Haemoglobin switching modulator SNPs rs5006884 is associated with increased HbA_2 in β-thalassaemia carriers

Cyril Cyrus[1], Chittibabu Vatte[1], Shahanas Chathoth[1], Abdul Azeez Sayed[2], J. Francis Borgio[2], Mohammed Abdullah Alrubaish[3], Rawan Alfalah[3], Jana Alsaikhan[3], Amein K. Al Ali[1]

[1]Department of Biochemistry, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[2]Department of Genetic Research, Institute for Research and Medical Consultation, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[3]King Fahd Hospital of the University, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Submitted: 28 March 2018; **Accepted:** 15 May 2018
Online publication: 18 July 2019

Arch Med Sci 2021; 17 (4): 1064–1074
DOI: https://doi.org/10.5114/aoms.2019.86705

Corresponding author:
Dr. Cyril Cyrus
Department
of Biochemistry
College of Medicine
Imam Abdulrahman Bin
Faisal University
P.O. Box 1982, Dammam 31441
Saudi Arabia
Phone: +966553241441
E-mail: ccyrus@iau.edu.sa

Abstract

Introduction: Haemoglobin A_2 (HbA_2), the tetramer of α- and δ-globin chains, is used as a diagnostic biomarker for β-thalassaemia carriers. The HbA_2 levels are regulated by the presence of HPFH, δ-thalassaemia, $HbA_{1/2}$ gene triplication, and variants of KLF1, β-globin gene, and HbF regulating QTLs. Saudi Arabia has a high incidence of borderline HbA_2 levels, thereby making it difficult to classify the haemoglobinopathies. This study aims to investigate the association of known HbF enhancer QTL gene SNPs with HbA_2 levels
Material and methods: 14 Specific SNPs in *BCL11A*, *HMIP*, *OR51B6*, *HBBP1*, and *HBG2* loci were genotyped in 164 Saudi β-thalassaemia carriers by TaqMan assay to validate their role as regulators of HbA_2 levels. HbA2 levels were determined using the Variant II β-Thalassemia Short Program Recorder kit. The non-random association of these SNPs was tested using HaploView software. Protein interaction was assessed using 3D structure modelling for OR51B6 (rs5006884), comparative energy minimisation, and root-mean-square deviation (RMSD) prediction.
Results: Elevated HbA_2 levels were associated with SNPs in *HBBP1*, *OR51B6*, and TCT haplotype from HBG2 promoter region. The bioinformatics modelling and prediction revealed that the exonic rs5006884 had RMSD value deviations and significantly varied binding energy minimisation. α-globin variations were found in 57.92% of individuals but were not associated with elevated HbA_2.
Conclusions: The haemoglobin switching modulators rs2071348, rs7482144, and rs5006884 are involved in regulation of HbA_2 level with rs5006884 influencing the tetramer formation. Screening for haemoglobinopathies should take these SNPs into consideration, specifically in borderline HbA_2 cases. Assiduous analysis of rs5006884 as HbA_2 modulator for amelioration of disease severity is recommended.

Key words: β-thalassaemia, haemoglobin A_2, SNPs, OR51B6, tetramer.

Introduction

Haemoglobin A_2 (HbA_2) is a tetramer of α- and δ-globin chains, which constitutes approximately 2-3% of the total circulating haemoglobin. The *HBD*-coded HbA_2 prevents the polymerisation of deoxy-sickle haemoglobin and is a reliable diagnostic biomarker for β-thalassaemia trait. The globin genes are arranged in the order of their expression and are activated sequentially as embryonic (HBE1, ε), foetal (HBG1/2, γ), and adult (HBD/HBB, δ/β) genes during development. The globin promoters initiate the molecular switching between these genes by gaining access to the upstream regulatory elements, namely the β locus control region (β-LCR) [1] and autonomous silencing of the preceding gene. Thus, the gene closest to the β-LCR is activated first (*HBE1*) and the furthest (*HBB*) is activated last. Inactivation of the HBB promoter results in extended interaction of β–LCR with the HBD and HBG promoters, thereby leading to elevated HbA_2 and HbF levels. The deficit of β-globin and profusion of α-globin results in the post-translational elevation of HbA_2 levels in β-thalassaemia carriers [2]. However, HbA_2 is not characteristically elevated in all β-thalassaemia carriers, making it difficult to classify haemoglobinopathies in patients with borderline HbA_2 levels. Saudi Arabia has a high incidence of borderline HbA_2 levels. HbF level-associated QTLs are differentially distributed among populations and a reciprocal relationship among the HbA_2 and HbF levels has been inferred.

Regulation of HbA_2 level which is heritable with genetic variations accounting for 42% of total HbA_2 variability [2], which is either attributed to a variation in the β-globin gene or its regulators, in addition to other factors such as the presence of δ-thalassaemia, triplication of $HbA_{1/2}$ genes, hereditary persistence of foetal haemoglobin (HPFH) and Krueppel-like factor 1 (*KLF1*) variants, all of which make it difficult to classify haemoglobinopathies [3]. *KLF1* variants are an important modulator of haemoglobin switching [4] by transcriptional repression of the δ-globin gene and activation of the HbA_2 gene (HBD), resulting in increased HbA_2 levels [5–8]. *KLF1* also increases HbF levels. Our recent study on *KLF1* variations reported no significant association with borderline HbA_2 among Saudis [9].

In northern Europeans, altered transcription within the HBB cluster and SNPs in HBS1L-MYB locus were associated with HbA_2 [2]. HbA_2 has an advantage over HbF as it is pancellularly distributed and expressed [10]. Therefore, in sickle cell disease (SCD), increased levels might have therapeutic potential compared with HbF [11]. Few studies have been conducted on the role of other HbF-regulating QTLs on the level of HbA_2. An increased level of HbA_2 may ameliorate the severity of SCD and β-thalassaemia in the same manner that an increased level of HbF does. Haematopoiesis and HbF level associated SNPs of the HBS1L-MYB interval are reported to induce higher HbF levels in SCD and are differentially distributed among populations [12, 13]. A reciprocal relationship among the HbA_2 and HbF levels has been inferred in acquired disorders [10].

