Influence of single nucleotide polymorphisms in the *BCL11A*, *HBS1L-MYB* intergenic region, and HBB gene cluster on the fetal hemoglobin levels in Bangladeshi patients with β-thalassemia/hemoglobin E disease

Jenjira Chalerm[1, 2], Mahmood Ahmed Chowdhury[3, 4], Chayanon Peerapittayamongkol[2], Kittiphong Paiboonsukwong[1], Suthat Fucharoen[1], Orapan Sripichai[1]*

[1]*Thalassemia Research Center, Institute of Molecular Biosciences, Mahidol University, Nakorn Pathom, Thailand.*
[2]*Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand*
[3]*Department of Paediatrics, Chattagram Maa-Shishu O General Hospital, Chittagong, Bangladesh*
[4]*Thalassaemia Welfare Centre - Bangladesh, Chittagong, Bangladesh*

**Corresponding author: orapan.sri@mahidol.ac.th*

ABSTRACT

Thalassemia is considered a major health burden in the Southeast Asian and Indian populations. Patients with β-thalassemia display disease heterogeneity ranging from nearly asymptomatic to severe anemia with complications. Increased levels of fetal hemoglobin (HbF) can ameliorate the clinical severity of β-thalassemia patients. The HbF production is influenced by many quantitative trait loci (QTL). Three major HbF-QTLs are the *BCL11A* gene, the *HBS1L-MYB* intergenic region (HMIR), and the β-globin (*HBB*) gene cluster. Therefore, this study aimed to evaluate the influence of these genetic modifiers (*BCL11A*, rs766432; HMIR, rs4895411; and *HBB* cluster, rs2071348) on the HbF levels in Bangladeshi patients with β-thalassemia/HbE disease. The cohort study comprised of 90 patients with β-thalassemia/HbE disease from in and around the area of Chittagong, Southeast Bangladesh. The HbF levels ranged from 7.9% to 59.1%. The results showed that levels of HbF were primarily influenced by alleles of the *HBB* cluster (rs2071348), and to a lesser extent by rs766432 *HBL11A* gene and HMIP (rs4895441) loci. The rs2071348 SNP explained 12.5% of the variation in the HbF levels, while 3.6% and 3.9% of trait variation were explained by rs766432 and rs4895441, respectively. In a case-control model of the low and high HbF analysis, we found that genotypes *AC* and *AA* ($p = 2.0\times10^{-4}$) and the allele *C* ($p = 2.0\times10^{-4}$) of rs2071348, and genotypes *AG* and *GG* ($p = 0.02$) and the allele *G* ($p = 0.05$) of rs4895441were associated with a significantly higher frequency with high HbF. However, the rs766432 did not exhibit such features. Our results suggest these three major HbF-QTLs as the influencing phenotypic factors of β-thalassemia in Bangladeshi β-thalassemia/HbE patients.

Keywords: β-thalassemia; HbE; fetal hemoglobin; HbF; SNPs

INTRODUCTION

β-Thalassemia/HbE is a monogenic disorder caused by a compound heterozygous mutation of the β-globin gene. Patients with β-thalassemia/HbE co-inherit the β-thalassemia allele with the structural variant hemoglobin E (HbE; c.79G>A (p.Glu27Lys)). The condition presents with exceptional and variable clinical symptoms ranging from a mild asymptomatic anemia to a transfusion-dependent thalassemia (Sripichai *et al*., 2008). The reduction of β-globin chain synthesis in β-thalassemia causes an excess of unbounded α-globin chains, which precipitate in erythroid cells, leading to ineffective erythropoiesis and a shortened lifespan of red blood cells (Fucharoen and Weatherall, 2016). Therefore, the degree of imbalance between the α-globin chains and the non-α-globin chains dictates the phenotypic severity of the disorder. Increased production of the γ-globin chains can thus modulate the β-thalassemia disease severity by reducing the pool of free α-chains (Weatherall, 2001). In particular, an increase in the levels of fetal hemoglobin (HbF; α2γ2) in many studies was shown to affect the clinical outcome of patients with β-thalassemia (Sripichai and Fucharoen, 2016). Remarkably, the diversity of HbF levels is strongly influenced by genetic factors.

