGK-puro-pA). Then, the EF-1 alpha promoter and reverse tetracycline transactivator (rtTA) were amplified from the vector provided in the Tet-On 3G Inducible Expression System, and cloned into AAVS1 SA-PGK-puro-pA by using the In-Fusion system. The KLF1 c.973G>A mutation was introduced into the targeting vector by site-directed mutagenesis using the In-Fusion system.

Sequencing analysis

For DNA sequencing, genomic DNA was isolated using a DNeasy Blood & Tissue Kit (Qiagen), and the targeted region was amplified by PCR. PCR products from CDA-iPSCs were sequenced by Eurofins Genomics (Tokyo, Japan). Plasmid vectors were directly sequenced by Eurofins Genomics.

Cytogenetic Analysis

CDA-iPSCs passaged in feeder-free conditions were harvested and G-banded, and metaphase cell analysis was supported by Nihon Gene Research Laboratories (Sendai, Japan).

Teratoma assay

Three NOD/Shi-scid, IL-2Rγnull (NOG) mice were anaesthetized and injected under the kidney capsule with 1×10^7 cells each of undifferentiated CDA-iPSCs mixed with Matrigel. Mice were euthanized after 14 weeks. Paraformaldehyde-fixed paraffin-embedded tumors were sectioned and stained with hematoxylin-eosin.

Mass spectrometry analysis

Cell samples containing erythroid cells were lysed using 1 ml of 10 mM Tris-HCl buffer (pH 7.8). After centrifugation, the protein concentrations of the supernatants obtained were evaluated using a BCA Protein Assay kit (Thermo Fisher Scientific). The globin proteins in the supernatants were digested using trypsin and quantified using a high resolution multiple reaction monitoring assay on a TripleTOF 5600 system coupled to a NanoLC Ultra system (SCIEX, Framingham, MA). The peak areas of fragment ions for each globin-specific peptide (ε: AAVTSLWSK, δ: EFTPQMQAAYQK, $\gamma 1$: MVTAVASALSSR, $\gamma 2$: MVTGVASALSSR, and β: VNVDEVGGEALGR) were calculated to determine the concentrations of globin proteins. Synthetic peptides corresponding to each sequence were used to generate calibration curves.

Quantitative real-time PCR

RNA was isolated using the RNeasy Plus Mini RNA Isolation kit (Qiagen, Hilden, Germany) and reverse-transcribed using ReverTra Ace (Toyobo, Osaka, Japan). The cDNA reaction was diluted in water and used in Taqman real-time PCR reactions (Thermo Fisher Scientific). Reactions were run in duplicate on a StepOne real-time PCR system (Thermo Fisher Scientific) according to the manufacturer's instructions. Predesigned TaqMan gene expression assays (Thermo Fisher Scientific) or PrimeTime Predesigned qPCR Assays (Integrated DNA Technologies) were used to determine the expression of genes. Values were normalized to those of glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and then compared to the normalized values for control cells.