The aim of this study was to investigate the association of HbA_2 level with 14 SNPs of the HbF enhancer QTL genes. The molecular interactions involved on the variations in the level of HbA_2 synthesis is not completely understood. Shedding light on the role of SNPs that act as triggering factors of HbA_2 may indicate that inducing elevated levels of HbA_2 could be of clinical importance.

Material and methods

Ethical approval

This study was approved by the Ethical Committee of Imam Abdulrahman Bin Faisal University in accordance with the 1964 Helsinki Declaration and its later amendments. Written, informed consent was obtained from each participant.

Pheno- and genotyping

This study was conducted on 164 β-thalassaemia carriers residing in the Eastern Province of Saudi Arabia. The blood samples were collected in EDTA anti-coagulated vacutainers for HbA_2 level estimation using Bio-Rad Variant II (Variant II β-Thalassaemia Short Program Recorder kit) and DNA extraction using blood mini kit (Qiagen, USA).

SNP genotyping of 14 SNPs, namely rs2071348, rs7482144, rs5006884 (*HBG2* promoter region), rs766432, rs11886868, rs4671393, rs7557939 (*BCL11A* region), rs28384513, rs9376090, rs9399137, rs4895441, rs9389269, rs9402686, and rs9494142 (*HMIP* region) was carried out by nuclease allelic discrimination assay. The target-specific forward and reverse primers along with TaqMan probes labelled with VIC and FAM for each allele (TaqMan Assay, Applied Biosystems) were amplified on the ABI 7500 real time PCR system (ABI, Foster City, USA) according to the manufacturer's instructions.

The α-3.7 deletion and variations were identified using ViennaLab StripAssays and PCR-Sequencing method. The globin gene was amplified as previously described [14]. In summary, the amplicons were purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany), and the purified PCR products were cycle sequenced using a Big Dye Terminator Cycle Sequencing Kit (Applied Biosystem, Foster City, CA, USA). The cycle-sequenced products were then purified and

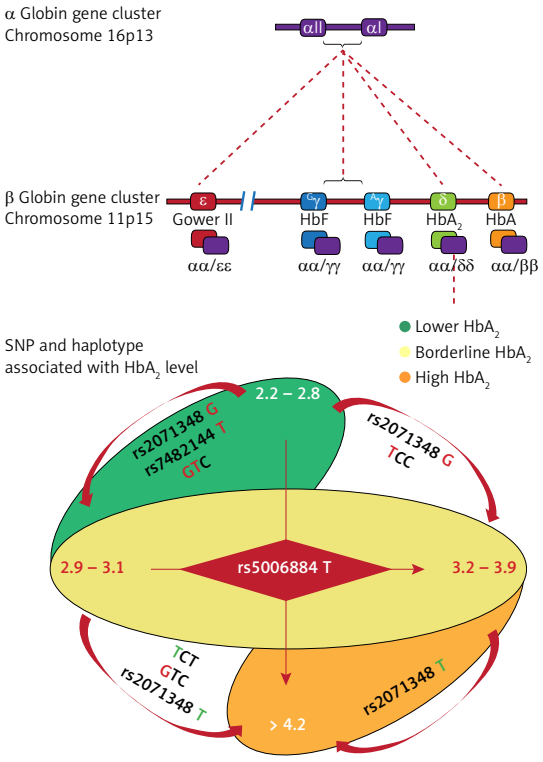


Figure 1. Diagram exhibiting the SNP and haplotype associated with the elevation of HbA_2 levels between low to borderline and borderline to high

electrophoresed in a Genetic Analyzer 3500 (Life Technologies Corporation, Carlsbad, CA, USA). Sequencing Analysis Software Version 5.4 (Applied Biosystem, USA) was used for data analysis.

Structure modelling and root-mean-square deviation (RMSD) prediction

Three-dimensional (3D) protein structure of OR51B6 native chain was designed using automated homology modelling [15] and the mutated OR51B6 protein using SWISS MODEL (http://swissmodel.expasy.org/). The 3D structure was generated based on template Protein Data Bank (PDB) ID: 5cxv with highest resolution of 2.7*Å*. PROCHECK (http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) online was used to validate the native and modelled structure [16]. The mutant models were generated using SWISS PDB Viewer Ver. 4.1 (http://spdbv.vital-it.ch/) [17]. The energy minimisation and its RMSD were checked for the native and mutants using the GROMACS 5.1.1 (http://www.gromacs.org/) program [18].

Protein-protein interaction

Protein-protein interaction was conducted using a PRISM server [19–21] because it predicts the interface structures between two interacting proteins. We submitted the haemoglobin proteins with both OR51B6 wild and mutant models to check the possible interaction variations between the wild and mutant proteins towards the haemoglobin proteins.

Statistical analysis

Hardy-Weinberg equilibrium was tested for all the SNPs, and χ^2 and odds ratio were determined by SPSS ver. 19 to evaluate allele association. Linkage disequilibrium (LD) test was carried out using HaploView 4.2 software [22] to identify the non-random association of these 14 SNPs. Haplotype blocks were constructed using HaploView 4.2 program. Haplotypes associated with the study subjects were inferred based on the partition-ligation approach through an EM algorithm. A p-value < 0.05 was considered significant for all statistical analyses.