Genome-wide association studies (GWAS) have identified quantitative trait loci (QTL) harboring single nucleotide polymorphisms (SNPs) showing strong associations with HbF levels. The major HbF-QTLs include the *BCL11A* gene on chromosome 2p16.2 (Menzel *et al*., 2007; Uda *et al*., 2008; Nuinoon *et al*., 2010), the *HBS1L-MYB* intergenic region (HMIR) on

chromosome 6q23.3 (Thein *et al*., 2007; Lettre *et al*., 2008; Nuinoon *et al*., 2010), and the β-like-globin gene cluster (*HBB* cluster) on chromosome 11p15.5 (Lettre *et al*., 2008; Thein *et al*., 2009; Nuinoon *et al*., 2010). Several genetic association studies in different populations have shown that the impact of each variant on the variation of HbF levels could be population-specific and is dependent on the genetic background and frequency of each HbF-associated allele (Menzel *et al*., 2007; Lettre *et al*., 2008; Galarneau *et al*., 2010; Thein, 2013)

However, the genetic basis of the β-thalassemia/HbE disease heterogeneity and the genotype-phenotype association are not entirely clear. Several HbF-QTLs previously identified in other populations have not yet been assessed in Bangladeshi individuals. Therefore, the aim of this study was to evaluate the associations between genetic variations in the major HbF-QTLs, namely *BCL11A*, HMIR and *HBB* cluster, and the HbF levels in Bangladeshi patients with β-thalassemia/HbE disease.

MATERIALS AND METHODS

Study population

A total of 90 unrelated β-thalassemia/HbE patients (38.9% were female) with different HbF levels were recruited. The patient ages were 1-43 years (mean age 11.2 ± 8.2 years; median age 9 years). The HbF levels were determined by high performance liquid chromatography (HPLC) using the VARIANT™ Hemoglobin Testing System (β-Thalassemia Short Program, Bio-Rad Laboratories, Hercules, CA, USA). The levels of HbF in this cohort ranged from 7.9% to 59.1%, the mean was 27.3% (SD 13.1) and the median was 24.6%. Forty-four subjects with HbF levels $\geq$ 25% were classified as high HbF whereas 46 subjects with HbF levels < 25% were categorized as low HbF. The study was approved by the Institutional Ethical Committee (IEC) in accordance with the Declaration of Helsinki; certificate of approval number 2010/010.0701. Written informed consent was obtained from all participants or their respective parents in the case of children.

Genotyping

Genomic DNA was extracted from peripheral blood using the Gentra Puregene blood kit (Qiagen Science, Germantown, MD, USA) following the manufacturer's instructions. The quality and quantity of DNA were measured by a Nano Drop spectrophotometer. The β-thalassemia and HbE mutations were confirmed by the GenoFlow Beta thalassemia array test kit (DiagCor Bioscience, Hong Kong, China). Three SNPs were chosen for this study based on the previous studies that reported genetic variants most strongly associated with increased HbF levels in β-thalassemia/HbE patients; *BCL11A* (rs766432), HMIP (rs4895441), and *HBB* cluster (rs2071348). SNPs were genotyped by a PCR-HRM assay using specific primers (oligonucleotide sequences available upon request). Allelic discrimination was carried out using SsoFast EvaGreen Supermix (Bio-rad Laboratories, Hercules, CA, USA) in a CFX Connect™ Real-Time PCR machine (Bio-rad Laboratories).