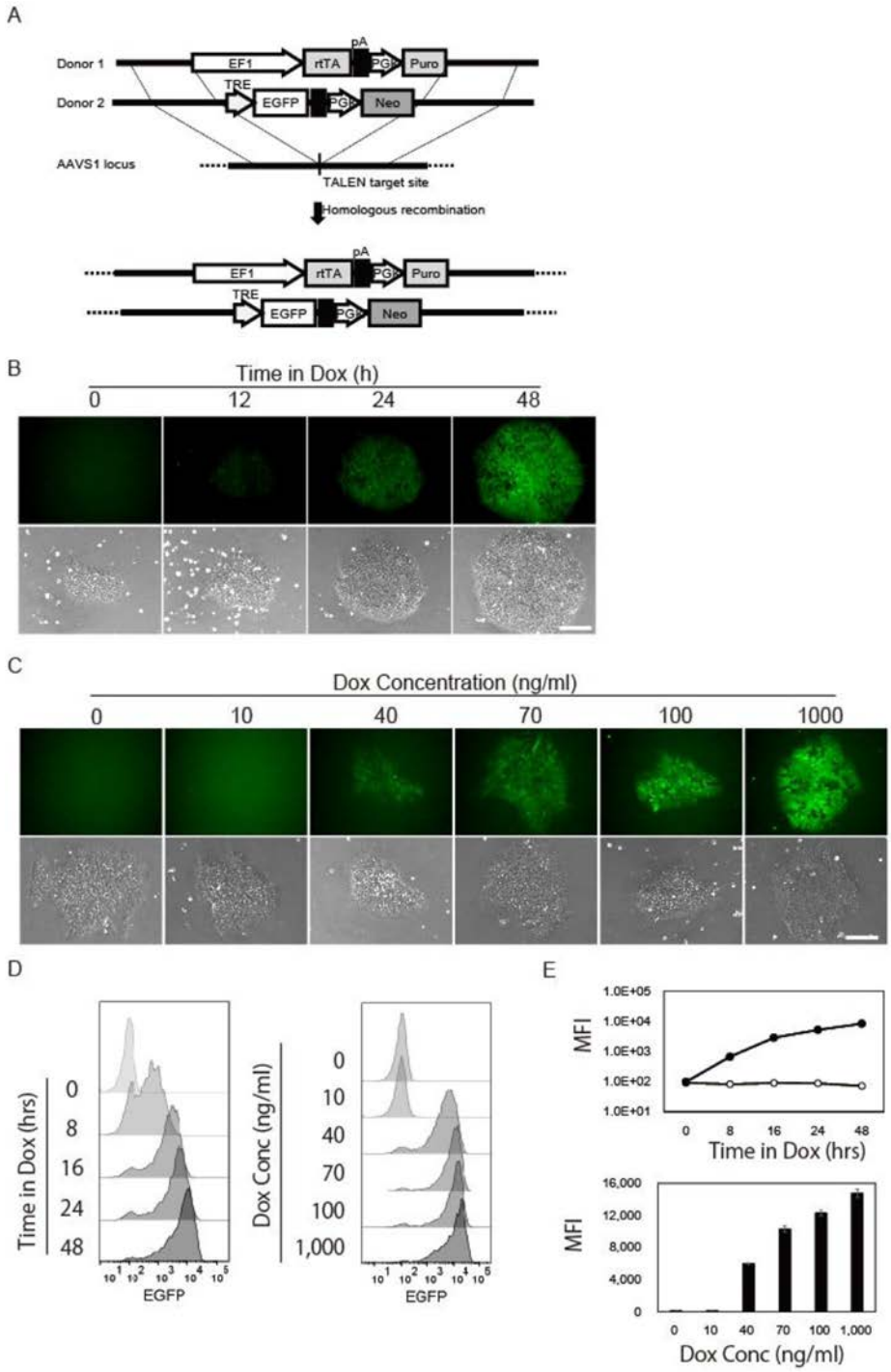

Figure S1. Quality check of tet-on inducible gene expression system in iPSCs. (A) Schematic illustration of AAVS1 gene targeting with two different donor vectors, i.e., reverse tetracycline trans-activator (rtTA) under control of the EF1 promoter and EGFP under control of the tetracycline-responsive element (TRE), for homology-directed repair of the double-strand break. (B) Fluorescent images and phase contrast images of iPSCs with the tet-on inducible EGFP expression cassette cultured for the indicated hours in the presence of 1,000 ng/ml doxycycline (Dox). Scale bar indicates 200 μm. (C) Fluorescence images and phase-contrast images of iPSCs with the tet-on inducible EGFP expression cassette cultured for 48 hours in the presence of Dox at the indicated concentrations. Scale bar indicates 200 μm. (D) Histograms of EGFP intensity in iPSCs with the tet-on inducible EGFP expression cassette cultured for the indicated hours in the presence of Dox at the indicated concentrations. Data are representative of three independent experiments. (E) Mean fluorescence intensity of EGFP in iPSCs with the tet-on inducible EGFP expression cassette cultured for the indicated hours in the presence of Dox at the indicated concentrations. Data are shown as the mean ± SD of three experimental replicates.

