A Dominant Mutation in the Gene Encoding the Erythroid Transcription Factor KLF1 Causes a Congenital Dyserythropoietic Anemia

Lionel Arnaud,[1,*] Carole Saison,[1] Virginie Helias,[1] Nicole Lucien,[1] Dominique Steschenko,[2] Marie-Catherine Giarratana,[3] Claude Prehu,[4] Bernard Foliguet,[5] Lory Montout,[1,6,7] Alexandre G. de Brevern,[1,6,7] Alain Francina,[8] Pierre Ripoche,[1,7] Odile Fenneteau,[9] Lydie Da Costa,[9,10,19] Thierry Peyrard,[1,11] Gail Coghlan,[12] Niels Illum,[13] Henrik Birgens,[14] Hannah Tamary,[15] Achille Iolascon,[16,17] Jean Delaunay,[18] Gil Tchernia,[10,20] and Jean-Pierre Cartron[1]

The congenital dyserythropoietic anemias (CDAs) are inherited red blood cell disorders whose hallmarks are ineffective erythropoiesis, hemolysis, and morphological abnormalities of erythroblasts in bone marrow. We have identified a missense mutation in *KLF1* of patients with a hitherto unclassified CDA. KLF1 is an erythroid transcription factor, and extensive studies in mouse models have shown that it plays a critical role in the expression of globin genes, but also in the expression of a wide spectrum of genes potentially essential for erythropoiesis. The unique features of this CDA confirm the key role of KLF1 during human erythroid differentiation. Furthermore, we show that the mutation has a dominant-negative effect on KLF1 transcriptional activity and unexpectedly abolishes the expression of the water channel AQP1 and the adhesion molecule CD44. Thus, the study of this disease-causing mutation in *KLF1* provides further insights into the roles of this transcription factor during erythropoiesis in humans.

The CDAs represent a heterogeneous group of rare congenital anemias predominantly caused by dyserythropoiesis in the bone marrow.[1] Three major (types I to III) and several minor subgroups have been differentiated, mainly according to the morphological abnormalities of erythroblast nuclei observed in bone marrow smears (e.g., chromatin bridges or double nuclei).[2,3] The gene responsible for CDA I (MIM 224120) was identified by positional cloning in 2002[4] and coined *CDAN1* (MIM 607465), but its function remains to be elucidated. The gene responsible for CDA II (MIM 224100) has recently been shown to encode SEC23B (MIM 610512), which was known to be involved in the vesicular transport between the endoplasmic reticulum and Golgi apparatus but whose erythroid-specific role was unsuspected.[5,6] Although CDA I and CDA II represent most cases, the identification of causative genetic defects in other CDA subgroups or in patients with unclassified CDA may offer further insights into the different pathways underlying erythropoiesis.

The first CDA patient investigated in this study (male patient ME) was born at 28 weeks of gestation to nonconsanguineous healthy parents in a context of acute fetal distress. Hydrops fetalis-associated anemia had been detected at 23 weeks of gestation and treated with two intrauterine transfusions; the karyotype of the fetus was normal. The neonatal examination revealed severe hyperbilirubinemia, hepatomegaly, hypertrophic cardiomyopathy, and several dysmorphic features (micropenis, hypospadia, large anterior fontanel, and hypertelorism). Anemia did not improve after birth and required transfusions. At 4 months of age, the analysis of bone marrow smears showed marked hyperplasia of the erythroid lineage, leading to a diagnosis of CDA, but the dysplastic changes in the erythroblasts did not clearly fit any classification of CDA (Figure S1). Despite treatments with erythropoietin or interferon-alpha, the hemolytic anemia persisted and required recurrent transfusions (at 2–3 week intervals) until a splenectomy was performed at 4 years of age (the enlarged spleen showed no pathologic features). Shortly thereafter, transfusion independence was achieved, and hemoglobin levels were stabilized at around 8.0 g/dl (Table S1). At 13 years of age, patient ME showed short

[1]National Institute of Blood Transfusion, 75015 Paris, France; [2]Hôpital d'Enfants de Brabois, Centre Hospitalier Universitaire de Nancy, 54511 Vandoeuvre-les-Nancy, France; [3]UMR_S 938, INSERM, Faculté de Médecine Pierre et Marie Curie, Université Pierre et Marie Curie, Paris 6, 75012 Paris, France; [4]Laboratoire de Biochimie et Génétique, Centre Hospitalier Universitaire Henri Mondor, 94010 Créteil, France; [5]Service commun de microscopie électronique, Faculté de Médecine, Université de Nancy, 54506 Vandoeuvre-les-Nancy, France; [6]Dynamique des Structures et Interactions des Macromolécules Biologiques, 75015 Paris, France; [7]UMR_S 665, INSERM, Institut National de la Transfusion Sanguine, Université Paris Diderot, Paris 7, 75015 Paris, France; [8]Hospices Civils de Lyon, Hôpital Edouard Herriot, Unité de Pathologie Moléculaire du Globule Rouge, 69003 Lyon, France; [9]Laboratoire d'Hématologie Biologique, Centre Hospitalier Universitaire Hôpital Robert Debré, 75019 Paris, France; [10]Laboratoire d'Hématologie, Hôpital de Bicêtre, 94275 Le Kremlin Bicêtre, France; [11]Centre National de Référence pour les Groupes Sanguins, 75011 Paris, France; [12]Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, MB R3A 1S1 Canada; [13]Hans Christian Andersen Children's Hospital, Odense University Hospital, 5000 Odense, Denmark; [14]Herlev Hospital, University of Copenhagen, 2730 Herlev, Denmark; [15]Pediatric Hematology Unit, Schneider Children's Medical Center of Israel, 49202 Petach Tikvah, Israel; [16]Department of Biochemistry and Medical Biotechnologies, University Federico II of Naples, 80131 Naples, Italy; [17]Centro Ingegneria Genetica, Biotecnologie Avanzate, 80145 Naples, Italy; [18]UMR_S 779, INSERM, Faculté de Médecine Paris-Sud, Université Paris-Sud, 94275 Le Kremlin-Bicêtre, France
[19]Present address: UMR 1009, INSERM, Institut de Cancérologie Gustave Roussy, Université Paris-Sud, Paris 11, 94805 Villejuif, France
[20]Present address: Ministère de la Santé, 75007 Paris, France
*Correspondence: larnaud@ints.fr
DOI 10.1016/j.ajhg.2010.10.010.

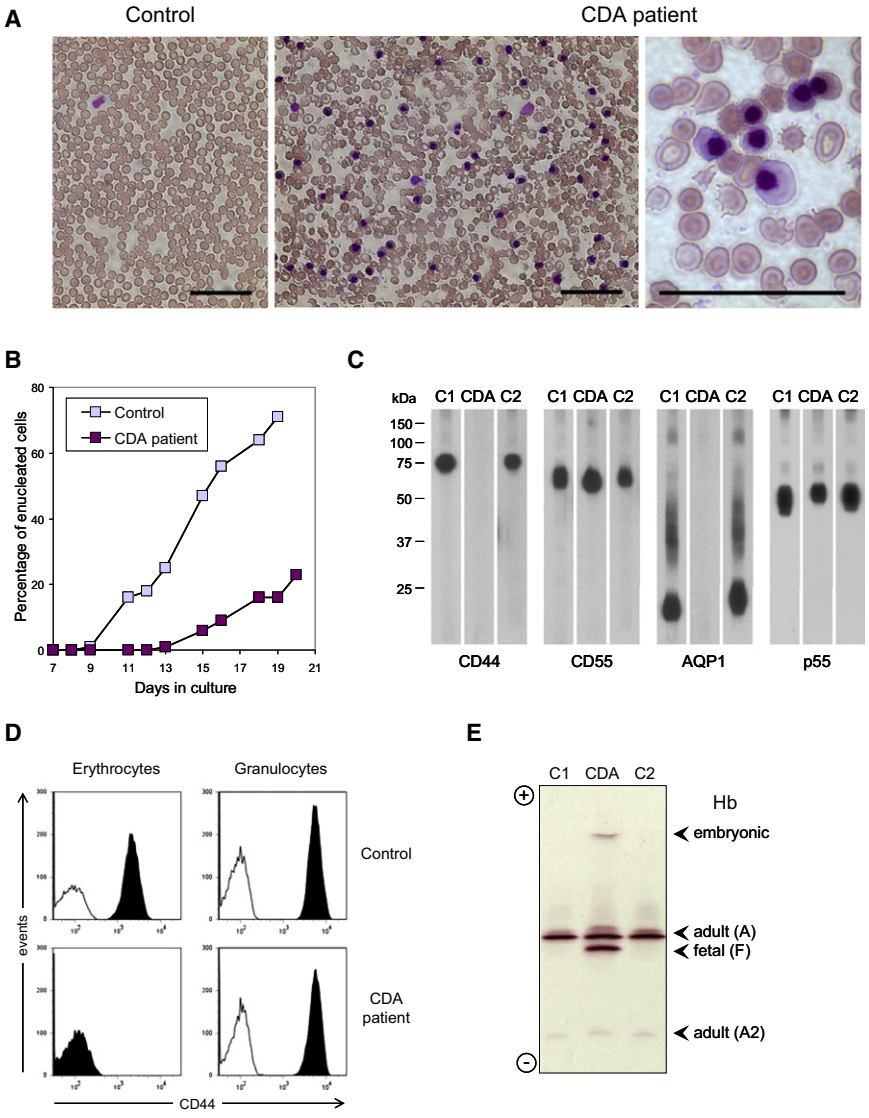


Figure 1. Analysis of the Peripheral Blood of CDA Patient ME Shows Unique Abnormalities

(A) Peripheral blood smears from the patient (right panels; sample taken on 10/28/2008) and a control (left panel; sample taken at the same time) stained with May-Grünwald Giemsa. Note the large number of circulating erythroblasts (purple nuclei) as well as poikilocytosis, anisocytosis and fragmented erythrocytes in the patient. The scale bars represent 40 μm.

(B) Study of the number of enucleated cells during in vitro erythroid differentiation of $CD34^+$ cells from the patient (purple; sample taken on 10/25/2005) or a control donor (blue). Note the markedly reduced enucleation capacity of the patient's cell culture.

(C) Immunoblot analysis of CD44, CD55, AQP1, and p55 in erythrocyte membrane lysates from the patient (CDA; sample taken on 6/7/2004), his healthy mother (C2), or a random control (C1). Note the combined deficiency of CD44 and AQP1 in the patient's erythrocytes.

(D) Flow-cytometry analysis of CD44 on mature erythrocytes (left panels) and granulocytes (right panels) from the patient (bottom panels; sample taken on 1/5/2010) and a control (top panels; sample taken at the same time). Note the erythroid-specific deficiency of CD44 in the patient. For the analysis of erythrocytes, whole blood samples were costained with fluorochrome-conjugated anti-CD44 (black histogram) or isotype control antibody (white histogram) and anti-CD71, and the mature erythrocytes were gated on FSS, SCC, and $CD71^-$ so that the reticulocytes and erythroblasts would be eliminated; for the analysis of granulocytes, nucleated blood cells were first isolated by hypotonic erythrocyte lysis, then costained as above and directly gated on FSC and SSC.

(E) Isoelectric focusing analysis of the different hemoglobin (Hb) variants in the patient (CDA; sample taken on 10/28/2008) and two controls (C1 and C2). Note the atypical globin expression in the patient; he exhibited very high levels of fetal Hb ($\alpha_2\gamma_2$ tetramer, 37.3%; normal range, less that 1%) as well as embryonic Hb Portland ($\zeta_2\gamma_2$ tetramer, 2.9%; normal range, absent) as ascertained by reverse-phase liquid chromatography; adult HbA and HbA2 were at 55.5% ($\alpha_2\beta_2$ tetramer, normal range: 90%–100%) and 1.2% ($\alpha_2\delta_2$ tetramer, normal range: 2%–3%), respectively. Extensive sequencing of patient ME's globin loci detected no gross abnormalities but a heterozygous mutation in the α1-globin gene (c.62_63insT, p.His21fsX36), which could not be responsible for the profound β-globin locus dysregulation and was indeed present in his healthy paternal aunt. Of note, patient SF was a carrier of a 4 bp deletion in the promoter of $^A\gamma$-globin gene, as was her healthy father.[14]

All analyses presented herein were performed on blood samples taken from patient ME after splenectomy and at least 6 months after transfusion.

stature (height −3 SD, weight −2 SD) despite growth-hormone therapy and treatment for hypothyroidism and thalassemic facies.