Results

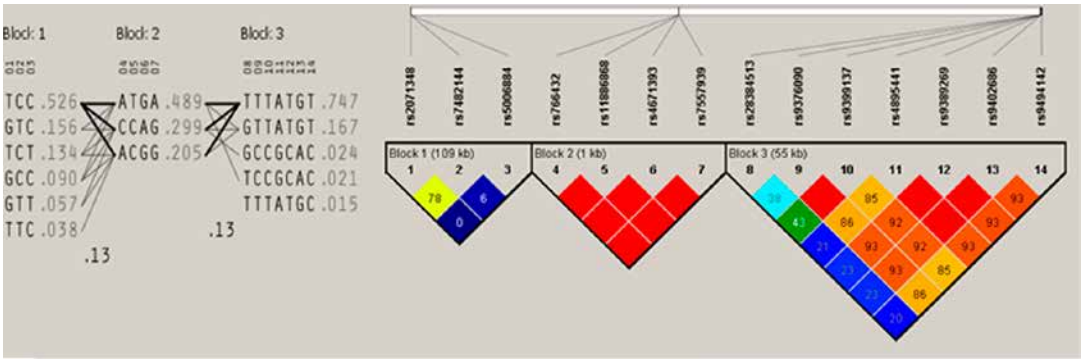
The independent segregation genotype for all the SNPs was in agreement with the Hardy-Weinberg equilibrium. Standard allelic association analysis of the 14 SNPs tested in the patient cohort showed that only three SNPs, namely rs2071348 (*HBBP1*), rs7482144 (*HBG2*), and rs5006884 (*OR51B6*), were significantly associated with borderline and higher HbA_2 levels. The rs5006884 (p = 0.02; χ^2 = 5.08) was the key SNP associated with elevated HbA_2 levels compared to the normal and borderline levels. There were no significant differences in allele frequencies of SNPs in the *HBS1L-MYB* region and *BCL11A* region between the normal and HbA_2-elevated cohorts (Table I, Figure 1).

The predominant haplotypes in the borderline and higher HbA_2 cohort consisted of GTC, TCC, and TCT in the HBG2 promoter region comprising SNPs rs2071348 (HBBP1 gene), rs7482144 (HBG2 gene), and rs5006884 (OR51B6), respectively (Table II, Figure 2). The TCT (rs2071348T, rs7482144C, rs5006884T) haplotype pattern was the most significant HbA_2 enhancer haplotype (p = 0.01, χ^2 = 7.8) (Table II).

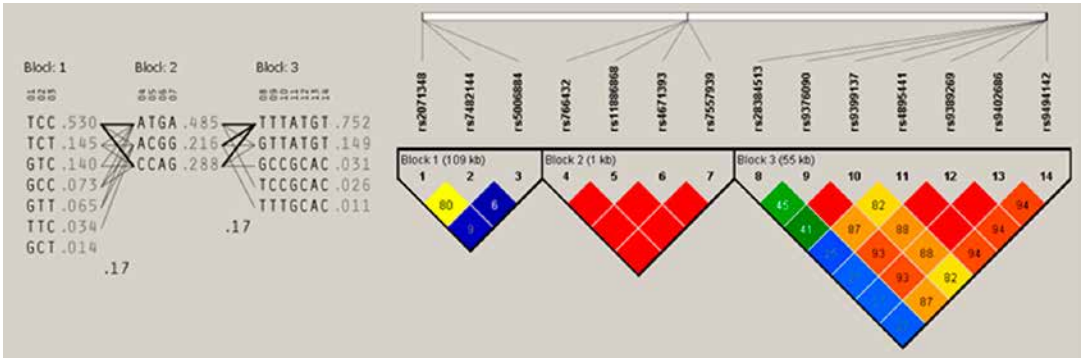
The α globin variations

The detailed study on the α globin genotype revealed that the $-\alpha_2^{3.7}/\alpha_1\alpha_2$ is the most frequent, with an overall prevalence of 57.92% (n = 95) of Saudi individuals included in the study. The other α globin variations included $\alpha_1^{-4.2}/\alpha_1\alpha_2$, $\alpha_1^{polyA-1}$, $\alpha_2/\alpha_1\alpha_2$, and $^{-3.7}\alpha_2/\alpha_1^{polyA-1}\alpha_2$. None of these variants were found to be significantly associated with the level of HbA_2.

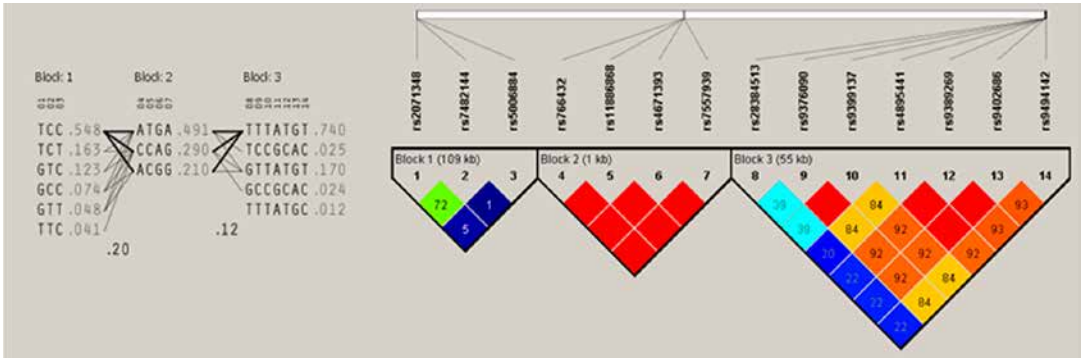
Normal vs. 2.9–3.1



Normal vs. 3.2–3.9



Normal vs. 4.2



2.9–3.1 vs. 3.2–3.9

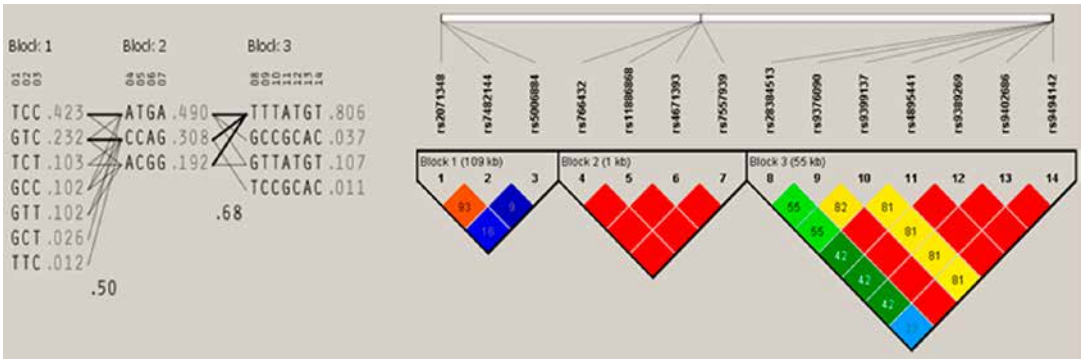


Figure 2. Haploview output of LD across 14 SNPs from the genotyping data in the Saudi population. Linkage Disequilibrium (LD) analysis patterns between 3 HBG2, 4 BCL11A, and 7 HMIP SNPs compared in thalassaemia patients. The pairwise correlation between the SNPs were measured as r^2 and shown (x100) in each diamond. Enhancer haplotypes are in red boxes and diminisher haplotypes are in blue boxes. Coordinates are according to the NCBI build dbSNP 144 *Homo sapiens* annotation release 107 (reference sequence NT_025741.16)