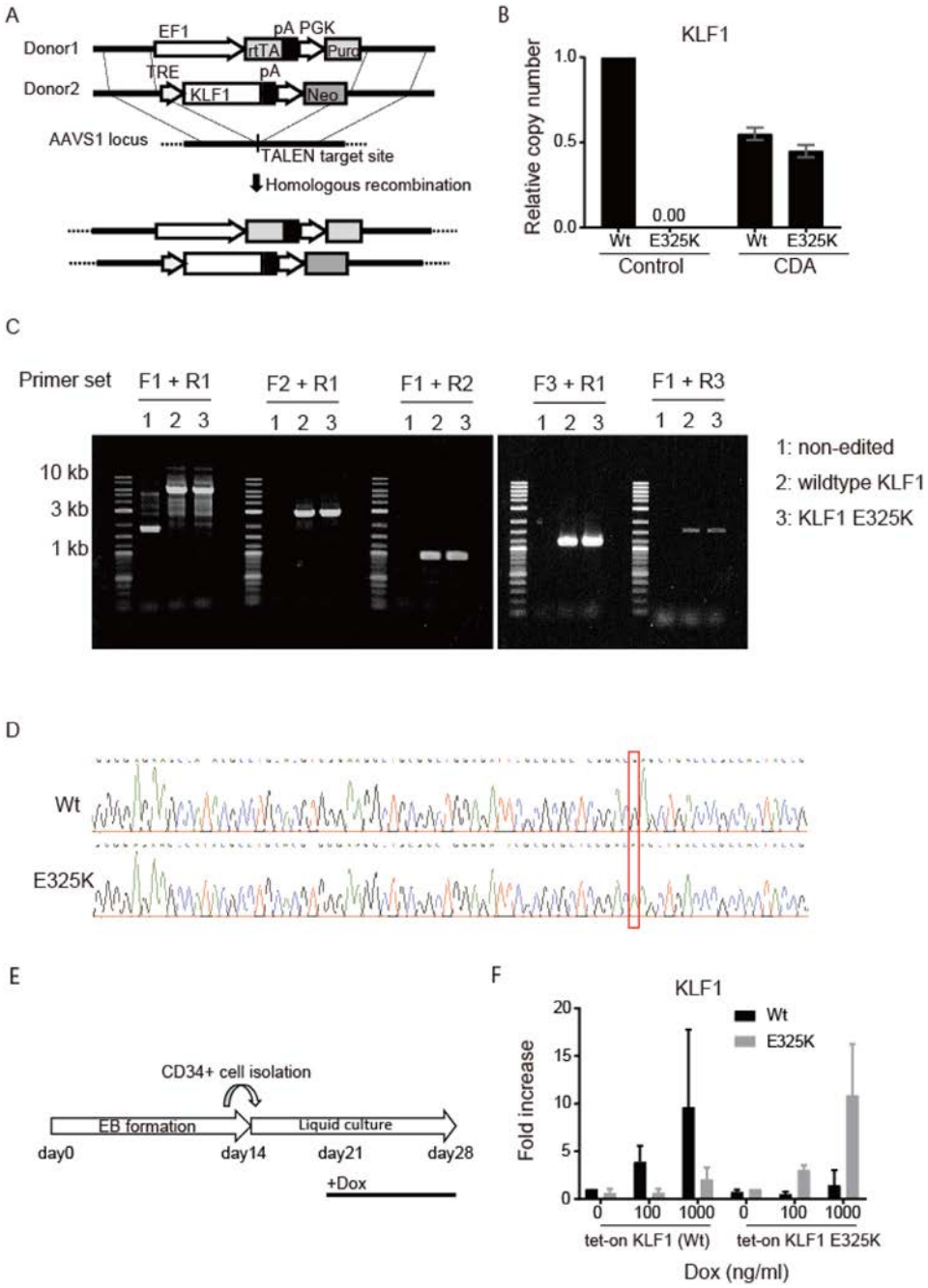

Figure S2. Quality check of iPSCs with the tet-on inducible KLF1 expression cassette. (A) Schematic illustration of AAVS1 gene targeting with two different donor vectors for homology-directed repair of the double-strand break. (B) Droplet digital PCR analysis of erythroblastic cells at day 28 of differentiation culture using probe sets to distinguish wildtype (Wt) and mutant (E325K) *KLF1* mRNA. Data are shown as the mean ± SD of three experimental replicates. (C) Gel images of PCR products amplified from genome-edited CDA-iPSCs with the primer set indicated in (A). (D) Genome DNA sequencing of genome-edited CDA-iPSCs, showing the G>A transition in the KLF1 gene inserted into the AAVS1 locus. (E) Schematic representation of the differentiation procedure. Embryoid bodies (EBs) were differentiated in the presence of cytokine cocktails. $CD34^+$ cells isolated from day14 embryoid bodies were cultured for an additional 14 days in the presence of EPO, IL-3, and SCF. Doxycycline (Dox) was added from day 21 to day 28. (F) Droplet digital PCR analysis of erythroblastic cells cultured in the presence of doxycycline (Dox) at the indicated concentration. The fold change in KLF1 expression was determined relative to the expression in erythroblastic cells cultured in the absence of Dox.