A striking feature of patient ME's CDA was the very large number of nucleated red blood cells in his peripheral blood (there were 210% the number of white blood cells before splenectomy and up to 1,000% thereafter; Figure 1A and Table S1). Most of these circulating nucleated red blood cells were orthochromatic erythroblasts, but only a few of them were enucleating, which suggested a failure of terminal erythroid differentiation. Analysis of these cells by electron microscopy revealed various ultrastructural abnormalities, especially atypical cytoplasmic inclusions and enlarged nuclear pores (Figure S2). The in vitro study of erythroid differentiation of $CD34^+$ cells[7] isolated from patient ME's peripheral blood showed normal proliferation and differentiation but impaired enucleation capacity (Figure 1B). Furthermore, when we analyzed a panel of markers on the surface of his erythrocytes by flow cytometry (Figure S3), we noticed the absence of CD44, which was confirmed by immunoblot analysis (Figure 1C), as well as reduced expression of two other adhesion

molecules, BCAM and ICAM4. CD44 was similarly absent from his mature erythrocytes and circulating erythroblasts, but it was present on his granulocytes and all his other leukocyte populations (Figure 1D and Figure S4), suggesting that only the erythroid lineage was affected, consistent with a CDA trait. We also found that patient ME's erythrocytes were deficient in the water channel AQP1 (Figure 1C) and, consequently, had a reduced water permeability similar to that of erythrocytes in the very rare *AQP1*$^{-/-}$ individuals[8] (Figure S5). Patient ME's CDA was unique and certainly different from CDA I and CDA II, as suggested by bone marrow analysis and later confirmed by the absence of mutations in *CDAN1* and *SEC23B* (data not shown).

In order to identify the causative genetic defect, we took fresh blood samples from patient ME and his relatives after obtaining a signed informed consent under an institutional-review-board-approved protocol, and we extracted genomic DNA. After unsuccessfully exploring the possibility of an inherited recessive mutation by performing a SNP-based genome-wide screen in patient ME's family (Affymetrix GeneChip Human Mapping 250K Nsp Array, data not shown), we thought that this unique CDA might be caused by a de novo mutation in a transcription factor essential for expression of *CD44* (MIM 107269) and *AQP1* (MIM 107776), among others. Because CD44 deficiency was apparently restricted to the erythroid lineage, we first focused our analysis on erythroid transcription factors such as GATA1 and KLF1 (also known as EKLF).[9,10] GATA1 represented an attractive candidate,[11,12] but no mutations in *GATA1* (MIM 305371) were detected (data not shown). Sequencing of *KLF1* (MIM 600599) (Table S2) in patient ME revealed the presence of two heterozygous mutations: a T-to-C transition in exon 2 (NM_006563.3:c.304T>C, NP_006554.1:p.Ser102Pro) and a G-to-A transition in exon 3 (NM_006563.3:c.973G>A, NP_006554.1:p.Glu325Lys) (Figure S6). A comparison with the NCBI dbSNP database (build 131) showed that c.304T>C mutation corresponded to a previously reported SNP (rs2072597) and was unlikely to be pathogenic; this was consistent with rs2072597 heterozygosity of his healthy mother (data not shown). In contrast, c.973G>A mutation had never been reported and was not present in 96 regular blood donors or in patient ME's relatives (Figure 2A and data not shown), suggesting it was the disease-causing mutation.

To verify that the *KLF1* mutation c.973G>A was responsible for this unique CDA, we searched for it in other patients. Female patient SF was extensively studied in the 90s, and her atypical CDA showed striking similarities with that of patient ME: combined deficiency of CD44 and AQP1 on erythrocytes,[13] circulating erythroblasts,[14] unique ultrastructural abnormalities in erythroblasts,[15] and increased electrophoretic mobility of the erythrocyte membrane protein Band 3[13] (Figure S7). When we sequenced *KLF1* in patient SF, we found the same heterozygous *KLF1* mutation, c.973G>A, as in patient ME (Figure 2A), verifying that this mutation was responsible for this type of CDA. Of note, and consistent with a de novo mutation, analysis of *KLF1* haplotypes of patients ME and SF does not support the hypothesis of a founder effect for the pathogenic *KLF1* mutation c.973G>A. In addition to the previously mentioned unique features of this CDA, patient SF was shown to express high levels of fetal hemoglobin, along with embryonic ζ-globin chain, in the majority of her erythrocytes.[14] When we analyzed patient ME's hemoglobin by isoelectric focusing (Figure 1E), we similarly found large amounts of fetal hemoglobin as well as an unusual hemoglobin, migrating like embryonic hemoglobin Portland ($\zeta_2\gamma_2$ tetramer), which was confirmed by reverse-phase liquid chromatography (Figure S8). Thus, the *KLF1* mutation c.973G>A was associated with a profound dysregulation of globin gene expression and provided in vivo evidence of the critical role of KLF1 in this complex transcriptional regulation in humans, as predicted by thalassemia-associated mutations in the proximal KLF1 binding sites of β-globin gene[16] and extensive studies in mouse models.[17–19] Interestingly, while this manuscript was being finalized, Borg et al.[20] described a large Maltese family in which a defective KLF1 allele segregates with the persistence of high levels of fetal hemoglobin in adults—a benign and usually asymptomatic condition known to alleviate the severity of β-thalassemia and sickle cell disease—and they further demonstrated that KLF1 indirectly downregulates fetal globin gene expression by activating *BCL11A* expression.

The pathogenic *KLF1* mutation c.973G>A results in the substitution of the evolutionarily conserved glutamate 325 by a lysine (E325K) in the second zinc finger (ZF2) of KLF1 (Figure 2B). As with arginine 322 and arginine 328, glutamate 325 is predicted to contact DNA[16] (Figure 2C). In order to investigate the possible structural effect of the E325K variant, we built a new structural model for the zinc-finger domain of KLF1 on the basis of the X-ray structures of Wilms' tumor protein in complex with DNA.[21] The currently used models of KLF1 are based on the structure of Zif268 bound to DNA,[22] but alignment of their respective zinc-finger domains requires a 2 amino acid gap in KLF1 ZF2, which significantly changes the local topology. According to the new structural model, glutamate 325 is located on a helix and is directed toward the DNA backbone (Figure 2D, upper left panel). The single amino acid variant E325K does not alter the overall topology of the KLF1 zinc-finger domain but mainly affects the side chain of residue 325 (Figure 2D, lower panels). Actually, this charge-reversal variant enhances the electrostatic interaction between KLF1 and the DNA and potentially creates a novel hydrogen bond (Figure 2D, upper right panel). Thus, the E325K variant is predicted to stabilize, rather than disrupt, the binding of KLF1 to its DNA target sequences.

Before testing the potential effect of the E325K variant on KLF1 transcriptional activity, we first wanted to check whether it affected the stability or cellular localization of the variant protein. For this purpose, we constructed vectors encoding either wild-type or E325K variant KLF1

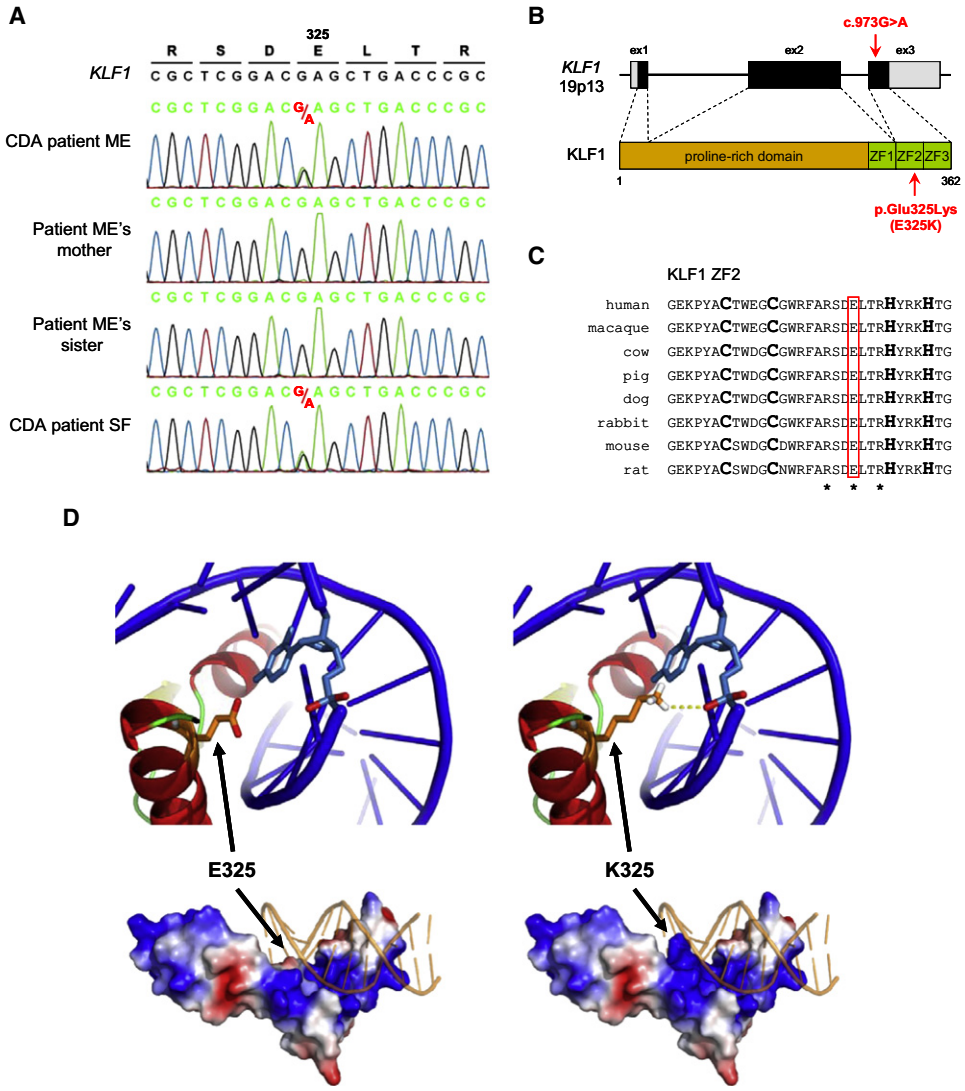


Figure 2. Identification of the *KLF1* Mutation Associated with this CDA and Structural Analysis of the Variant Transcription Factor
(A) A detail of *KLF1* sequencing in patient ME, his unaffected mother and sister, and unrelated patient SF show the same heterozygous *KLF1* mutation in both CDA patients. The experimental sequences were aligned with the NCBI reference sequence of *KLF1* (NG_013087.1). Genomic DNA samples from patient ME's father and patient SF's parents were not available.
(B) A diagram shows the structure of *KLF1* (based on NG_013087.1) and its products and highlights the localization of the mutation c.973G>A, p.Glu325Lys (E325K) was found in patients ME and SF. *KLF1* consists of three exons (boxes; black represents coding regions, and gray represents untranslated regions) and encodes a 362 amino acid peptide with a proline-rich domain at the amino-terminus (the transactivation domain is in brown) and three C_2H_2-type zinc fingers (ZF) at the carboxy-terminus (the DNA-binding domain is in green). The pathogenic mutation is located in the third exon and encodes a single amino acid change in the second zinc finger of the transcription factor.
(C) A sequence alignment of the second zinc finger of KLF1 from various mammalian species shows the conservation of a glutamate at amino acid position 325 (red box). The two cysteines and the two histidines contacting Zn^{2+} are indicated in bold, and the three conserved residues contacting the DNA are indicated by stars.
(D) A modeling structure of wild-type (left) and variant E325K (right) zinc-finger domain of KLF1 shows the enhanced electrostatic interaction between the E325K variant and the DNA backbone. The change of glutamate to lysine at position 325 exerts a double effect by reversing the side-chain charge from negative to positive and extending the side-chain length toward the negatively charged DNA backbone; the putative hydrogen bond created by the E325K variant is shown as a yellow dashed line. The bottom panels show the overall structure of the protein-DNA complexes, as well as the electrostatic potential surfaces (negative is in red, positive is in blue) of wild-type and variant proteins. The top panels show a close-up view of a cartoon representation of the region around residue 325, highlighting as sticks the side chain of residue 325 (orange) and the closest nucleotide (light blue); of note, residue 325 is not oriented toward the nucleotide base but toward the phosphate.

tagged with a Flag epitope at the N terminus to allow its detection and then transfected them into human erythroid K-562 cells. Flag-tagged KLF1 E325K had the same expression level as the wild-type protein by immunoblot analysis (Figure 3A) and showed the same nuclear localization by immunofluorescence analysis (Figure 3B). The combined CD44 and AQP1 deficiencies observed in this unique CDA suggested that *AQP1* and *CD44*

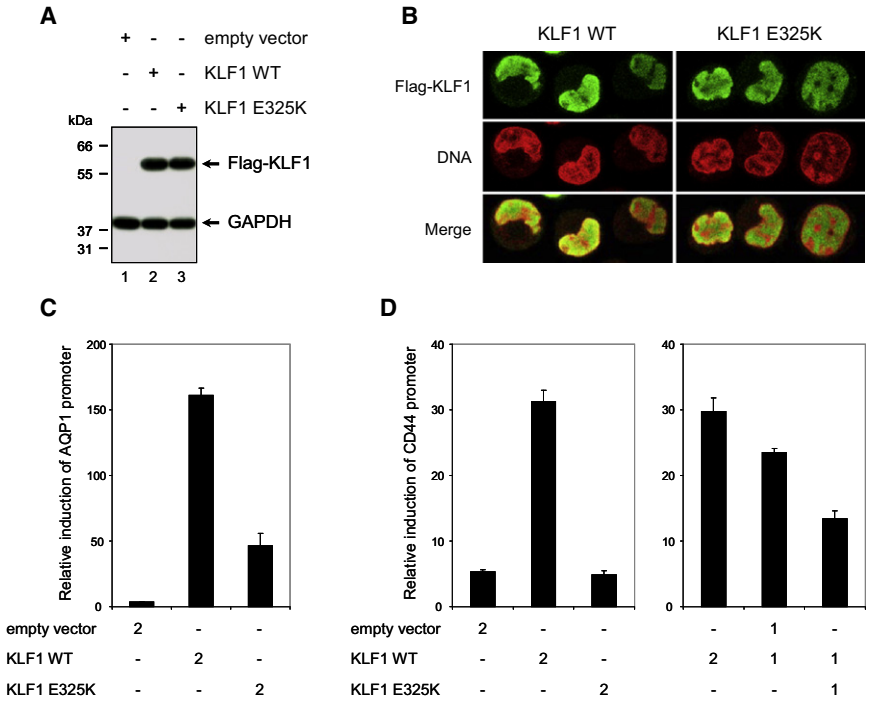


Figure 3. Characterization of the Effect of the E325K Variant on KLF1 Function

(A) Immunoblot analysis of transfected K-562 cells shows that the E325K variant does not affect the stability of KLF1. Constructs encoding Flag-tagged KLF1 wild-type (WT), variant E325K, or empty vector were transfected in K-562 cells, and protein expression was analyzed after 24 hr by immunoblot with anti-Flag and anti-GAPDH (loading control).

