

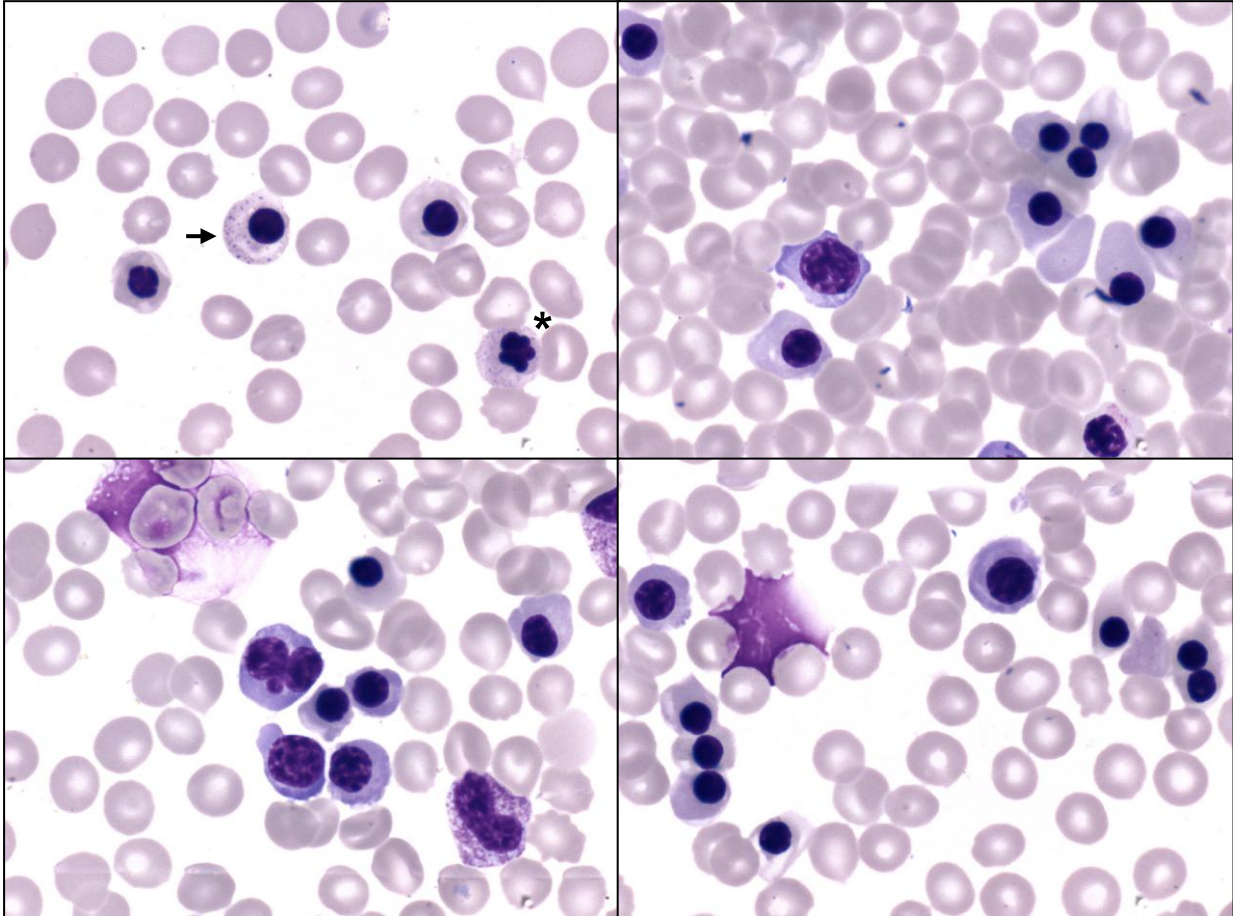
The American Journal of Human Genetics Volume 87

**Supplemental Data**

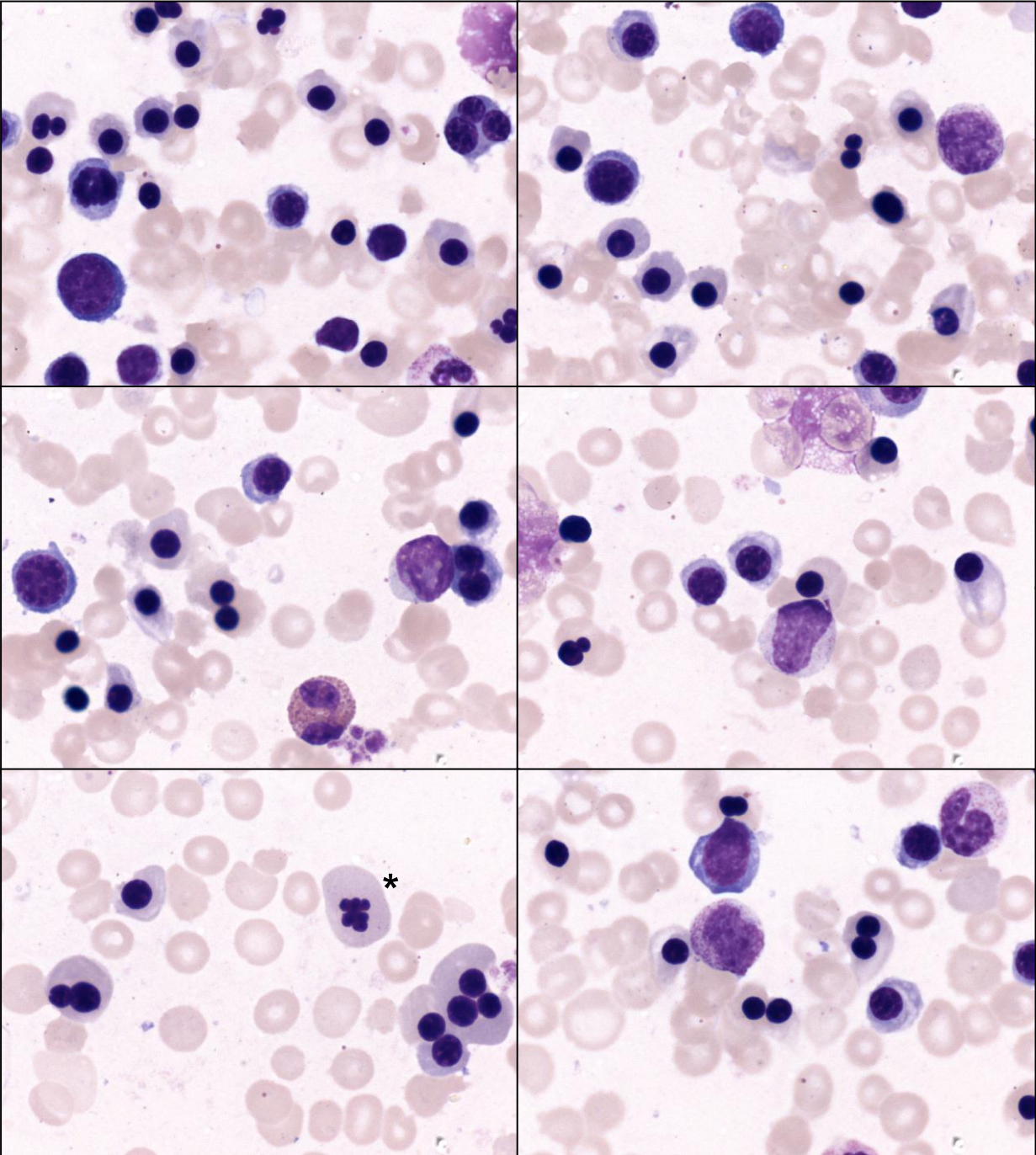
**A Dominant Mutation in the Gene Encoding  
the Erythroid Transcription Factor KLF1 Causes  
a Congenital Dyserythropoietic Anemia**

