



HAL
open science

Role of the Cold Shock Domain Protein A in the transcriptional regulation of gamma-globin gene expression.

Raffaella Petruzzelli, Sara Gaudino, Giovanni Amendola, Raffaele Sessa, Stella Puzone, Rosanna Di Concilio, Giovanna d'Urzo, Maria Amendolara, Paola Izzo, Michela Grosso

► To cite this version:

Raffaella Petruzzelli, Sara Gaudino, Giovanni Amendola, Raffaele Sessa, Stella Puzone, et al.. Role of the Cold Shock Domain Protein A in the transcriptional regulation of gamma-globin gene expression.. British Journal of Haematology, 2010, 150 (6), pp.689. 10.1111/j.1365-2141.2010.08303.x . hal-00556389

HAL Id: hal-00556389

<https://hal.science/hal-00556389v1>

Submitted on 16 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Role of the Cold Shock Domain Protein A in the
transcriptional regulation of gamma-globin gene
expression.**

Journal:	<i>British Journal of Haematology</i>
Manuscript ID:	BJH-2010-00179.R1
Manuscript Type:	Ordinary Papers
Date Submitted by the Author:	22-Apr-2010
Complete List of Authors:	<p>Petruzzelli, Raffaella; University of Naples Federico II, Dipartimento di Biochimica e Biotecnologie Mediche Gaudino, Sara; University of Naples Federico II, Dipartimento di Biochimica e Biotecnologie Mediche Amendola, Giovanni; Ospedale Umberto I, U.O.C. Pediatria-TIN and Ematologia-Oncologia Pediatrica Sessa, Raffaele; University of Naples Federico II, Dipartimento di Biochimica e Biotecnologie Mediche Puzone, Stella; University of Naples Federico II, Dipartimento di Biochimica e Biotecnologie Mediche Di Concilio, Rosanna; Ospedale Umberto I, U.O.C. Pediatria-TIN and Ematologia-Oncologia Pediatrica D'Urzo, Giovanna; Ospedale Umberto I, U.O.C. Pediatria-TIN and Ematologia-Oncologia Pediatrica Amendolara, Maria; Ospedale Umberto I, U.O.C. Pediatria-TIN and Ematologia-Oncologia Pediatrica Izzo, Paola; University of Naples Federico II, Dipartimento di Biochimica e Biotecnologie Mediche Grosso, Michela; University of Naples Federico II, Biochimica e Biotecnologie Mediche</p>
Key Words:	Cold Shock Domain, dbpA, HPFH, globin gene expression, beta-thalassemia intermedia, mRNA differential display

1
2
3 1 **Role of the Cold Shock Domain Protein A in the transcriptional regulation of α -globin**
4
5 2 **gene expression.**
6
7
8
9

10 4 Raffaella Petruzzelli,^{1*} Sara Gaudino,^{1*} Giovanni Amendola,² Raffaele Sessa,¹ Stella Puzone,¹
11
12 5 Rosanna Di Concilio,² Giovanna d'Urzo,² Maria Amendolara,² Paola Izzo,¹ and Michela
13
14 6 Grosso¹.

17 * who equally contributed to this work

18
19
20 ¹Dipartimento di Biochimica e Biotecnologie Mediche, Università degli Studi di Napoli
21
22 ²“Federico II”, Naples, 80131, Italy; ²U.O.C. Pediatria-TIN and Ematologia-Oncologia
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

12 Running title: Role of CSDA in modulation of globin gene expression

14 Corresponding author: Dr. Michela Grosso, Dipartimento di Biochimica e Biotecnologie
15 Mediche, Università degli Studi di Napoli “Federico II”, 80131 Napoli, Italy;
16 e-mail: grosso@dbbm.unina.it; Tel: +39-081-7463140; Fax: +39-081-7464359

18 **Acknowledgements**

19 This work was supported by grants from the University of Naples Federico II, from San
20 Paolo-Banco di Napoli and from Regione Campania. We are deeply indebted to Prof. P.
21 Costanzo for helpful discussions and critical reading of the manuscript, to Dr. F. Romano for
22 flow cytometric analysis and to Dr. L. Prossomariti and Dr. L. Pagano for providing blood
23 samples from thalassemic controls. We are also grateful to Jean Ann Gilder for editing the
24 text.

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

1 Summary

2 Impaired switching from fetal haemoglobin (HbF) to adult globin gene expression leads to
3 hereditary persistence of fetal haemoglobin (HPFH) in adult life. This is of prime interest
4 because elevated HbF levels ameliorate α -thalassaemia and sickle cell anaemia. Fetal
5 haemoglobin levels are regulated by complex mechanisms involving factors linked or not to
6 the α -globin gene locus. To search for factors putatively involved in α -globin gene expression,
7 we examined the reticulocyte transcriptome of three siblings who had different HbF levels
8 and different degrees of α -thalassaemia severity although they had the same α - and β -globin
9 gene cluster genotypes. By mRNA differential display we isolated the cDNA coding for the
10 cold shock domain protein A (CSDA), also known as dbpA, previously reported to interact *in*
11 *vitro* with the α -globin gene promoter. Expression studies performed in K562 and in primary
12 erythroid cells showed an inverse relationship between α -globin and CSDA expression levels.
13 Functional studies performed by Chromatin Immunoprecipitation and reporter gene assays in
14 K562 cells demonstrated that CSDA is able to bind the promoter of the α -globin gene and to
15 suppress its expression. Therefore, our study demonstrates that CSDA is a trans-acting
16 repressor factor of α -globin gene expression and contributes to modulate the HPFH
17 phenotype.

18
19
20 Keywords: Cold Shock Domain, dbpA, HPFH, globin gene expression, α -thalassaemia
21 intermedia. mRNA differential display.

1 Introduction

2 The switch from fetal to adult globin gene expression occurs around birth when fetal globin
3 genes are progressively silenced thereby leading to a gradual decline of fetal haemoglobin
4 (HbF) (Stamatoyannopoulos & Gronsveld, 2001; Oneal *et al*, 2006). Impaired haemoglobin
5 switching leads to the persistence of the expression of fetal globin genes throughout
6 adulthood. This condition is without clinical relevance except when co-inherited with
7 haemoglobinopathies. In fact, HPFH has great therapeutic potential because high levels of
8 HbF reduce the α /non α globin chain imbalance and thus ameliorate the thalassaemic and
9 sickle cell disease phenotypes (Garner *et al*, 2003; Thein, 2008). Consequently, there is a race
10 to clarify the molecular basis of haemoglobin switching and persistence of high HbF levels in
11 the hope of identifying new therapeutic tools for thalassaemia and sickle cell disease.
12 Together with epigenetic factors such as age and gender (Chang *et al*, 1995), several
13 genetic determinants *in cis* to the α -globin gene cluster as well as HPFH quantitative trait
14 loci (QTL) unlinked to the α -globin gene cluster (6q23, 8p, Xp22.2-23) influence HbF
15 levels (Dover *et al*, 1992; Craig *et al*, 1996; Garner *et al*, 2000; Garner *et al*, 2002;
16 Garner *et al*, 2004). However, little is known about the *trans*-acting factors mapping on
17 such external loci, although experimental evidences indicate that HBS1L and MYB,
18 whose genes map on chromosome 6q23, and BCL11A, a zinc-finger protein encoded on
19 chromosome 2p15, may modulate HbF production in adult life (Jiang *et al*, 2006; Menzel
20 *et al*, 2007; Thein *et al*, 2007; Lettre *et al*, 2008; Uda *et al*, 2008, Wahlberg *et al*, 2009)
21 However, the molecular mechanisms by which these factors affect HbF levels are still
22 poorly understood, although recent studies in K562 cells suggest that BCL11A binds to
23 the promoter region of the α -globin gene to form a repressor complex (Chen *et al*, 2009)

1 This study was aimed at identifying other factors putatively involved in modulation of α -
2 globin gene expression. Our findings provide experimental evidence that CSDA expression
3 levels modulate α -globin gene expression in adult life and could lead to innovative gene
4 therapy approaches in haemoglobinopathies.

5 **Materials and Methods**

6 *Patients and hematological data*

7 *Thalassaemia intermedia* was diagnosed in three adult siblings. The eldest brother (patient I-
8 1) underwent splenectomy at the age of 33 years and was on chronic transfusion therapy since
9 his thirties. His two sisters (patients I-2 and I-3) received blood transfusions only occasionally
10 during their pregnancies. Three other thalassaemic patients were enrolled as control in this
11 study (patients C1-C3). Table I lists the haematological and clinical data of all the examined
12 patients.

13 *Genetic analysis of α - and β -globin gene clusters*

14 Screening for α - and β -globin gene mutations and polymorphisms was carried out on DNA
15 obtained from peripheral leucocytes by reverse dot-blot, sequencing, Southern blot or **MLPA**
16 **(multiplex-ligation-dependent probe amplification, MRC-Holland, Amsterdam,**
17 **Nederland, P140-B2 HBA probemix)** analysis as reported elsewhere after local Ethics
18 Committee approval and written informed consent were obtained. The α -globin gene cluster
19 haplotypes were determined as reported previously (Sutton *et al*, 1989).

20
21
22

1 *Genetic analysis of polymorphisms in the BCL11A and MYB/HBS1L regions*

2 **Genotyping of polymorphisms in the BCL11A (SNP rs11886868) (Uda *et al*, 2008) and**

3 **MYB/HBS1L (SNP rs4895441) (Thein *et al*, 2007) regions was carried out on genomic**

4 **DNA from peripheral leucocytes by sequence analysis of amplicons generated with the**

5 **following primers: SNP rs11886868 forward 5'-GAACACGTCACCAGTCTAGA-3',**

6 **reverse 5'-GTAGCACTGTTTCATAGTGGTC-3'; SNP rs4895441 forward 5'-**

7 **GTTGGGATATAGGCCATAGAC-3', reverse 5'-CACGTTCTTATTGTGGTCTAC-3'.**

8
9 *RNA isolation, differential mRNA display, quantitative real-time RT-PCR and Western blot*

10 Total RNA was isolated from peripheral reticulocytes according to a recently described

11 procedure (Petruzzelli *et al*, 2009). Differential mRNA display analysis was performed using

12 the Delta Differential Display kit (BD Biosciences-Clontech, Palo Alto, CA, USA) according

13 to the manufacturer's instructions. Selected bands were cloned in a pGEM T-vector (Promega,

14 Madison, WI, USA) and sequenced on a ABI-Prism 3730 automatic sequencer (Applied

15 Biosystems, Foster City, CA, USA). Comparative sequence searches were performed using

16 the BLAST algorithm (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). cDNA was synthesized using

17 an MMLV reverse transcriptase (Invitrogen, Carlsbad, CA, USA) from 1 μ g of total RNA.

18 Quantitative real-time RT-PCR for *CSDA* and β -globin gene mRNAs was carried out with an

19 iCycler instrument (Bio-Rad Laboratories, Hercules, CA, USA) using the SyBR Green master

20 mix (Bio-Rad Laboratories) and the manufacturer's protocol. Primers were designed using the

21 Primer Express 2.0 program (sequences available on request). All data were normalized using

22 endogenous β_2 -microglobulin mRNA as control (Livak & Schmittgen, 2001; Howe *et al*,

23 2004).

1 CSDA protein levels were evaluated by Western blot with the 4D9 monoclonal antibody
2 against CSDA (Sigma-Aldrich, Saint Louis, MO, USA) using a procedure previously
3 described (Medugno *et al*, 2005).

4 5 *Cell cultures*

6 K562 cells were grown in RPMI 1640 medium supplemented with 10% fetal calf serum
7 (FCS) plus 4 mM glutamine, 10 U/mL penicillin and 10 mg/mL streptomycin at 37°C in a
8 humidified 5% CO₂-containing atmosphere. Cell cultures were kept sub-confluent and
9 transfected for CSDA RNA interference (RNAi) or over-expression analysis.