Statistical analysis

The genotypic distribution frequencies of all polymorphisms were analyzed for Hardy-Weinberg equilibrium (HWE). The association between the HbF levels and the SNPs was analyzed by linear regression under an additive genetic model and by a case-control model of low and high HbF levels (the cutoff value was set up by the median at 25% HbF). Odds ratio (OR) and 95% confidence intervals (CIs) were calculated for association of high and low HbF phenotypes with genotypic and allelic tests. All analyses were conducted using the statistical package R, version 2.5.1. A P-value below 0.05 was considered statistically significant.

RESULTS

The minor allele frequencies (MAF) and HWE observed for the three HbF-associated polymorphisms in the total sample were displayed in Table 1. The minor allele of the *HBB* cluster SNP rs2071348 was *C*, accounting for 34% of the total chromosomes. By contrast, the HMIP SNP, rs4895441 only had 10% of the minor allele *G*, with a *GG* homozygosity percentage of 1%. Meanwhile, the *C* allele of the *BCL11A* SNP rs766432 constituted 1.4% of the whole chromosomes examined, whereas homozygous *CC* was not present in this cohort. The genotype distributions of rs766432 and rs4895441were in agreement with the Hardy-Weinberg equilibrium (HWE; P > 0.05), but the rs2071348 was not in HWE (P = 0.0007).

The distribution of HbF levels within each genotype group of the three HbF- associated SNPs was examined, as shows in box-plots in Figure 1. β-Thalassemia/HbE patients who were homozygous for the major ancestral alleles, *AA* for rs766432, *AA* for rs4895441 and *AA* for rs2071348, showed lower HbF values when compared to those with the derived alleles. The *CC* and *AC* genotypes of rs2071348 in the *HBB* gene cluster showed the higher mean HbF levels

Table 1. HbF association results for the SNPs *BCL11A* rs766432, HMIR rs4895441 and *HBB* gene cluster rs2071348 in Bangladeshi patients with β-thalassemia/HbE disease.

Chr:position	Gene	SNP	Variant	MAF (allele)	P-HWE	β (SE)	Variance explained (%)	P
2:60492835	*BCL11A*	rs766432	A/C	0.14 (C)	0.463	6.23 (2.99)	3.63	0.040
6:135105435	HMIP	rs4895441	A/G	0.10 (G)	1.000	6.82 (3.17)	3.94	0.034
11:5242916	*HBB* cluster	rs2071348	A/C	0.34 (C)	0.0007	5.71 (1.54)	12.47	0.0004

Abbreviations: MAF: minor allele frequency; p-HWE: P-value for Hardy–Weinberg Equilibrium; HMIR: HBS1L-MYB intergenic region. The table includes the effect sizes of the minor allele (regression coefficient beta, β), standard error (SE) and P-values for the HbF levels using a linear regression model. Chromosome position (Chr:position) is according Ensembl.

(33.1±11.6% and 33.4±15.2%, respectively) than the *AA* genotype (22.8±11.2%). The *AC* genotype of *BCL11A* SNP rs766432 showed high HbF levels of 31.7±13.1 % as compared to the *AA* genotype (25.5±12.7%). Similarly, the *AG* genotype of HMIP SNP rs4895441 showed higher HbF levels (35.3±13.0%) in comparison to the *AA* genotype (25.7±12.6%). Regression analysis under an additive genetic model was carried out to examine the impact of SNPs in modifying levels of HbF in the patients. The results showed that rs766432, rs4895441 and rs2071348 polymorphisms displayed similar correlation coefficient values (6.23%, 6.82% and 5.71%, respectively) in increasing HbF level. The HbF phenotypic variation was explained by the three loci from 23.5% to 30.1%. These three SNPs explained more variation in the HbF levels than the cumulative sum of the phenotypic variance explained by the SNPs individually (7.3% compared to 6.8%), although the difference was not statistically significant. The rs2071348 SNP was more strongly associated with the HbF levels ($P = 4.0\times10^{-4}$) than rs766432 and rs4895441 (P = 0.04 and 0.03, respectively) and accounted for 12.5% of the trait variance. The effect of these three HbF-associated SNPs could explain 20% of phenotypic variation in the HbF levels.