(B) Immunofluorescence analysis of transfected K-562 cells shows that the E325K variant does not affect the nuclear localization of KLF1. K-562 cells transfected as in (a) were analyzed after 24 hr by immunofluorescence with anti-Flag (green) along with propidium iodide for DNA staining (red). No anti-Flag staining was detected in K-562 cells transfected with the empty vector (data not shown).

(C) An *AQP1* promoter-reporter assay in K-562 cells shows that the E325K variant affects the transcriptional activity of KLF1. K-562 cells were cotransfected with an *AQP1* promoter-Photinus luciferase (Pluc) construct and a *HSV-TK* promoter-Renilla luciferase (Rluc) construct for normalization, along with 2 μg of the indicated *KLF1* constructs, and the luciferase activities were analyzed after 24 hr; the results are shown as means ± SD (n = 3) of Pluc activity normalized by Rluc activity.

(D) A *CD44* promoter-reporter assay in K-562 cells shows that the E325K variant has a dominant-negative effect on the transcription activity of KLF1. Not only does KLF1 E325K have a markedly reduced transcriptional activity, but it is also able to impede the transcriptional activity of coexpressed KLF1 wild-type. K-562 cells were cotransfected with a *CD44* promoter-Pluc construct and a *HSV-TK* promoter-Rluc construct for normalization, along with 2 or 1 μg of the *KLF1* constructs as indicated, and the luciferase activities were analyzed after 24 hr; the results are shown as means ± SD (n = 3) of Pluc activity normalized by Rluc activity.

were direct targets of KLF1, which was consistent with the presence of several potential KLF1 binding sites (CCNCNCCCN)[16] upstream of their respective initiating codons. Therefore, we studied the effect of the E235K variant on KLF1 transcriptional activity with *CD44* and *AQP1* promoter-reporter assays in K-562 cells. Flag-tagged KLF1 wild-type was able to activate *CD44* and *AQP1* promoters, whereas KLF1 E235K showed markedly reduced transcriptional activity (Figures 3C and 3D). Furthermore, when we coexpressed KLF1 E325K with KLF1 wild-type in order to mimic heterozygosity of patients ME and SF, we observed that KLF1 E325K was able to inhibit the activation of *CD44* promoter induced by KLF1 wild-type (Figure 3D). On the basis of these data, we conclude that the E325K variant has a dominant-negative effect on the transcriptional activity of KLF1, in total agreement with the phenotype of heterozygous patients.

This study describes a missense *KLF1* mutation responsible for a human pathology, a hitherto unclassified CDA. Interestingly, while this manuscript was being finalized, a new CDA patient with *KLF1* mutation c.973G>A was identified (female patient SE; A.I., unpublished data), suggesting that this type of CDA might be less rare than expected. The previously reported *KLF1* mutations, found in the heterozygous state too, are responsible for the blood-group phenotype In(Lu)[23] (MIM 111150), which is mainly characterized by a reduced expression of Lu (also known as BCAM) on erythrocytes but which is not associated with any pathology. Of note, patient SF does not show a reduced expression of Lu (Figure S9), indicating that the *KLF1* mutation c.973G>A is not directly responsible for the reduced Lu expression observed in patient ME (Figure S3). We assume that distinct *KLF1* mutations might differently affect the gene repertoire of this transcription factor and thus lead to different phenotypes, as observed with *GATA1* mutations.[11,12] Fortuitously, while this manuscript was being finalized, Siatecka et al.[24] reported that the mouse mutation *Nan*, a dominant ethylnitrosourea-induced mutation causing hemolytic anemia, corresponds to a missense *Klf1* mutation encoding the E339D variant and affecting the expression of only a subset of *Klf1* target genes, and they further showed that the E339D variant indeed alters Klf1 DNA binding to only a subset of its target sequences. Of note, glutamate 339 in mouse Klf1 is the equivalent of glutamate 325 in human KLF1; however, *Klf1* variant E339D is intrinsically different from *KLF1* variant E325K because of the opposite charge of the variant residue (see structure prediction above), which is consistent with the distinct resulting pathologies in mice and humans. Extensive studies with mouse models suggested that KLF1 played a global role during erythropoiesis by regulating a wide spectrum of genes[19,25,26] but also

organizing nuclear hubs for efficient and coordinated transcription of genes coregulated with the β-globin gene.[27,28] This study provides in vivo evidence that human KLF1 plays a critical role in the regulation of fetal globin genes, as recently demonstrated by Borg et al.[20] (see above), but also of other unexpected genes such as *AQP1* and *CD44*. Although AQP1 deficiency in humans doesn't cause dyserythropoiesis,[8] we cannot exclude that CD44 deficiency contributes, even though *CD44* is dispensable for mouse erythropoiesis.[29] Future studies will need to identify the KLF1 target genes whose expression is essential for human erythropoiesis but that are deficient in this form of CDA. Finally, we propose that *KLF1* should be systematically sequenced as a novel candidate gene in all CDA cases with unknown genetic cause; such sequencing might eventually lead to the discovery of other pathogenic *KLF1* mutations.

Supplemental Data

Supplemental Data include nine figures and two tables and can be found with this article online at http://www.cell.com/AJHG.

Acknowledgments

First, we would like to thank the patients and patient ME's relatives for providing blood samples for this study. We would like also to thank D. Sommelet, J. Buisine, L. Douay, M. Goossens, C. Etchebest, Y. Colin, T. Zelinski, P. Gane, T. Cynober, M. Feneant-Thibault, C. Schmitt, R. Russo, M.R. Esposito, C. Menanteau, S. Kappler-Gratias, B.-N. Pham, P.-Y. Le Pennec, E. Vera, C. Verheyde, G. Nicolas, B.A. Ballif, and M. Le Gall for their contributions to this study. L.A. and J.-P.C. were supported by the National Institute of Blood Transfusion (INTS), the National Institute for Health and Medical Research (INSERM), and Paris Diderot University (Paris 7). G.C. was supported by an operating grant from the Winnipeg Rh Institute Foundation to T. Zelinski. H.T. was supported by an Israel Science Foundation grant (No. 699/_03-18.4), by an Israeli Ministry of Science, Culture, and Sport grant in the framework of the Israel-France Program, and by an Israeli Ministry of Science-Eshkol Fellowship. A.I. was supported by the Italian Ministry of University and Research (grant MUR-PS 35-126/Ind), the Italian Telethon Foundation (project GGP09044), and Regione Campania (DGRC 1901/2009).

Received: August 13, 2010
Revised: October 7, 2010
Accepted: October 13, 2010
Published online: November 4, 2010

Web Resources

The URLs for data presented herein are as follows:

Entrez Gene browser, http://www.ncbi.nlm.nih.gov/gene/
Entrez Single Nucleotide Polymorphism (SNP) browser, http://www.ncbi.nlm.nih.gov/snp/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/omim/
Structure modeling of KLF1, http://www.dsimb.inserm.fr/~debrevern/KLF1

References

1. Renella, R., and Wood, W.G. (2009). The congenital dyserythropoietic anemias. Hematol. Oncol. Clin. North Am. *23*, 283–306.
2. Heimpel, H., Wendt, F., Klemm, D., Schubothe, H., and Heilmeyer, L. (1968). Congenital dyserythropoietic anemia. Arch. Klin. Med. *215*, 174–194.
3. Heimpel, H., Kellermann, K., Neuschwander, N., Högel, J., and Schwarz, K. (2010). The morphological diagnosis of congenital dyserythropoietic anemia: Results of a quantitative analysis of peripheral blood and bone marrow cells. Haematologica *95*, 1034–1036.
4. Dgany, O., Avidan, N., Delaunay, J., Krasnov, T., Shalmon, L., Shalev, H., Eidelitz-Markus, T., Kapelushnik, J., Cattan, D., Pariente, A., et al. (2002). Congenital dyserythropoietic anemia type I is caused by mutations in codanin-1. Am. J. Hum. Genet. *71*, 1467–1474.
5. Schwarz, K., Iolascon, A., Verissimo, F., Trede, N.S., Horsley, W., Chen, W., Paw, B.H., Hopfner, K.P., Holzmann, K., Russo, R., et al. (2009). Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. Nat. Genet. *41*, 936–940.
6. Bianchi, P., Fermo, E., Vercellati, C., Boschetti, C., Barcellini, W., Iurlo, A., Marcello, A.P., Righetti, P.G., and Zanella, A. (2009). Congenital dyserythropoietic anemia type II (CDAII) is caused by mutations in the SEC23B gene. Hum. Mutat. *30*, 1292–1298.
7. Giarratana, M.C., Kobari, L., Lapillonne, H., Chalmers, D., Kiger, L., Cynober, T., Marden, M.C., Wajcman, H., and Douay, L. (2005). Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. Nat. Biotechnol. *23*, 69–74.
8. Preston, G.M., Smith, B.L., Zeidel, M.L., Moulds, J.J., and Agre, P. (1994). Mutations in aquaporin-1 in phenotypically normal humans without functional CHIP water channels. Science *265*, 1585–1587.
9. Kim, S.I., and Bresnick, E.H. (2007). Transcriptional control of erythropoiesis: Emerging mechanisms and principles. Oncogene *26*, 6777–6794.
10. Sankaran, V.G., Xu, J., and Orkin, S.H. (2010). Advances in the understanding of haemoglobin switching. Br. J. Haematol. *149*, 181–194.
11. Hollanda, L.M., Lima, C.S., Cunha, A.F., Albuquerque, D.M., Vassallo, J., Ozelo, M.C., Joazeiro, P.P., Saad, S.T., and Costa, F.F. (2006). An inherited mutation leading to production of only the short isoform of GATA-1 is associated with impaired erythropoiesis. Nat. Genet. *38*, 807–812.
12. Nichols, K.E., Crispino, J.D., Poncz, M., White, J.G., Orkin, S.H., Maris, J.M., and Weiss, M.J. (2000). Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat. Genet. *24*, 266–270.
13. Parsons, S.F., Jones, J., Anstee, D.J., Judson, P.A., Gardner, B., Wiener, E., Poole, J., Illum, N., and Wickramasinghe, S.N. (1994). A novel form of congenital dyserythropoietic anemia associated with deficiency of erythroid CD44 and a unique blood group phenotype [In(a-b-), Co(a-b-)]. Blood *83*, 860–868.
14. Tang, W., Cai, S.P., Eng, B., Poon, M.C., Waye, J.S., Illum, N., and Chui, D.H. (1993). Expression of embryonic zeta-globin and epsilon-globin chains in a 10-year-old girl with congenital anemia. Blood *81*, 1636–1640.