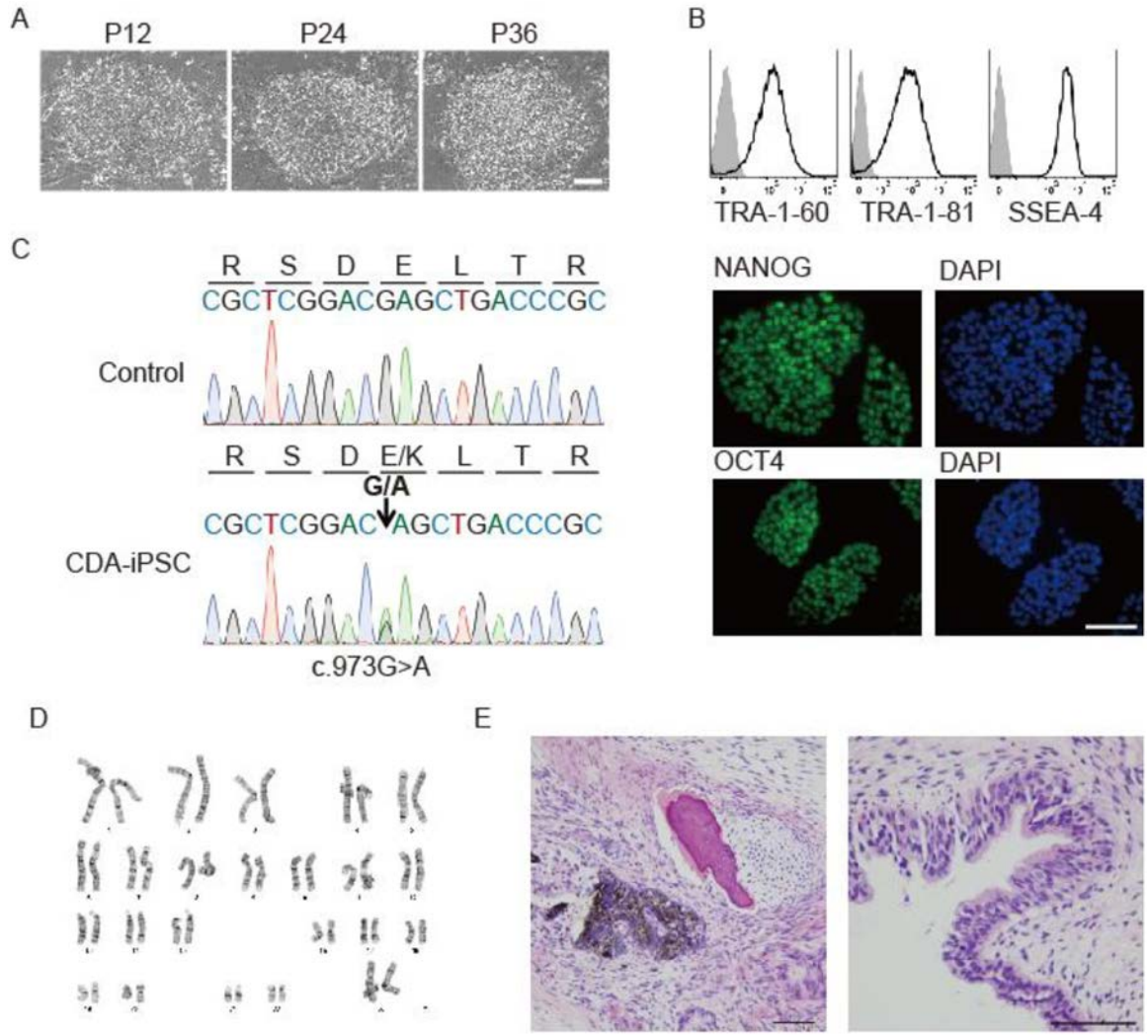

Figure S3. Generation of iPS cells from peripheral blood T lymphocytes in a CDA type IV patient. (A) Phase contrast images of CDA-iPSCs maintained on MEF feeder layers at three different passages: P12, P24 and P36. Scale bar indicates 100 μm. (B) Pluripotent stem cell markers detected by flow cytometry (TRA-1-60, TRA-1-81, and SSEA-4) and by immunofluorescence microscopy (NANOG and OCT4). Scale bar indicates 100 μm. (C) Genome DNA sequencing of CDA-iPSCs and control iPSCs, showing the G>A transition at exon 2. (D) G-band staining of CDA-iPSCs showing normal female karyotype (46, XX). (E) Histological analysis of tissues including three germ layers in teratoma induced from CDA-iPSCs. Scale bar indicates 200 μm.

