bjh research paper

The genetic basis of asymptomatic codon 8 frame-shift (*HBB*: c25_26delAA) β^0-thalassaemia homozygotes

Zhihua Jiang,[1,*] Hong-yuan Luo,[1,*] Shengwen Huang,[1,†] John J. Farrell,[1] Lance Davis,[1] Roger Théberge,[2] Katherine A. Benson,[1] Suchada Riolueang,[3] Vip Viprakasit,[3] Nasir A.S. Al-Allawi,[4] Sule Ünal,[5] Fatma Gümrük,[5] Nejat Akar,[6] A. Nazli Başak,[7] Leonor Osorio,[8] Catherine Badens,[9] Serge Pissard,[10] Philippe Joly,[11] Andrew D. Campbell,[12] Patrick G. Gallagher,[13] Martin H. Steinberg,[1] Bernard G. Forget[14,‡] and David H.K. Chui[1]

[1]Departments of Medicine, Pathology and Laboratory Medicine, Boston University School of Medicine, [2]Center for Biomedical Mass Spectrometry, Boston University School of Medicine, Boston, MA, USA, [3]Department of Paediatrics and Thalassaemia Centre, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, [4]Faculty of Medical Sciences, Scientific Research Centre, University of Duhok, Duhok, Iraq, [5]Division of Paediatric Haematology, Department of Paediatrics, Hacettepe University, [6]Department of Paediatric Molecular Genetics, Medical School, Ankara University, Ankara, [7]Department of Molecular Biology and Genetics, Neurodegeneration Research Laboratory, Boğaziçi University, Istanbul, Turkey, [8]Laboratório de Genética Molecular, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal, [9]Laboratoire de génétique moléculaire, Centre de référence Thalassémies, APHM, Hôpital d'enfants de la Timone, Marseille, [10]Departement de genetique, GHU Henri Mondor, Créteil, [11]Laboratoire de Biochimie et de Biologie Moléculaire, Hôpital Edouard Herriot, Unité de Pathologie Moléculaire du Globule Rouge, Hospices Civils & Université Claude Bernard-Lyon 1, Lyon, France, [12]Department of Pediatrics, University of Michigan School of Medicine, Ann Arbor, MI, [13]Departments of Pediatrics, Genetics, Yale University School of Medicine, and [14]Departments of Medicine, Genetics, Yale University School of Medicine, New Haven, CT, USA

Summary

Two 21-year old dizygotic twin men of Iraqi descent were homozygous for *HBB* codon 8, deletion of two nucleotides (–AA) frame-shift β^0-thalassaemia mutation (FSC8; *HBB*:c25_26delAA). Both were clinically well, had splenomegaly, and were never transfused. They had mild microcytic anaemia (Hb 120-130 g/l) and 98% of their haemoglobin was fetal haemoglobin (HbF). Both were carriers of Hph α-thalassaemia mutation. On the three major HbF quantitative trait loci (QTL), the twins were homozygous for G>A *HBG2 Xmn1* site at single nucleotide polymorphism (SNP) rs7482144, homozygous for 3-bp deletion *HBS1L-MYB* intergenic polymorphism (HMIP) at rs66650371, and heterozygous for the A>C *BCL11A* intron 2 polymorphism at rs766432. These findings were compared with those found in 22 other FSC8 homozygote patients: four presented with thalassaemia intermedia phenotype, and 18 were transfusion dependent. The inheritance of homozygosity for HMIP 3-bp deletion at rs66650371 and heterozygosity for Hph α-thalassaemia mutation was found in the twins and not found in any of the other 22 patients. Further studies are needed to uncover likely additional genetic variants that could contribute to the exceptionally high HbF levels and mild phenotype in these twins.

Keywords: β^0-Thalassaemia, fetal haemoglobin, HbF quantitative trait loci, *HBS1L-MYB* intergenic polymorphism, α-Thalassaemia.

First published online 13 January 2016
doi: 10.1111/bjh.13909

Received 27 August 2015; accepted for publication 9 November 2015
Correspondence: Dr David H.K. Chui, Department of Medicine, Boston University School of Medicine, 72 East Concord Street, Evans 248, Boston, MA 02118, USA.
E-mail: david.chui@bmc.org

*ZJ and HL contributed equally to this work.
†Present address: Department of Laboratory, Guizhou Provincial People's Hospital, 83 East Zhongshan Road, Guiyang 550002, Guizhou, China
‡Deceased.

Patients who are either homozygous or compound heterozygous for β-thalassaemia mutations have intracellular accumulation of free HbA that precipitates and leads to erythroid cell apoptosis, ineffective erythropoiesis and haemolysis. Most of these patients are severely anaemic from an early age and require chronic red blood cell (RBC) transfusions (Weatherall & Clegg, 2001).

Increased HbF production is found in adults with hereditary persistence of fetal haemoglobin (HPFH) point mutations in *HBG* promoters or with HPFH, (*HBD HBB*)^0- or (*HBG1 HBD HBB*)^0-thalassaemia deletions. In addition, three major HbF quantitative trait loci (QTL) mediate increased HbF production (Thein & Menzel, 2009). They are defined by the *HBG2 Xmn1* polymorphism at Chr11p15 (rs7482144), the *HBS1L-MYB* intergenic polymorphism (HMIP) at 6q23 (rs66650371) and the *BCL11A* intron 2 polymorphism at 2p16 (rs766432). Their alternate or minor alleles are associated with elevated HbF levels.