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.

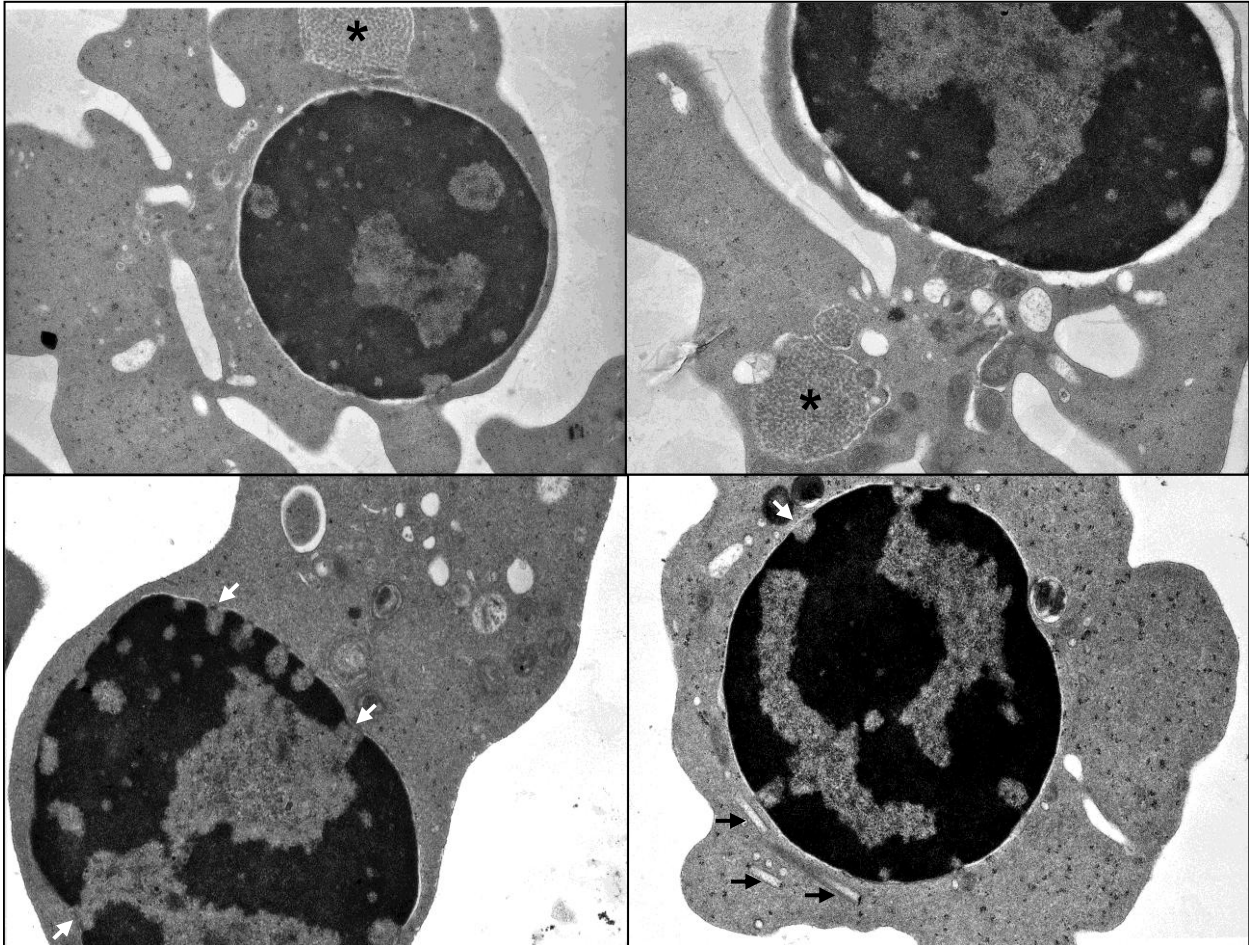
**Figure S1. (A) Bone marrow smear from the CDA patient ME before splenectomy** (sample taken on 1997/07/14) showing marked erythroid hyperplasia (77%) with a majority of acidophilic erythroblasts (58%). Binucleated erythroblasts (1.0%), multinucleated erythroblasts (1.5%) and few karyorrhexis (star)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