10 11 *Isolation and culture of primary erythroid cells*

12 CD34⁺ cells were isolated in high purity using MACS magnetic separation system (Miltenyi
13 Biotec GmbH, Bergisch, Germany), according to manufacturer's instructions, from buffy
14 coats of healthy blood donors at our local Transfusion Centre, after informed consent had
15 been obtained. To induce erythroid differentiation, cells were kept for one week in IMDM
16 medium (Mascia Brunelli, Milan, Italy) supplemented with 20% FCS (HyClone, Thermo
17 Scientific, Logan, UT, USA) plus 10 ng/ml Stem Cell Factor (Peprotech, Hamburg,
18 Germany), 1u/ml erythropoietin (Sigma-Aldrich, St. Louis, MO, USA), 1ng/ml interleukin 3
19 (Peprotech), 10 U/mL penicillin (BioWhittaker, Basel, Switzerland) and 10 mg/mL
20 streptomycin (BioWhittaker) at 37°C in a humidified 5% CO₂-containing atmosphere,
21 according to procedures previously described (Migliaccio *et al*, 2002; Ronzoni *et al*, 2008).
22 Cells were stained with anti CD34 phycoeritrin (PE)-conjugated (Pharmigen/Becton
23 Dickinson, San Diego, CA) and anti CD71 fluorescein isothiocyanate (FITC)-conjugated

1 (Pharmigen/Becton Dickinson) for 20 min at room temperature. Then cells were washed and
2 analyzed with a flow cytometer (FACScan Becton Dickinson) (Supporting Information,
3 **Figure 1S**).

4

5 *CSDA RNA interference and over-expression in K562 cell line and in primary erythroid cells*

6 Transfection of K562 cells was performed with Lipofectamine 2000 (Invitrogen) as
7 transfectant agent. The day before transfection, cells were plated into 12- or 6-well plates at a
8 density of 2.5×10^5 /ml in Optimem medium (Invitrogen) according to the manufacturer's
9 instructions. Primary erythroid cells were transfected at day 7 of culture with HiPerFect
10 Transfection Reagent (Qiagen) according to the manufacturer's instructions. The day before
11 transfection, cells were diluted at a density of 3×10^5 /ml in the culture medium containing
12 serum and antibiotics and incubated under normal growth conditions. On the day of
13 transfection cells were plated into 24-well plates at a density of 2×10^5 /ml in the culture
14 medium. *CSDA* RNAi was achieved by transiently transfecting cells with double-stranded
15 small interfering RNAs (stealth™ siRNAs) from Invitrogen. Two different siRNA oligos
16 were used as follows (sense strand shown): siRNA 1
17 GCCUUACCACGUGGGACAGACCUUU; siRNA 2
18 ACCUUUGACCGUCGCUCACGGGUCU; an RNAi negative control (High GC,
19 Invitrogen), certified by the manufacturer to have no matches with mRNAs sequences in
20 GenBank, was used as mock control. Cells were collected 48 h after transfection for total
21 protein or RNA isolation procedures, as previously described, in order to perform Western
22 blot or real-time RT-PCR analysis (Medugno *et al*, 2005; Petruzzelli *et al*, 2009). *CSDA*
23 silencing was verified at the protein level by Western blot with the 4D9 anti-*CSDA* antibody,

1 whereas quantitative real-time RT-PCR analyses were carried out to evaluate variations of
2 *CSDA* and γ -globin gene mRNA levels, using the procedure described above.

3 For over-expression experiments, the full-length cDNA of *CSDA* spanning exons 1-9
4 (*CSDA* isoform *a*, Genbank accession number: NM_003651.4), was cloned in a p3xFLAG-
5 CMV 7.1 plasmid vector (Sigma-Aldrich). K562 cells were transiently transfected with 2 μ g
6 of expression vector p3xFLAG-*CSDA*. Cells were collected 48 h after transfection for total
7 protein or RNA purification in order to perform Western blot or real-time RT-PCR analysis.
8 Western blot analysis was carried out with **the anti-FLAG or with the 4D9 anti-*CSDA***
9 **antibodies** (Sigma-Aldrich) to check exogenous *CSDA* over-expression, whereas quantitative
10 real-time RT-PCR analysis for γ -globin gene mRNAs was performed as above described in
11 order to evaluate effects on fetal globin gene expression.

12 *Plasmid constructs and reporter gene assays*

13 The recombinant plasmid -268 γ -luc containing the firefly luciferase reporter gene was
14 obtained by cloning the -268/-1 bp fragment of the proximal γ -globin gene promoter in a
15 pGL4 vector (Promega). The **HPFH point mutation G γ -202 C \rightarrow G as well as a deletion**
16 **mutant of the homopyrimidine tract (del-204/-194) reported as the putative recognition**
17 **site for *CSDA* were generated by mutagenesis of the wild-type recombinant plasmid -**
18 **268 γ -luc using the QuikChange Site-Directed Mutagenesis kit (Stratagene, Cedar Creek,**
19 **TX, USA), according to manufacturer's recommendations.** To normalize the luciferase
20 assay, 0.1 μ g of the pRL-CMV vector (Promega) coding for the *Renilla* luciferase was
21 transiently co-transfected with 0.9 μ g of the -268 γ -luc constructs in K562 cells at the density
22 of 1.5×10^5 cells per well in 24-well dishes. The pGL4 and p3XFlag empty vectors (0.9 μ g)

1 were used as negative controls whereas the pCMV-luc (0.9 μ g) was the positive control for
2 the assay. To evaluate the transcriptional role of CSDA, cells were co-transfected with 0.5, 1
3 or 2 μ g of the expression vector p3xFLAG-CSDA described above. Luciferase activity was
4 measured 48 h after transfection with the Dual-Luciferase Reporter Assay System (Promega)
5 on a 20/20ⁿ luminometer (Turner Biosystems, Sunnyvale, CA, USA), according to the
6 protocols of the manufacturers.

8 *Chromatin ImmunoPrecipitation (ChIP) analysis*

9 Chromatin from K562 cells (12.5×10^6) was purified and immunoprecipitated with anti-CSDA
10 and immunoglobulin G (IgG) antibodies (Sigma) as described previously (Medugno *et al*,
11 2005). Antigen-DNA complexes cross-links were reversed over-night at 65°C after addition
12 of RNase (10 ng/ μ l) in a 100 μ l final volume. SDS (0,5% final concentration) and proteinase
13 K (0,6 μ g/ μ l final concentration) were then added and the mix was incubated at 50°C for 3h
14 followed by phenol-chloroform extraction and EtOH precipitation. DNA pellets were
15 resuspended in 30 μ l of distilled water. Real-time PCR detection of the proximal G β -globin
16 gene promoter region was performed in a 20 μ l reaction mix containing 1X SYBR Green I
17 PCR Master mix (Bio-Rad Laboratories), 20 μ M of each primer mix (gamma For: 5'-
18 ACTACAGGCCTCACTGGAG-3' and gamma Rev: 5'-TGGAAGTCTGAAGGGTGC-3'),
19 1/10 volume of purified DNA and nuclease-free water on an iCycler instrument (Bio-Rad
20 Laboratories). The thermal profile consisted of 1 cycle at 95°C for 3 minutes followed by 40
21 cycles at 95°C for 15 seconds, 58°C for 30 seconds, 72°C for 20 seconds. As a negative
22 control, PCR was done using *GAPDH*-specific primers (GAPDH For: 5'-
23 GGTCGTATTGGGCGCCTGGTCACCA-3' and GAPDH Rev: 5'-

1 CACACCCATGACGAACATGGGGGC-3') to show that the anti-CSDA 4D9 antibody did
2 not immunoprecipitate nonspecific genomic DNA. Real time PCR data analysis followed the
3 methodology previously described (Chakrabarti *et al*, 2002).

4 5 **Results**

6 *Genetic analysis of α - and β -globin gene clusters*

7 We identified three siblings affected by different severity degrees of β -thalassaemia
8 *intermedia*. Characterization of β -thalassaemic defects revealed β^+ IVSI-6 (T→C) in
9 homozygosis in all of them. For both alleles, the mutation was associated to Orkin's haplotype
10 VI (Orkin *et al*, 1982). To search for genetic determinants of different levels of β -globin gene
11 expression we first carried out an extensive sequence analysis of putative regulatory regions
12 within the β -globin gene cluster potentially involved in these mechanisms. We examined
13 sequence variations in binding sites for transcriptional factors or in polymorphic sequences
14 within the hypersensitive site-2 (HS-2) of the locus control region (LCR) (Öner *et al*, 1992)
15 and the promoter region of the G β and β -globin genes (Berg *et al*, 1991; Chang *et al*, 1995;
16 Pissard & Beuzard, 1994; Lu & Steinberg 1996; Chan *et al*, 2007), all of which had been
17 reported to be involved in modulation of globin gene expression (Table II). We found the
18 same β -globin gene cluster genotype in the three patients and were able to exclude known
19 HPFH mutations and hypothetical parental germ line rearrangements within this locus.

20 **In all the three siblings Southern blot and MLPA analysis excluded deletional forms of**
21 **β -thalassemia as well as triplication or quadruplication within the β -globin gene cluster**
22 **(Supporting information, Table 1S). Non-deletional β -thalassemic mutations were**
23 **excluded by sequence analysis of the genomic regions encompassing the entire α_1 - and α_2 -**

1 **globin genes (data not shown). Therefore, the variability of the clinical phenotype could**
2 **not be associated with α -globin gene imbalance for any of these patients.**

3
4 *Genetic analysis of polymorphisms in the BCL11A and MYB/HBS1L regions*

5 **Analysis of the polymorphisms in the BCL11A and MYB/HBS1L associated with high**
6 **HbF levels revealed the same heterozygous genotype (C/T for SNP rs11886868 and A/G**
7 **for SNP rs4895441) for all the three siblings, thus excluding involvement of these two**
8 **loci in the clinical variability of their α -thalassemic conditions (Supporting information,**
9 **Figure 2S) .**

10
11 *Identification of CSDA as a putative modulator for α -globin gene expression*

12 Based on the foregoing data, it was conceivable that HbF level variations were mostly
13 associated to genetic determinants not linked to the α -globin gene cluster. To explore this
14 hypothesis, we analyzed the reticulocyte transcriptome from the three patients using a
15 differential mRNA display approach. We found several bands that were differentially
16 expressed in the sample from patient I-1, the more severely affected sibling, compared with
17 his two sisters (Fig. 1). Selected bands were cloned in a pGEM T-vector and sequenced. A
18 complete homology was found between clones originated from two different up-regulated
19 bands of patient I-1 and the cDNA sequence corresponding to the Cold Shock Domain Protein
20 A (CSDA) also known as DNA-binding protein A (dbpA), reported to act as a repressor factor
21 for several genes (Coles *et al*, 1996; Coles *et al*, 2002). To confirm the differential display
22 results CSDA expression levels were evaluated by real time PCR analysis in the three siblings
23 and three other unrelated homozygotes for the IVSI-6 mutation affected by varying degrees of

1 clinical conditions, as controls (Table I, patients C1-C3) (Fig. 2A). Also, protein levels of
2 CSDA were found consistent with these findings (Fig. 2B). Interestingly, evaluation of α -
3 globin gene mRNA levels revealed an inverse relationship with the CSDA expression levels
4 (Fig. 2C) comparable to that found for peripheral blood values of HbF (Fig. 2D).

5
6 *Effects of CSDA knock-down and over-expression on α -globin gene expression in K562 and*
7 *primary erythroid cells*

8 In order to examine the role played by CSDA in the regulation of α -globin gene expression we
9 first performed transient RNAi in the K562 cell line with two siRNAs specific for CSDA
10 mRNA sequences. Cells were collected 48 h after transfection and subjected to total protein
11 and RNA extraction procedures. Knock-down of CSDA was verified at protein and mRNA
12 levels by Western blot and real-time PCR procedures, respectively (Fig. 3). The effects of
13 CSDA silencing on α -globin gene expression were evaluated by quantitative real-time PCR in
14 cells treated with specific or mock siRNAs. The expression level of the α -globin gene was on
15 average two-fold higher when CSDA mRNA levels were reduced to about 40-50% (Fig. 3A
16 and 3B).