Table I. Allelic association of 14 SNPs related to *BCL11A*, *HBS1L-MYB*, and *HBG2* promoter region in cohorts with normal, borderline and high HbA_2 levels

HbA2 vs. SNP		*HBG2* promoter region			*BCL11A* region				HMIP region						
		rs2071348	**rs7482144**	**rs5006884**	**rs766432**	**rs11886868**	**rs4671393**	**rs7557939**	**rs28384513**	**rs9376090**	**rs9399137**	**rs4895441**	**rs9389269**	**rs9402686**	**rs9494142**
Normal vs. 2.9–3.1	Associated allele	G	T	C	C	T	A	A	T	T	T	A	T	G	T
	χ^2	10.30	5.23	0.78	0.24	0.03	0.39	0.034	0.41	0.62	1.80	0.62	0.98	0.98	0.49
	P-value	**0.001***	**0.02***	0.38	0.63	0.85	0.53	0.854	0.52	0.43	0.18	0.43	0.32	0.32	0.49
Normal vs. 3.2–3.9	Associated allele	G	T	T	A	T	G	G	T	C	C	G	C	A	C
	χ^2	6.57	2.25	3.31	0.07	0.03	0.02	0.006	1.94	0.14	0.73	0.73	0.35	0.35	0.13
	P-value	**0.01***	0.13	0.07	0.79	0.86	0.89	0.939	0.16	0.71	0.39	0.39	0.55	0.55	0.72
Normal vs. 4.2	Associated allele	T	C	T	C	T	A	A	T	T	T	A	T	G	T
	χ^2	3.41	0.81	5.08	0.00	0.50	0.01	0.497	0.00	0.70	0.70	0.70	0.81	0.81	0.92
	P-value	0.06	0.37	**0.02***	1.00	0.48	0.94	0.481	0.97	0.40	0.40	0.40	0.37	0.37	0.34
2.9–3.1 vs. 3.2–3.9	Associated allele	T	C	T	A	–	G	G	T	C	C	G	C	A	C
	χ^2	0.18	0.29	4.32	0.35	0.00	0.35	0.042	0.46	0.88	3.17	1.64	1.64	1.64	0.72
	P-value	0.67	0.59	**0.04***	0.56	1.00	0.56	0.838	0.50	0.35	0.08	0.20	0.20	0.20	0.40
2.9–3.1 vs. 4.2	Associated allele	G	T	C	C	C	A	G	T	C	C	G	C	A	C
	χ^2	8.11	2.80	6.45	0.04	0.34	0.04	0.338	0.07	0.38	0.19	0.38	0.38	0.38	0.58
	P-value	**0.004***	0.09	**0.01***	0.84	0.56	0.84	0.561	0.80	0.54	0.66	0.54	0.54	0.54	0.45
3.2–3.9 vs. 4.2	Associated allele	T	C	T	C	T	A	A	G	T	T	A	T	G	T
	χ^2	6.95	1.99	1.19	0.02	0.33	0.02	0.48	0.46	0.86	1.09	1.09	1.09	1.09	1.09
	P-value	**0.01***	0.16	0.28	0.90	0.56	0.90	0.488	0.50	0.35	0.30	0.30	0.30	0.30	0.30

Table II. Frequency of haplotypes of SNPs in HBG2, BCL11A and HBS1L-MYB compared between normal, borderline and high HbA_2 level cohorts

Block	Haplotype	Normal vs. 2.9–3.1			Normal vs. 3.2–3.9			Normal vs. 4.2			2.9–3.1 vs. 3.2–3.9			2.9–3.1 vs. 4.2			3.2–3.9 vs. 4.2		
		Freq.	χ^2	*p*	Freq.	χ^2	*p*	Freq.	χ^2	*p*	Freq.	χ^2	*p*	Freq.	χ^2	*p*	Freq.	χ^2	*p*
HBG2 promoter region	TCC	0.526	3.02	0.08	0.53	4.54	**0.03**	0.548	0.70	0.40	0.423	0.16	0.69	0.355	0.16	0.69	0.396	0.55	0.46
	GTC	0.156	6.52	**0.01***	0.14	1.27	0.26	0.123	1.46	0.23	0.232	0.72	0.40	0.29	4.85	**0.03***	0.23	3.59	0.06
	TCT	0.134	2.24	0.13	0.145	0.12	0.73	0.163	7.80	**0.01***	0.103	2.80	0.09	0.201	6.15	**0.01***	0.217	2.35	0.13
	GCC	0.09	1.83	0.18	0.073	0.29	0.59	0.074	0.84	0.36	0.102	0.53	0.47	0.116	1.55	0.21	0.024	0.30	0.59
	GTT	0.057	0.49	0.48	0.065	3.67	0.06	0.048	0.52	0.47	0.102	0.43	0.51				0.033	0.40	0.53
	TTC	0.038	0.26	0.61	0.034	1.53	0.22	0.041	0.42	0.52	0.012	0.80	0.37	0.036	1.01	0.31			
	GCT				0.014	1.95	0.16				0.026	1.11	0.29				0.08	1.04	0.31
	TTT																0.02	3.80	0.05
BCL11A region	ATGA	0.489	0.03	0.85	0.485	0.01	0.94	0.491	0.50	0.48	0.49	0.04	0.84	0.516	0.34	0.56	0.5	0.48	0.49
	CCAG	0.299	0.39	0.53	0.288	0.02	0.89	0.29	0.01	0.94	0.308	0.35	0.56	0.328	0.04	0.84	0.283	0.02	0.90
	ACGG	0.205	0.62	0.43	0.216	0.01	0.94	0.21	0.76	0.38	0.192	0.48	0.49	0.156	0.28	0.59	0.2	0.75	0.39
	ATGG																0.017	0.20	0.65
HMIP region	TTTATGT	0.747	0.75	0.39	0.752	1.15	0.28	0.74	0.19	0.66	0.806	0.08	0.78	0.797	0.00	0.98	0.815	0.02	0.90
	GTTATGT	0.167	0.12	0.73	0.149	3.15	0.08	0.17	0.07	0.80	0.107	2.03	0.15	0.156	0.17	0.68	0.085	2.02	0.15
	TTTATGC	0.015	0.06	0.81				0.012	0.12	0.73				0.016	0.19	0.66			
	GCCGCAC	0.024	0.20	0.66	0.031	0.87	0.35	0.024	0.26	0.61	0.037	1.16	0.28	0.016	0.19	0.66	0.048	0.61	0.44
	TCCGCAC	0.021	1.06	0.30	0.026	0.02	0.90	0.025	0.27	0.61	0.011	1.09	0.30				0.019	0.23	0.63
	TTTGCAC				0.011	0.41	0.52										0.017	0.20	0.65
	TTCATGT																0.017	0.20	0.65
	TCTGCAC													0.016	0.19	0.66			