15. Wickramasinghe, S.N., Illum, N., and Wimberley, P.D. (1991). Congenital dyserythropoietic anaemia with novel intra-erythroblastic and intra-erythrocytic inclusions. Br. J. Haematol. *79*, 322–330.
16. Feng, W.C., Southwood, C.M., and Bieker, J.J. (1994). Analyses of beta-thalassemia mutant DNA interactions with erythroid Krüppel-like factor (EKLF), an erythroid cell-specific transcription factor. J. Biol. Chem. *269*, 1493–1500.
17. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R., and Grosveld, F. (1995). Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. Nature *375*, 316–318.
18. Perkins, A.C., Sharpe, A.H., and Orkin, S.H. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature *375*, 318–322.
19. Drissen, R., von Lindern, M., Kolbus, A., Driegen, S., Steinlein, P., Beug, H., Grosveld, F., and Philipsen, S. (2005). The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. Mol. Cell. Biol. *25*, 5205–5214.
20. Borg, J., Papadopoulos, P., Georgitsi, M., Gutiérrez, L., Grech, G., Fanis, P., Phylactides, M., Verkerk, A.J., van der Spek, P.J., Scerri, C.A., et al. (2010). Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nat. Genet. *42*, 801–805.
21. Stoll, R., Lee, B.M., Debler, E.W., Laity, J.H., Wilson, I.A., Dyson, H.J., and Wright, P.E. (2007). Structure of the Wilms tumor suppressor protein zinc finger domain bound to DNA. J. Mol. Biol. *372*, 1227–1245.
22. Elrod-Erickson, M., Benson, T.E., and Pabo, C.O. (1998). High-resolution structures of variant Zif268-DNA complexes: Implications for understanding zinc finger-DNA recognition. Structure *6*, 451–464.
23. Singleton, B.K., Burton, N.M., Green, C., Brady, R.L., and Anstee, D.J. (2008). Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. Blood *112*, 2081–2088.
24. Siatecka, M., Sahr, K.E., Andersen, S.G., Mezei, M., Bieker, J.J., and Peters, L.L. (2010). Severe anemia in the Nan mutant mouse caused by sequence-selective disruption of erythroid Kruppel-like factor. Proc. Natl. Acad. Sci. USA *107*, 15151–15156.
25. Hodge, D., Coghill, E., Keys, J., Maguire, T., Hartmann, B., McDowall, A., Weiss, M., Grimmond, S., and Perkins, A. (2006). A global role for EKLF in definitive and primitive erythropoiesis. Blood *107*, 3359–3370.
26. Tallack, M.R., Whitington, T., Yuen, W.S., Wainwright, E.N., Keys, J.R., Gardiner, B.B., Nourbakhsh, E., Cloonan, N., Grimmond, S.M., Bailey, T.L., and Perkins, A.C. (2010). A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. Genome Res. *20*, 1052–1063.
27. Patrinos, G.P., de Krom, M., de Boer, E., Langeveld, A., Imam, A.M., Strouboulis, J., de Laat, W., and Grosveld, F.G. (2004). Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. Genes Dev. *18*, 1495–1509.
28. Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., et al. (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat. Genet. *42*, 53–61.
29. Schmits, R., Filmus, J., Gerwin, N., Senaldi, G., Kiefer, F., Kundig, T., Wakeham, A., Shahinian, A., Catzavelos, C., Rak, J., et al. (1997). CD44 regulates hematopoietic progenitor distribution, granuloma formation, and tumorigenicity. Blood *90*, 2217–2233.

The American Journal of Human Genetics Volume *87*

Supplemental Data

A Dominant Mutation in the Gene Encoding the Erythroid Transcription Factor KLF1 Causes a Congenital Dyserythropoietic Anemia

Lionel Arnaud, Carole Saison, Virginie Helias, Nicole Lucien, Dominique Steschenko, Marie-Catherine Giarratana, Claude Prehu, Bernard Foliguet, Lory Montout, Alexandre G. de Brevern, Alain Francina, Pierre Ripoche, Odile Fenneteau, Lydie Da Costa, Thierry Peyrard, Gail Coghlan, Niels Illum, Henrik Birgens, Hannah Tamary, Achille Iolascon, Jean Delaunay, Gil Tchernia, and Jean-Pierre Cartron.

Figure S1. (A) Bone marrow smear from the CDA patient ME before splenectomy (sample taken on 1997/07/14) showing marked erythroid hyperplasia (77%) with a majority of acidophilic erythroblasts (58%). Binucleated erythroblasts (1.0%), multinucleated erythroblasts (1.5%) and few karyorrhexis (star) are observed. In some erythroblasts, the cytoplasm contains fine or coarse basophilic stippling of pale grayish-blue (arrow).

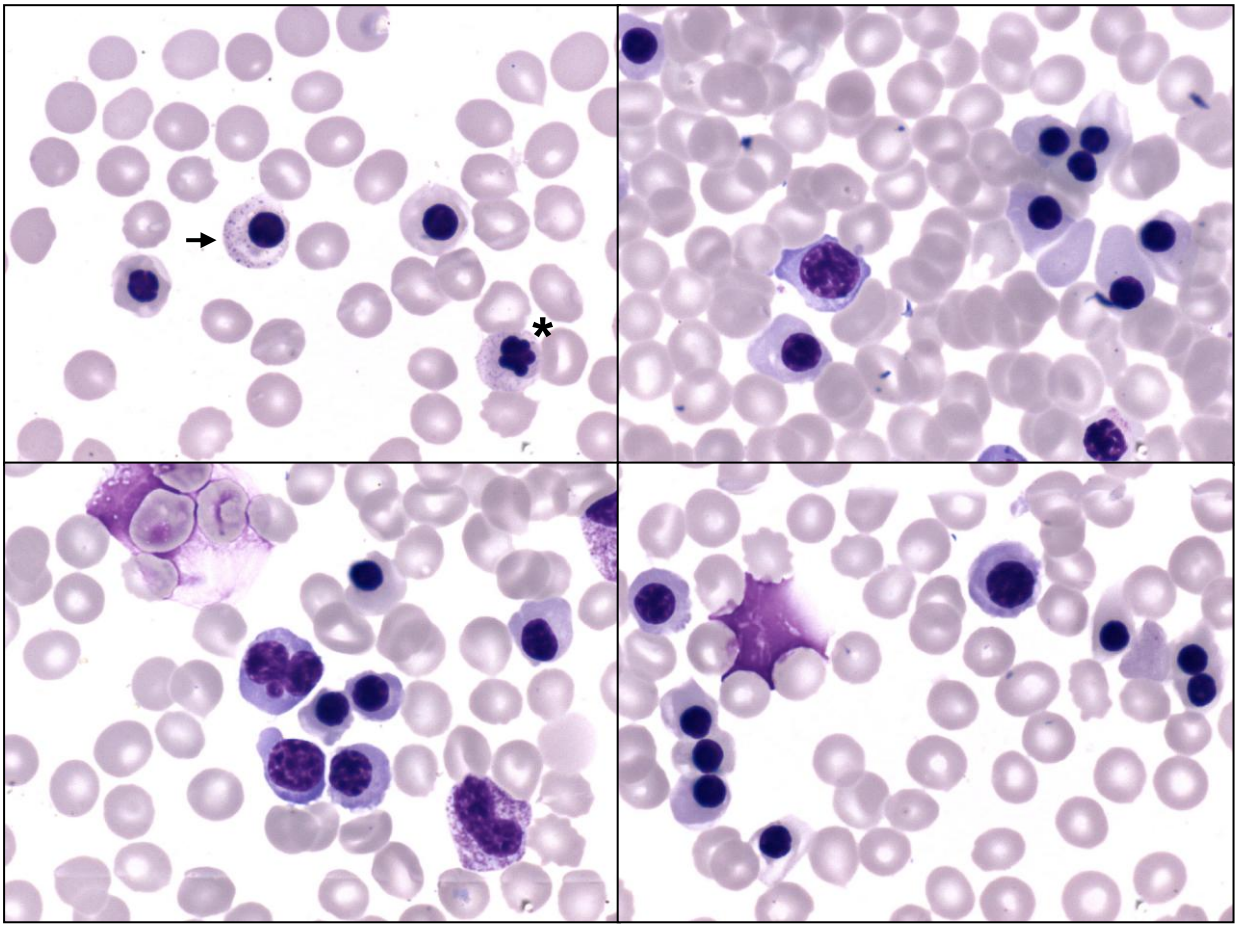


Figure S1. (B) Bone marrow smear from the CDA patient ME after splenectomy (sample taken on 2001/02/07) showing marked erythroid hyperplasia (80%) with a majority of acidophilic erythroblasts (55%). Binucleated erythroblasts (4.5%), multinucleated erythroblasts (2.5%) and few karyorrhexis (star) are observed.

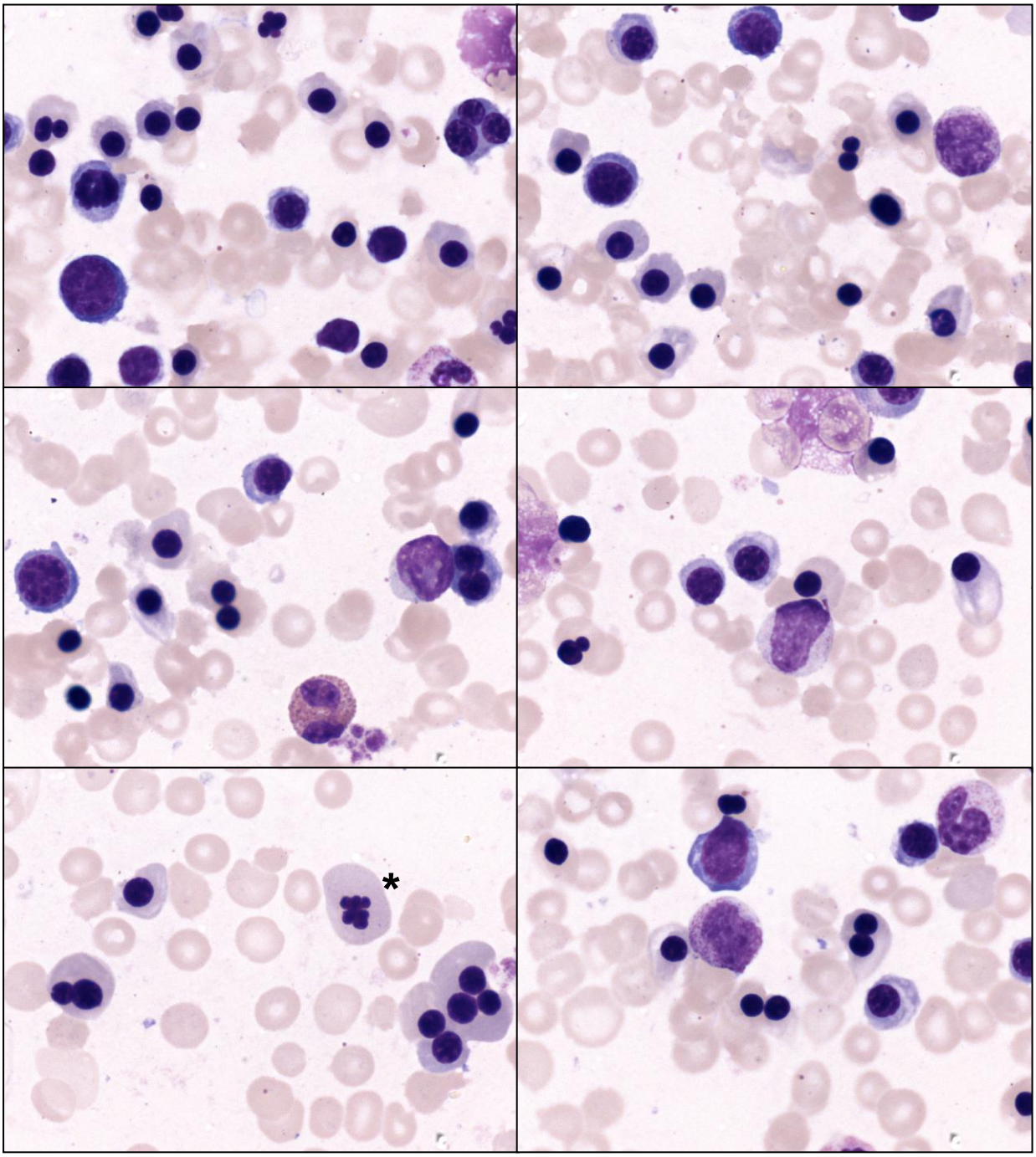


Figure S2. (A) Transmission electron microscopy of circulating erythroblasts of the CDA patient ME showing tubulosaccular inclusions (stars) and elongated cylindrical structures (black arrows) in the cytoplasm, and abnormally large nuclear pores (white arrows). Original magnification ×10,000.