Two dizygotic twin young men were homozygous for FSC8 mutation (*HBB*:c25_26delAA). Both were clinically well, had mild microcytic anaemia and almost all haemoglobin produced was HbF. In contrast, among 22 other FSC8 homozygote patients, four had thalassaemia intermedia phenotype and 18 were transfusion dependent. We now report genetic determinants found in both twins pertaining to their capacity to produce large amounts of HbF. Further studies are needed to uncover likely additional genetic variants that can up-regulate HbF expression in these patients.

Methods

Haematological analyses

Blood counts were done using an automated haematology analyser. Haemoglobin analyses were carried out with Variant II cation high-performance liquid chromatography (HPLC) system (Bio-Rad, Hercules, CA, USA).

Molecular diagnoses

Genomic DNA was extracted from peripheral blood leucocytes by QIAGEN BioRobot EZ1 workstation with EZ1 DNA kit (Qiagen, Valencia, CA, USA). Globin genes and other genetic loci of interest were amplified by polymerase chain reaction (PCR), and the purified amplicons were subjected to Sanger bidirectional nucleotide sequencing using the ABI Big Dye Terminator Cycle Sequencing kit (Applied BioSystems, Foster City, CA, USA) and ABI 3730 DNA Analyser. *HBA* deletions were assessed by gap-PCR (Tan *et al*, 2001).

Sentinel single nucleotide polymorphisms (SNPs), rs66650371, rs9399137 and rs766432 in known HbF QTL were determined by PCR-based amplification refractory mutation system (ARMS) (Farrell *et al*, 2011; Table SI).

Multiplex ligation-dependent probe amplification (MLPA) analysis was done on the *HBB* gene cluster including its locus control region (LCR) using the SALSA MLPA KIT P102 HBB kit (MRC Holland, Amsterdam, the Netherlands) according to manufacturer's instructions.

Haplotype analysis of the HBB *globin gene cluster*

The haplotype of the *HBB* globin gene cluster found in Mediterranean populations as described by Orkin *et al* (1982) was determined by the polymorphic profile of seven restriction sites (*HBE*-Hind*II*, *HBG2*-Hind*III*, *HBG1*-Hind*III*, 5′ and 3′ *HBBP1*-Hind*II*, *HBB*-Ava*II* and *HBB*-Hinf*I*) using PCR-restriction fragment length polymorphism procedures as previously described (Ekwattanakit *et al*, 2012).

Results

Case report and family study

A 21-year old Iraqi-American man was found to have splenomegaly. He was otherwise well, and had never been transfused. He had mild microcytic anaemia (Hb 123 g/l; MCV

Table I. Haematological findings and globin genotypes.

Relationship	Father	Mother	Sister	Proband	Twin brother
Age (years)	53	50	24	21	21
Hb (g/l)	127	129	124	123	133
Hct (%)	41·8	42·1	40·8	38·9	40·8
MCV (fl)	65	76	74	72	73
MCH (pg)	19·8	23·4	22·6	22·9	23·9
MCHC (g/l)	304	306	304	316	326
Reticulocyte count	0·87%	1·14%	0·38%	1·92%	1·52%
WBC (10^9/l)	7·1	8·1	5·6	12·5	10·7
Platelets (10^9/l)	145	236	197	147	143
Hb A	90%	90%	88%	—	—
Hb A_2	5·1%	5·0%	5·1%	2·2%	2·6%
Hb F	3·3%	4·3%	5·4%	97·8%	97·4%
β-Globin genotype	Codon 8 (-AA) heterozygous	Codon 8 (-AA) heterozygous	Codon 8 (-AA) heterozygous	Codon 8 (-AA) HOMOZYGOUS	Codon 8 (-AA) HOMOZYGOUS
α-Globin genotype	(α α/α α)	(α^{HpH} α/α α)	(α^{HpH} α/α α)	(α^{HpH} α/α α)	(α^{HpH} α/α α)
HBB globin gene cluster haplotype	I/IV	I/IV	I/IV	IV/IV	IV/IV

Hb, haemoglobin concentration; Hct, haematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular haemoglobin; MCHC, mean cell haemoglobin concentration; WBC, white blood cells.

72 fl) with slightly increased reticulocyte count (1·9%). His mean RBC corpuscular haemoglobin (MCH) was 22·9 pg, similar to β-thalassaemia trait. Haemoglobin analysis by cation exchange HPLC showed HbF 97·8% and HbA_2 2·2% (Table I). Matrix-assisted laser desorption ionization (MALDI) mass spectrometry revealed the proportion of HBG2/HBG1 at approximately 3/1 to 4/1, and HBE was not detected (data not shown). Nucleotide sequencing of his *HBB, HBA2, HBA1, HBG2* and *HBG1* was done. The patient was homozygous for FSC8 β^0-thalassaemia mutation (Fig S1A), heterozygous for Hph α-thalassaemia mutation in his *HBA2* (*HBA2*:c.95 + 2_95 + 6delTGAGG), and homozygous for the G>A *HBG2* promoter *Xmn1* polymorphism at rs7482144 (*HBG2*:c.-211G>A). Known HPFH mutations were not found in his *HBG2* and *HBG1* promoters. There was no deletion or duplication involving his *HBB* locus, as shown by MLPA (Fig S1B). His *KLF1* gene was sequenced, and no mutation was found (data not shown).