17 To verify the quantitative effects of CSDA levels on the regulation of α -globin gene
18 expression, we over-expressed a recombinant FLAG-tagged full-length CSDA protein (40
19 kDa) in K562 cells. CSDA over-expression was evaluated in total protein extracts by Western
20 blot with anti-FLAG **and anti-CSDA (4M9) antibodies**, whereas the effects on α -globin gene
21 expression were evaluated by real-time PCR analysis. As shown in Figure 4A and 4B,
22 increased CSDA levels resulted in a substantial reduction (around 50%) of α -globin gene
23 mRNA levels versus untreated cells.

24 Effects of CSDA knock-down and over-expression on α -globin gene expression were also

1 evaluated in primary erythroid cells in which similar results were found (Fig. 3C, 3D, 4C and
2 4D).

3
4 *CSDA interacts with the region at -200 bp of the γ -globin gene promoter*

5 To verify that CSDA plays a transcriptional role on the γ -globin gene promoter, we performed
6 reporter gene assays by transiently transfecting into K562 cells a plasmid construct containing
7 the firefly luciferase gene driven by the 268-bp fragment of the γ -globin gene promoter (nt
8 -268/-1) where a putative binding site for CSDA had already been described (Horwitz *et al*,
9 1994). When K562 cells were co-transfected with increasing amounts of a CSDA expression
10 vector (p3XFLAG-CSDA), the luciferase activities were decreased in a dose-dependent
11 manner (Fig. 5), thus providing further evidence that CSDA acts as a repressor factor of γ -
12 globin gene expression. **We also examined the effects of CSDA expression levels on γ -
13 globin gene promoter constructs bearing the -202 C→G HPFH point mutation as well as
14 on a deletion mutant of the homopyrimidine tract (del-204/-194), previously identified as
15 a putative binding site of CSDA (Ulrich *et al*, 1992; Horwitz *et al*, 1994; Bacolla *et al*,
16 1995). Luciferase activities of these constructs were significantly higher compared to
17 those found for the wild-type construct, thus confirming that this region is involved in
18 the CSDA repression activity on the γ -globin gene promoter (Fig. 5).**

19 To verify that CSDA is capable of binding to the proximal promoter of the γ -globin gene we
20 performed CHIP analysis on chromatin from K562 cells immunoprecipitated with anti-CSDA
21 antibody. Primers specific for the γ -promoter revealed a strong enrichment compared with
22 IgG controls (Fig. 6). These data provide experimental evidence that CSDA interacts *in vivo*
23 with the -200 bp region of the γ -globin gene.

1 Taken as a whole, our results demonstrate that CSDA suppresses β -globin gene expression at
2 least in part at the transcriptional level.

4 **Discussion**

5 We examined the reticulocyte transcriptome of three siblings who differed in HbF levels and
6 in β -thalassaemia severity although they had the same α - and β -globin gene cluster
7 genotypes. The aim of the study was to identify factors that modulate the level of β -globin
8 gene expression. The two sisters of the family were affected by a milder form of *thalassaemia*
9 *intermedia* and had higher levels of HbF than their elder brother who was affected by a more
10 severe form of the disorder and was transfusion-dependent. Nevertheless, all the three siblings
11 were homozygous for the α^+ IVS I-6 (T→C) mutation and all had a normal set of β -globin
12 genes. Therefore, the different severity of *thalassaemia intermedia* was clearly related to the
13 different HbF levels.

14 Variability in the severity of the disease has been extensively reported in α^+ IVS I-6 (T→C)
15 homozygotes, mainly relied on variations in β -globin chain production. Molecular studies also
16 indicate that such variations cannot be attributed to regulatory elements within the β -globin
17 gene cluster, suggesting a major role for transacting factors able to modulate β -globin gene
18 expression (Efremov *et al*, 1994; Ho *et al*, 1998). Therefore, this type of β -thalassaemia
19 condition provides an interesting experimental model to evaluate the effects of genetic
20 molecular modifiers of disease severity in β -thalassaemia syndromes.

21 In our study, analysis of genetic loci linked to the β -globin gene cluster and putatively
22 involved in regulation of globin gene expression showed the same genetic background for all
23 the three siblings, and this drew our attention to factors not linked to the β -globin gene cluster.

1 **Firstly, we genotyped two polymorphisms mapping in the BCL11A gene region or in the**
2 **region between MYB and HBS1L, respectively, and reported to be associated with**
3 **elevated HbF levels in adulthood. Results showed the same allelic sequence variations in**
4 **the three siblings, thus excluding involvement of these two loci in the clinical variability**
5 **of their α -thalassemic conditions.**

6 **Therefore, in order to search for other potential genetic modifiers, we examined the**
7 **reticulocyte transcriptome of the three patients using a mRNA differential display approach.**

8 Among bands differentially expressed our attention was focused on those corresponding to
9 factors potentially involved in gene expression or in signal transduction mechanisms.

10 Particularly, in clones originated from two different up-regulated bands of the more severely
11 affected patient (subject I-1) we found homology with the cDNA of CSDA, a cold shock
12 domain (CSD) protein.

13 CSD proteins have been reported to be involved in transcriptional activation and
14 repression, as well as in post-transcriptional mechanisms of gene expression regulation,
15 including mRNA packaging, transport, localization and stability (Landsman, 1992; Coles *et al*
16 *al*, 1996; Kudo *et al*, 1995; Diamond *et al*, 2001; Kloks *et al*, 2002; Coles *et al*, 2002, Coles *et*
17 *al*, 2004).

18 Although a general consensus binding site has not been established for CSD proteins, they
19 are generally considered to bind to CT-rich sequences that can generate H-DNA structures.

20 (Kloks *et al*, 2002). HPFH point mutations (G α -202 C \rightarrow G or C \rightarrow T) occurring in a region with
21 homopyrimidine tracts reported to be critical for α -globin gene repression are thought to
22 abolish the high-affinity binding sites for CSDA by disrupting the formation of the intra-
23 molecular triplex and thus the H-DNA structure (Ulrich *et al*, 1992; Horwitz *et al*, 1994;

1 Bacolla *et al*, 1995). It was thus proposed that reduced binding of CSDA in the -200 G \square
2 promoter region could lead to persistent expression of \square -globin genes in adult life (Horwitz *et*
3 *al*, 1994).

4 Our differential display results were in agreement with this model but also indicated that
5 modulation of CSDA expression levels could contribute to regulate fetal globin gene
6 expression. The role of CSDA on \square -globin gene regulation was therefore investigated by
7 expression studies in human erythroleukemia K562 cells and in primary CD71⁺ erythroid cells
8 and confirmed that CSDA plays a repression role in the transcriptional regulation of \square -globin
9 gene expression. In fact, in both these cell systems, down- and up-modulation of CSDA levels
10 consistently corresponded to variations of \square -globin gene expression levels: CSDA knock-
11 down induced by RNAi resulted in significantly increased expression of \square -globin genes,
12 whereas its over-expression was associated with reduced \square -globin gene mRNA levels.
13 The putative mechanism by which CSDA modulates \square -globin gene expression was
14 investigated in K562 cells by gene reporter assays **on wild-type and mutant constructs of**
15 **the \square -globin gene promoter region suspected to bind CSDA, providing experimental**
16 **evidence that CSDA acts as repressor of \square -globin gene expression. Furthermore,**
17 **chromatin immunoprecipitation analysis on K562 cells showed that CSDA interacts *in***
18 ***vivo* with this promoter region. In this way we were able to demonstrate that CSDA**
19 **modulates \square -globin gene expression at least in part at the transcriptional level.**

20 The results of CSDA knockdown both in K562 and in primary erythroid cells drive
21 attention on CSDA as a potential molecular target for treatment of haemoglobinopathies
22 because it is conceivable that this treatment could induce persistence expression of HbF in
23 adult life. However, experiments of stable interference of CSDA expression are required to

1 explore the feasibility of such approach and to evaluate the long-term effects of this treatment
2 in an erythroid environment.

3 Taken altogether our results indicate that CSDA contributes to regulation of fetal globin
4 gene expression and sheds further light on the complex molecular mechanisms involved in
5 globin gene switching. In fact, CSDA dose-dependent effects on γ -globin gene expression
6 suggest that other factors associated with HPFH could indirectly affect HbF production
7 through modulation of CSDA expression levels.

8 Our data from the differential mRNA display analysis also shed light on other factors
9 potentially involved in modulation of HbF levels. In fact, among bands differentially
10 displayed, we identified the cDNA for a GDP-dissociation inhibitor (GDI) for Rho GTP-
11 binding proteins, a class of proteins that play a primary role in modulating the activity of
12 GTPases (Etienne-Manneville & Hall, 2002; DerMardirossian & Bokoch, 2005; Heasman &
13 Ridley 2008). Expression levels for this factor resulted to be increased in the more severely
14 affected patient (subject I-1) respect to his two sisters (data not shown). Although preliminary,
15 these data appear quite intriguing, particularly at the light of recent reports indicating that
16 HBS1L, a putative member of the GTPases superfamily, could contribute to modulate HbF
17 levels in adults (Thein *et al*, 2007; Lettre *et al*, 2008; Wahlberg *et al*, 2009).

18 In conclusion, our study helps to elucidate the complex network of factors regulating
19 globin gene expression and could eventually pave the way to promising novel RNA-
20 interference-based therapeutics for haemoglobinopathies.

21

22

23

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 18 -

1 **Authorship contribution:** R.P., S.G., S. P. and R.S. performed research and collected and
2 analyzed data; G.A. acquired patients, collected and analyzed data and revised the manuscript;
3 R.D.C., G.dU. and M.A. collected and analyzed data; P.I. applied for funding, supervised
4 research and revised the manuscript; M.G. designed research, analyzed data and wrote and
5 revised the manuscript.

For Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 **References**

- 2 1. Bacolla, A., Ulrich, M.J., Larson, J.E., Ley, T.J. & Wells, R.D. (1995) An
3 intramolecular triplex in the human α -globin 5'-flanking region is altered by point
4 mutations associated with hereditary persistence of fetal hemoglobin. *Journal of*
5 *Biological Chemistry*, **270**, 24556-24563.
- 6 2. Berg, P.E., Mittelman, M., Elion, J., Labie, D. & Schechter, A.N. (1991) Increased
7 protein binding to a -530 mutation of the human α -globin gene associated with
8 decreased α -globin synthesis. *American Journal of Hematology*, **36**, 42-47.
- 9 3. Chakrabarti, S.K., James, J.C. & Mirmira, R.G. (2002) Quantitative assessment of
10 gene targeting in vitro and in vivo by the pancreatic transcription factor, Pdx1.
11 *Journal of Biological Chemistry*, **277**, 13286-13293.
- 12 4. Chan, P.K., Ma, E.S., Philipsen, S. & Tan-Un, K.C. (2007) The study of sequence
13 configuration and functional impact of the (AC)_n(AT)_xT_y motif in human α -globin
14 gene promoter. *American Journal of Hematology*, **82**, 342-348.
- 15 5. Chang, Y.C., Smith, K.D., Moore, R.D., Serjeant, G.R. & Dover, G.J. (1995) An
16 analysis of fetal hemoglobin variations in sickle cell disease: the relative contributions
17 of the X-linked factor, α -globin haplotypes, α -globin gene number, gender and age.
18 *Blood*, **85**, 1111-1117.
- 19 6. Chen, Z., Luo, H., Steinberg, M.H. & Chui, D.H.K. (2009). BCL11A represses HBG
20 transcription in K562 cells. *Blood Cells Molecules and Diseases*, **42**, 144-149.
- 21 7. Coles, L.S., Diamond, P., Occhiodoro, F., Vadas, M.A. & Shannon, M.F. (1996) Cold
22 shock domain proteins repress transcription from the GM-CSF promoter. *Nucleic*
23 *Acids Research*, **24**, 2311-2317.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 1 8. Coles, L.S., Diamond, P., Lambrusco, L., Hunter, J., Burrows, J. & Mathew, A.
2 (2002). A novel mechanism of repression of the vascular endothelial growth factor
3 promoter, by single strand DNA binding cold shock (Y-box) proteins in normoxic
4 fibroblasts. *Nucleic Acids Research*, 30, 4845-4854.
 - 5 9. Coles, L.S., Bartley, M.A., Bert, A., Hunter, J., polyak, S., Diamond, P., Vadas, M.A.
6 & Goodall, G.J. (2004) A multi-protein complex containing cold shock domain (Y-
7 box) and polypyrimidine tract binding proteins forms on the vascular endothelial
8 growth factor mRNA. Potential role in mRNA stabilization. *European Journal of*
9 *Biochemistry*, **271**, 648-660.
 - 10 10. Craig, J.E., Rochette, J., Fisher, C.A., Weatherall, D.J., Marc, S., Lathrop, G.M.,
11 Demanais, F. & Thein, S.L. (1996) Dissecting the loci controlling fetal haemoglobin
12 production on chromosomes 11p and 6q by the regressive approach. *Nature Genetics*,
13 **12**, 58-64.
 - 14 11. DerMardirossian, C. & Bokoch, G.M. (2005) GDIs: central regulatory molecules in
15 Rho GTPase activation. *Trends in Cell Biology*, **15**, 356-363.
 - 16 12. Diamond, P., Shannon, M.F., Vadas, M.A. & Coles, L.S. (2001) Cold Shock Domain
17 factors activate the granulocyte-macrophage colony-stimulating factor promoter in
18 stimulated Jurkat T cells. *Journal of Biological Chemistry*, **276**, 7943-7951.
 - 19 13. Dover, G.J., Smith, K.D., Chang, Y.C., Purvis, S., Mays, A., Meyers, D.A., Sheils &
20 C., Serjeant, G. (1992) Fetal hemoglobin levels in sickle cell disease and normal
21 individuals are partially controlled by an X-linked gene located at Xp 22.2. *Blood*, **80**,
22 816-824.

- 1
2
3 14. Efremov, D.G., Dimovski, A.J., Baysal, E., Ye, Z., Adekile, A.D., Ribeiro, M.L.S.,
4
5 2 Schirilo, G., Altay, C., Gürgey, A., Efremov, G.D. & Huisman, T.H.J. (1994) Possible
6
7 3 factors influencing the haemoglobin and fetal haemoglobin levels in patients with α -
8
9 4 thalassemia due to a homozygosity for the IVS-I-6 (T→C) mutation. *British Journal of*
10
11 5 *Haematology*, **86**, 824-830.
12
13
14
15 6 15. Etienne-Manneville, S. & Hall, A. (2002) Rho GTPases in cell biology. *Nature*, **420**,
16
17 7 629-635.
18
19
20 8 16. Garner, C., Tatu, T., Reittie, J.E., Littlewood, T., Darley, J., Cervino, S., Farrall, M.,
21
22 9 Kelly, P., Spector, T.D. & Thein, S.L. (2000) Genetic influences on F cells and other
23
24 10 hematologic variables: a twin heritability study. *Blood*, **95**, 342-346.
25
26
27 11 17. Garner, C.P., Tatu, T., Best, S., Creary, L. & Thein, S.L. (2002) Evidence of genetic
28
29 12 interaction between the α -globin complex and chromosome 8q in the expression of
30
31 13 fetal hemoglobin. *American Journal of Human Genetics*, **70**, 793-799.
32
33
34 14 18. Garner, C., Dew, T.K., Sherwood, R., Rees, D. & Thein, S.L. (2003) Heterocellular
35
36 15 hereditary persistence of fetal haemoglobin affects the haematological parameters of
37
38 16 α -thalassemia trait. *British Journal of Haematology*, **123**, 353-358.
39
40
41 17 19. Garner, C., Silver, N., Best, S., Menzel, S., Martin, C., Spector, T.D. & Thein, S.L.
42
43 18 (2004) Quantitative trait locus on chromosome 8q influences the switch from fetal to
44
45 19 adult hemoglobin. *Blood*, **104**, 2184-2186.
46
47
48 20 20. Heasman, S.J. & Ridley, A.J. (2008) Mammalian Rho GTPases: new insights into
49
50 21 their functions from in vivo studies. *Nature Reviews Molecular Cell Biology*, **9**, 690-
51
52 22 701.
53
54
55
56
57
58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- 1 21. Ho, P.J., Hall, G.W., Luo, L.Y., Weatherall, D.J. & Thein, S.L. (1998) Beta-
2 thalassemia intermedia: is it possible consistently to predict phenotype from
3 genotype? *British Journal of Haematology*, **100**, 70-78.
- 4 22. Horwitz, E.M., Maloney, K.A. & Ley, T.J. (1994) A human protein containing a
5 “cold shock” domain binds specifically to H-DNA upstream from the human β -globin
6 genes. *Journal of Biological Chemistry*, **269**, 14130-14139.
- 7 23. Howe, J.G., Crouch, J., Cooper, D. & Smith, B.R. (2004) Real-time quantitative
8 reverse transcription-PCR for cyclin D1 mRNA in blood, marrow and tissue
9 specimens for diagnosis of mantle cell lymphoma. *Clinical Chemistry*, **50**, 80-87.
- 10 24. Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector T.D. &
11 Thein, S.L. (2006) cMYB is involved in the regulation of fetal haemoglobin
12 production in adults. *Blood*, **108**, 1077-1083.
- 13 25. Kloks, C.P., Spronk, C.A., Lasonder, E., Hoffmann, A., Vuister, G.W., Grzesiek, S.
14 & Hilbers, C.W. (2002) The solution structure and DNA-binding properties of the
15 cold-shock domain of the human Y-box protein YB-1. *Journal of Molecular Biology*,
16 **316**, 317-26.
- 17 26. Kudo, S., Mattei, M.G. & Fukuda, M. (1995) Characterization of the gene for dbpA,
18 a family member of the nucleic-acid-binding proteins containing a cold-shock domain.
19 *European Journal of Biochemistry*, **231**, 72-82.
- 20 27. Landsman, D. (1992) RNP-1, an RNA binding motif, is conserved in the DNA-
21 binding cold shock domain. *Nucleic Acids Research*, **20**, 2861-2864.
- 22 28. Lettre, G., Sankaran, V.G., Bezerra, M.A., Araújo A.S., Uda, M., Sanna, S., Cao, A.,
23 Schlessinger, D., Costa, F.F. Hirschhorn, J.N. & Orkin, S.H. (2008) DNA

- 1 polymorphisms at the BCL11A, HBS1L-MYB, and γ -globin loci associate with fetal
2 hemoglobin levels and pain crises in sickle cell disease. *Proceedings of the National*
3 *Academy of Sciences USA*, **105**, 11869-11874.
- 4 29. Livak, K.J. & Schmittgen, T.D. (2001) Analysis of relative gene expression data
5 using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, **25**, 402-408.
- 6 30. Lu, Z.H. & Steinberg, M.H. (1996) Fetal hemoglobin in sickle cell anemia: relation
7 to regulatory sequences cis to the γ -globin gene. *Blood*, **87**, 1604-1611.
- 8 31. Medugno, L., Florio, F., De Cegli, R., Grosso, M., Lupo, A., Costanzo, P. & Izzo, P.
9 (2005) The Kruppel-like zinc-finger protein ZNF224 represses aldolase A gene
10 transcription by interacting with the KAP-1 co-repressor protein. *Gene*, **359**, 35-43.
- 11 32. Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Health, S., Foglio, M.,
12 Zelenika, D., Boland, A., Rooks, H., Best, S., Spector, T.D., Farrall, M., Lathrop, M.
13 & Thein, S.L. (2007) A QTL influencing F cell production maps to a gene encoding a
14 zinc-finger protein on chromosome 2p15. *Nature Genetics*, **39**, 1197-1199.
- 15 33. Migliaccio, G., Di Pietro, R., di Giacomo, V., Di Baldassarre, A., Migliaccio, A.R.,
16 Maccioni, L., Galanello, R. & Papayannopoulou, T. (2002) In vitro mass production
17 of human erythroid cells from the blood of normal donors and of thalassemic patients.
18 *Blood Cells Molecules and Diseases*, **28**, 169-180.
- 19 34. Oneal, P.A., Gantt, N.M., Schwartz, J.D., Bhanu, N.V., Lee, Y.T., Moroney, J.W.,
20 Reed, C.H., Schechter, A.N., Luban, N.L. & Miller, J.L. (2006) Fetal hemoglobin
21 silencing in humans. *Blood*, **108**, 2081-2086.
- 22 35. Öner, C., Dimovski, A.J., Altay, Ç., Gürgey, A., Gu, Y.C., Huisman, T.H. &
23 Lanclos, K.D. (1992) Sequence variations in the 5' hypersensitive site-2 of the locus

- 1 control region of α^S chromosomes are associated with different levels of fetal globin in
2 hemoglobin S homozygotes. *Blood*, **79**, 813-819.
- 3 36. Orkin, S.H., Kazazian, H.H., Antonarakis, S.E., Goff, S.C., Boehm, C.D., Waber,
4 P.G., Sexton, J.P., Ostrer, H., Fairbanks, V.F. & Chakravarti, A. (1982) Linkage of α -
5 thalassemia mutations and α -globin gene polymorphisms with DNA polymorphisms in
6 human α -globin gene cluster. *Nature*, **296**, 627-631.
- 7 37. Petruzzelli, R., Palumbo, I., Gaudino, S., Scalia, G., del Vecchio, L., Izzo, P. &
8 Grosso, M. (2009) An optimized method to separate reticulocytes from peripheral
9 blood for molecular analysis. *International Journal of Laboratory Hematology*, **31**,
10 359-363.
- 11 38. Pissard, S. & Beuzard, Y. (1994) A potential regulatory region for the expression of
12 fetal hemoglobin in sickle cell disease. *Blood*, **84**, 331-338.
- 13 39. Ronzoni, L., Bonara, P., Rusconi, D., Frugoni, C., Libani, I. & Cappellini, M.D.
14 (2008) Erythroid differentiation and maturation from peripheral CD34⁺ cells in liquid
15 culture: Cellular and molecular characterization. *Blood Cells Molecules and Diseases*,
16 **40**, 148-155.
- 17 40. Sutton, M., Bouhassira, E.E. & Nagel, R.L. (1989) Polymerase chain reaction
18 amplification applied to the determination of α -like globin cluster haplotype.
19 *American Journal of Hematology*, **32**, 66-69.
- 20 41. Stamatoyannopoulos G. & Gronsveld, F. (2001) Hemoglobin switching. In: *The*
21 *Molecular Basis of Blood Diseases* (ed. by G. Stamatoyannopoulos, P.W. Majerus,
22 R.M. Perlmutter & H. Varmus), Saunders, Philadelphia, PA.