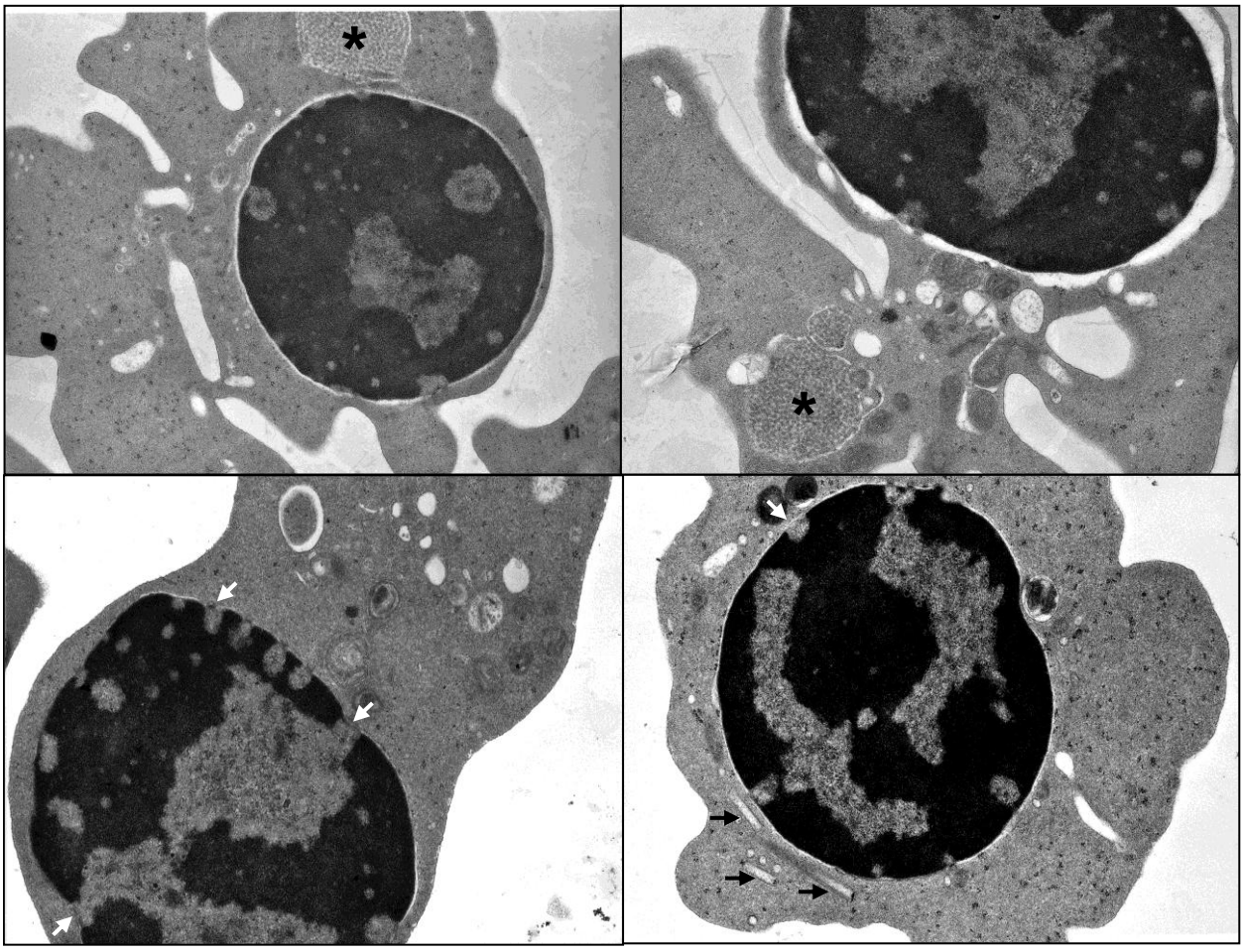


Figure S2. (B) Transmission electron microscopy of circulating erythroblasts of the CDA patient ME showing a wide area of endoreplication of the nuclear membrane. Original magnification ×8,000 (top panel) and ×35,000 (bottom panel).

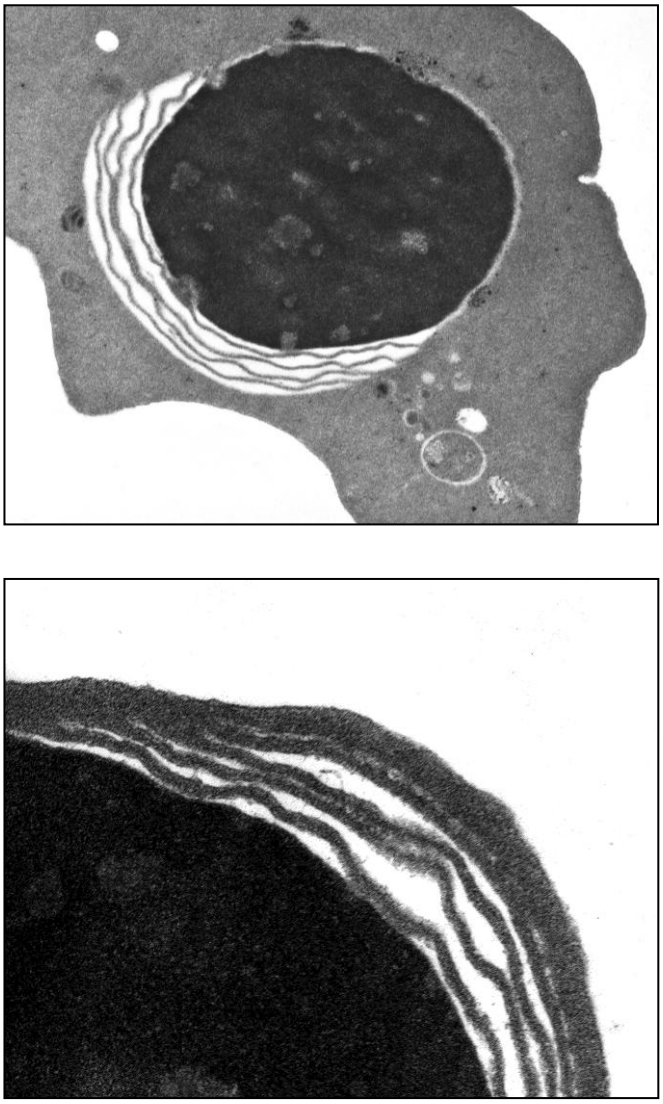


Figure S2. (C) Scanning electron microscopy of blood of the CDA patient ME showing severe anisopoikilocytosis (top panel, original magnification ×1,930) and small invaginations in erythrocyte membrane (stars, bottom panel, original magnification ×4,590). The scale bars represent 10 µm.

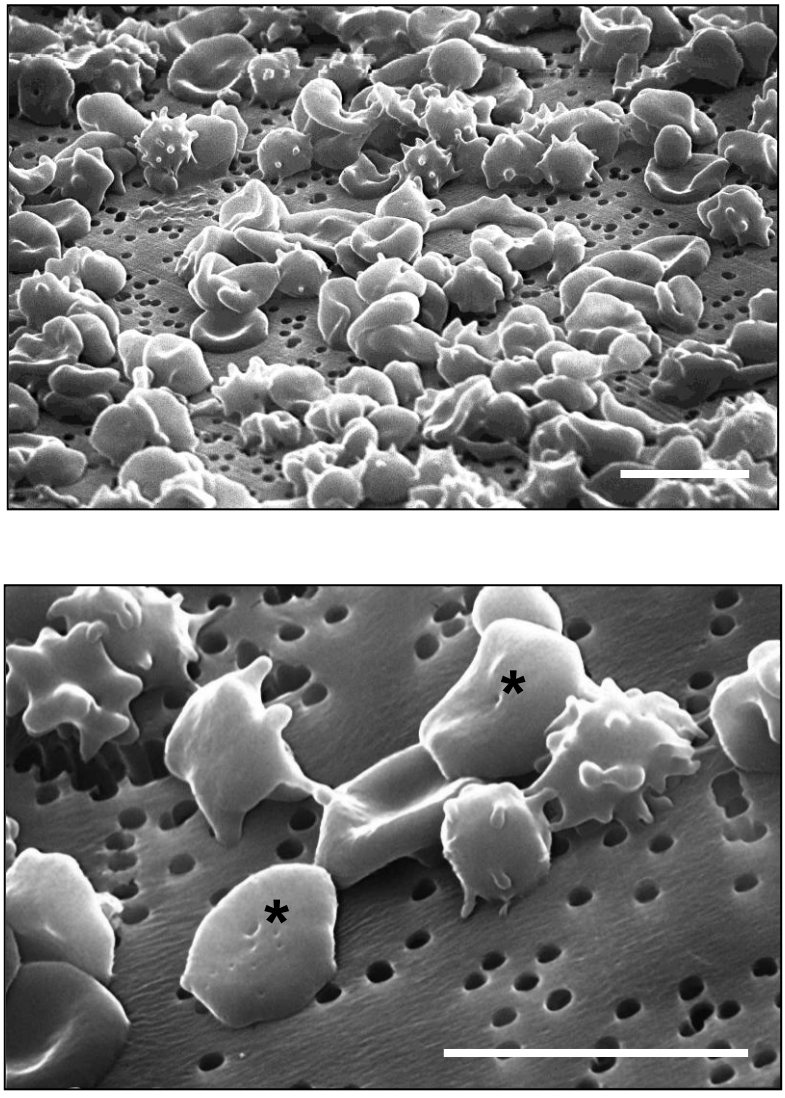


Figure S3. (A) Flow cytometry analysis of cell markers CD44, CD47, GYPA, BCAM/Lu, ICAM4/LW, CD29 and DARC on erythrocytes of the CDA patient ME (right histograms, sample taken on 2010/01/05) or his healthy mother (left histograms, sample taken at the same time). Erythrocytes were stained with mouse monoclonal antibodies (profiles in color) or an isotype control antibody (profile in black) and phycoerythrin-conjugated goat F(ab')2 anti-mouse IgG(H+L), and were gated on FSC and SSC (density dot plots). This blood sample from patient ME was taken 4 years after his most recent transfusion, and thus was free of exogenous erythrocytes. Patient ME's erythrocytes abnormally showed no expression of CD44, reduced expression of BCAM and ICAM4, and slightly increased expression of CD29. Similar results have consistently been observed on samples taken since 2003.

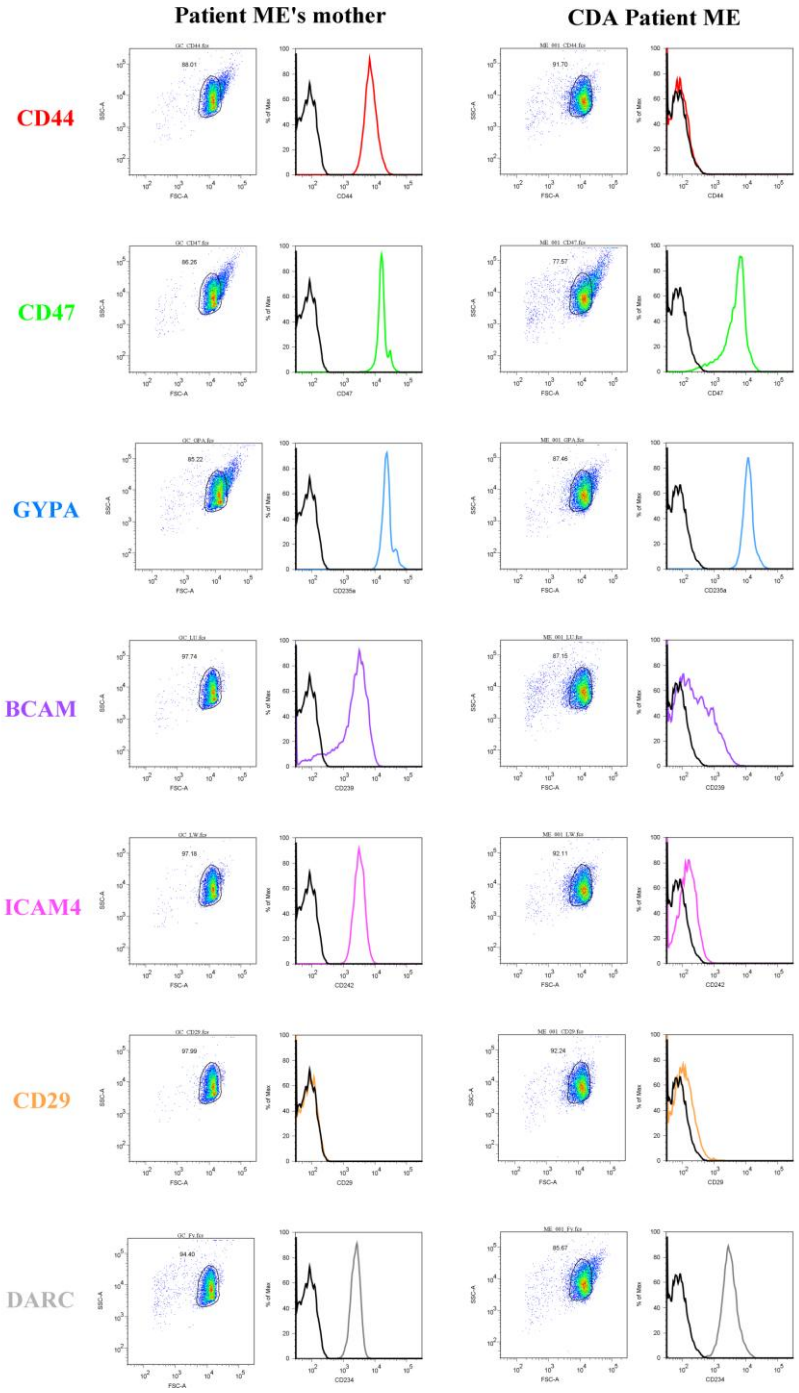


Figure S3. (B) Flow cytometry analysis of cell markers CD55, CD58, CD59, CD99, CD151, KEL and AChE on erythrocytes of the CDA patient ME (right histograms, sample taken on 2010/01/05) or his healthy mother (left histograms, sample taken at the same time). Erythrocytes were stained with mouse monoclonal antibodies (profiles in color) or an isotype control antibody (profile in black) and phycoerythrin-conjugated goat F(ab')2 anti-mouse IgG(H+L), and were gated on FSC and SSC (density dot plots). This blood sample from patient ME was taken 4 years after his most recent transfusion, and thus was free of exogenous erythrocytes. The expression of CD99 is highly variable and patient ME's profile is not abnormal.