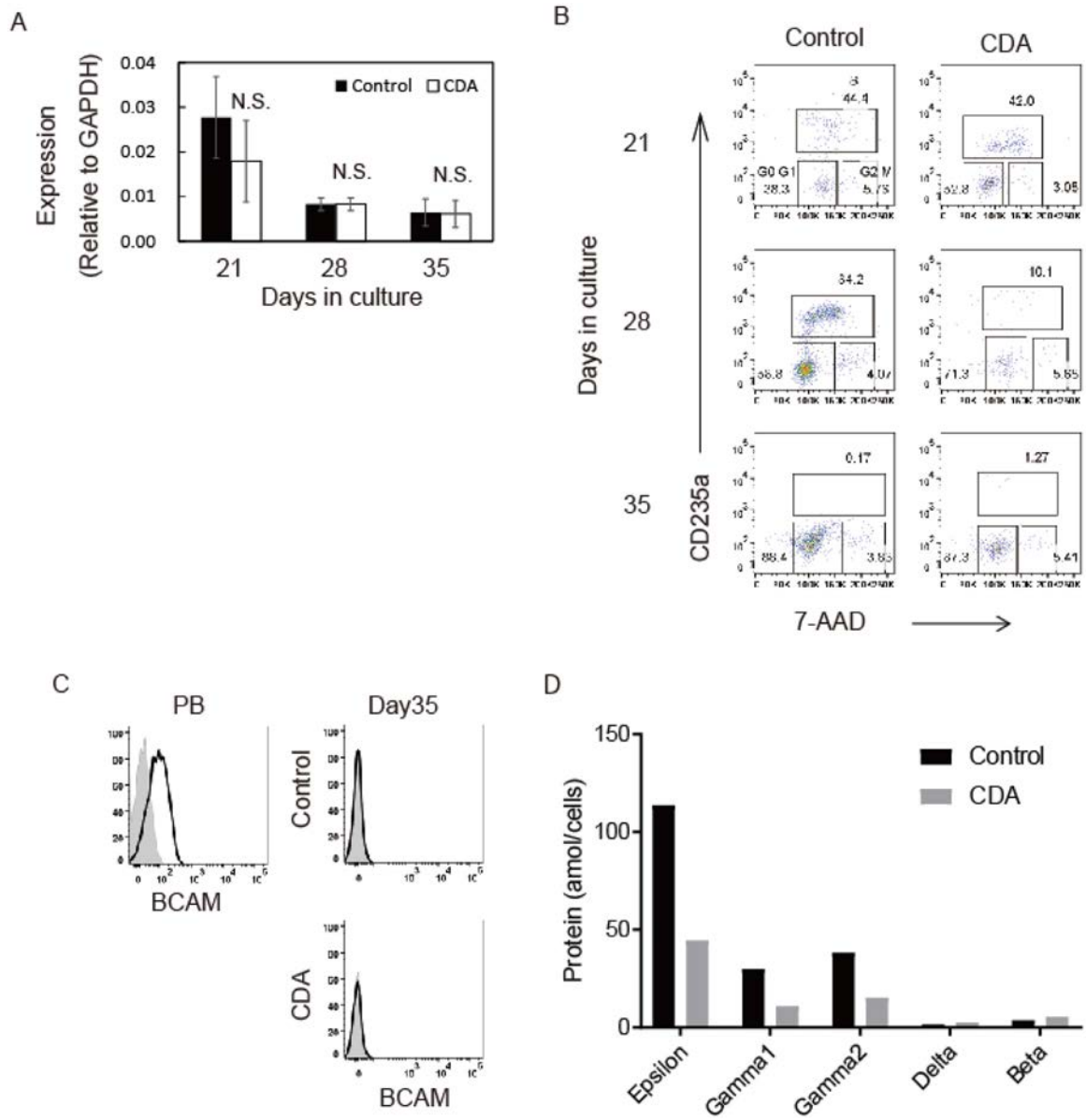

Figure S4. Erythropoiesis from CD34$^+$ cells derived from CDA-iPSCs. (A) qRT-PCR analysis of KLF1 in erythroblastic cells during erythropoietic liquid culture started from CD34$^+$ cells derived from CDA-iPSCs and control iPSCs. Probes designed to detect the boundary between exon 1 and exon 2 were used. Data are shown as the mean ± SD of three experimental replicates. (B) Representative flow cytometry dot plots of BrdU-labeled CD235a$^+$CD71$^+$ cells during erythropoietic liquid culture started from CD34$^+$ cells derived from CDA-iPSCs and control iPSCs. The numbers indicate the percentage of gated cells. (C) BCAM expression on peripheral blood erythrocytes of healthy donor and CD235a$^+$CD71$^+$ cells derived from control and CDA-iPSCs detected by flow cytometry. (D) Amount of globin proteins analyzed by mass spectrometry. Data are shown as the mean of two experimental replicates.