The proband had a dizygotic twin brother, documented by non-identical SNP genotypes at 7 different loci on chromosomes 7, 13, 16 and 19 (Table SII). Table I tabulates the clinical laboratory results of the proband, his dizygotic twin brother, his older sister and both his parents. The other twin also was homozygous for FSC8 mutation, and their sister and parents were heterozygous for the same mutation.

Both twins were homozygous for *HBB* cluster haplotype IV, which is found in only 1% of the Mediterranean population (Orkin *et al*, 1982). Except for *HBB*-Ava*II* polymorphism, haplotype IV is identical to Senegal haplotype found in sickle cell anaemia patients who have relatively high HbF levels (Akinsheye *et al*, 2011). The twins' parents and sister had haplotypes I/IV.

HBB cluster locus

HBB locus control region (LCR) DNase I hypersensitive site 2 (HS2) core in the proband was sequenced from Chr11:5 301 778–5 302 235. No mutation or deletion was found. He had identical purine/pyrimidine repetitive sequences [$(TA)_9$ $(CA)_2$ $(TA)_2$ CG $(TA)_{10}$] downstream of LCR HS2 core as in sickle cell anaemia patients with Senegal haplotype (Öner *et al*, 1992).

Within the *HBD-HBG1* intergenic region, there are multiple *cis*-regulatory elements capable of modulating *HBG* expression (Fig S2). Overlapping amplicons covering the entire 14·5 kb intergenic region in the proband were obtained by PCR, and Sanger nucleotide sequencing was done. No deletion was found. The twin was homozygous for alternate alleles of rs10128556 and rs2071348 (Table II, Fig S2), which were previously shown to be associated with increased HbF in sickle cell anaemia patients and in patients with HbE/β^0-thalassaemia mutation, respectively (Galarneau *et al*, 2010; Nuinoon *et al*, 2010). He was homozygous for the alternate alleles of 19 SNPs, none of which resides within known functional motifs (Fig S2, Table SIII).

The transcription factor BP1 (DLX4) binding site at 530 bp 5′ to the *HBB* in the proband was found to be homozygous for the common reference sequence, [$(AC)_2$ $(AT)_7$ T_7] (Zoueva *et al*, 2008).

HbF quantitative trait loci (QTL)

XmnI polymorphism in the HBG2 promoter on Chr. 11p15. Both twins were homozygous for the alternate allele A at rs7482144, also known as the *XmnI* polymorphism at nt –158 in the *HBG2* promoter (Table II).

Table II. SNP Genotypes at 3 HbF QTL.

HbF QTL	Gene locus	SNP ID	Position (hg19)	Reference allele	Alternate allele	Father	Mother	Sister	Proband	Twin brother
HBB cluster on Chr.11p15	*HBP1* intron 2	rs10128556	5 263 683	C	T				T/T	
		rs2071348	5 264 146	T	G				G/G	
	HBG2 promoter	rs7482144	5 276 169	G	A	G/A	G/A	G/A	A/A	A/A
HMIP on Chr.6q23	*HBS1L – MYB* intergenic region	rs66650371	135 418 633 –135 418 635	TAC	Deletion	del/del	del/del	del/del	del/del	del/del
		rs9399137	135 419 018	T	C	C/C	C/C	C/C	C/C	C/C
BCL11A intron 2 on Chr.2p16	DHS +62	rs1427407	60 718 043	G	T				G/T	
		rs1896293	60 718 848	G	T				G/T	
		rs766432	60 719 970	A	C	A/C	C/C	C/C	A/C	A/C
	DHS +58	rs6706648	60 722 040	C	T				C/C	
		rs6738440	60 722 241	A	G				A/A	
		rs10539208	60 724 764 –60 724 766	ATA	Deletion				del/del	
	DHS +55	rs7606173	60 725 451	G	C				G/C	

SNP ID, single nucleotide polymorphism identification; HMIP, *HBS1L-MYB* intergenic polymorphism.

HMIP polymorphism on Chr. 6q23. Both twins and their three other immediate family members were homozygous for the alternate allele C at the sentinel SNP rs9399137 in HMIP (Table II). Furthermore, all five family members were homozygous for the 3-bp deletion, rs66650371, which is in linkage disequilibrium with rs9399137 and primarily responsible for mediating this QTL's effect upon HbF expression (Farrell *et al*, 2011; Stadhouders *et al*, 2014).

BCL11A intron 2 polymorphism on Chr. 2p16. Both twins and their father were heterozygous for the alternate allele of *BCL11A* intronic SNP rs766432 (Table II) (Menzel *et al*, 2007a). There are three erythroid specific DNase I hypersensitive sites (DHS) within *BCL11A* intron 2, known as +55, +58 and +62 based on the distances in kb between the DHS motif and *BCL11A* transcription start site (Fig 1) (Bauer *et al*, 2013). The three DHS sites in the proband were sequenced. The proband was found to be heterozygous for alternate alleles among three other SNPs located within DHS +62 and +55, rs1427407, rs1896293 and rs7606173, all known to be associated with elevated HbF levels in sickle cell anaemia patients (Fig 1, Table II) (Bauer *et al*, 2013).