Structure modelling and RMSD prediction

Of the three SNPs (rs2071348, rs7482144, and rs5006884) that were significantly associated with borderline and higher HbA_2 levels, only rs5006884 was an exonic polymorphism that made an amino acid change. The three-dimensional (3D) structural modelling for rs5006884 native using a Protein Data Bank template and mutant models using SWISS MODEL and PROCHECK [16, 23–25] were studied to elucidate the effects of the amino acid changes, and were also compared for the energy minimisation and its RMSD using the GROMACS 5.1.1 program [18]. Binding energy minimisation varied significantly, and the total energy values for rs5006884 (7529.6 kJ/mol) variegated proteins were deviated from the wild (7227.9 kJ/mol) protein. In addition, the RMSD values, which are directly proportional to structural deviations with RMSD –0.01 Å, play a significant role during disease initiation. The superimposed position of wild OR51B6 (red colour) and mutant L172F (green colour) model was generated by SWISS PDB Viewer, as shown in Figure 3. This observation suggests that the variation rs5006884 at the *OR51B6* gene affects the protein-protein interaction.

Protein-protein interaction

OR51B6 wild and mutant models were checked for possible interaction variations towards the globin proteins (HBB, HBD, HBE, HBA1, HBA2, and HBG) using PRISM server. Based on the template and target protein interface, model complex was generated by prism, with multiple prediction of interaction of molecules bound with free binding energies. Only the lowest free binding energy complexes were selected. The influence of the mutated *OR51B6* gene in the binding energy of the OR51B6-target protein complex varied significantly