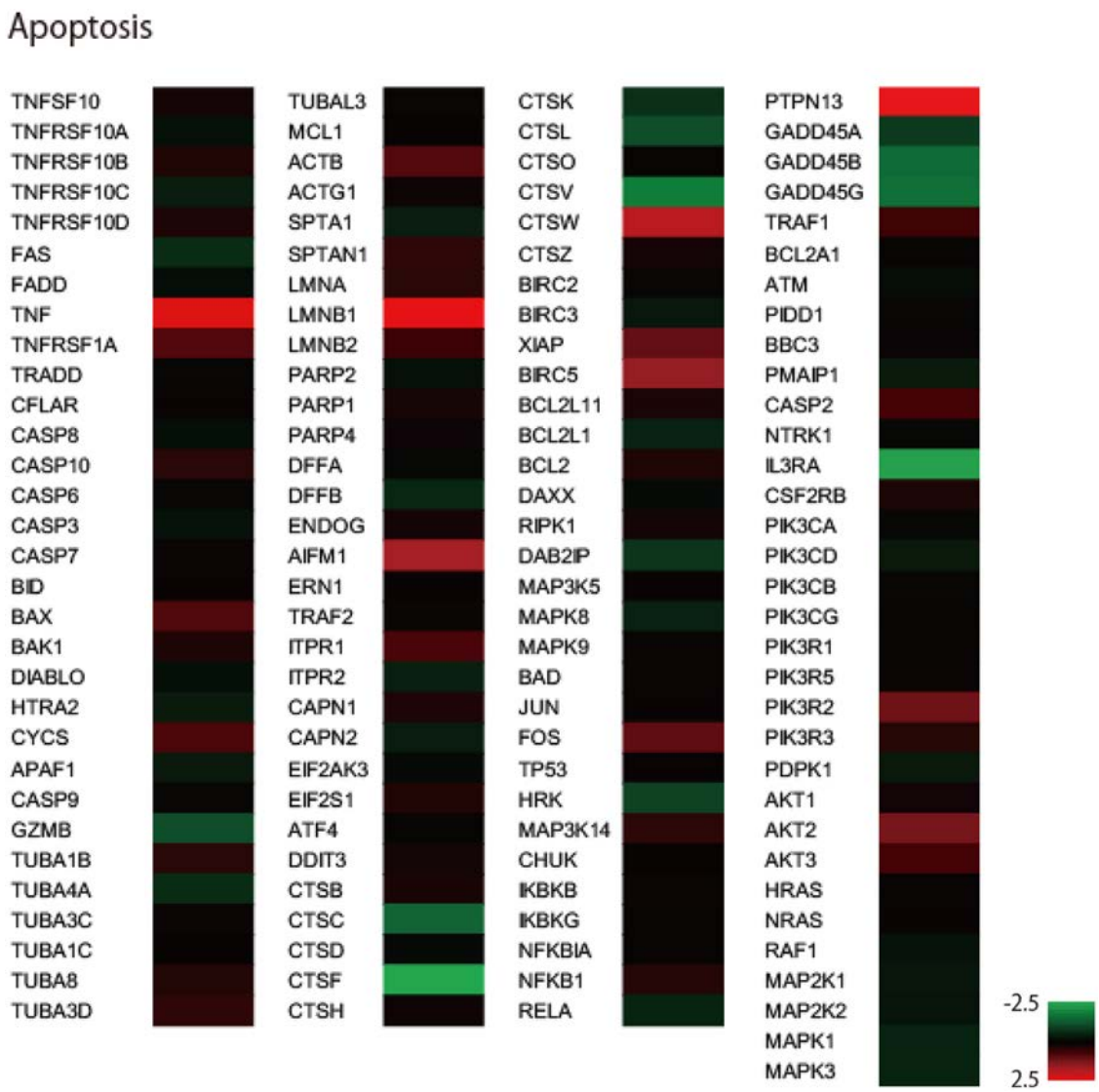

Figure S5. Apoptosis-related genes in CDA-iPSC-derived erythroblasts. The subset of KLF1 target genes expressed in cells under erythroid differentiation culture is indicated in the heat map.