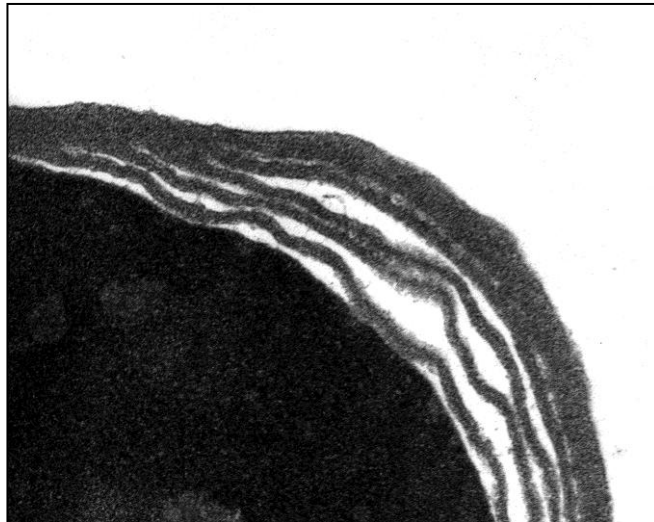
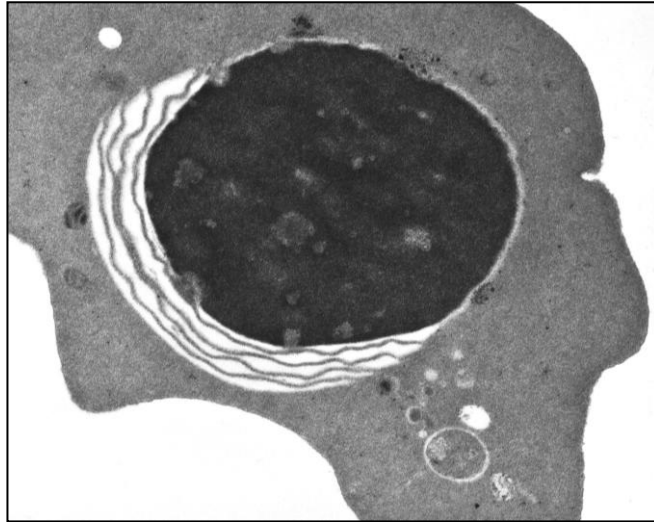
**Figure S1. (B) Bone marrow smear from the CDA patient ME after splenectomy** (sample taken on 2001/02/07) showing marked erythroid hyperplasia (80%) with a majority of acidophilic erythroblasts (55%). Binucleated erythroblasts (4.5%), multinucleated erythroblasts (2.5%) and few karyorrhexis (star)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



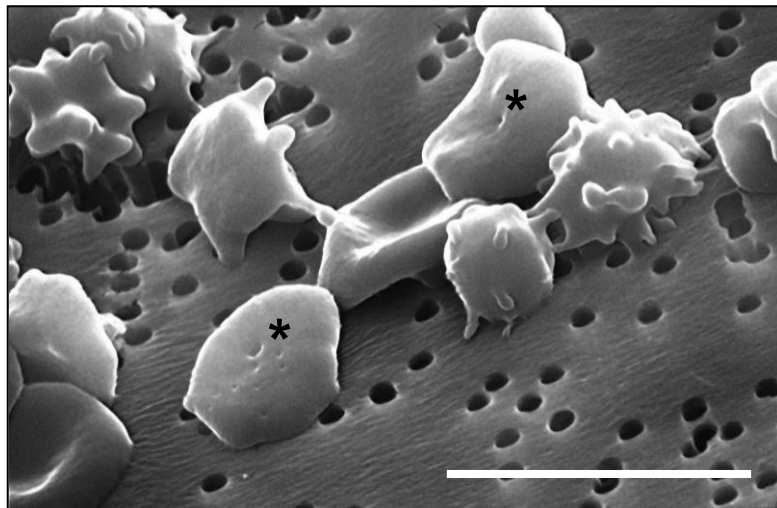
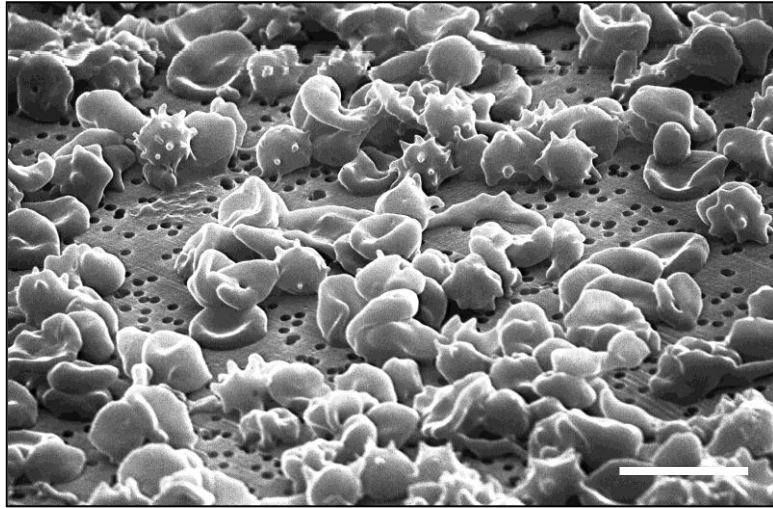
**Figure S2. (A) Transmission electron microscopy of circulating erythroblasts of the CDA patient ME showing tubulosaccular inclusions (stars) and elongated cylindrical structures (black arrows) in the cytoplasm, and abnormally large nuclear pores (white arrows). Original magnification  $\times 10,000$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



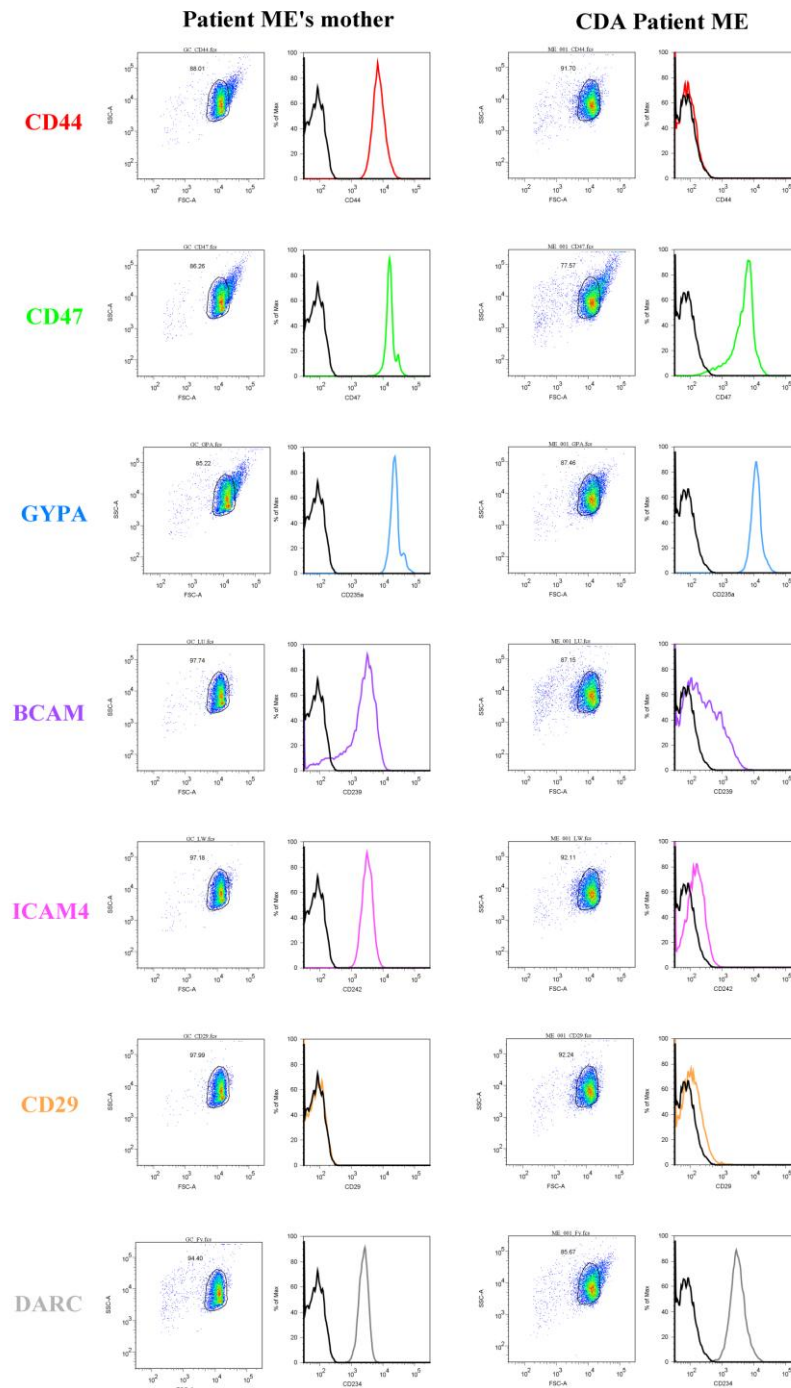
**Figure S2. (B) Transmission electron microscopy of circulating erythroblasts of the CDA patient ME showing a wide area of endoreplication of the nuclear membrane. Original magnification  $\times 8,000$  (top panel) and  $\times 35,000$  (bottom panel).**