Other FSC8 homozygote patients

Table SIV presents findings on four Turkish FSC8 homozygotes. Two sisters, aged 30 and 32 years, were transfused sparingly before splenectomy. No transfusions have been given since splenectomy, at age 8 and 11 years, respectively. Their Hb was 86 and 93 g/l. Two boys, aged 7 and 13 years, had a Hb of 77 and 98 g/l, respectively, and neither had been transfused.

Table SV presents findings on 18 additional transfusion-dependent FSC8 homozygotes. Six were Iraqi, 6 Turkish, 2 Moroccan and 4 or were classed as North African (Moroccan/Algerian). Ten had undergone splenectomy. Fifteen of these transfusion dependent patients were homozygous for *HBB* cluster haplotype IV/IV, and homozygous for *XmnI* polymorphism in the *HBG2* promoter, identical to that found in the asymptomatic twins. Three other patients, 1 Turkish and 2 North African, were homozygous for haplotype VII/VII, and they did not have the *XmnI* polymorphism in their *HBG2* promoter.

Table III summarizes the HbF QTL findings among all FSC8 homozygotes. Both dizygotic twins were homozygous for the 3-bp deletion at rs66650371 in HMIP (Table II). Among the four patients with "mild" phenotype, three were heterozygous for the 3-bp deletion, and 1 did not have the deletion (Table SIV). Among the 18 transfusion-dependent patients, eight were heterozygous for the 3-bp deletion, and the other 10 did not have the deletion (Table SV). The differences in the allele frequencies among these three groups of patients are statistically highly significant ($P = 0{\cdot}0035$).

Discussion

The FSC8 β^0-thalassaemia mutation (*HBB*:c25_26delAA) was first described in a Turkish patient (Orkin & Goff, 1981). This deletion leads to a premature termination codon TGA at codon 21 upstream of the first exon-intron junction. The aberrant *HBB* mRNA transcripts are subjected to nonsense-mediated decay resulting in β^0-thalassaemia. Homozygotes for this mutation usually are transfusion-dependent, and are found in Turkey, Northern Iraq, Azerbaijan, Morocco, and their neighbouring regions (Gurgey *et al*, 1989; Cürük *et al*, 1992; Kuliev *et al*, 1994; Lemsaddek *et al*, 2003, 2004; Al-Allawi *et al*, 2010; Hamamy & Al-Allawi, 2013).

Gurgey *et al* (1989) reported 8 Turkish patients homozygous for FSC8 mutation who presented as thalassaemia

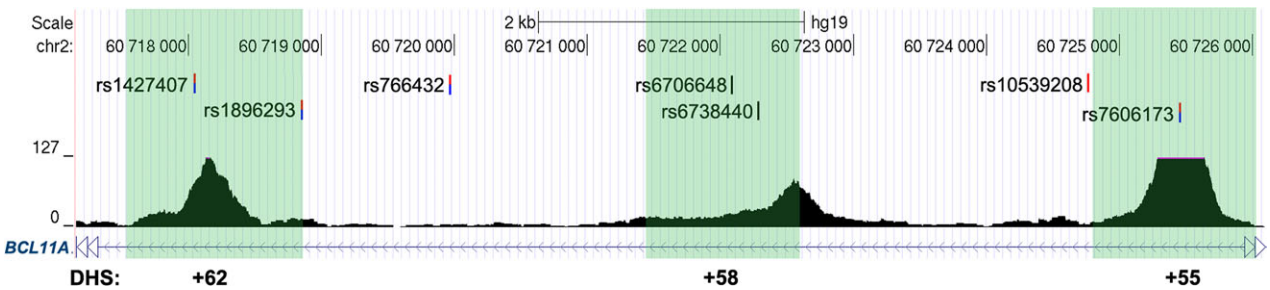

Fig 1. The figure shows ~8·5 kb of *BCL11A* intron 2 region encompassing the 3 erythroid-specific DHS (Bauer *et al*, 2013). The numbers on the first row represent hg19 Chr2 positions (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/). On the second and third rows, each perpendicular bar represents a single SNP found in the proband. Red and blue bars represent heterozygosity for the SNP alternate allele. The black bars represent homozygosity for the SNP reference allele. The red bar represents homozygosity for the SNP alternate allele (see Table II). The green shaded areas represent the 3 DHS, and the numbers on the bottom row represent distances in kb from *BCL11A* transcription start site (Bauer *et al*, 2013). They are DHS +62, Chr2:60 717 492–60 718 860; DHS +58, Chr2:60 721 411–60 722 674; and DHS +55, Chr2:60 724 802–60 726 084. The black areas represent Chip-seq signals of GATA1 in human peripheral blood-derived erythroblasts, from ENCODE data (https://www.encodeproject.org/experiments/ENCSR000EXP/). DHS, DNase 1 hypersensitive site; SNP, single nucleotide polymorphism.

intermedia. They were homozygous for *HBB* haplotype IV, associated with the *Xmn1* polymorphism in *HBG2* promoter. Their age ranged between 6–35 years with Hb 97 ± 10 g/l, MCV 73 ± 2 fl and HbF 98·6 ± 0·2%. Five had mild thalassaemic facies; all had splenomegaly; six had undergone splenectomy.