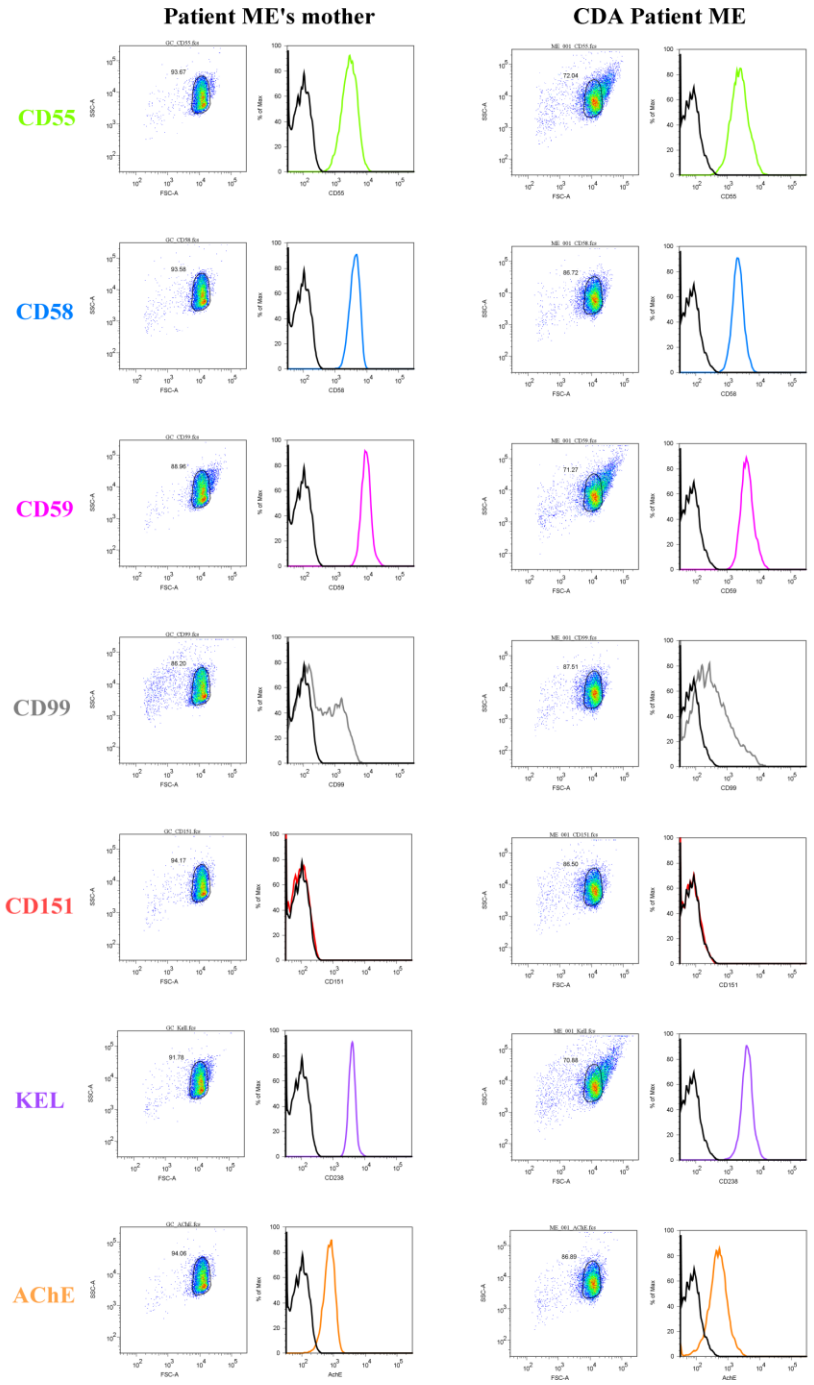


Figure S4. Flow cytometry analysis of CD44 expression in the nucleated cells present in the CDA patient ME's peripheral blood, showing absence of CD44 on his erythroblasts but presence on his leukocyte populations. Nucleated blood cells were isolated by hypotonic erythrocyte lysis and triple stained with anti-CD44-APC or isotype antibody-APC, anti-CD71-PE and anti-CD45-FITC; granulocytes (panel 1) and monocytes (panel 2) were directly gated on FSC and SSC (top left contour plot), while lymphocytes ($CD71^-/CD45^+$, panel 3) and erythroblasts ($CD71^+/CD45^-$, panel 4) were further gated on CD71 and CD45 expression (bottom left contour plot); the labeling corresponding to anti-CD44-APC (colored histograms) or isotype antibody-APC (grey histograms) was analyzed with the same voltage for all cell populations.

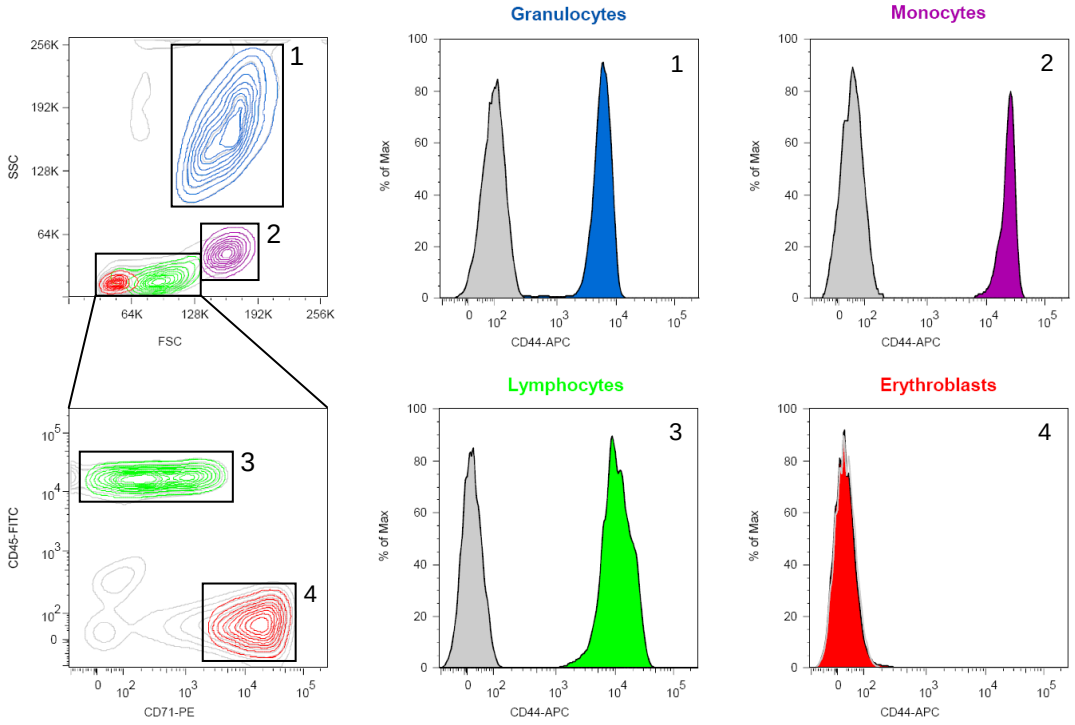


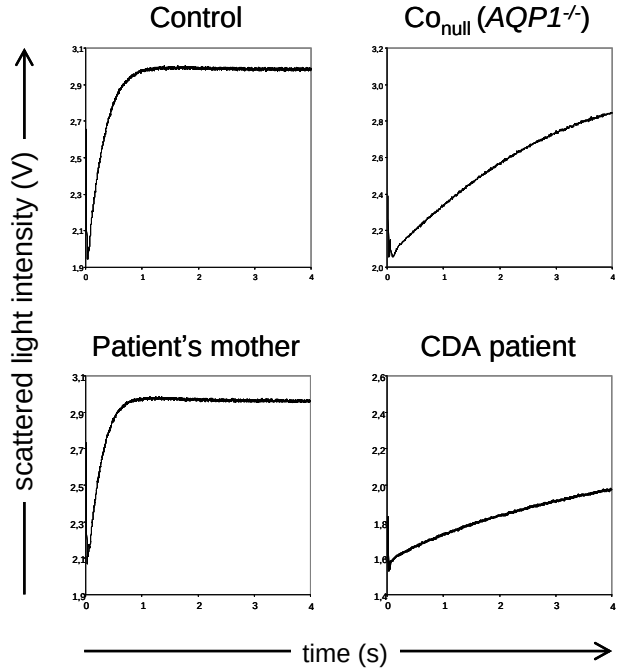
Figure S5. Analysis of the water permeability of the CDA patient ME's erythrocytes confirming his AQP1 deficiency.

(A) Stopped-flow light scattering measurements of erythrocytes from a control donor (top left panel), an *AQP1*$^{-/-}$ person (also known as Co_{null} blood group phenotype, top right panel), the CDA patient ME (bottom right panel) and his healthy mother (bottom left panel) in response to a 150 mOsm inwardly directed gradient of mannitol at 15°C; the rapid increase in light scattering results from the shrinkage of erythrocytes due to water efflux, allowing to determine their osmotic water permeability.

(B) Osmotic water permeability coefficients (*Pf*) of indicated erythrocytes. As the extremely rare *AQP1*$^{-/-}$ individuals, patient ME's erythrocytes showed severely reduced water permeability. Of note, AQP1 is a major erythrocyte membrane protein that carries the Colton blood group antigens as well as ABO blood group antigens.

Due to AQP1 deficiency of patient ME's erythrocytes, their Colton phenotype is expected to be Co(a-b-) while blood group genotyping (BioArraySolutions HEA BeadChip Kit) predicts the following extended phenotype: C+E–c+e+, K–k+, Kp(a–b+), Js(a–b+), Fy(a+b+), Jk(a+b–), M+N–S–s+, Lu(a–b+), Di(a–b+), Co(a+b–), Do(a–b+), Hy+Jo(a+), LW(a+b–), Sc:1,–2.

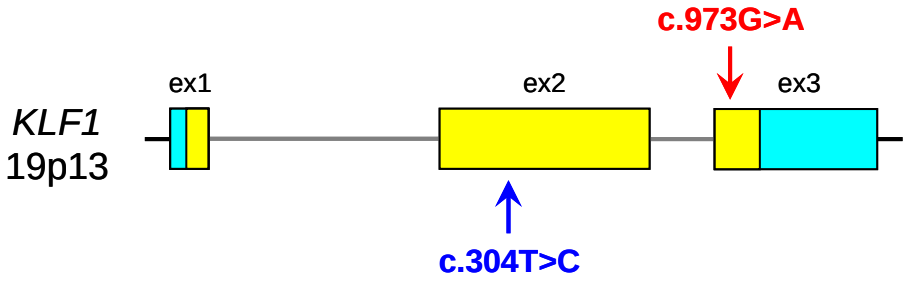
A



B

Samples	ABO type	*Pf* (cm.s^{-1}) at 15°C	+ $HgCl_2$
CDA patient	A	0.002485	0.001635
Patient's mother	A	0.036400	0.003250
Control	A	0.031300	0.003760
Co_{null} (*AQP1*$^{-/-}$)	O	0.004185	0.003930
Controls	O	0.030025	0.003441
(n=4)		± 0.004470	± 0.000343

Figure S6. (A) Experimental sequence data of *KLF1* from the CDA patient ME compared with genomic NCBI Reference Sequence NC_000019.8.