To further analyze the distribution frequencies of genotypes and alleles, the β-thalassemia/HbE patients were divided into two groups, a high HbF group (n = 44; range 25.5% to 59.1%, mean HbF = 37.8±10.2%), and a low HbF group (n = 46; range 7.9% to 24.7%, mean HbF = 17.2±5.1%) according to the cutoff value of 25% HbF (Table 2). The mean HbF levels differ significantly between the two groups ($P = 1.2\times10^{-17}$). The case-control study showed significant associations between levels of HbF and the *HBB* cluster SNP rs2071348 (OR = 5.56; 95% CI, 2.27–14.29; $P = 2.0\times10^{-4}$) and HMIP SNP rs4895441 (OR = 4.35; 95% CI, 1.41–16.67; P = 0.02) (Table 2). We found that the *C* allele of rs2071348 and the *G* allele of rs4895441 had a significantly higher frequency in the high HbF group than the low HbF group ($P = 7.0\times10^{-4}$ and 0.05, respectively). Although, the linear regression models revealed a statistical significant association between *BCL11A* SNP rs766432 and levels of HbF, this SNP was not found to correlate with the HbF levels in the case-control model of low and high HbF analysis.

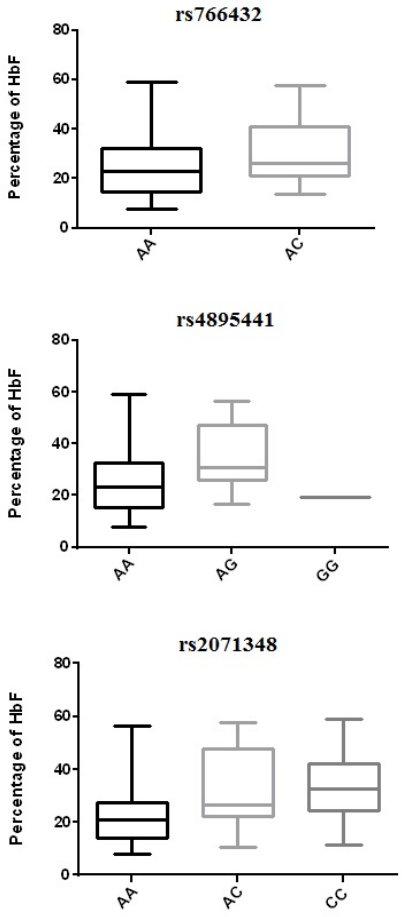

Figure 1 Box plots showing the distribution of HbF levels within genotypes of the SNPs *BCL11A* rs766432, HMIR rs4895441 and *HBB* gene cluster rs2071348 in Bangladeshi patients with β-thalassemia/HbE disease. Each rectangle represents the data between the 25th and 75th quartiles, and the bar within each rectangle is the median value for HbF.

Table 2. Distribution frequencies of genotypes and alleles for the SNPs *BCL11A* rs766432, HMIR rs4895441 and *HBB* gene cluster rs2071348 in the high and low HbF groups of Bangladeshi β-thalassemia/HbE patients.

SNPs	HbF Group	Genotypes		P	OR (95% CI)	Alternative allele	Reference allele	P	OR (95% CI)
rs766432		AC+CC	AA			C	A		
	High	14	30	0.545	1.32 (0.53-3.33)	14	74	0.549	1.32 (0.53-3.34)
	Low	12	34			8	80		
rs4895441		AG+GG	AA			G	A		
	High	13	31	0.017	4.35 (1.41-16.67)	13	75	0.047	3.05 (1.08-10.14)
	Low	4	42			5	87		
rs2071357		AC+CC	AA			C	A		
	High	28	16	0.0002	5.56 (2.27-14.29)	44	44	0.0007	2.64 (1.54-4.79)
	Low	11	35			17	75		

Abbreviations: SNPs: single nucleotide polymorphisms; HbF: fetal hemoglobin; OR (95% CI): odds ratio (95% confidence interval).