- 1
2
3 1 42. Thein, S.L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., Silver, N., Gerovasilli,
4
5 2 A., Ping, C., Yamaguchi, M., Wahlberg, K., Ulug, P., Spector, T.D., Garner, C.,
6
7 3 Matsuda, F., Farrall, M. & Lathrop, M. (2007) Intergenic variants of HBS1L-MYB are
8
9 4 responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal
10
11 5 hemoglobin levels in adults. *Proceedings of the National Academy of Sciences USA*,
12
13 6 **104**, 11346-11351.
14
15
16
17 7 43. Thein, S.L. (2008) Genetic modifiers of the α -haemoglobinopathies. *British Journal*
18
19 8 *of Haematology*, **141**, 357-366.
20
21
22 9 44. Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G.,
23
24 10 Busonero, F., Maschio, A., Albai, G., Piras, M.G., Sestu, N., Lai, S., Dei, M., Mulas,
25
26 11 A., Crisponi, L., Naitza, S., Asunis, I., Deiana, M., Nagaraja, R., Perseu, L., Satta, S.,
27
28 12 Cipollina, M.D., Sollaino, C., Moi, P., Hirschhorn, J.N., Orkin, S.H., Abecasis, G.R.,
29
30 13 Schlessinger, D., & Cao, A. (2008) Genome-wide association study shows BCL11A
31
32 14 associated with persistent fetal hemoglobin and amelioration of the phenotype of α -
33
34 15 thalassemia. *Proceedings of the National Academy of Sciences USA*, **105**, 1620-1625.
35
36
37 16 45. Ulrich, M.J., Gray, W.J. & Ley, T.J. (1992) An intramolecular DNA triplex is
38
39 17 disrupted by point mutations associated with hereditary persistence of fetal
40
41 18 hemoglobin. *Journal of Biological Chemistry*, **267**, 18649-18658.
42
43
44 19 46. Wahlberg, K., Jiang, J., Rooks, H., Jawaid, K., Matsuda, F., Yamaguchi, M.,
45
46 20 Lathrop, M., Thein S.L. & Best, S. (2009) The HBS1L-MYB intergenic interval
47
48 21 associated with elevated HbF levels shows characteristics of a distal regulatory region
49
50 22 in erythroid cells. *Blood*, **114**, 1254-1262.
51
52
53
54
55
56
57
58
59
60

1
2
3 **1 Figure Titles and Legends**
4

5
6 2

7
8 3 **Figure 1: Identification of transcripts differently expressed in peripheral**
9
10 4 **reticulocytes isolated from the examined patients.** Silver stained differential mRNA
11
12 5 display electropherogram and family tree. The arrows indicate bands differently
13
14 6 expressed among the three patients. **The asterisk indicates the two bands from which**
15
16 7 **clones corresponding to the cDNA of CSDA were originated.**
17
18
19

20 8
21
22 9 **Figure 2: Correlation between CSDA mRNA levels and fetal globin gene expression.**
23

24
25 10 (A) Quantitative real-time PCR of CSDA mRNA levels isolated from peripheral
26
27 11 reticulocytes; (B) Western blot analysis with anti-CSDA or anti- β -actin antibodies on
28
29 12 protein extracts isolated from peripheral reticulocytes; (C) Quantitative real-time PCR of
30
31 13 β -globin gene mRNA levels isolated from peripheral reticulocytes; (D) HbF values
32
33 14 detected in the peripheral blood. Quantitative real-time PCR data are represented as mean
34
35 15 \pm standard error of the mean (SEM).
36
37
38
39

40 16
41
42 17 **Figure 3: Knock-down of CSDA in K562 and in primary erythroid cells and**
43
44 18 **evaluation of the corresponding variations of β -globin mRNA levels.** (A) Quantitative
45
46 19 real-time PCR analysis of CSDA and β -globin gene mRNA levels in K562 cells
47
48 20 transfected with CSDA siRNA 1, CSDA siRNA 2 or a siRNA negative control (mock);
49
50 21 (B) Western blot analysis with anti-CSDA or anti-tubulin antibodies on protein extracts
51
52 22 from transfected K562 cells; (C) Quantitative real-time PCR analysis of CSDA and β -
53
54 23 globin gene mRNA levels in primary erythroid cells transfected with CSDA siRNA 1 or a
55
56
57
58
59
60

1 siRNA negative control (mock); (D) Western blot analysis with anti-CSDA or anti- β -
2 actin antibodies on protein extracts from transfected primary erythroid cell. Data are
3 expressed as mean \pm SEM.

4
5 **Figure 4: CSDA over-expression in K562 and in primary erythroid cells and**
6 **evaluation of the corresponding variations of β -globin mRNA levels.** (A) Quantitative
7 real-time PCR analysis of β -globin gene mRNA levels in K562 cells transfected with a
8 3xFLAG-CSDA expression plasmid vector (p3xFLAG-CSDA) or with a 3xFLAG empty
9 plasmid vector (p3xFLAG) used as negative control; (B) Western blot analysis with anti-
10 FLAG, **anti-CSDA** or anti- β -actin antibodies on protein extracts from transfected K562
11 cells; (C) Quantitative real-time PCR analysis of β -globin gene mRNA levels in primary
12 erythroid cells transfected with a 3xFLAG-CSDA expression plasmid vector (p3xFLAG-
13 CSDA) or with a 3xFLAG empty plasmid vector (p3xFLAG) used as negative control;
14 (D) Western blot analysis with anti-FLAG, **anti-CSDA** or anti- β -actin antibodies on
15 protein extracts from transfected primary erythroid cells. Data are represented as mean \pm
16 SEM.

17
18 **Figure 5: CSDA suppresses transcriptional activity of the β -globin gene in K562 cells.**
19 **(A) Schematic representation of the reporter plasmid constructs in which position of**
20 **the homopyrimidine tract (pyr) is indicated with a white rectangle: -268 β -luc (wild-**
21 **type β -globin gene promoter), -202 C \rightarrow G β -luc (asterisk indicates position of the -202**
22 **C \rightarrow G HPFH mutation), del-204/194 β -luc (which lacks the entire homopyrimidine**
23 **tract); (B) Relative luciferase activities measured 48 h after cotransfection in K562**

1 cells of reporter plasmid constructs with a p3XFlag-empty vector (lane 1) and with
2 increasing amounts (0.5, 1 or 2 μ g) of the expression vector p3XFlag-CSDA (lanes 2-
3 4). pCMV-Luc and pGL4 were used as positive and negative controls, respectively;
4 pRL-CMV plasmid was used to normalize the results. All luciferase results are
5 reported as relative light units (RLU). Data are represented as mean \pm SEM.

6
7 Figure 6: **CSDA interacts with the proximal promoter region of β -globin gene in**
8 **K562 cells.** (A) Schematic representation of the β -globin gene locus and of the β -globin
9 gene promoter fragment examined in this study; (B) 1.5% ethidium bromide-stained
10 agarose gel-electrophoresis of β -globin gene promoter and GAPDH amplicons,
11 respectively, from K562 DNA immunoprecipitated with anti-CSDA or IgG antibodies;
12 (C) Quantitative real-time PCR of the immunoprecipitated samples compared to
13 normalized IgG control values. Data are represented as mean \pm SEM.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