The dizygotic twins in this report were homozygous for the FSC8 mutation. Their phenotype was the mildest among those reported in the literature, and among 22 other patients we have studied. They have inherited *HBB* haplotype IV/IV associated with *XmnI* polymorphism in *HBG2* promoter, similar to Senegal haplotype found in sickle cell anaemia patients with increased HbF. Fifteen other transfusion-dependent patients also were homozygous for haplotype IV (Table III). Therefore these findings could not account for the twins' exceptionally mild phenotype.

Genotyping done on skeletal remains of an 8-year old child found in an archaeological site in Israel revealed FSC8 homozygosity (Filon *et al*, 1995). The site was dated to the Ottoman period between 16th and 19th centuries. The authors surmised that the child must have had inherently high HbF production to sustain her survival (Filon *et al*, 1995).

Stress erythropoiesis and HBG mRNA accumulation and translation

Stress erythropoiesis plays an important role in enhanced HbF production. For example, sickle cell anaemia patients from Eastern Provinces in Saudi Arabia have markedly elevated HbF levels, yet their HbS carrier parents, with the same Arab-Indian haplotype and no haemolysis, have normal HbF (Miller *et al*, 1986). The twins had splenomegaly and minimal reticulocytosis consistent with mild stress erythropoiesis.

In studies on the Corfu δβ-thalassaemia deletion in conjunction with β IVS I-5 G>A (*HBB*:c.92 + 5G>A), a severe β^+-thalassaemia mutation, robust *HBG* mRNA accumulation and translation resulting in elevated HbF levels occurred only when *HBB* mRNA fell below a critical threshold (Chakalova *et al*, 2005). FSC8 homozygotes have absence of *HBB* mRNA due to nonsense-mediated decay. This might account for the propensity of all FSC8 homozygotes to produce increased HbF. Alone, this would not explain the twins' asymptomatic phenotype compared with other patients.

Genetic determinants of HbF expression

There are three major HbF QTL, on Chr.11p15, 6q23 and 2p16. Together they account for approximately 20–50% of HbF variance in different populations. However genetic determinants accounting for more than half of HbF variance in all populations are yet to be discovered. The twins were (i) homozygous for *HBB* haplotype IV; (ii) homozygous for rs7482144 alternate allele on chr.11p15; (iii) homozygous for HMIP 3-bp deletion on chr.6q23; (iv) heterozygous for *BCL11A* intron 2 QTL on chr.2p16; and (v) heterozygous for Hph α-thalassaemia mutation. These genetic variants must be partially responsible for increased HbF expression in the twins and in other FSC8 homozygotes.

Among patients of African descent with sickle cell anaemia, the Senegal haplotype, characterized by rs7482144, have the highest HbF. The twins were homozygous for *HBB* haplotype IV, which shares six identical and consecutive polymorphisms from *HBE* to *HBBP1* with the Senegal haplotype. The proband has inherited the same purine/pyrimidine repetitive sequence within LCR HS2 as in the Senegal haplotype (Öner *et al*, 1992), and rs2071348 and rs10128556 within the *HBD-HBG1* intergenic region known to be associated with elevated *HBG* expression (Galarneau *et al*, 2010; Nuinoon *et al*, 2010). There were 3–4 times more HBG2 than HBG1 globin chains in both twins, as predicted by the presence of rs7482144. The pathways whereby haplotype IV

Table III. Summary of HbF QTL in FSC8 Homozygotes.

HbF QTL	SNP genotype	Asymptomatic dizygotic twins ($N = 2$)	Patients with mild disease ($N = 4$)	Patients with severe disease ($N = 18$)	*P*-value
Xmn I on Chr. 11p15 rs7482144	A/A	2	4	15	0·69
	G/A	0	0	0	
	G/G	0	0	3	
HMIP on Chr. 6q23 rs66650371 or 3-bp deletion	del/del	2	0	0	0·0035
	WT/del	0	3	8	
	WT/WT	0	1	10	
BCL11A intron 2 on Chr. 2p16 rs766432	C/C	0	0	2	0·62
	A/C	2	2	7	
	A/A	0	2	9	

SNP, single nucleotide polymorphism; HMIP, *HBS1L-MYB* intergenic polymorphism; QTL, quantitative trait loci.
Pearson Chi-square test was used to compare HbF QTL genotype frequencies among 3 groups of patients. Statistical analyses were performed using R Statistical Software (http://www.r-project.org/). *P*-values were obtained from the simulation. SNP alternate alleles are: A for rs7482144, del for rs66650371, and C for rs766432.

is associated with HbF up-regulation are presently not known.