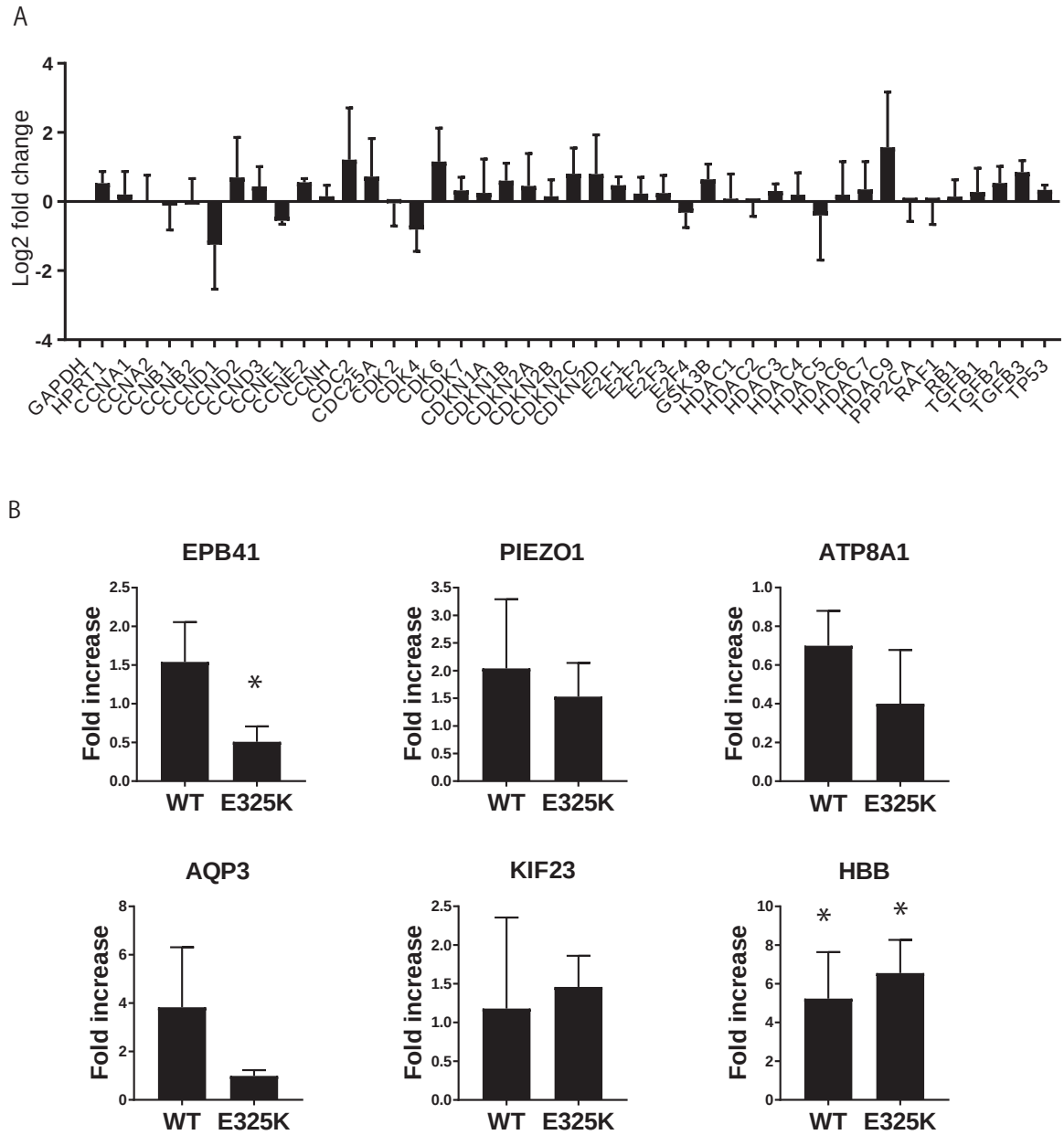

Figure S6. Disruption of transcriptional control by inducible expression of KLF1 E325K. (A) qRT-PCR analysis of cell cycle-related genes in erythroblastic cells cultured in the presence of 1,000 ng/ml Dox. Results were indicated as log2-fold change relative to controls withoud Dox. Data are shown as the mean ± SD of three experimental replicates. (B) qRT-PCR analysis of anemia-associated genes in erythroblastic cells cultured in the presence of 1,000 ng/ml Dox. Data are shown as the mean ± SD of three experimental replicates. * $P \leq 0.05$; ** $P \leq 0.02$ for controls without Dox.

Table S1. Subset of genes with KLF1 binding motifs

Gene ID	Gene Symbol	Z Score	CDA/Control Ratio	Control Signal	CDA Signal
53335	BCL11A	-1.94	0.34	413.10	139.32
358	AQP1	-2.00	0.43	3498.09	1516.72
960	CD44	-0.78	0.58	60.64	35.23
3386	ICAM4	0.16	1.05	17680.48	18575.66
1026	CDKN1A	0.15	1.05	2660.68	2802.51