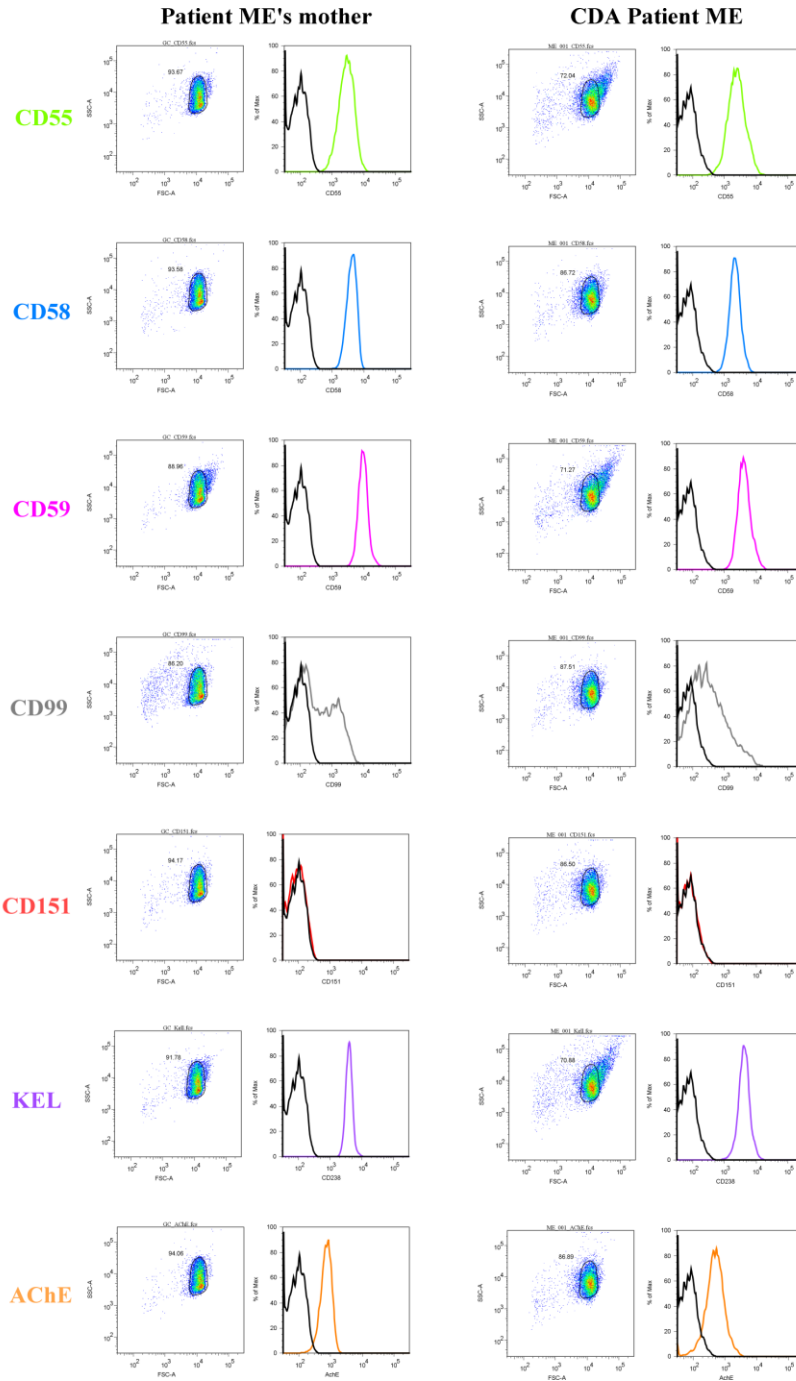
**Figure S2. (C) Scanning electron microscopy of blood of the CDA patient ME showing severe anisopoikilocytosis (top panel, original magnification  $\times 1,930$ ) and small invaginations in erythrocyte membrane (stars, bottom panel, original magnification  $\times 4,590$ ). The scale bars represent  $10\ \mu\text{m}$ .**



**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')<sub>2</sub> 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.**