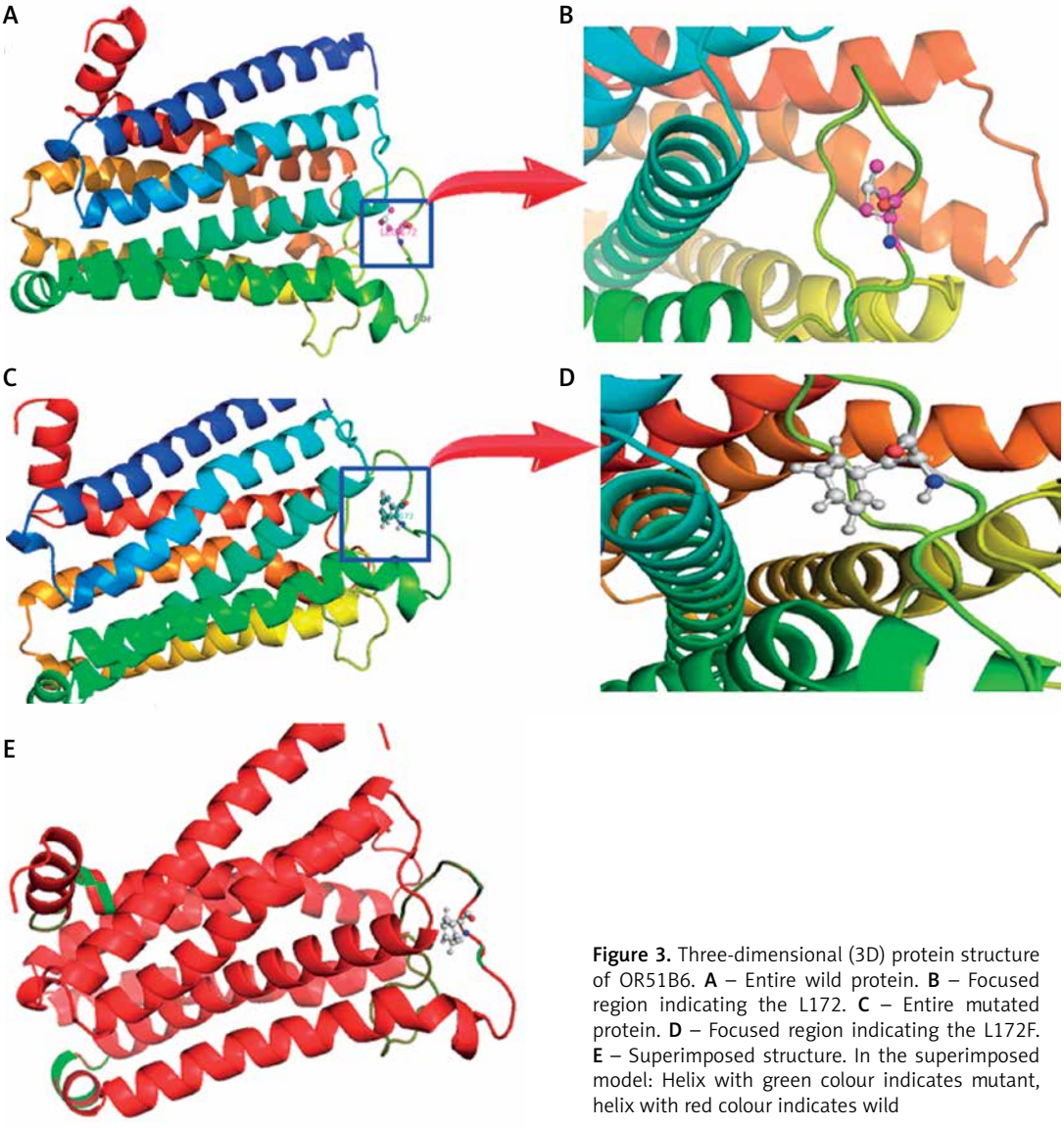


Figure 3. Three-dimensional (3D) protein structure of OR51B6. **A** – Entire wild protein. **B** – Focused region indicating the L172. **C** – Entire mutated protein. **D** – Focused region indicating the L172F. **E** – Superimposed structure. In the superimposed model: Helix with green colour indicates mutant, helix with red colour indicates wild

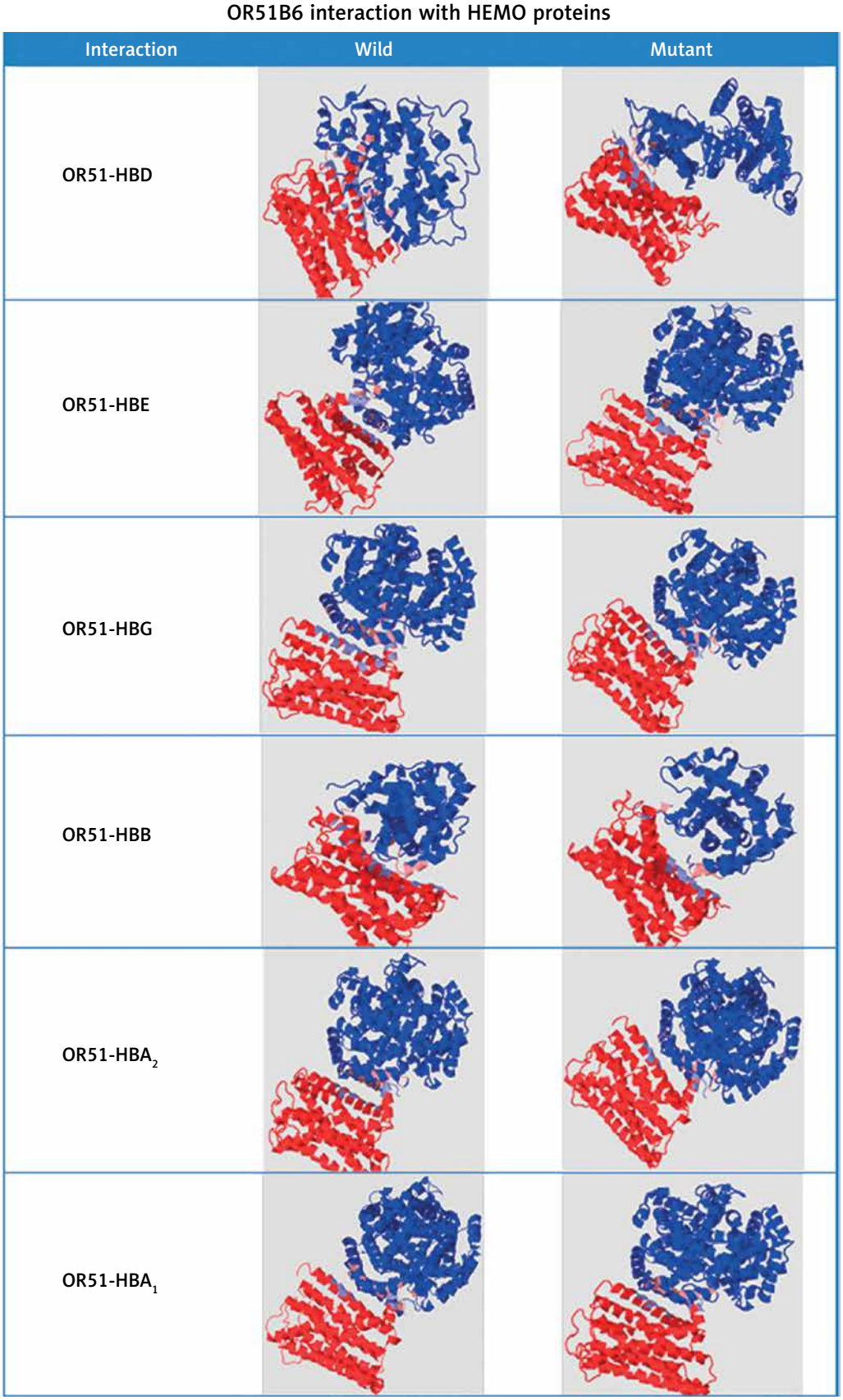


Figure 4. OR51B6 protein and globin protein interaction. Column 1 indicates the template OR51B6 protein and interacting globin protein; column 2 denotes the target globin protein and wild OR51B6 protein. Column 3 denotes the target globin protein and mutant OR51B6 protein

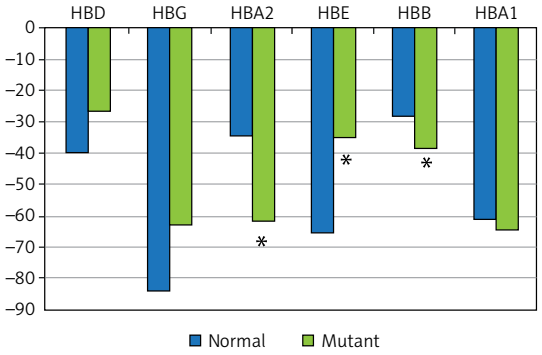


Figure 5. Influence of the mutated OR51B6 gene in the binding energy of the OR51-target protein complex

**Significantly varied binding energy at 25% compared with the wild type.*

with other haemoglobins (Figure 4). This observation suggests that the variation rs5006884 at the *OR51B6* gene affects the protein-protein interaction. Binding energies varied significantly between native and mutant OR51B6 with respect to globin protein, with the greatest shift between OR51B6/HBA2 and the least between OR51B6/HBA1. The influence and sensitivity of the binding energy upon amino acid substitution is significant at a threshold of 10% variation in the binding energy [26]. In the present study, the threshold to predict the most influential effect of the amino acid substitution on the binding energy is 25% variation, which indicated the most deleterious effect of the amino acid substitution among the tested combinations (Figure 5).