DISCUSSION

The reactivation of γ-globin expression in β-thalassemia patients is a response to the reduced level of β-globin synthesis due to the β-thalassemia mutation. The high levels of HbF production contribute to the amelioration of the β-thalassemia clinical manifestations and disease severity. In this study of 90 Bangladeshi β-thalassemia/HbE patient cohort, the range of HbF levels was 7.9–59.1%. According to a systematic review, this level was in accordance with previous studies where HbF levels were found to be 7–70% (Sripichai and Fucharoen, 2016). Genetic variants in three modifier loci; the *BCL11A* gene, the HBS1L-MYB intergenic region and the β-globin like gene cluster, have been proved to be the major factors that can regulate the expression of HbF in different ethnic populations. Similarly, SNPs rs766432 in the *BCL11A* gene, rs4895441 in the HMIR and rs2071348 in the HBB cluster were identified to be significantly associated with the HbF levels in our cohort. This is the first report on the association between these modifier loci and HbF values in Bangladeshi population.

The rs2071348 *A/C* polymorphism, located at the intergenic region between the *HBPB* and the *HBD* genes in the *HBB* cluster, has shown the strongest association with the HbF levels in the cohort of 618 Thai patients with β-thalassemia/HbE (Nuinoon *et al*., 2010). Another study in Indonesian patients also showed that rs2071348 was associated with HbF levels (Nuinoon *et al*., 2010). High-risk genotypes and alleles of SNP rs2071348 (*CC* and *AC*, allele *C*) were determined to be susceptible to having high HbF levels in this study, which demonstrated that there was no difference in racial difference for this variant. In addition, the *C* allele of rs2071348 was significantly more frequent in mildly affected β-thalassemia/HbE patients than in severely affected patients, suggesting that this SNP may help to predict the disease severity (Sherva *et al*., 2010).

Studies in populations in Europe and Asia showed that several polymorphisms in the *HBS1L-MYB* intergenic region had a strong relationship with high HbF levels in both β-hemoglobinopathy patients and normal individuals (Thein *et al*., 2007; Sherva *et al*., 2010; Makani *et al*., 2011). In this study, the rs4895441 SNP was selected to evaluate the effect on HbF levels, and the significant effect was observed by the *G* allele. We found the rs4895441 SNP had significant correlation with the HbF levels in both regression analysis and the case-control study, replicating previous findings in Thai and Indonesian patients. The frequency of the *G* allele in our study was 10%, which is lower than approximately 18% in Thai and Indonesian. However, this study is carried out in a small cohort of Bangladeshi β-thalassemia/HbE patients. Disruption in the intergenic region between the *HBS1L* and *MYB* leads to elevated HbF expression (Farrell *et al*., 2011; Suzuki *et al*., 2013) through the reduction of transcription factor bindings that affect long-range interactions with *MYB*, resulting in reduced *MYB* expression (Stadhouders *et al*., 2014). MYB is an erythroid transcription factor, whose down-regulation in adult erythroid cell was reported to result in increased γ-globin expression with reduced cell expansion and accelerated erythroid differentiation (Jiang *et al*., 2006; Soza-Ried *et al*., 2010), suggesting that levels of MYB might affect HbF production through its influence on the cell cycle. HBS1L is a G-protein/elongation factor involved in a variety of cellular processes. But overexpression of *HBS1L* in K562 erythroid cells did not affect the γ-globin expression (Jiang *et al*., 2006).