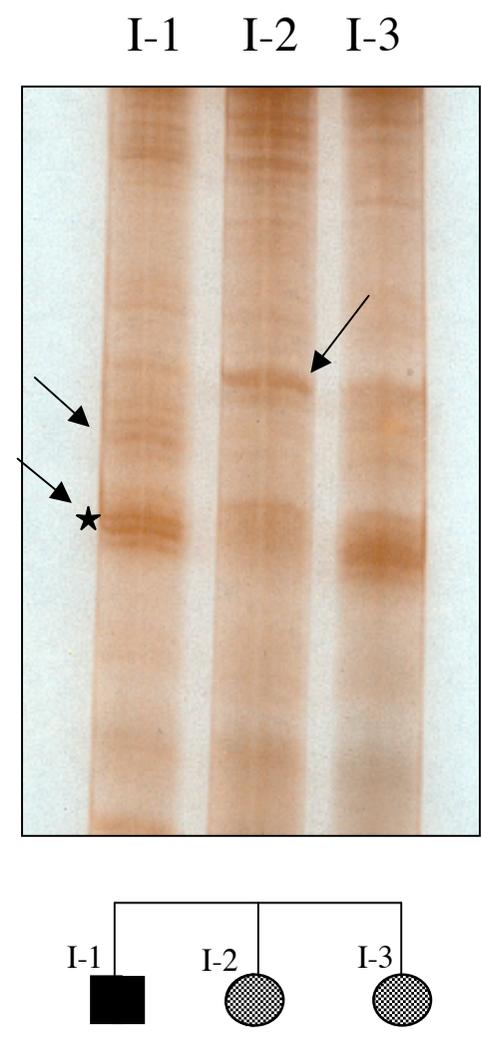


Figure 1

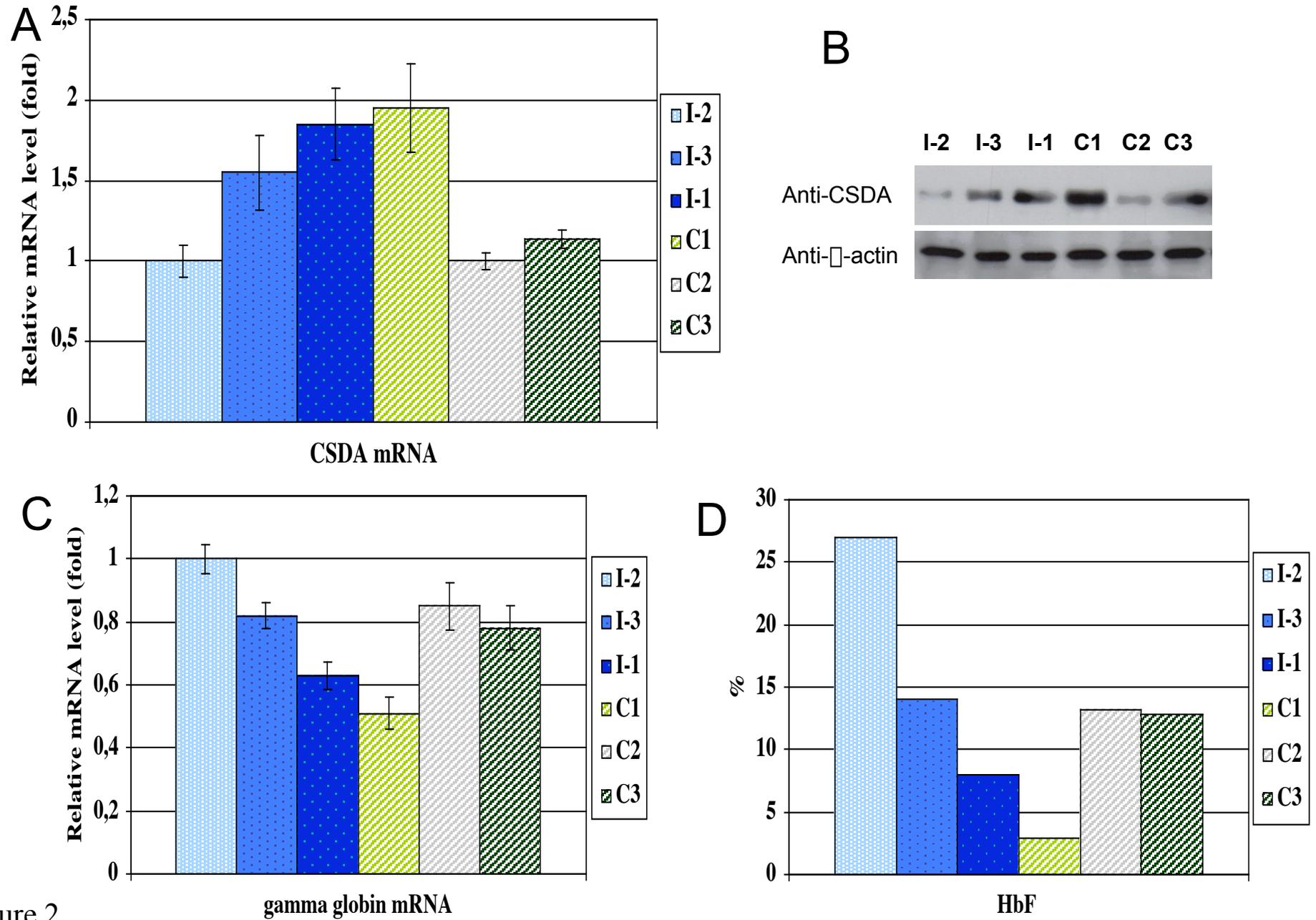


Figure 2

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

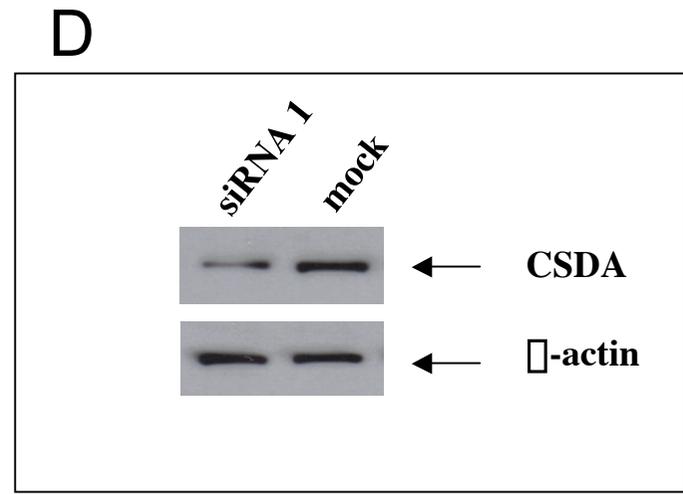
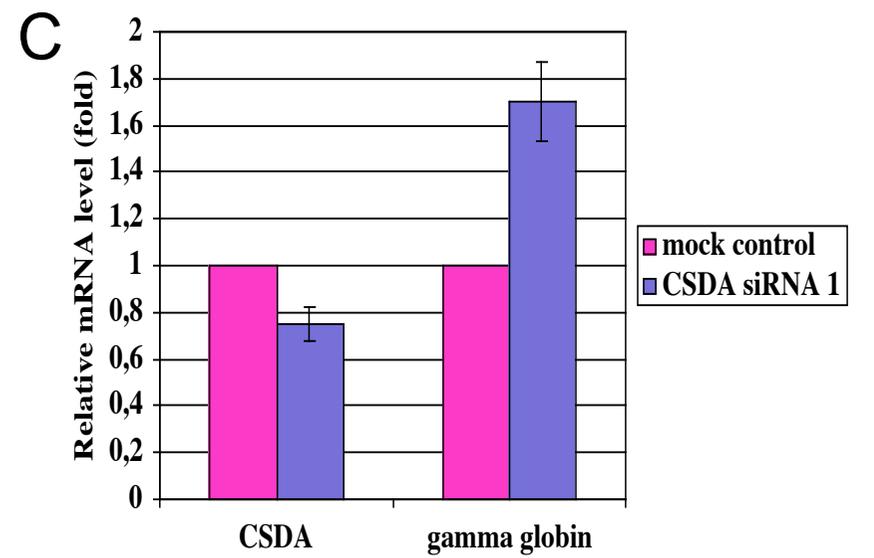
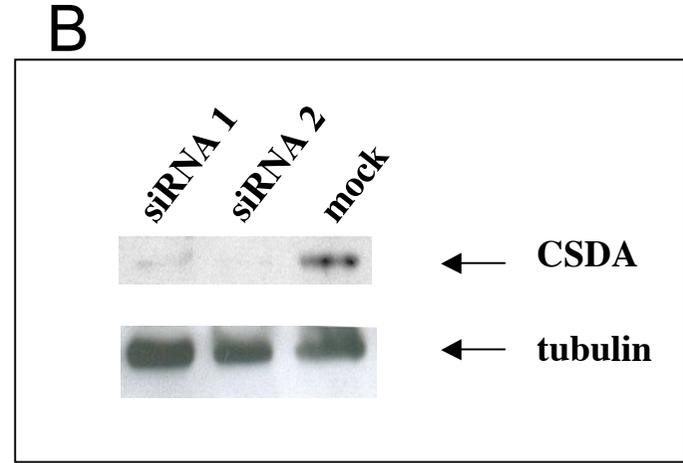
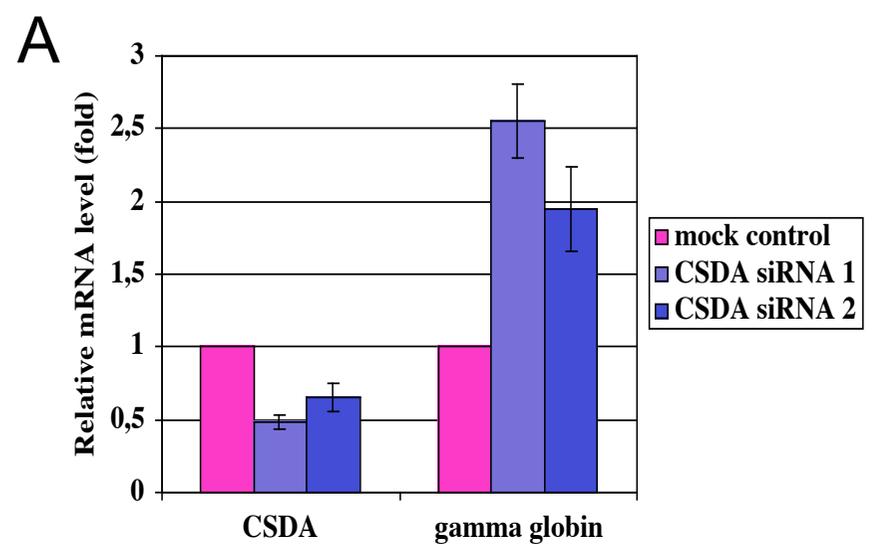


Figure 3

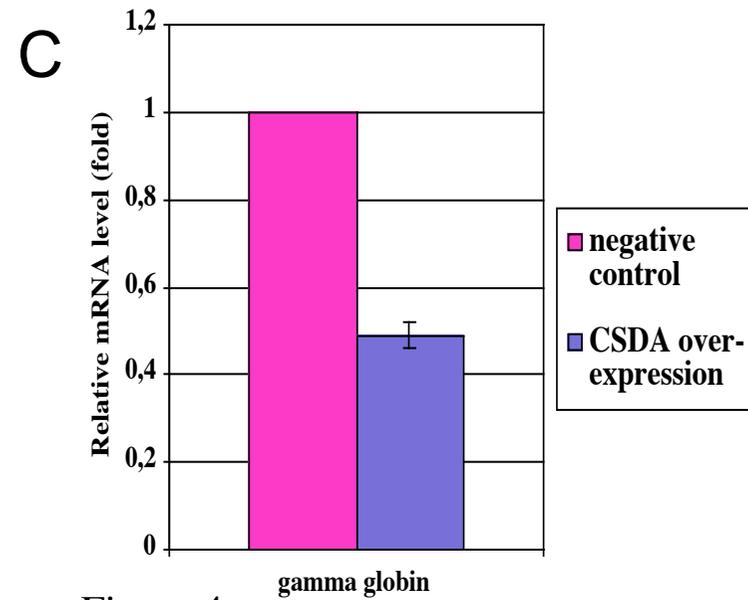
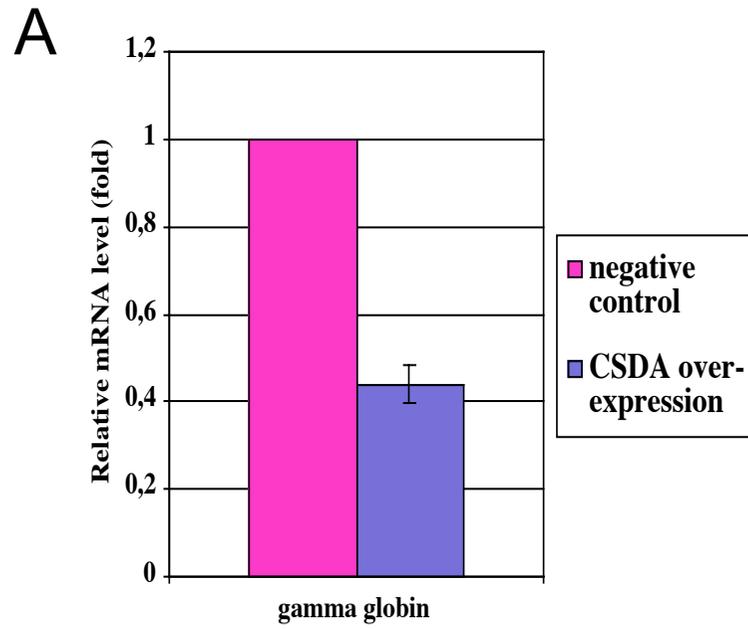
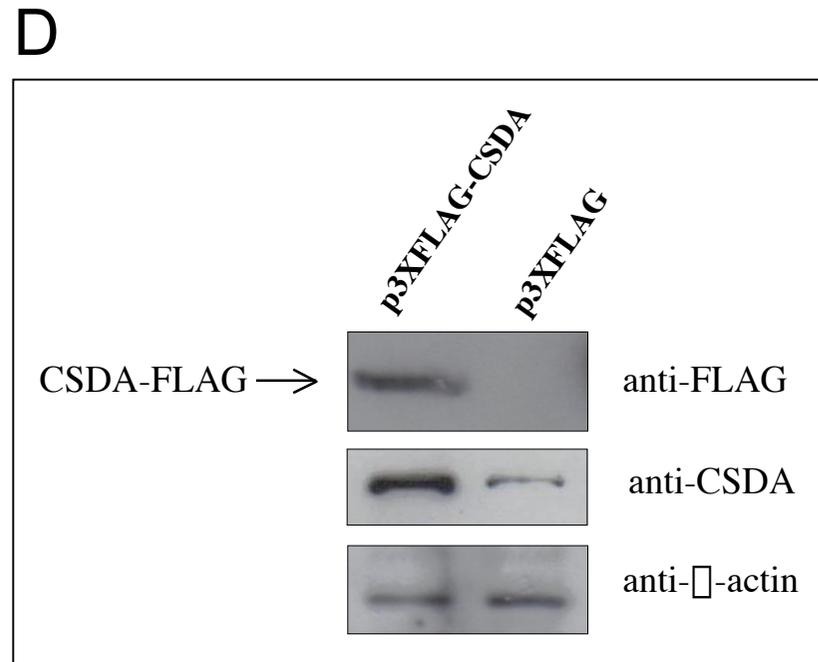
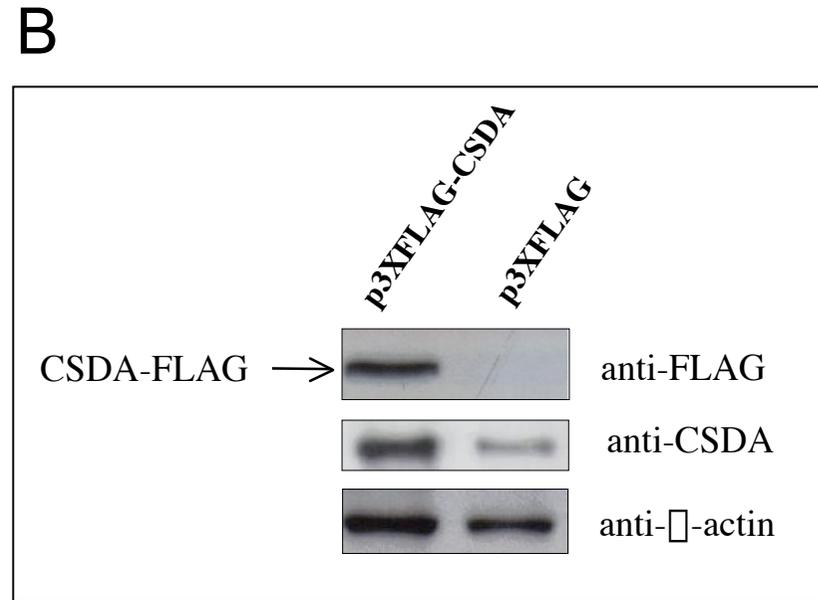
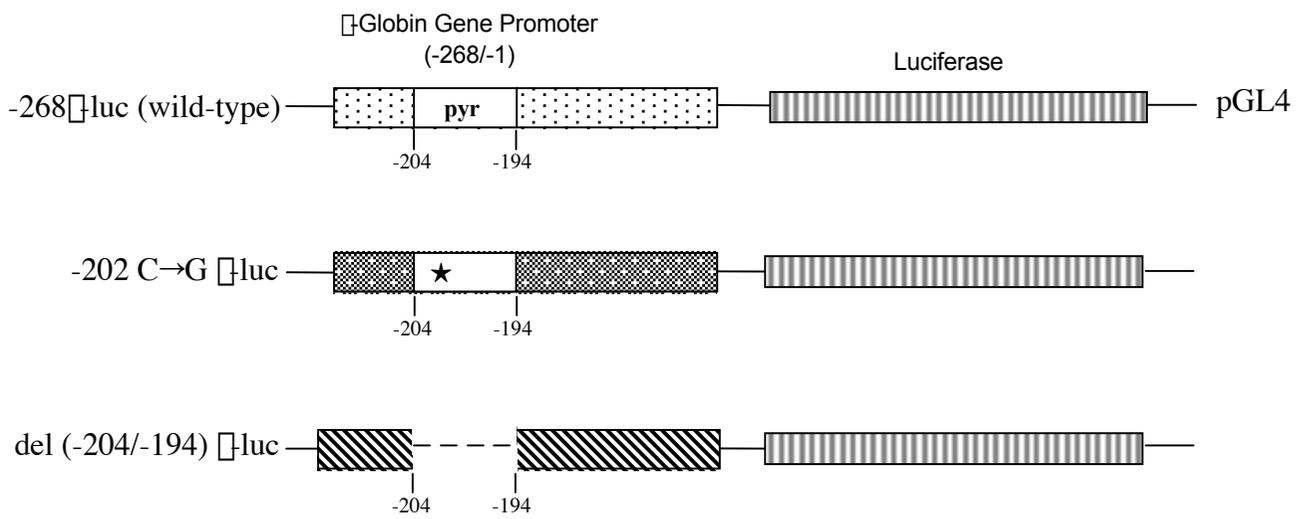