Discussion

Screening for β-thalassaemia trait is important in genetic and pre-marital counselling. This screening is based on the determination of the level of HbA_2, which is usually elevated in subjects heterozygous for the β-thalassaemia mutation. However, there are several factors that may influence this level, such as coinheritance of α-thalassaemia or δ-thalassaemia, where the HbA_2 concentration is reduced to normal and remains < 3.5%. Also, the presence of silent HBB mutations projects normal RBC indices and normal or borderline HbA_2 levels [27]. HbA_2 levels are lower in β+-thalassaemia mutation carriers than in β^0-thalassaemia mutation carriers, and normal HbA_2 levels are reported in a small subset of traits with non-mutant δ-globin gene [28, 29]. These studies with divergent results entrench the confusion about the reliability of the HbA_2 level [30].

A genome-wide association study identified rs5006884 SNP in the olfactory genes *OR51B5/OR51B6* locus on chromosome 11 to exceed the stringent genome-wide significance threshold of 10^{-8}, and is considered to be the most significant SNP explaining 5.6% of the variability in HbF level [31]. The effect of rs5006884 on HbF levels was independent of the Xmn1 site, sickle cell haplotype, β-globin gene-like complex, and LCR region. A recent study also showed a lack of association between the rs5006884 SNP with HbF in β-thalassaemia among the Saudi population [32].

An additional regulatory region modulating HbF expression was reported to be centromeric to the β-globin gene cluster, with genetic modifiers rs2071348 and rs7482144 influencing disease severity [33]. The rs2071348 polymorphism on the HBB pseudogene (*HBBP1*) was previously significantly associated with a milder disease phenotype in Asian β(0)-thalassaemia/haemoglobin E patients [34], but causality due to the rs7482144 marker represented by the Xmn1-HBG2 site has not been fully demonstrated [35]. Detailed analysis of the association of HbF QTL – HBG2 promoter region, BCL11A region, and HMIP region genes on Saudis with various levels of HbA2 revealed that these three SNPs, rs2071348, rs7482144, and rs5006884, were significantly associated with borderline and higher HbA_2 levels in the present study, which are reported as insignificant SNPs for HbF level in Saudis [32] confirming the reciprocal relationship among the HbA_2 and HbF levels [10]. The predominant haplotypes in the borderline and higher HbA_2 cohort consisted of GTC, TCC, and TCT in the HBG2 promoter region comprising the same SNPs: rs2071348, rs7482144, and rs5006884, respectively.

The bioinformatics tool-based, systematic, in silico approach to ascertain the impact of genetic variation rs5006884C>T on the structure of OR51B6 protein and its significant impact on the interaction with globin proteins were predicted. The single amino acid change, rs5006884 in the OR51B6, resulted in structural modification with RMSD –0.01 Å that suggests it plays a significant role during disease initiation. The binding energy minimisation also varied significantly, indicating that the variation rs5006884 affects the protein-protein interaction.

Molchanova *et al.* [36] reported the α-2 expression level to be twice the level seen in α-1, with less efficient translation of the α-2-mRNA, to maintain a relative level of α-2 protein. The protein interface energy between HBB/HBA2 and HBD/HBA2 is almost the same, but among haemoglobin proteins with variants of OR51B6, the greatest shift is between OR51B6/HBA2 (78.7%) and the least between OR51B6/HBA1 (5.07%) (Figure 5). It was evident from the interaction of OR51B6 and HBA_2 that a functional modification due to the presence of rs5006884C>T on the structure of OR51B6 that has doubled the binding

energy might be the reason for the slight increase in the level of HBA_2. Previous studies on the HBA_2 protein interaction reported a significant variation in the interface energy between mutant-template and wild-template protein complex [37]. The OR51B6 can be considered as a HbA_2 modulator for large-scale studies.

In conclusion, the haemoglobin switching modulator SNPs in the *HBG2* gene region in general and *OR51B6* (rs5006884) variation in particular might intervene in the tetramer formation of α2 with β and δ globin, thereby playing a significant role in the regulation of HbA_2 level in β-thalassaemia carriers. Therefore, these variations need to be considered when screening haemoglobinopathies. Furthermore, the *OR51B6* can be studied in detail for the HbA_2 modulator to ameliorate disease severity.

Acknowledgments

The authors extend their gratitude to King Abdulaziz City for Science and Technology, Saudi Arabia (To Dr. Cyril Cyrus, Grant No: LGP 31-34) for the funding of this research. We also thank Mr. Geoffrey James Tam Moro and Mr. Mohammed Al Shamlan for their technical and administrative support.

Conflict of interest

The authors declare no conflict of interest.