The most important functional motif for HMIP is the 3-bp deletion, rs66650371 in an enhancer capable of regulating *MYB* expression (Farrell *et al*, 2011; Stadhouders *et al*, 2014). In the immediate vicinity of the 3-bp deletion polymorphism are binding sites for TAL1, GATA2, and RUNX1. The 3-bp deletion interferes with normal spatial orientation for DNA-protein binding and/or protein-protein interactions, leading to decreased enhancer activity and down-regulation of *MYB* expression. MYB can transactivate *KLF1* expression, and also plays an important role in cellular proliferation and haematopoiesis (Jiang *et al*, 2006; Menzel *et al*, 2007b; Ganesh *et al*, 2009; Bianchi *et al*, 2010). Both twins were homozygous for this 3-bp deletion (Tables II and III). This has to be one of the pivotal genetic determinants accounting for the twin's heightened HbF production.

BCL11A is an essential transcriptional repressor of *HBG* expression. Both twins were heterozygous for this QTL in *BCL11A* intron 2. Their mother and sister were homozygous for the same QTL, and yet their HbF levels were merely 4–5% (Table I). Two patients (7 and 22 in Table SV) were also homozygous for this QTL, and yet they were transfusion dependent. Therefore, *BCL11A* polymorphisms are not responsible for the twins' exceptionally high HbF.

The twins were heterozygous for a deletion of 5-bp (–TGAGG) in the IVS-1 donor splice site in *HBA2* (*HBA2*: c.95 + 2_95 + 6delTGAGG), the more highly expressed of the two *HBA*s. In β-thalassaemia patients, the co-inheritance of α-thalassaemia reduces HBA to non-HBA synthetic imbalance, and results in less unpaired HBA that can precipitate and damage the erythroid cell. The inheritance of α-thalassaemia in the twins in conjunction with their HbF QTL alternate alleles should aid in the preferential survival of many more RBC with high HbF expression. Of note is Patient 12 who had single *HBA* deletion and severe disease phenotype (Table SV).

To summarize, homozygosity for the 3-bp deletion in HMIP and heterozygosity for the Hph α-thalassaemia mutation in *HBA2* were the two genetic determinants found only in the asymptomatic twins and not found in any of the other 22 FSC8 homozygote patients. These two genetic variants are most likely among the essential genetic factors contributing to the twins' uniquely high HbF expression. Future investigations by genetic approaches, such as whole genome sequencing and transcriptome expression, to study and compare FSC8 homozygotes with markedly disparate phenotypes should identify additional genetic variants that are important in regulating *HBG* expression.

Acknowledgements

We are grateful to Dr. Peter Fraser for his insightful discussions, and Dr. Paola Sebastiani for her expert advice on statistical analyses. We wish to dedicate this manuscript to the memory of the late Professor Bernard G. Forget who died on November 6, 2015. We are grateful for his wisdom, guidance, and generosity throughout this work.

Author contributions

Z. Jiang, H-Y. Luo, S. Huang, L. Davis, R. Théberge, K.A. Benson, S. Riolueang performed the research and analysed the data. N.A.S. Al-Allawi, S. Ünal, F. Gümrük, N. Akar, A.N. Başak, L. Osorio, C. Badens, S. Pissard, P. Joly, and A.D. Campbell provided patient samples and analysed their clinical and laboratory findings. J.J. Farrell, V. Viprakasit, A.D. Campbell, P.G. Gallagher, M.H. Steinberg, B.G. Forget and D.H.K. Chui designed the research and analysed the data. M.H. Steinberg, B.G. Forget and D.H.K. Chui wrote

the manuscript. All authors reviewed, revised and approved of the final version of the manuscript.

Disclosure of conflicts of interest

All authors have no conflicts of interest to declare.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig S1. The top figure shows the nucleotide sequences from a normal individual beginning from the *HBB* initiation codon ATG to codon 13. Codon 8 is AAG. The lower figure shows the nucleotide sequences from the proband. Codon 8 is GTC, as a result of homozygous deletion of dinucleotides AA from the normal codon 8 sequences.

Fig S2. MLPA on proband's *HBB* gene cluster showing equal signal strength of probes from LCR to 9 kb downstream from *HBB*.

Fig S3. The figure shows ~14·5 kb *HBG1 – HBD* intergenic region. The numbers on the first row represent hg19 Chr11 positions. Each bar represents a single SNP found in the proband, all homozygous for the alternate alleles (Table SIII). Also shown are: enhancer 3' to *HBG1* (Bodine, Ley, EMBO J 1987; 6:2997-3004), BGLT3 template for a long non-coding RNA (Kiefer *et al*, Blood 2011; 118:6200–6208), Corfu deletion (Kulozik *et al*, Blood 1988; 71:457–462), PYR complex DNA-binding site (Bank, Blood 2006; 107:435–443), and two purported BCL11A binding sites, at 3 kb downstream of *HBG1*, and 1 kb upstream of *HBD* (Sankaran *et al*, N Engl J Med 2011; 365:807–814). SNP rs10128556 is associated with elevated HbF in sickle cell anemia patients (Galarneau *et al*, Nat Genet 2010; 42:1049–1051). SNP rs2071348 is associated with elevated HbF and milder disease severity in patients with HbE/β^0-thalassemia (Nuinoon *et al*, Hum Genet 2010; 127:303–314).

Table SI. ARMS tests to determine HbF QTL genotypes.

Table SII. Non-identical SNP genotypes at 7 loci found in two dizygotic twins.

Table SIII. 21 SNPs present within proband's *HBD* to *hbg1* intergenic region.