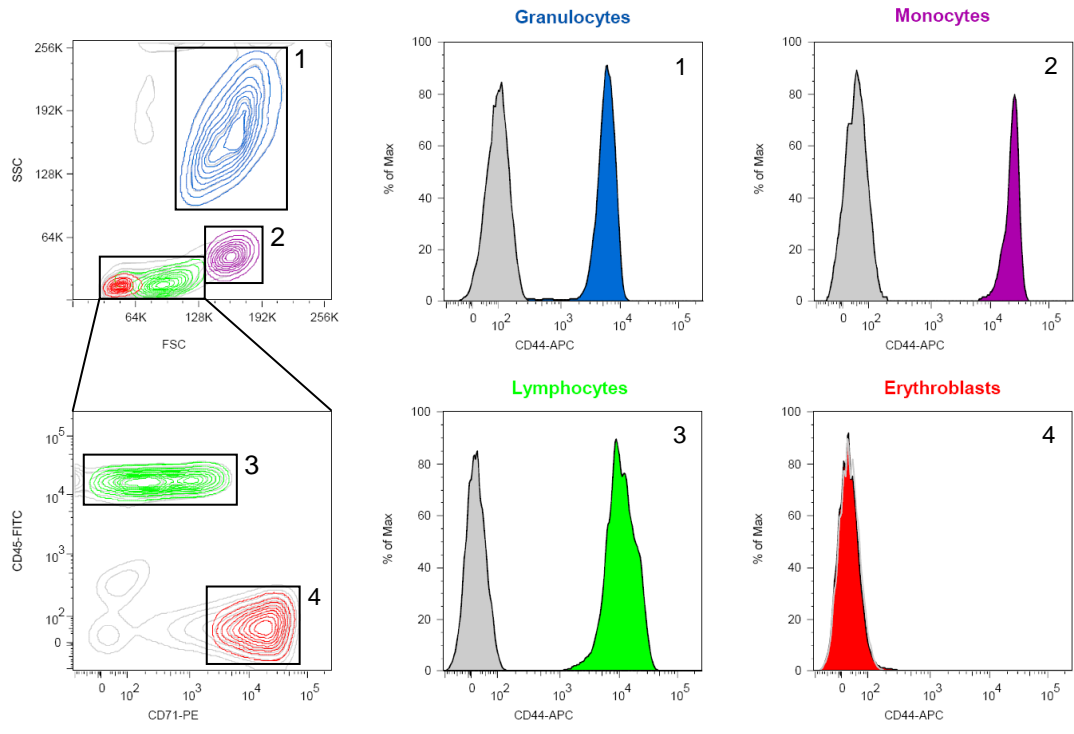


**Figure S3. (B) Flow cytometry analysis of cell markers CD55, CD58, CD59, CD99, CD151, KEL and AChE on erythrocytes of the CDA patient ME (right histograms, sample taken on 2010/01/05) or his healthy mother (left histograms, sample taken at the same time). Erythrocytes were stained with mouse monoclonal antibodies (profiles in color) or an isotype control antibody (profile in black) and phycoerythrin-conjugated goat F(ab')<sub>2</sub> 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<sup>+</sup>/CD45<sup>+</sup>, panel 3) and erythroblasts (CD71<sup>+</sup>/CD45<sup>-</sup>, 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.

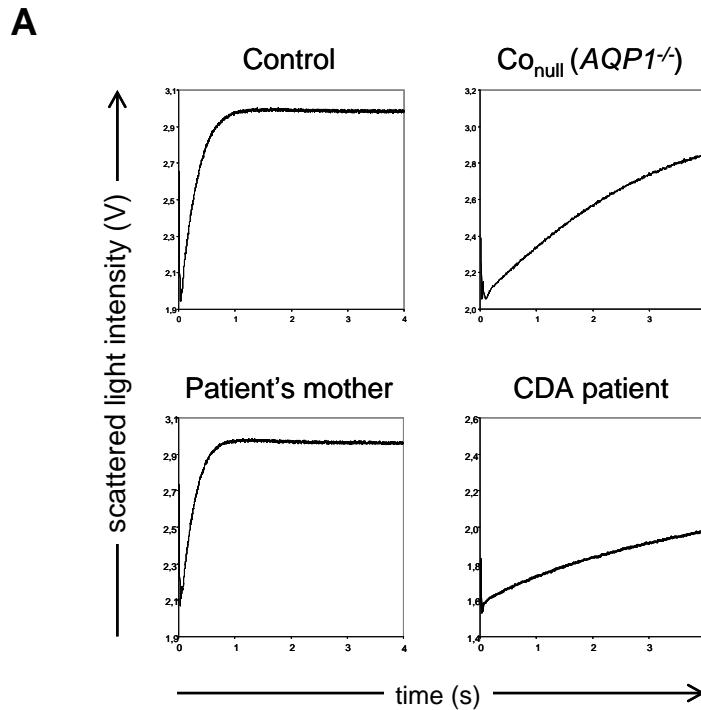


**Figure S5. Analysis of the water permeability of the CDA patient ME's erythrocytes confirming his AQP1 deficiency.**

(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.

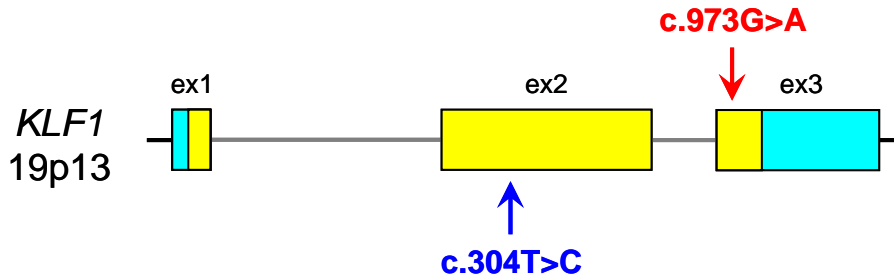
Due to AQP1 deficiency of patient ME's erythrocytes, their Colton phenotype is expected to be  $Co(a-b-)$  while blood group genotyping (BioArraySolutions HEA BeadChip Kit) predicts the following extended phenotype: C+E-c+e+, K-k+, Kp(a-b+), Js(a-b+), Fy(a+b+), Jk(a+b-), M+N-S-s+, Lu(a-b+), Di(a-b+),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



**B**

Samples	ABO type	$Pf$ (cm.s <sup>-1</sup> ) at 15°C	
			+ HgCl <sub>2</sub>
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.**