NNN coding sequence
NNN untranslated sequence
NNN intronic sequence

GCCGCCTTGCCTTGCTTTGCCTTATCAGAGGCTGCAGCCAATCAGCTAAGGACAGAGAGGAGCCCTCGAAGGGGCTATCACAG
CCTCAGAGTTCACGAGGCAGCCGAGGAAGAGGAGGCTTGAGGCCCAGGGTGGGCACCAGCCAGCCATGGCCACAGCCGAGACC
GCCTTGCCCTCCATCAGCACACTGACCGCCCTGGGCCCCTTCCCGGACACACAGGATGACTTCCTCAAGGTGGGGCCTAGAAG
GTGGGGTCTAGGTGGGCTGGCTGGAATCCAGGGCCACAGTCACAGATCTTGGGGTCCAGACCTGCATCTTGACCTGAAATCAA
GAGACTTAACCAGGACTGAGGTACGCTCAGTCCAGGAGAGGAGATCTCAGCTTAGTCTGGCAGGGGGTGAGGAGGGTGGTCTA
GGGGTTTGAGGTTCTAAGTGTGATCTATTTCGGTAATAGAAAACGAAGGTAGCCTGGGCAACATGGTGAAACCCTATCTCTAC
AAAAAATACCAAAAACATTAGGCCAGGCATGGGGGCGTGTGCCTGTAGTCCCAGGTACTCCGTAGGCTGATGCAGGAGGATCA
TTAGAGCCCAGGAGATTAAGGATACAGTGAGCTGCACCACTGCACTCCAGCCTGGGCAAAAGAGTAAGACCCTATCTCAAGAA
AAAAAAAAAAAAAAAGGAACGAGATCTAGGCTCACAGACAATCTTCCAGATATCAGCGGGAAATGATGAGTGTGTCTGGGGGA
CATCCAAAATTTCGGAATTAATCTTGTTTTGGGAGACAGGGAAGGAGAGGGATGTTCTGGGGGAAAACTAAGTCAAGGCTGGC
ATCCTCTCCCCCGTCCCCTGCCAGTTTTCCATCTCCAGCAGCTCTGCTACCCCTTCCCCCATCCCCGAGTGTGGTTCCAGATA
GTGGAAGTCTTATCTCCTGTCTCCAGCCAGACCTGATCGGTTTCTGTCCCTGGAGCTGGGGGGGGAGCGGGGAGAGGGGCGGT
TAGAGGGGCAGTGTTGGGGAAGTGGGACAGACAGACAGGCAAACAAGACCCCTTTCCAAAGCCTCTGCGTCAGAGTGTCCAGC
CCGCGATGTCCCTGGGCAGGGCACCCCAGTGTCCACCGAACCTCGAGCTGCCTGCTCCCTCCCGCAGTGGTGGCGCTCCGAAG
AGGCGCAGGACATGGGCCCGGGTCCTCCTGACCCCACGGAGCCGCCCCTCCACGTGAAGTCTGAGGACCAGCCCGGGGAGGAA
GAGGACGATGAGAGGGGCGCGGACGCCACCTGGGACCTGGATCTCCTCCTCACCAACTTCTCGGGCCCGGAGCCCGGTGGCGC
GCCCCAGACCTGCGCTCTGGCGCCCAGCGAGGCCYCCGGGGCGCAATATCCGCCGCCGCCCGAGACTCTGGGCGCATATGCTG
GCGGCCCGGGGCTGGTGGCTGGGCTTTTGGGTTCGGAGGATCACTCGGGTTGGGTGCGCCCTGCCCTGCGAGCCCGGGCTCCC
GACGCCTTCGTGGGCCCAGCCCTGGCTCCAGCCCCGGCCCCCGAGCCCAAGGCGCTGGCGCTGCAACCGGTGTACCCGGGGCC
CGGCGCCGGCTCCTCGGGTGGCTACTTCCCGCGGACCGGGCTTTCAGTGCCTGCGGCGTCGGGCGCCCCCTACGGGCTACTGT
CCGGGTACCCCGCGATGTACCCGGCGCCTCAGTACCAAGGGCACTTCCAGCTCTTCCGCGGGCTCCAGGGACCCGCGCCCGGT
CCCGCCACGTCCCCCTCCTTCCTGAGTTGTTTGGGACCCGGGACGGTGGGCACTGGACTCGGGGGGACTGCAGAGGATCCAGG
TGTGATAGCCGAGACCGCGCCATCCAAGCGAGGCCGACGTTCGTGGGCGCGCAAGAGGCAGGCAGCGCACACGTGCGCGCACC
CGGGTTGCGGCAAGAGCTACACCAAGAGCTCCCACCTGAAGGCGCATCTGCGCACGCACACAGGTGAGGGGGCGGGGCCCCGG
ACATGAGAAAGGGCGCGGCGCCCGCTGTAGTTACAGGGGAAGAAGGGTTGCAGAGGGCGGGACTTGGACTTGGCTGGCCTCTG
AGAGTGAGTGCCTCCTTAAATTTTGTGCCCTAGGGGCCTCACTTTGTTCATCCTAGTCCCAGCCCAGGCTGAGTAAAGGGGTG
TGGCCAGATGCAGGGGACCCGGGGACATGACTGGGCAGACAGTGGCGCTTATGGCTTCCTTGTCCCCTAGGGGAGAAGCCATA
CGCCTGCACGTGGGAAGGCTGCGGCTGGAGATTCGCGCGCTCGGACRAGCTGACCCGCCACTACCGGAAACACACGGGGCAGC
GCCCCTTCCGCTGCCAGCTCTGCCCACGTGCTTTTTCGCGCTCTGACCACCTGGCCTTGCACATGAAGCGCCACCTTTGAGCC
CTGCCCTGGCACTTGGACTCTCCTAGTGACTGGGGATGGGACAAGAAGCCTGTTTGGTGGTCTCTTCACACGGACGCGCGTGA
CACAATGCTGGGTGGTTTTCCCACGAATGGACCCTCTCCTGGACTCGCGTTCCCAAAGATCCACCCAAATATCAAACACGGAC
CCATAGACAGCCCTGGGGGAGCCTCTTACGGAAAATCCGACAAGCCTTCAGCCACAGGGAGCCACACAGAGATGTCCAAACTG
TCGTGCAAACCCAGTGAGACAGACCGCCAAATAAACGGACTCAGTGGACACTCAGACCAGCTCCCAGATGGCCCTGGACAGCA
GGAGAGGGTGTGGGATGAGGCTTCCCAGAGACCCTGGGTCTAGAAAGCGGCTCCTGAAGGTCCCTTATTGTGGCTGATATTAA
CTGTCAATGGTTATGGGTCCTATAAAAATGCCCCTCCCAGATAAAGCTTCAGCGTTGGCCTGAATTTTGGGTTTTGGAGGCTG
GAAGATGGCAGGAGGCTGTAAACCTAGAGGTCATATTAGAGGG

Y (heterozygous T/C) for NM_006563.3:c.304T>C, NP_006554.1:p.Ser102Pro (rs2072597)
R (heterozygous G/A) for NM_006563.3:c.973G>A, NP_006554.1:p.Glu325Lys

Figure S6. (B) Analysis of heterozygous KLF1 mutation c.973G>A (E325K) from the CDA patient ME in both directions with Mutation Surveyor software (SoftGenetics).

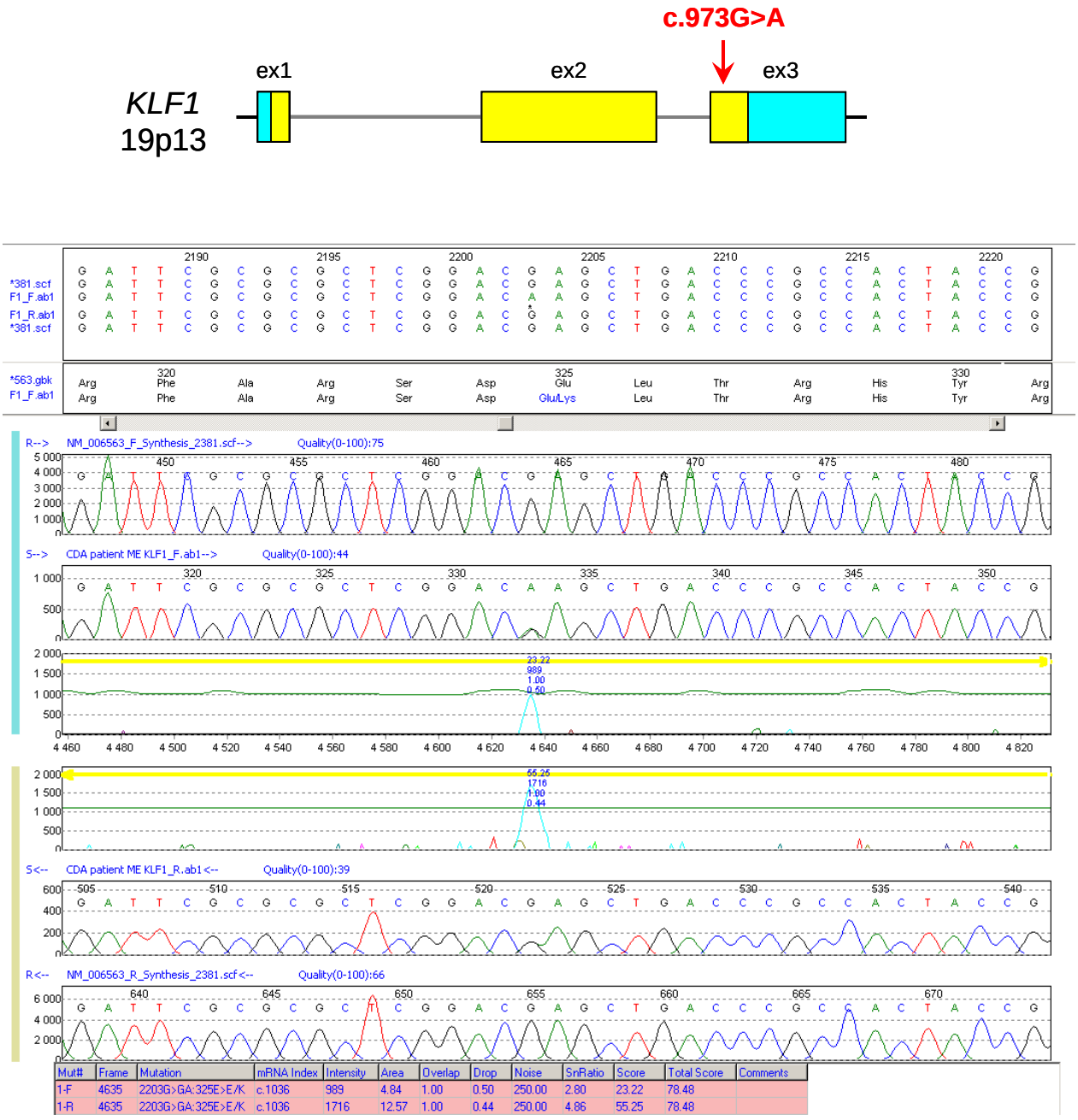


Figure S7. Analysis of the erythrocyte membrane proteins from the CDA patient ME (CDA, sample taken on 2010/01/05) or a control (CTL) showing abnormal migration of Band 3 (arrows) as usually observed in CDA II but also slight reduction of spectrins (stars). Proteins were resolved either by Laemmli's (left gel) or Fairbanks' (right gel) electrophoretic method, and stained with Coomassie Blue.

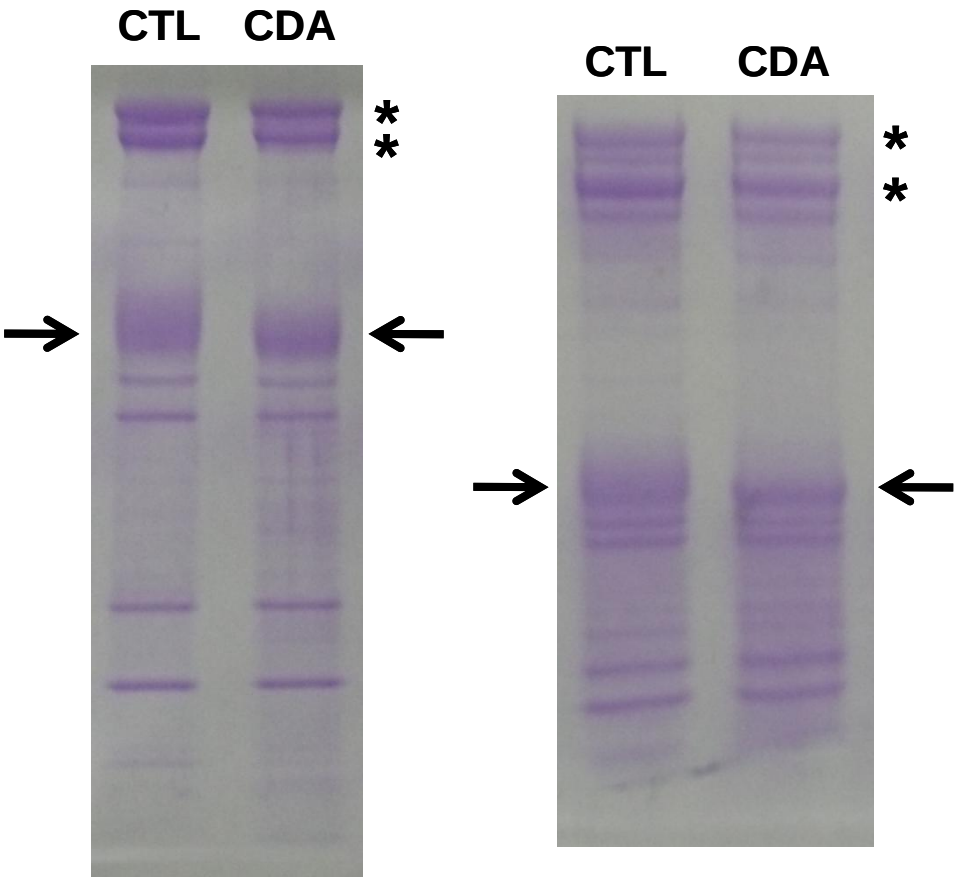


Figure S8. Reverse phase liquid chromatography (RPLC) elution pattern of the hemolysate from the CDA patient ME showing the presence of the hydrophobic ζ-globin chain (retention time 84 min). This confirmed that the fast migrating hemoglobin observed in isoelectric focusing (IEF) was indeed embryonic hemoglobin Portland ($\zeta_2\gamma_2$ tetramer). The peaks and the corresponding retention times of heme, β-, $^{G}\gamma$-, $^{A}\gamma$- and α-globin chains are indicated.