Table S2. Subset of cell cycle regulators

Gene ID	Gene Symbol	Z Score	CDA/Control Ratio	Control Signal	CDA Signal
891	CCNB1	1.34	1.74	880.53	1,532.89
9133	CCNB2	1.59	1.93	1181.80	2,279.69
894	CCND2	1.57	1.91	1155.31	2,212.05
896	CCND3	1.23	1.67	1548.47	2,578.21
1869	E2F1	1.25	2.02	186.03	376.58
1870	E2F2	0.75	1.53	525.42	801.37
1028	CDKN1C	-4.43	0.22	8396.35	1,824.95
1029	CDKN2A	1.52	1.88	1329.70	2,494.10
1031	CDKN2C	1.52	1.88	545.90	1,025.54

Table S3. KLF1 target genes down-regulated

Gene ID	Gene Symbol	Z Score	CDA/Control Ratio	Control Signal	CDA Signal
7286	TUFT1	-2.98	0.36	4480.68	1602.45
9903	KLHL21	-2.01	0.51	20235.32	10261.10

Table S4. KLF1 target genes up-regulated

Gene ID	Gene Symbol	Z Score	CDA/Control Ratio	Control Signal	CDA Signal
340024	SLC6A19	2.15	3.34	101.34	338.00
1958	EGR1	2.38	2.27	1,915.72	4,344.69
7504	XK	2.73	2.55	3,348.91	8,555.67
7036	TFR2	2.93	5.19	115.10	597.15
55859	BEX1	3.05	2.85	5,044.25	14,397.65