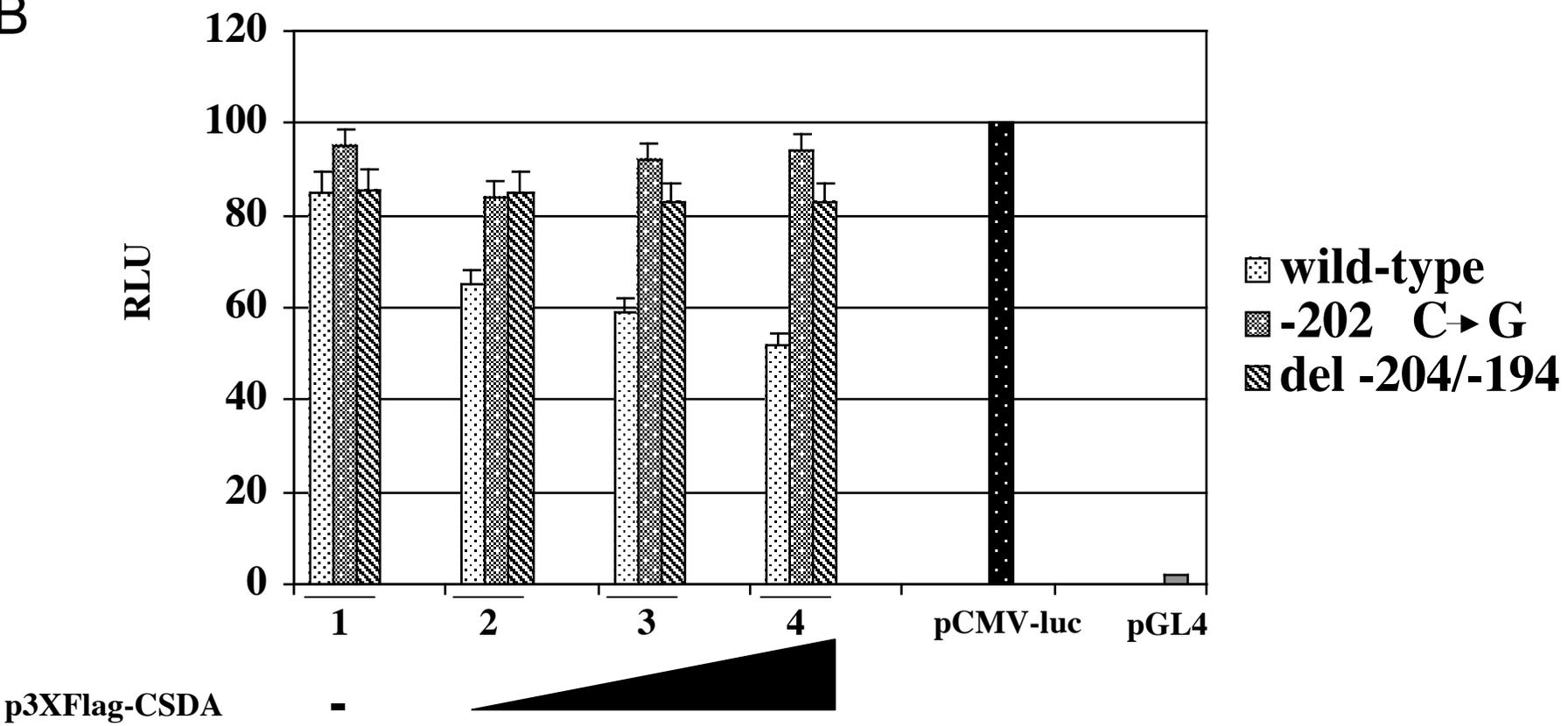
Figure 4



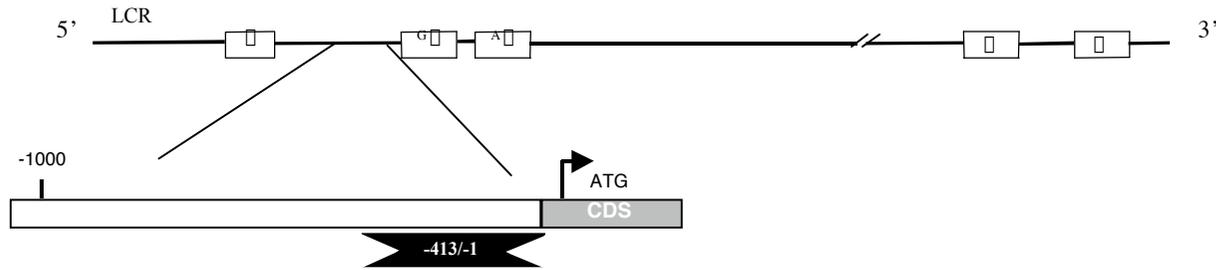
A



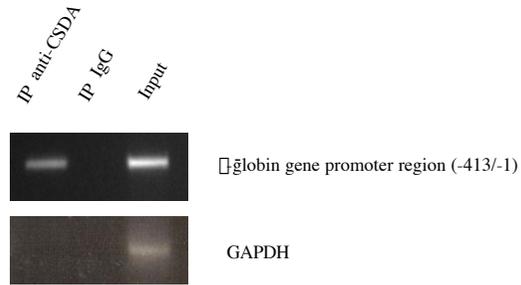
B



A



B



C

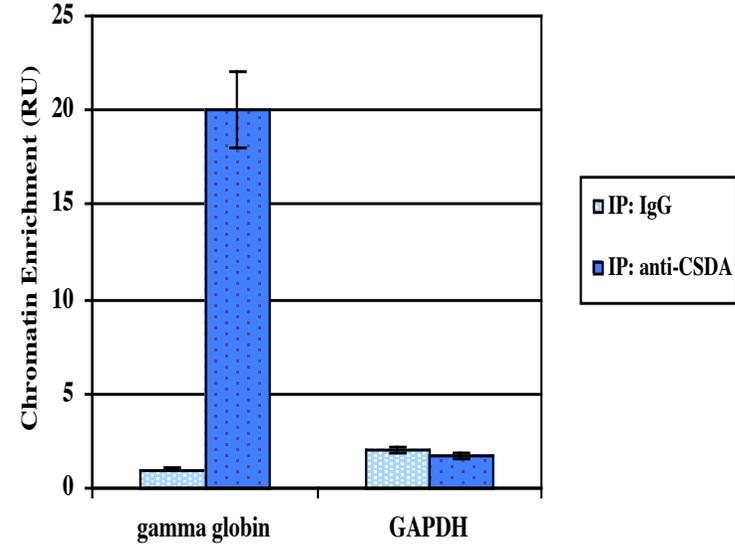


Figure 6

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Table I: Clinical, hematological data at diagnosis and β -globin gene genotypes of the patients examined in this study.

Patient	Gender	Age (yr)	β -thalassemia genotype	Transfusion therapy	Chelation therapy	Hb (g/dL)	HbF (%)	Hb A2 (%)	Splenectomy/Cholecystectomy	β -thalassemia phenotype
I-1	M	58	IVSI-6/IVSI-6	Occasionally transfused until his thirties, transfusion dependent for the last 20 years	Yes	6.5	6.0	5.9	yes/yes	severe intermedia
I-2	F	40	IVSI-6/IVSI-6	Occasionally	No	8.9	27.6	6.0	no/no	intermedia
I-3	F	38	IVSI-6/IVSI-6	Occasionally	No	8.6	14.5	6.8	no/yes	intermedia
C1	M	40	IVSI-6/IVSI-6	Transfusion-dependent since his childhood	Yes	6.2	2.9	7.2	yes/yes	severe intermedia
C2	M	42	IVSI-6/IVSI-6	Occasionally	No	8.1	13.2	6.2	no/no	intermedia
C3	M	37	IVSI-6/IVSI-6	Occasionally	No	8.5	12.8	6.3	no/no	intermedia

Table II: α - and β -globin gene cluster genotypes in the three siblings affected by *thalassemia intermedia*. For each examined putative regulatory transcription site within the β -globin gene cluster, the nucleotide position, the consensus sequences or the polymorphic variants and the corresponding patient genotypes are reported (the described polymorphic nucleotides are underlined). The N or N' sequences at position -10623/-10570 of HS2 are as reported elsewhere (Lu & Steinberg, 1996).