Table SIV. FSC8 Homozygotes, with no or rare transfusion. Patients #3 and #4 were sisters. They were originally reported in 1989 by Gurney *et al*. Patients #6 and #16 (Table SV) were siblings. SNP alternate alleles are: A for rs7482144, C for rs9399137, del for rs66650371, and C for rs766432.

Table SV. FSC8 Homozygotes, transfusion dependent. Patients #13 and #14 were siblings. Patients #16 and #6 (Table SIV) were siblings. Patients #23 and #24 were siblings. Patient #23 underwent bone marrow transplantation in 2011 at age 3.

References

Akinsheye, I., Alsultan, A., Solovieff, N., Ngo, D., Baldwin, C.T., Sebastiani, P., Chui, D.H.K. & Steinberg, M.H. (2011) Fetal hemoglobin in sickle cell anemia. *Blood*, **118**, 19–27.

Al-Allawi, N.A.S., Hassan, K.M., Sheikha, A.K., Nerweiy, F.F., Dawood, R.S. & Jubrael, J. (2010) β-Thalassemia mutations among transfusion–dependent thalassemia major patients in Northern Iraq. *Molecular Biology International*, **2010**, 479282.

Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L., Sabo, P.J., Vierstra, J., Voit, R.A., Yuan, G.C., Porteus, M.H., Stamatoyannopoulos, J.A., Lettre, G. & Orkin, S.H. (2013) An erythroid enhancer of *BCL11A* subject to genetic variation determines fetal hemoglobin level. *Science*, **342**, 253–257.

Bianchi, E., Zini, R., Salati, S., Tenedini, E., Norfo, R., Tagliafico, E., Manfredini, R. & Ferrari, S. (2010) c-MYB supports erythropoiesis through the transactivation of KLF1 and LMO2 expression. *Blood*, **116**, e99–e110.

Chakalova, L., Osborne, C.S., Dai, Y.F., Goyenechea, B., Metaxotou-Mavromati, A., Kattamis, A., Kattamis, C. & Fraser, P. (2005) The Corfu δβ thalassemia deletion disrupts γ globin gene silencing and reveals post-transcriptional regulation of HbF expression. *Blood*, **105**, 2154–2160.

Cürük, M.A., Yüregir, G.T., Asadov, C.D., Dadasova, T., Gu, L.H., Baysal, E., Gu, Y.C., Ribeiro, M.L. & Huisman, T.H.J. (1992) Molecular characterization of β-thalassemia in Azerbaijan. *Human Genetics*, **90**, 417–419.

Ekwattanakit, S., Monteerarat, Y., Riolueang, S., Tachavanich, K. & Viprakasit, V. (2012) Association of *Xmn I* polymorphism and hemoglobin E haplotypes on postnatal γ globin gene expression in homozygous hemoglobin E. *Advances in Hematology*, **2012**, 528075.

Farrell, J.J., Sherva, R.M., Chen, Z.Y., Luo, H.Y., Chu, B.F., Ha, S.Y., Li, C.K., Lee, A.C., Li, R.C., Li, C.K., Yuen, H.L., So, J.C.C., Ma, E.S.K., Chan, L.C., Chan, V., Sebastiani, P., Farrer, L.A., Baldwin, C.T., Steinberg, M.H. & Chui, D.H.K. (2011) A 3-bp deletion in the *HBS1L-MYB* intergenic region on chromosome 6q23 is associated with HbF expression. *Blood*, **117**, 4935–4945.

Filon, D., Faerman, M., Smith, P. & Oppenheim, A. (1995) Sequence analysis reveals a β-thalassemia mutation in the DNA of skeletal remains from the archaeological site of Akhziv, Israel. *Nature Genetics*, **9**, 365–368.

Galarneau, G., Palmer, C.D., Sankaran, V.G., Orkin, S.H., Hirschhorn, J.N. & Lettre, G. (2010) Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nature Genetics*, **42**, 1049–1051.

Ganesh, S.K., Zakai, N.A., van Rooij, F.J., Soranzo, N., Smith, A.V., Nalls, M.A., Chen, M.H., Kottgen, A., Glazer, N.L., Dehghan, A., Kuhnel, B., Aspelund, T., Yang, Q., Tanaka, T., Jaffe, A., Bis, J.C., Verwoert, G.C., Teumer, A., Fox, C.S., Guralnik, J.M., Ehret, G.B., Rice, K., Felix, J.F., Rendon, A., Eiriksdottir, G., Levy, D., Patel, K.V., Boerwinkle, E., Rotter, J.I., Hofman, A., Sambrook, J.G., Hernandez, D.G., Zheng, G., Bandinelli, S., Singleton, A.B., Coresh, J., Lumley, T., Uitterlinden, A.G., Vangils, J.M., Launer, L.J., Cupples, L.A., Oostra, B.A., Zwaginga, J.J., Ouwehand, W.H., Thein, S.L., Meisinger, C., Deloukas, P., Nauck, M., Spector, T.D., Gieger, C., Gudnason, V., van Duijn, C.M., Psaty, B.M., Ferrucci, L., Chakravarti, A., Greinacher, A., O'Donnell, C.J., Witteman, J.C., Furth, S., Cushman, M., Harris, T.B. & Lin, J.P. (2009) Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nature Genetics*, **41**, 1191–1198.