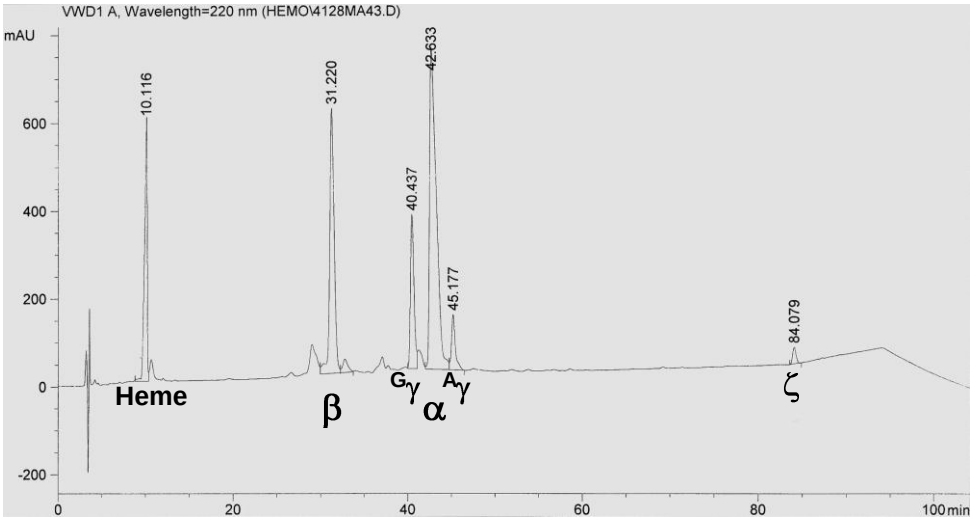


Figure S9. Flow cytometry analysis of cell markers CD44, CD47, GYPA, BCAM/Lu, ICAM4/LW, CD29 and DARC on erythrocytes of the CDA patient SF (right histograms, sample taken on 2009/06/29) or a healthy control (left histograms). Erythrocytes were stained with mouse monoclonal antibodies (profiles in color) or an isotype control antibody (profile in black) and phycoerythrin-conjugated goat F(ab')2 anti-mouse IgG(H+L), and were gated on FSC and SSC (density dot plots). Patient SF's erythrocytes abnormally showed no expression of CD44 and reduced expression of ICAM4; the small $CD44^{+}$ population (9.5%, star) likely corresponded to the residual erythrocytes of the most recent transfusion (2009/04/20) and should be taken into account for the analysis of the other markers, especially ICAM4.

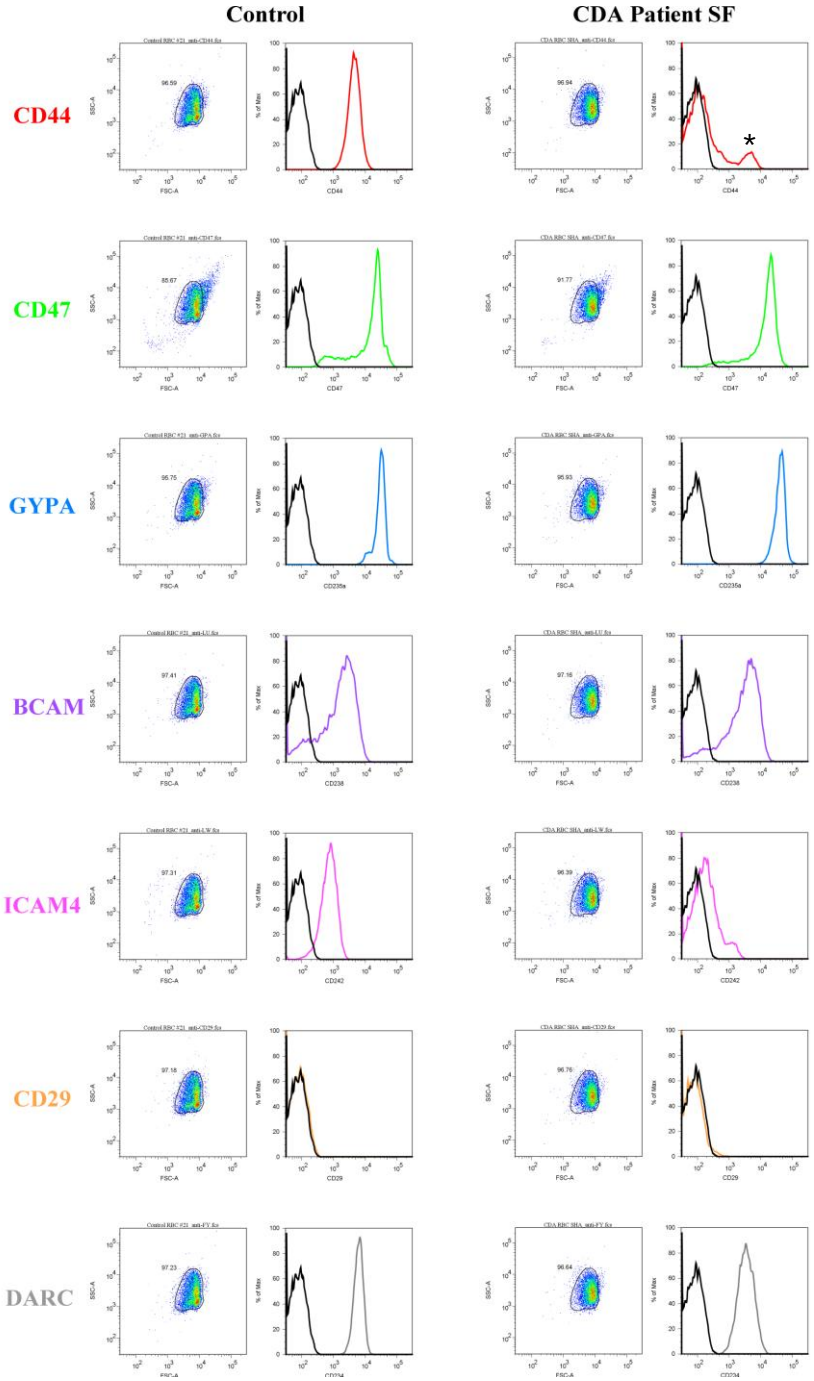


Table S1. Essential hematologic parameters of CDA patient ME from birth, before and after splenectomy.

Red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean cellular volume (MCV), mean cellular hemoglobin (MCH), mean cellular hemoglobin concentration (MCHC), red-blood-cell distribution width (RDW), platelet count (PLT), mean platelet volume (MPV), white blood cell count (WBC). Macrocytosis, severe thrombocytosis and extremely high erythroblastosis appeared after splenectomy. The three last RBC transfusions were on January 2004 (PT-Hb 7.2 g/dL), February 2005 (PT-Hb 7.2 g/dL) and October 2006 (PT-Hb 7.3 g/dL).

Parameters	Date / Age (Hospital) / Units	12/12/1996 1 day (Strasbourg)	14/02/1997 2 months (Strasbourg)	14/03/1997 3 months (Strasbourg)	05/05/1997 5 months (Nancy)	16/09/1997 9 months (Nancy)	29/01/1998 1 year (Nancy)	20/07/1998 1 year (Nancy)	20/05/1999 2 years (Nancy)	17/08/1999 2 years (Nancy)	29/03/2000 3 years (Nancy)	03/07/2000 3 years (Nancy)	20/04/2001 4 years (Paris)	14/05/2002 5 years (Paris)	25/10/2005 8 years (Nancy)	28/10/2008 11 years (Nancy)	10/02/2009 12 years (Nancy)	05/01/2010 13 years (Nancy)
RBC	10^{12}/L	4,0	2,7	3,8	3,3	2,2	2,7	2,7	2,4	2,7	1,7	1,9	2,5	2,5	2,3	2,7	2,9	2,6
HGB	g/dL	11,5	7,5	11,4	9,4	6,5	7,4	7,6	6,7	7,6	4,7	5,4	8,0	8,4	7,5	8,4	9,1	8,0
HCT	%	34,1	22,5	33,9	27,9	19,4	21,9	22,8	20,0	23,5	14,2	16,7	26,0	27,0	23,5	27,6	30,1	26,7
MCV	fL	86,3	85,0	89,0	83,6	87,7	82,3	84,8	83,8	87,4	85,3	85,9	106,9	108,9	101,0	102,4	102,8	103,4
MCH	pg	29,1	28,1	29,9	2,8	29,4	27,8	28,2	28,0	28,2	28,3	27,8	32,2	33,8	32,2	31,1	31,1	31,0
MCHC	g/dL	33,7	33,0	33,6	33,7	33,5	33,8	33,3	33,5	32,3	33,1	32,3	30,1	31,0	31,9	30,4	30,2	30,0
RDW	%	18,4	16,7	14,6	14,1	14,6	17,5	13,9	20,9	18,2	20,6	24,0						
Reticulocytes	10^9/L	52,5	13,9	62,9	9,0	80,7	134,0	21,5					838	362	422	351	469	304
PLT	10^9/L	406,0	393,0	229,0	380,0	447,0	339,0	223,0	230,0	211,0	173,0	83,0	526,0	612,0	1020,0	1101,0	1036,0	1264,0
MPV	fL	10,4	12,0	10,3	8,9	7,9	8,4	8,6	8,1	8,5	7,8	7,3	10,6	9,3	7,5	9,2	8,1	8,3
WBC	10^9/L	9,8	20,4	8,2	8,1	13,0	24,2	11,9	15,9	16,0	19,2	27,8	22,7	24,9	15,9	13,9	16,3	16,4
Neutrophils	10^9/L	3,2	13,5	2,1	1,5	3,7	5,6	3,3	5,6	3,5	11,9	18,6	7,1	7,7	7,8	5,5	9,8	9,2
Eosinophils	10^9/L	0,4	0,6	0,2	0,3	0,2	0,2	0,1	0,3	0,1	0,2	0,2	1,4	0,8	0,3	0,3	0,3	0,2
Basophils	10^9/L	0,0	0,0	0,0	0,1	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,2	0,0
Lymphocytes	10^9/L	3,6	3,3	4,8	4,8	7,4	16,5	6,3	8,4	9,8	6,3	7,2	10,9	14,7	6,4	6,1	4,7	4,8
Monocytes	10^9/L	1,1	2,0	1,2	1,2	1,0	1,9	2,1	1,6	2,6	0,8	1,7	3,4	1,5	1,1	1,8	1,3	1,8
Neutrophils	%	33,0	66,0	25,0	18,5	28,4	23,0	28,0	35,0	22,0	62,0	67,0	31,0	34,0	34,3	24,4	42,9	40,5
Eosinophils	%	4,0	3,0	2,0	3,8	1,7	1,0	1,0	2,0	0,4	0,8	0,6	6,0	3,3	1,4	1,2	1,5	0,7
Basophils	%	0,0	0,0	0,0	0,7	1,4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,6	0,7	0,0
Lymphocytes	%	37,0	16,0	58,0	59,1	56,9	68,0	53,0	53,0	61,0	33,0	26,0	48,0	64,7	28,0	26,8	20,7	20,9
Monocytes	%	11,0	10,0	14,0	14,4	7,7	8,0	18,0	10,0	16,0	4,0	6,0	15,0	6,6	4,8	7,9	5,7	8,0
Erythroblasts	/100 WBCs	470	12	128	few	12	30	few	100	110	80	210	920	594	1000	1000	500	500

12/12/1996 BIRTH (28 weeks' gestation) — 06/09/2000 SPLENECTOMY (4 years)

Table S2. PCR and sequencing primers used to sequence *KLF1*.
Location and direction of the primers are based on NCBI *KLF1* mRNA Reference Sequence NM_006563.3 while position is based on NCBI human chromosome 19 genomic contig NT_011295.11. Complete *KLF1* was sequenced after PCR amplification of two overlapping fragments using high fidelity DNA polymerase as follows. A 5' *KLF1* fragment was amplified by semi-nested PCR with primers KLF1-1, KLF1-2 and KLF1-11, and sequenced with primers KLF1-2, KLF1-6 and KLF1-15. A 3' *KLF1* fragment was amplified by semi-nested PCR with primers KLF1-7, KLF1-3 and KLF1-4, and sequenced with primers KLF1-7, KLF1-8 and KLF1-4. PCR products were sequenced with ABI BigDye terminator chemistry (GATC Biotech) after gel purification (Macherey-Nagel). Sequencing data were analyzed with DNA Workbench software (CLC bio) and Mutation Surveyor software (SoftGenetics).

Name	Sequence	Location	Direction	Position
KLF1-1	TAGTCTTTAACCCCAGCCCCAG	5' side of exon 1	Sense	4260969-4260948
KLF1-2	ACGTGAAGTTTGTGCCCCAG	5' side of exon 1	Sense	4260939-4260920
KLF1-3	AGGGGGCTTGTTGGAGTTGAG	3' side of exon 3	Antisense	4257897-4257917
KLF1-4	AGGAGATGAGGGTGTGTAAGG	3' side of exon 3	Antisense	4257935-4257955
KLF1-6	TGATCGGTTTCTGTCCCTGG	Intron 1	Sense	4259958-4259939
KLF1-7	AGGATCACTCGGGTTGGGT	Exon 2	Sense	4259457-4259439
KLF1-8	TACACCAAGAGCTCCCACC	Exon 2	Sense	4258978-4258960
KLF1-11	CGGGTCCCAAACAACTCAG	Exon 2	Antisense	4259122-4259140
KLF1-15	GCTTTGGAAAGGGGTCTTG	Intron 1	Antisense	4259847-4259865