Our data in β-thalassemia/HbE patients showed nominal statistically significant associations

between *BCL11A* SNP rs766432 and the HbF levels. rs766432 is known to correlate with the levels of HbF in various cases of β-hemoglobinopathies in Thai, Chinese, Middle East, European, and African populations (Sedgewick *et al*., 2008; Nuinoon *et al*., 2010; Akinsheye *et al*., 2012; Pule *et al*., 2015). Thus, our finding is in line with previous studies, confirming that the rs766432 SNP has a significant correlation with increasing HbF levels among thalassemia patients. However, case-control models that assess the correlation of variants in patients with HbF levels between high and low HbF (the considered cutoff point of 25%) did not show significant association between HbF and rs766432. The absence of significant association between rs766432 and the HbF levels may be explained by the low population frequency of this variant, with a 1.4% frequency for the minor *C*-allele in the cohort. BCL11A is a regulator responsible for γ-globin to β-globin gene switching (Sankaran *et al*., 2008). BCL11A interacts with other repressor factors to form a repressor complex, thereby silencing the γ-globin gene in adult erythroid cells. The rs766432 SNP is located within an erythroid-specific enhancer of *BCL11A* and likely to influence *BCL11A* expression (Bauer *et al*., 2013).

In conclusion, our results suggest that the level of HbF in Bangladeshi β-thalassemia/HbE patients is strongly associated with the *HBB* gene cluster rs2071348, but also with *BCL11A* rs766432 and HMIR rs4895441. The effect of these three HbF-associated SNPs explained for 20% of the variance in HbF levels in this cohort. A larger sample size is needed and the mechanistic function(s) of these modifier loci should be explored in future studies.

ACKNOWLEDGEMENTS

This work was supported by Mahidol University and the Research Chair Grant, National Science and Technology Development Agency (NSTDA), Thailand. Jenjira Chalerm received the Siriraj Graduate Scholarship, Faculty of Medicine Siriraj Hospital, Mahidol University.

REFERENCES

Akinsheye I, Solovieff N, Ngo D, Malek A, Sebastiani P, Steinberg MH, Chui DH (2012) Fetal hemoglobin in sickle cell anemia: molecular characterization of the unusually high fetal hemoglobin phenotype in African Americans. Am J Hematol 87: 217–219.

Bauer DE, Kamran SC, Lessard S, Xu J, Fujiwara Y, Lin C, Shao Z, Canver MC, Smith EC, Pinello L, *et al*. (2013) An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science 342: 253–257.

Farrell JJ, Sherva RM, Chen Z, Luo H, Chu BF, Ha SY, Li CK, Lee ACW, Li RCH, Li CK, *et al*. (2011) A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood 117: 4935–4945.

Fucharoen S, Weatherall DJ (2016) Progress toward the control and management of the thalassemias. Hematol Oncol Clin North Am 30: 359–371.

Galarneau G, Palmer CD, Sankaran VG, Orkin SH, Hirschhorn JN, Lettre G (2010). Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. Nat Genet 42: 1049–1051.

Jiang J, Best S, Menzel S, Silver N, Lai MI, Surdulescu GL, Spector TD, Thein SL (2006) cMYB is involved in the regulation of fetal hemoglobin production in adults. Blood 108: 1077–1083.

Lettre G, Sankaran VG, Bezerra MA, Arau'jo AS, Uda M, Sanna S, Cao A, Schlessinger D, Costa FF, Hirschhorn JN, *et al*. (2008) DNA polymorphism at the BCL11A, HBS1L-MYB, and β-globin loci associates with fetal haemoglobin levels and pain crises in sickle cell disease. Proc Natl Acad Sci USA 105: 11869–11874.

Makani J, Menzel S, Nkya S, Cox SE, Drasar E, Soka D, Komba AN, Mgaya J, Rooks H, Vasavda N, *et al*. (2011) Genetics of fetal hemoglobin in Tanzanian and British patients with sickle cell anemia. Blood 117: 1390–1392.

Menzel S, Garner C, Gut I, Matsuda F, Yamaguchi M, Heath S, Foglio M, Zelenika D, Boland A, Rooks H, *et al*. (2007) A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet 39: 1197–1199.