Gurgey, A., Altay, C., Diaz-Chico, J.C., Kutlar, F., Kutlar, A. & Huisman, T.H.J. (1989) Molecular heterogeneity of β-thalassemia intermedia in Turkey. *Acta Haematologica*, **81**, 22–27.

Hamamy, H.A. & Al-Allawi, N.A.S. (2013) Epideminological profile of common haemoglobinopathies in Arab countries. *Journal of Community Genetics*, **4**, 147–167.

Jiang, J., Best, S., Menzel, S., Silver, N., Lai, M.I., Surdulescu, G.L., Spector, T.D. & Thein, S.L. (2006) cMYB is involved in the regulation of fetal hemoglobin production in adults. *Blood*, **108**, 1077–1083.

Kuliev, A.M., Rasulov, I.M., Dadasheva, T., Schwarz, E.I., Rosatelli, C., Saba, L., Meloni, A., Gemidjioglu, E., Petrou, M. & Modell, B. (1994) Thalassemia in Azerbaijan. *Journal of Medical Genetics*, **31**, 209–212.

Lemsaddek, W., Picanco, I., Seuanes, F., Mahmal, L., Benchekroun, S., Khattab, M., Nogueira, P. & Osório-Almeida, L. (2003) Spectrum of β-thalassemia mutations and HbF levels in the heterozygous Moroccan population. *American Journal of Hematology*, **73**, 161–168.

Lemsaddek, W., Picanco, I., Seuanes, F., Nogueira, P., Mahmal, L., Benchekroun, S., Khattab, M. & Osório-Almeida, L. (2004) The β-thalassemia mutation/haplotype distribution in the Moroccan population. *Hemoglobin*, **28**, 25–37.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., Best, S., Spector, T.D., Farrall, M., Lathrop, M. & Thein, S.L. (2007a) A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nature Genetics*, **39**, 1197–1199.

Menzel, S., Jiang, J., Silver, N., Gallagher, J., Cunningham, J., Surdulescu, G., Lathrop, M., Farrall, M., Spector, T.D. & Thein, S.L. (2007b) The *HBS1L-MYB* intergenic region on chromosome 6q23.3 influences erythrocyte, platelet, and monocyte counts in humans. *Blood*, **110**, 3624–3626.

Miller, B.A., Salameh, M., Ahmed, M., Wainscoat, J., Antognetti, G., Orkin, S.H., Weatherall, D.J. & Nathan, D.G. (1986) High fetal hemoglobin production in sickle cell anemia in the eastern province of Saudi Arabia is genetically determined. *Blood*, **67**, 1404–1410.

Nuinoon, M., Makarasara, W., Mushiroda, T., Setianingsih, I., Wahidiyat, P.A., Sripichai, O., Kumasaka, N., Takahashi, A., Svasti, S., Munkongdee, T., Mahasirimongkol, S., Peerapittayamongkol, C., Viprakasit, V., Kamatani, N., Winichagoon, P., Kubo, M., Nakamura, Y. & Fucharoen, S. (2010) A genome-wide association identified the common genetic variants influence disease severity in ß0-thalassemia/hemoglobin E. *Human Genetics* **127**, 303–314.

Öner, C., Dimovski, A.J., Altay, C., Gurgey, A., Gu, Y.C., Huisman, T.H.J. & Lanclos, K.D. (1992) Sequence variations in the 5′ hypersensitive site-2 of the locus control region of β^S chromosomes are associated with different levels of fetal globin in hemoglobin S homozygotes. *Blood*, **79**, 813–819.

Orkin, S.H. & Goff, S.C. (1981) Nonsense and frameshift mutations in β-thalassemia detected in cloned β-globin genes. *Journal of Biological Chemistry*, **256**, 9782–9784.

Orkin, S.H., Kazazian, H.H. Jr, Antonarakis, S.E., Goff, S.C., Boehm, C.D., Sexton, J.P., Waber, P.G. & Giardina, P.J. (1982) Linkage of β-thalassaemia mutations and β-globin gene polymorphisms with DNA polymorphisms in human β-globin gene cluster. *Nature*, **296**, 627–631.

Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van Ijcken, W., Lenhard, B., Rooks, H., Best, S., Menzel, S., Grosveld, F., Thein, S.L. & Soler, E. (2014) *HBS1L-MYB* intergenic variants moderate fetal hemoglobin via long-range MYB enhancers. *Journal of Clinical Investigation*, **124**, 1699–1710.

Tan, A.S., Quah, T.C., Low, P.S. & Chong, S.S. (2001) A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for α-thalassemia. *Blood*, **98**, 250–251.

Thein, S.L. & Menzel, S. (2009) Discovering the genetics underlying foetal haemoglobin production in adults. *British Journal of Haematology*, **145**, 455–467.

Weatherall, D.J. & Clegg, J.B. (2001) The Thalassaemia Syndromes, 4th edn. Blackwell Science, Oxford.

Zoueva, O.P., Garrett, L.J., Bodine, D. & Rodgers, G.P. (2008) BP1 motif in the human β-globin promoter affects β-globin expression during embryonic/fetal erythropoiesis in transgenic mice bearing the human β-globin gene. *Blood Cells, Molecules, and Diseases*, **41**, 244–251.