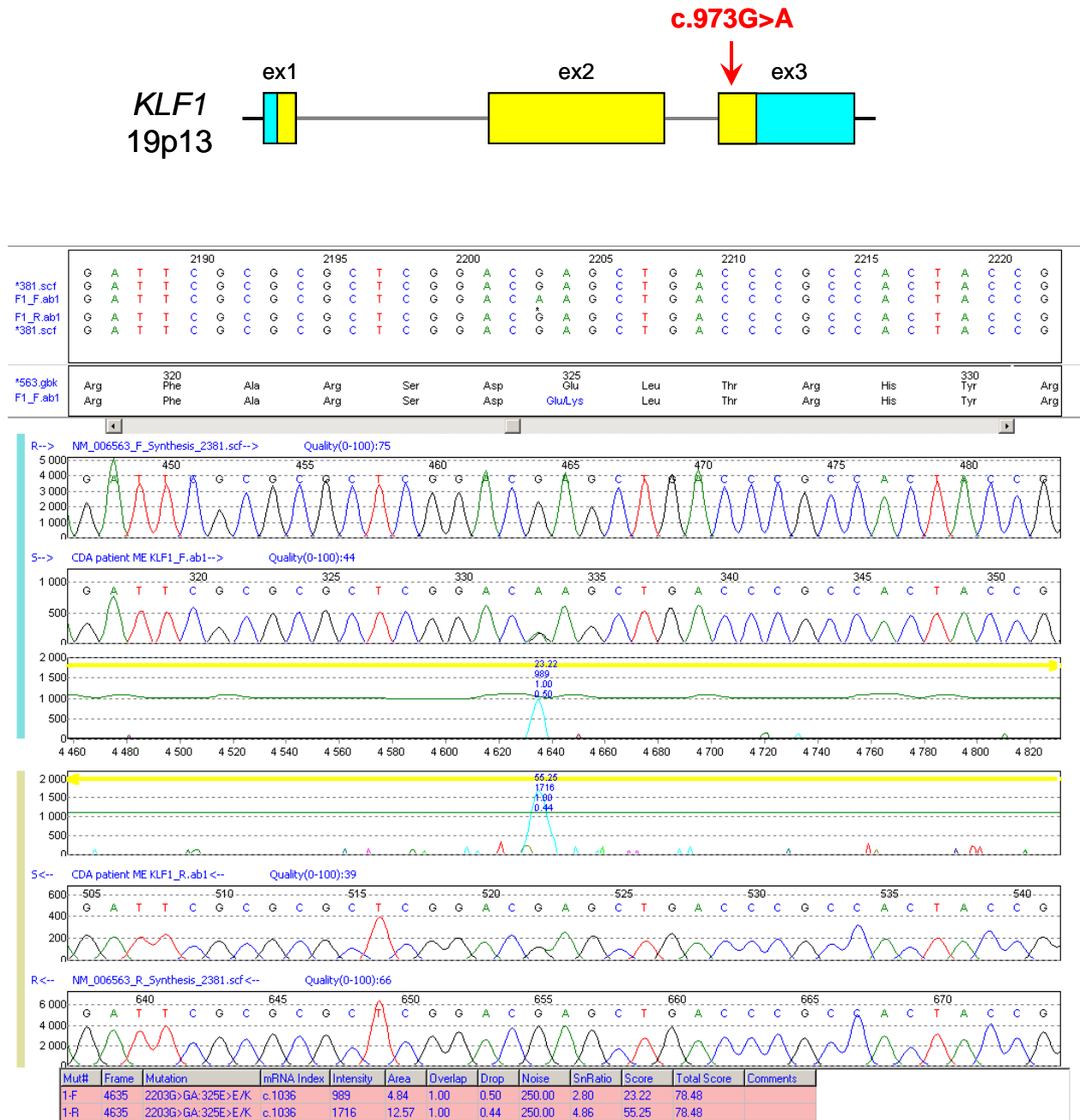
NNN coding sequence  
NNN untranslated sequence  
NNN intronic sequence

```

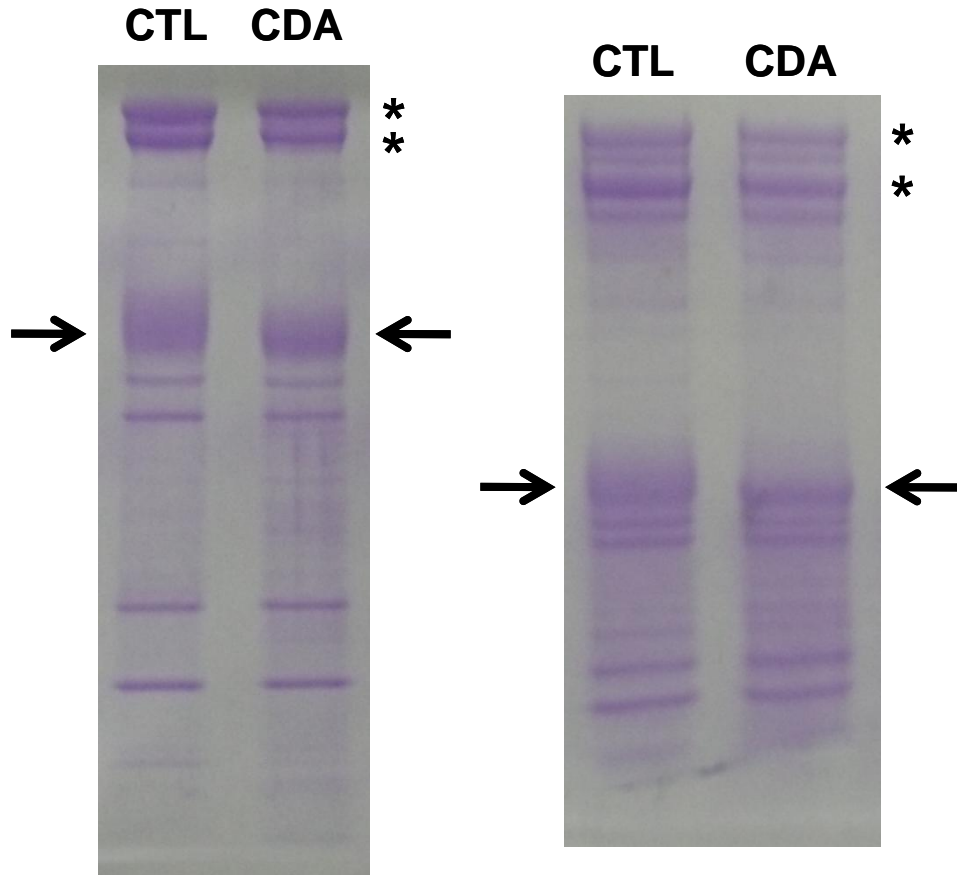
GCCGCTTGCCTTGCCTTGCCTTATCAGAGGCTGCAGCCAATCAGCTAAGGACAGAGAGGAGCCCTCGAAGGGGCTATCACAG
CCTCAGAGTTACAGAGGCAGCCGAGGAAGAGGAGCCTTGAGGCCACAGGGTGGGCACCAGCCAGCCATGGCCACAGCCGAGACC
GCCTTGCCTCCATCAGCACACTGACCGCCCTGGGCCCTTCCCGGACACACAGGATGACTTCCCTCAAGGTGGGGCCTAGAAG
GTGGGGTCTAGGTGGGCTGGCTGGAATCCAGGGCCACAGTCACAGATCTTGGGGTCCAGACCTGCATCTTGACCTGAAATCAA
GAGACTTAACCAGGACTGAGGTACGCTCAGTCCAGGAGAGGAGATCTCAGCTTAGTCTGGCAGGGGGTGAAGAGGGTGGTCTA
GGGTTTGGAGTTCTAAGTGTGATCTATTTCCGTAATAGAAAACGAAGGTAGCCTGGGCAACATGGTGAACCCCTATCTCTAC
AAAAATACAAAAACATTAGGCCAGGCATGGGGGGTGTGCCTGTAGTCCCAGGTAAGTCCGTAGGCTGATGCAGGAGGATCA
TTAGAGCCAGGAGATTAAGGATACAGTGCAGTGCACACTGCACCTCCAGCCTGGGCAAAAGAGTAAGACCCTATCTCAAGAA
AAAAAAAAAAAAAAAAAGGAACGAGATCTAGGCTCACAGACAATCTCCAGATATCAGCGGAAATGATGAGTGTGTCTGGGGGA
CATCCAAAATTTCCGAATTAATCTTGTTTTGGGAGACAGGGAAGGAGAGGGATGTTCTGGGGGAAAATAAGTCAAGGCTGGC
ATCCTCTCCCCCTGCCAGTTTCCATCTCCAGACTCTGCAGTCTGTACCCCTCCCCATCCCCGAGTGTGGTTCAGATA
GTGAAGTCTTATCTCCTGTCTCCAGCCAGACTGATCGGTTTCTGTCCCTGGAGCTGGGGGGGAGCGGGGAGAGGGGCGGT
TAGAGGGGAGTGTGGGGAAAGTGGGACAGACAGAGCAAGCAAGACCCCTTTCCAAAGCCTCTGCGTCAGAGTGTCCAGC
CCGCGATGTCCCTGGGCAGGGCACCCAGTGTCCACCGAACCTCGAGCTGCCTGCCTCCCTCCCGCAGTGGTGGCGCTCCGAAG
AGGCGCAGGACATGGGCCCGGTCTCCTGACCCACGGAGCCCGCCCTCCACGTGAAGTCTGAGGACCAGCCCGGGGAGGAA
GAGGACGATGAGAGGGGCGGGACGCCACCTGGGACCTGGATCTCCTCCTACCAACTTCTCGGGCCGGAGCCCGGTGGCGC
GCCCAGACCTGCCTCTGGCGCCAGCGAGGCCYCCGGGGCGCAATATCCGCCGCGCCCGAGACTCTGGGCGCATATGCTG
GCGCCCCGGGGCTGGTGGCTGGGCTTTTGGGTTCCGAGGATCACTCGGGTGGGTGCGCCCTGCCCTGCGAGCCCGGGCTCCC
GACGCTTTCGTGGGCCAGCCCTGGCTCCAGCCCCGGCCCCGAGCCCAAGGCGCTGGCGCTGCAACCGGTGTACCCGGGGCC
CGGCGCGGCTCCTCGGGTGGCTACTTCCCGCGGACCGGGCTTTCAGTGCCTGCGGGCTCGGGCGCCCCCTACGGGCTACTGT
CCGGTACCCCGCGATGTACCCGGCGCCTCAGTACCAAGGGCATTCCAGCTCTTCCGCGGGCTCCAGGACCCCGCGCCCGT
CCGCCACGTCCCTCCTTCTGAGTTGTTTGGGACCCCGGACGTTGGGCACCTGGACTCGGGGGGACTGCAGAGGATCCAGG
TGTGATAGCCGAGACCGCCATCCAAAGCAGGGCCAGTTCGTGGGCGCGCAAGAGGCAGGCAGCGCACACAGTGC
CGGTTGCGGCAAGAGCTACACCAAGAGCTCCCACTGAAGGCGCATCTGCGCACGCACACAGGTGAGGGGGCGGGGCCCGG
ACATGAGAAAGGGCGGGCGCCCGTGTAGTTACAGGGGAAGAAGGGTTCAGAGGGGCGGGACTTGGACTTGGCTGGCCTCTG
AGAGTGAAGTGCCTCTTAAATTTTGTGCCCTAGGGGCTCACTTTGTTCATCCTAGTCCCAGCCAGGCTGAGTAAAGGGGTG
TGCCAGATGCAGGGGACCCGGGACATGACTGGGACAGCAGTGGCGCTTATGGCTTCCCTGTCCCTAGGGGAGAAGCCATA
CGCTGCACGTGGGAAGGCTGCGGCTGGAGATTCGCGCGCTCGGACRAGCTGACCCGCCACTACCGGAAACACACGGGGCAGC
GCCCTTCCGCTGCCAGCTCTGCCACGTGCTTTTTCGCGCTCTGACCACCTGGCCTTGACATGAAGCGCCACCTTTGAGCC
CTGCCCTGGCACTTGGACTCTCCTAGTGAAGTGGGATGGGACAAGAAGCCTGTTTGGTGGTCTCTTACACGGACGCGGTGA
CACAAATGCTGGTGGTGTTCACGAATGGACCCCTCCTTGACTCGGCTTCCAAAGATCCACCAATATCAACACGGAC
CCTAGACGACCCCTGGGGAGCCTTACGGAAATCCGACAAGCCTTCAGCCACAGGGAGCCACACAGAGATGTCCAACTG
TCGTGCAAAACCCAGTGCAGACAGACCGCAAAATAACGGACTCAGTGGACACTCAGACCAGCTCCAGATGGCCCTGGACAGCA
GGAGAGGGTGTGGGATGAGGCTTCCAGAGACCCGGGTCTAGAAAGCGGCTCCTGAAGGTCCCTTATGTGGCTGATATTA
CTGTCAATGGTTATGGGCTCTATAAAAAATGCCCTCCAGATAAAGCTTACAGCTTGGCCTGAATTTGGGTTTGGAGGCTG
GAAGATGGCAGGAGGCTGTAACCTAGAGGTCATATTAGAGGG
    
```