Patient	α -globin gene genotype	β -globin gene genotype	HS2			5'-G γ				5'- β		
			-10924 CC <u>IGT</u> E	-1090 CC <u>ACC</u> Sp1	-10623/-10570 (AT) _x N(AT) _y Hox2H	-1280 TGATAG GATA1	-1225 TGAC <u>GTC</u> A CRE	-202 C/T/G	-158 C/T (Xmn)	-551 C/T	-530 (AT) _x (T) _y BP1	-521 C/T
I-1	$\alpha\alpha/\alpha\alpha$	IVS16/IVS6	T/T	A/G	(AT) ₈ N(AT) ₁₃ / (AT) ₈ N'(AT) ₁ 3	G/G	G/G	C/C	C/C (-/-)	T/T	(AT) ₆ (T) ₈ / (AT) ₆ (T) ₈	C/C
I-2	$\alpha\alpha/\alpha\alpha$	IVS16/IVS6	T/T	A/G	(AT) ₈ N(AT) ₁₃ / (AT) ₈ N'(AT) ₁ 3	G/G	G/G	C/C	C/C (-/-)	TT	(AT) ₆ (T) ₈ / (AT) ₆ (T) ₈	C/C
I-3	$\alpha\alpha/\alpha\alpha$	IVS16/IVS6	T/T	A/G	(AT) ₈ N(AT) ₁₃ / (AT) ₈ N'(AT) ₁ 3	G/G	G/G	C/C	C/C (-/-)	T/T	(AT) ₆ (T) ₈ / (AT) ₆ (T) ₈	C/C

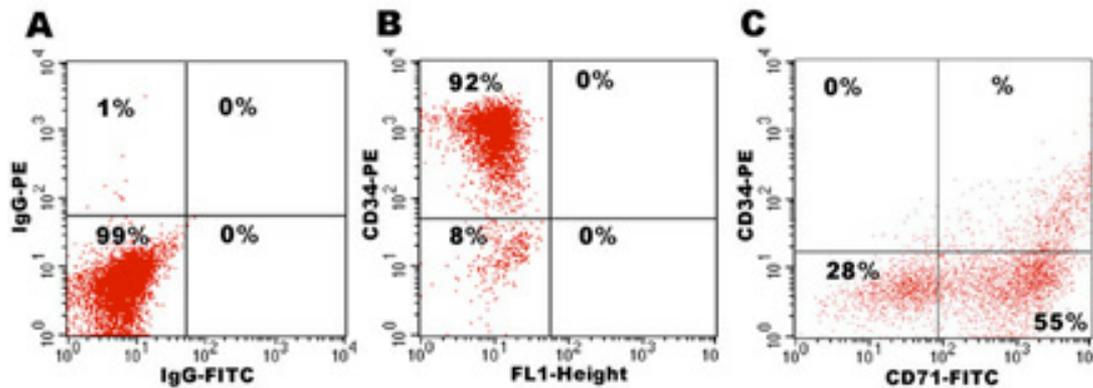
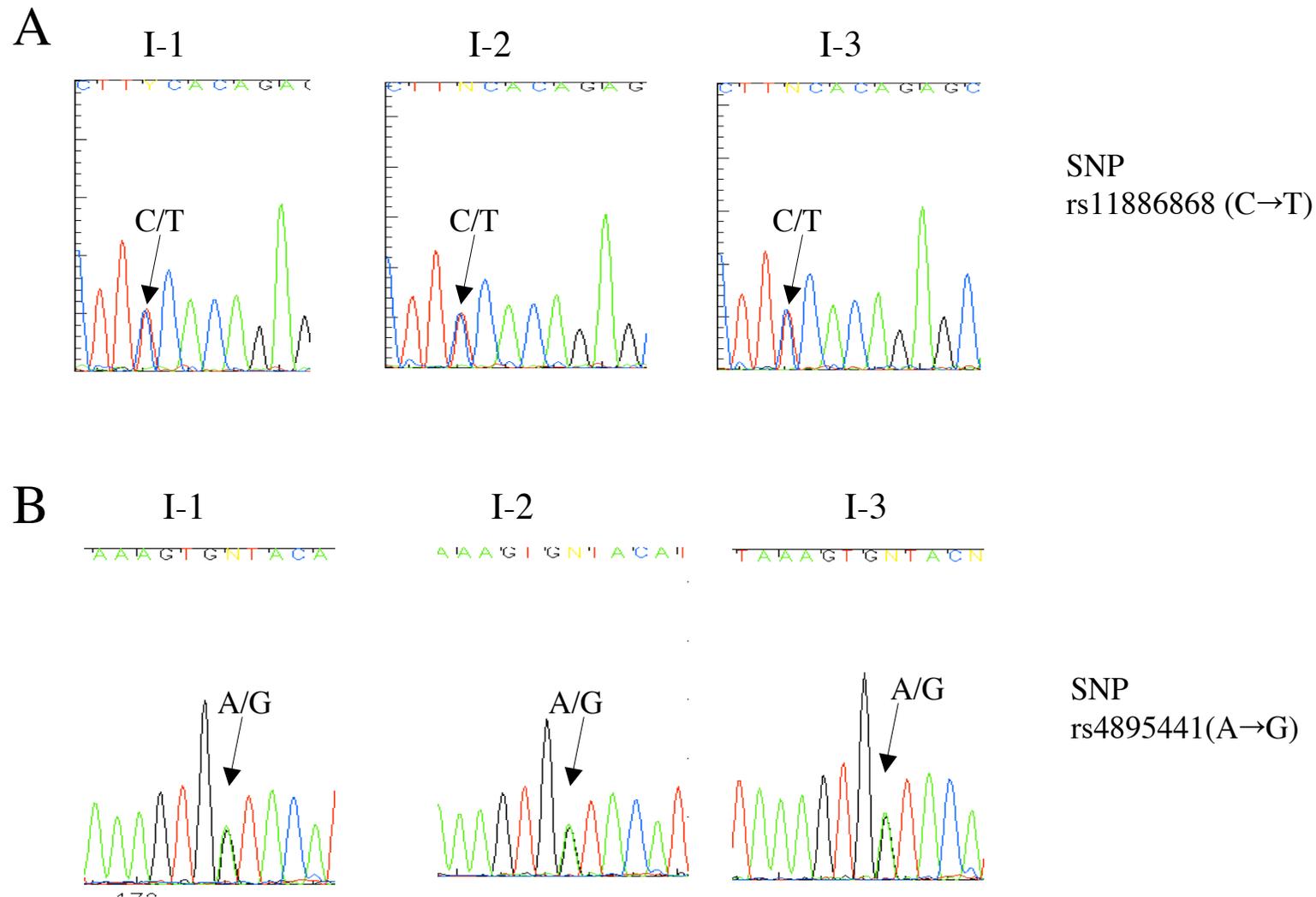


Figure 1S: Flow cytometric histograms of CD34 and CD71 staining. (A) Staining with control IgG. (B) Staining with anti CD34 after cell purification with magnetic beads (CD34= 92%). (C) Double-staining with anti CD34 and CD71 for control of erythroblastic differentiation at day 7 of culture (CD 71+/CD34+ =17%; CD 71+/CD34- =55%; total CD71 = 72%). Details on isolation, culture and flow cytometric analysis of erythroid cells are reported in the Materials and Methods Section.



38
39
40
41
42
43
44
45
46
47

Figure 2S: Sequencing analysis of rs11886868 (C→T) and rs4895441(A→G) SNPs reported to be associated with HPFH. Results show that the three siblings (patients I-1, I-2 and I-3) share the same heterozygous genotypes at both these polymorphic loci.

Gene	Chr pos	Length (bp)	MV36	I-1	I-2	I-3
POLR3K probe 4913-L1316	16p13.3	236	16p13, 60 Kb telomeric of HS40	Normal (0.86)	Normal (0.83)	Normal (0.86)
HS-40 probe 4799-L4797	16p13.3	178	HS-40	Normal (0.93)	Normal (1.12)	Normal (0.93)
HS-40 probe 4800-L4175	16p13.3	382	HS-40	Normal (0.88)	Normal (0.93)	Normal (0.88)
HBA region probe 4926-L4017	16p13.3	364	9.3 kb upstream HBZ	Normal (0.89)	Normal (1.01)	Normal (0.89)
HBA region probe 4622-L4001	16p13.3	346	3.5 kb upstream HBZ	Normal (0.98)	Normal (1.08)	Normal (0.98)
HBA region probe 4624-L4004	16p13.3	292	Between HBZ and HBZP	Normal (1.16)	Normal (0.95)	Normal (1.16)
HBA region probe 4625-L4005	16p13.3	318	Between HBZ and HBZP	Normal (1.03)	Normal (0.95)	Normal (1.03)
HBA region probe 4637-L4018	16p13.3	184	Between HBA2P and HBA1P	Normal (0.89)	Normal (0.97)	Normal (0.89)
HBA region probe 8488-L8410	16p13.3	373	Between HBA1P and HBA2	Normal (0.96)	Normal (1.01)	Normal (0.96)
HBA region probe 4627-L4007	16p13.3	201	Between HBA1P and HBA2	Normal (0.82)	Normal (0.77)	Normal (0.82)
HBA region probe 8492-L8415	16p13.3	214	Between HBA1P and HBA2	Normal (0.89)	Normal (0.82)	Normal (0.89)
HBA region probe 4628-L4008	16p13.3	229	Between HBA1P and HBA2	Normal (0.84)	Normal (0.79)	Normal (0.84)
HBA region probe 4630-L4011	16p13.3	142	HBA1 +HBA2 exon 1	Normal (1.17)	Normal (1.15)	Normal (1.17)
HBA region probe 8498-L8422	16p13.3	160	HBA2 intron 2	Normal (1)	Normal (1.11)	Normal (1)
HBA region probe 4633-L6249	16p13.3	240	HBA2 intron 2	Normal (0.93)	Normal (0.96)	Normal (0.93)
HBA region probe 4632-L6292	16p13.3	166	HBA1+HBA2 exon 3	Normal (1)	Normal (1.19)	Normal (1)
HBA region probe 8491-L8414	16p13.3	196	End of HBA2 exon 3	Normal (0.94)	Normal (0.93)	Normal (0.94)
HBA region probe 4626-L4740	16p13.3	190	Between HBA2 and HBA1	Normal (0.95)	Normal (0.97)	Normal (0.95)
HBA region probe 8493-L8416	16p13.3	220	Between HBA2 and HBA1	Normal (0.84)	Normal (0.77)	Normal (0.84)
HBA region probe 8494-L8417	16p13.3	256	Between HBA2 and HBA1	Normal (0.9)	Normal (0.85)	Normal (0.9)
HBA region probe 8497-L8420	16p13.3	339	Between HBA2 and HBA1	Normal (1.01)	Normal (1.04)	Normal (1.01)
HBA region probe 8499-L8423	16p13.3	154	0.2 kb downstream HBA1	Normal (0.98)	Normal (1)	Normal (0.98)
HBA region probe 4638-L4019	16p13.3	283	0.5 kb downstream HBA1	Normal (1.07)	Normal (0.86)	Normal (1.07)
HBA region probe 4639-L4020	16p13.3	310	2.4 kb downstream HBA1	Normal (1.13)	Normal (0.96)	Normal (1.13)
HBA region probe 6707-L6294	16p13.3	402	3.7 kb downstream HBA1	Normal (0.92)	Normal (0.97)	Normal (0.92)
Reference probe 2869-L2336	01p21	247	c	Normal (0.96)	Normal (0.99)	Normal (0.96)
Reference probe 3272-L2709	03q29	409	c	Normal (0.91)	Normal (0.92)	Normal (0.91)
Reference probe 3075-L2475	05p15	265	c	Normal (1.08)	Normal (0.9)	Normal (1.08)
Reference probe 1542-L0985	05q22	274	c	Normal (1.03)	Normal (0.85)	Normal (1.03)
Reference probe 1701-L1469	05q22	391	c	Normal (0.88)	Normal (0.98)	Normal (0.88)
Reference probe 0797-L0463	05q31	130	c	Normal (1.04)	Normal (1.12)	Normal (1.04)
Reference probe 3707-L3161	09q22	148	c	Normal (0.98)	Normal (1.07)	Normal (0.98)
Reference probe 3255-L2692	11p13	301	c	Normal (1.22)	Normal (1.01)	Normal (1.22)
Reference probe 0547-L0116	11q22	355	c	Normal (1.02)	Normal (1.13)	Normal (1.02)
Reference probe 2020-L1539	15q11	172	c	Normal (0.91)	Normal (1.1)	Normal (0.91)
Reference probe 3085-L4948	16p13	325	c	Normal (1.07)	Normal (1.01)	Normal (1.07)
Reference probe 2813-L2242	17q21	208	c	Normal (0.9)	Normal (0.81)	Normal (0.9)

Table 1S: Results of Multiplex Ligation-dependent Probe Amplification (MLPA) analysis of the α -globin gene cluster which excluded in the three siblings (patients I-1, I-2 and I-3) the presence of large deletions and duplications or other rearrangements within the cluster.