References

1. Palstra RJ, de Laat W, Grosveld F. Beta-globin regulation and long-range interactions. Adv Genet 2008; 61: 107-42.
2. Menzel S, Garner C, Rooks H, Spector TD, Thein SL. HbA2 levels in normal adults are influenced by two distinct genetic mechanisms. Br J Haematol 2013; 160: 101-5.
3. Steinberg MH, Rodgers GP. HbA2: biology, clinical relevance and a possible target for ameliorating sickle cell disease. Br J Hematol 2015; 170: 781-7.
4. Siatecka M, Bieker JJ. The multifunctional role of EKLF/KLF1 during erythropoiesis. Blood 2011; 118: 2044-54.
5. Perseu L, Satta S, Moi P, et al. KLF1 gene mutations cause borderline HbA2. Blood 2011; 118: 4454-8.
6. Gallienne AE, Dreau HM, Schuh A, Old JM, Henderson S. Ten novel mutations in the erythroid transcription factor KLF1 gene associated with increased fetal hemoglobin levels in adults. Haematologica 2012; 97: 340-3.
7. Tepakhan W, Yamsri S, Fucharoen G, Sanchaisuriya K, Fucharoen S. Kruppel like factor 1 mutations and expression of hemoglobins F and A in homozygous hemoglobin E syndrome. Ann Hematol 2015; 94: 1093-8.
8. Yu LH, Liu D, Cai R, et al. Changes in hematological parameters in alphathalassemia individuals co-inherited with erythroid Kruppel-like factor mutations. Clin Genet 2015; 88: 56-61.
9. Borgio JF, AbdulAzeez S, Al-Muslami AM, et al. KLF1 gene and borderline hemoglobin A2 in Saudi population. Arch Med Sci 2018; 14: 230-6.
10. Griffin PJ, Sebastiani P, Edward H, et al. The genetics of hemoglobin A2 regulation in sickle cell anemia. Am J Hematol 2014; 89: 1019-23.
11. Zhu J, Chin K, Aerbajinai W, Trainor C, Gao P, Rodgers GP. Recombinant erythroid Kruppel-like factor fused to GATA1 up-regulates delta- and gamma-globin expression in erythroid cells. Blood 2011; 117: 3045-52.
12. Farrell JJ, Sherva RM, Chen ZY, et al. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood 2011; 117: 4935-45.
13. Akinsheye I, Alsultan A, Solovieff N, et al. Fetal hemoglobin in sickle cell anemia. Blood 2011; 118: 19-27.
14. Al-Nafie AN, Borgio JF, AbdulAzeez S, et al. Co-inheritance of novel ATRX gene mutation and globin (alpha and beta) gene mutations in transfusion dependent beta-thalassemia patients. Blood Cells Mol Dis 2015; 55: 27-9.
15. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006; 22: 195-201.
16. Laskowski RA, Macarthur MW, Moss D, Thornton JM. PROCHECK – a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993; 26: 283-91.
17. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Electrophoresis 1997; 18: 2714-23.
18. Lindahl E, Azuara C, Koehl P, Delarue M. NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. Nucleic Acids Res 2006; 34: 52-56.
19. Ogmen U, Keskin O, Aytuna AS, Nussinov R, Gursoy A. PRISM: protein interactions by structural matching. Nucleic Acids Res 2005; 33: W331-6.
20. Keskin O, Nussinov R, Gursoy A. PRISM: protein-protein interaction prediction by structural matching. Methods Mol Biol 2008; 484: 505-21.
21. Tuncbag N, Gursoy A, Nussinov R, Keskin O. Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM. Nat Protoc 2011; 6: 1341-54.
22. Barrett JC, Fry B, Maller J, Daly MJ. Haploview analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21: 263-5.
23. Kiefer F, Arnold K, Künzli M, et al. The SWISS-MODEL repository and associated resources. Nucl Acids Res 2009; 37: D387-92.
24. Berman HM, Westbrook J, Feng Z. The protein data bank. Nucleic Acids Res 2000; 28: 235-42.
25. Feng L, Gell DA, Zhou S, et al. Molecular mechanism of AHSP-mediated stabilization of alpha-hemoglobin. Cell 2004; 119: 629-40.
26. Chen DE, Willick DL, Ruckel JB, Floriano WB. Principal component analysis of binding energies for single-point mutants of hT2R16 bound to an agonist correlate with experimental mutant cell response. J Comp Biol 2015; 22: 37-53.
27. Origa R. Editors In: M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, H.C. Mefford, K. Stephens, A. Amemiya, N. Ledbetter, editors. SourceGeneReviews®. Seattle (WA): University of Washington, Seattle; 1993-2017.2000 Sep 28 [updated 2015 May 14].
28. Galanello R, Barella S, Ideo A, et al. Genotype of subjects with borderline hemoglobin A2 levels: implication for beta-thalassemia carrier screening. Am J Hematol 1994; 46: 79-81.

29. Fucharoen S, Fucharoen G, Sanchaisuriya K, Pengjam Y. Molecular analysis of a thai beta-thalassaemia heterozygote with normal haemoglobin A2 level: implication for population screening. Ann Clin Biochem 2002; 39: 44-9.
30. Verhovsek M, So CC, O'Shea T, et al. Is HbA2 level a reliable diagnostic measurement for beta-thalassemia trait in people with iron deficiency? Am J Hematol 2012; 87: 114-6.
31. Solovieff N, Milton JN, Hartley SW, et al. Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5 olfactory receptor gene cluster. Blood 2010; 115: 1815-22.
32. Cyrus C, Vatte C, Borgio JF, et al. Existence of HbF enhancer haplotypes at HBS1L-MYB intergenic region in transfusion-dependent Saudi beta-thalassemia patients. Biomed Res Int 2017; 2017: 1972429.
33. Sherva R, Sripichai O, Abel K, et al. Genetic modifiers of Hb E/beta0 thalassemia identified by a two-stage genome-wide association study. BMC Med Genet 2010; 11: 51.
34. Giannopoulou E, Bartsakoulia M, Tafrali C, et al. A single nucleotide polymorphism in the HBBP1 gene in the human beta-globin locus is associated with a mild beta-thalassemia disease phenotype. Hemoglobin 2012; 36: 433-45.
35. Danjou F, Francavilla M, Anni F, et al. A genetic score for the prediction of beta-thalassemia severity. Haematologica 2015; 100: 452-7.
36. Molchanova TP, Pobedimskaya DD, Huisman THJ. The differences in quantities of alpha2-and alpha1-globin gene variants in heterozygotes. Br J Haematol 1994; 88: 300-6.
37. Borgio JF, Al-Madan MS, AbdulAzeez S, et al. Mutation near the binding interfaces at alpha-hemoglobin stabilizing protein is highly pathogenic. Am J Transl Res 2016; 8: 4224-32.