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

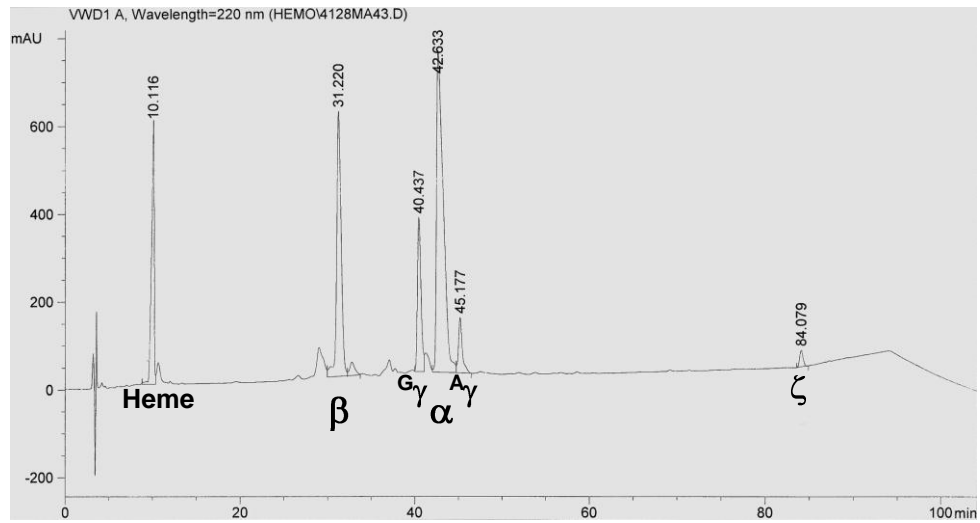
**Figure S6. (B) Analysis of heterozygous KLF1 mutation c.973G>A (E325K) from the CDA patient ME in both directions with Mutation Surveyor software (SoftGenetics).**