Nuinoon M, Makarasara W, Mushiroda T, Setianingsih I, Wahidiyat PA, Sripichai O, Kumasaka N, Takahashi A, Svasti S, Munkongdee T, *et al*. (2010) A genome-wide association identified the common genetic variants influence disease severity in β-thalassemia/ hemoglobin E. Hum Genet 127: 303–314.

Pule GD, Ngo Bitoungui VJ, Chetcha CB, Kengne AP, Antonarakis S, Wonkam A (2015) Association between variants at BCL11A erythroid-specific enhancer and fetal hemoglobin levels among sickle cell disease patients in cameroon: implications for future therapeutic interventions. Omics 19: 627–631.

Sankaran VG, Menne TF, Xu J, Akie TE, Lettre G, Van Handel B, Mikkola HK, Hirschhorn JN, Cantor AB, Orkin SH (2008) Human fetal hemoglobin expression

is regulated by the developmental stage-specific repressor BCL11A. Science 322: 1839–1842.

Sedgewick AE, Timofeev N, Sebastiani P, So JC, Ma ES, Chan LC, Fucharoen G, Fucharoen S, Barbosa CG, Vardarajan BN, *et al.* (2008) BCL11A is a major HbF quantitative trait locus in three different populations with beta-hemoglobinopathies. Blood Cells Mol Dis 41: 255–258.

Sherva R, Sripichai O, Abel K, Ma Q, Whitacre J, Angkachatchai V, Makarasara W, Winichagoon P, Svasti S, Fucharoen S, *et al.* (2010) Genetic modifiers of Hb E/beta0 thalassemia identified by a two-stage genome-wide association study. BMC Med Genet 11: 51–59.

Soza-Ried C, Hess I, Netuschil N, Schorpp M, Boehm T (2010) Essential role of c-myb in definitive hematopoiesis is evolutionarily conserved. Proc Natl Acad Sci USA 107: 17304–17308.

Sripichai O, Fucharoen S (2016) Fetal hemoglobin regulation in β-thalassemia: heterogeneity, modifiers and therapeutic approaches. Expert Rev Hematol 9: 1129–1137.

Sripichai O, Makarasara W, Munkongdee T, Kumkhaek C, Nuchprayoon I, Chuansumrit A, Chuncharunee S, Chantrakoon N, Boonmongkol P, Winichagoon P, *et al.* (2008) A scoring system for the classification of beta-thalassemia/Hb E disease severity. Am J Hematol 83: 482–484.

Stadhouders R, Aktuna S, Thongjuea S, Aghajanirefah A, Pourfarzad F, van Ijcken W, Lenhard B, Rooks H, Best S, Menzel S, *et al.* (2014) HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. J Clin Invest 124: 1699–1710.

Suzuki M, Yamazaki H, Mukai HY, Motohashi H, Shi L, Tanabe O, Engel JD, Yamamoto M (2013) Disruption of the Hbs1l-Myb locus causes hereditary persistence of fetal hemoglobin in a mouse model. Mol Cell Biol 33: 1687–1695.

Thein SL, Menzel S, Peng X, Best S, Jiang J, Close J, Silver N, Gerovasilli A, Ping C, Yamaguchi M, *et al.* (2007) Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. Proc Natl Acad Sci USA 104: 11346–11351.

Thein SL, Menzel S (2009) Discovering the genetics underlying foetal haemoglobin production in adults. Br J Haematol 145: 455–467.

Thein SL (2013) Genetic association studies in β-hemoglobinopathies. Hematology Am Soc Hematol Educ Program 2013: 354–361.

Uda M, Galanello R, Sanna S, Lettre G, Sankaran VG, Chen W, Usala G, Busonero F, Maschio A, Albai G, *et al.* (2008) Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta thalassemia. Proc Natl Acad Sci USA 105: 1620–1625.

Weatherall DJ (2001) Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. Nat Rev Genet 2: 245–255.