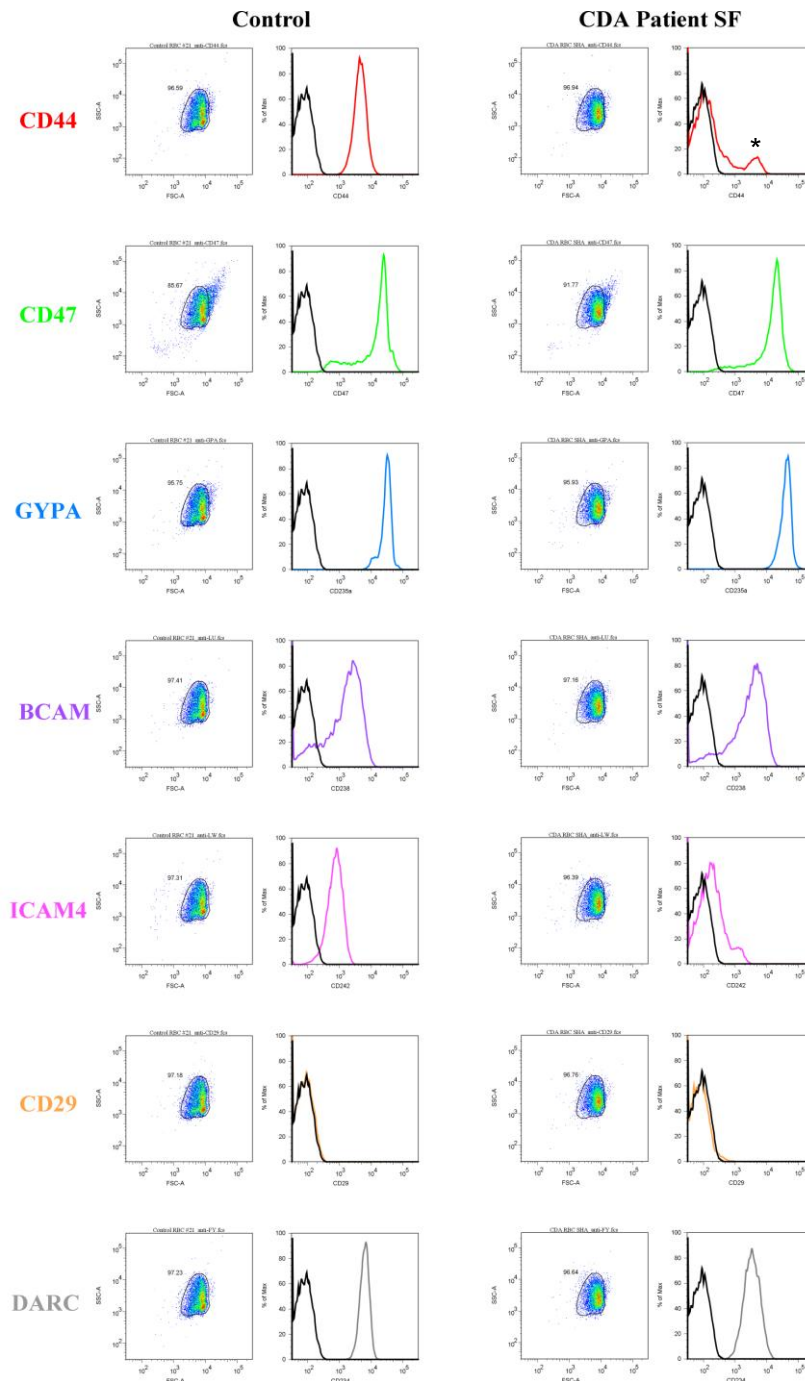
**Figure S7. Analysis of the erythrocyte membrane proteins from the CDA patient ME** (CDA, sample taken on 2010/01/05) or a control (CTL) showing abnormal migration of Band 3 (arrows) as usually observed in CDA II but also slight reduction of spectrins (stars). Proteins were resolved either by Laemmli's (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  $\zeta$ -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,  $\beta$ -,  $\zeta$ - $\gamma$ -,  $\alpha$ - $\gamma$ - and  $\alpha$ -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')<sub>2</sub> 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<sup>+</sup> 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	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)	25/01/1998 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 <sup>12</sup> /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	28.9	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	33.6	33.8	33.3	33.5	32.9	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 <sup>9</sup> /L	52.5	13.9	62.9	9.0	80.7	134.0	21.5					838	362	422	351	469	304
PLT	10 <sup>9</sup> /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 <sup>9</sup> /L	9.8	20.4	8.2	8.1	13.0	24.2	11.9	15.9	16.0	19.2	27.4	22.7	24.9	15.9	13.9	16.3	16.4
Neutrophils	10 <sup>9</sup> /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 <sup>9</sup> /L	0.4	0.6	0.2	0.3	0.2	0.1	0.3	0.1	0.2	0.1	0.2	1.4	0.8	0.3	0.3	0.3	0.2
Basophils	10 <sup>9</sup> /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 <sup>9</sup> /L	3.6	3.3	4.8	4.8	7.4	16.5	6.3	8.4	9.8	6.3	7.3	10.9	14.7	6.4	6.1	4.7	4.8
Monocytes	10 <sup>9</sup> /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



**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	TAGTCTTTAACCCAGCCCCAG	5' side of exon 1	Sense	4260969-4260948
KLF1-2	ACGTGAAGTTTGTGCCCCAG	5' side of exon 1	Sense	4260939-4260920
KLF1-3	AGGGGGCTTGTGGAGTTGAG	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	CGGGTCCCAAACAACACTCAG	Exon 2	Antisense	4259122-4259140
KLF1-15	GCTTTGGAAAGGGGTCTTG	Intron 1	Antisense	4259847-4259865