Dissecting the role of genetic modifiers of hemoglobinopathies: A futuristic approach towards precision medicine

Priya Hariharan
National Institute of Immunohaematology

Manju Gorivale
National Institute of Immunohaematology

Pratibha Sawant
National Institute of Immunohaematology

Pallavi Mehta
National Institute of Immunohaematology

Anita Nadkarni (✉ anitahnadkarni@yahoo.com)
National Institute of Immunohaematology

Research Article

Abstract

Introduction:

Hemoglobinopathies though a monogenic disorder, show phenotypic variability. Hence, understanding the genetics underlying the heritable sub-phenotypes of hemoglobinopathies, specific to each population, would be prognostically useful and could inform personalized therapeutics. This study aimed to evaluate the role of genetic modifiers leading to higher HbF production with cumulative impact of the modifiers on disease severity.

Materials and methods:

200 patients [100 β-thalassemia homozygotes,100 Sickle Cell Anemia], and 50 healthy controls were recruited. Primary screening followed with molecular analysis for confirming the β-hemoglobinopathy was performed. Co-existing α-thalassemia and the polymorphisms located in 3 genetic loci linked to HbF regulation were screened.

Results:

The most remarkable result was the association of SNPs with clinically relevant phenotypic groups. The γ-globin gene promoter polymorphisms [-158 C→T,+25 G→A],BCL11A rs1427407 G→T,-3 bp HBS1L-MYB rs66650371 and rs9399137 T→C polymorphisms were correlated with higher HbF, in group that has lower disease severity score (P<0.00001), milder clinical presentation, and a significant delay in the age of the first transfusion.

Conclusion:

Our study emphasizes the complex genetic interactions underlying the disease phenotype that may be a prognostic marker for predicting the clinical severity and assist in disease management.

Key Points:

- The present study expands the knowledge of the frequency of the genetic modifiers in the Indian hemoglobinopathy patients.
- A strong independent association of the SNPs located in the γ-globin promoter region: BCL11A, HBS1L-MYB intergenic region *with the HbF levels* were observed among the milder patient groups(P<0.05).
- The sickle cell anemia group inheriting a greater number of HbF modulating alleles showed a pronounced negative correlation with the disease severity score.
- The predictions based on genetic modifiers thus can foresee the severity of β-thalassemia and SCA, thus may help in the efficient management of the disease.

1. Introduction:

β-thalassemia and sickle cell anemia (SCA) form a major health burden in India as they cause a high degree of morbidity, moderate to severe hemolytic anemia with the carrier frequency varying from 3 to 17% among different population groups of India [1]. Both the diseases are caused by mutations in the β-globin gene and are inherited as an autosomal recessive single gene disorder. However, despite this apparent genetic simplicity, both the disorders display a remarkable spectrum of phenotypic severity. In β-thalassemics, the primary determinant of disease severity is the type of β-globin gene mutation that the patient inherits. Severe β-thalassemia mutation (β^0/β^+) may completely down-regulate β-globin chain synthesis. However, milder β^{++} mutations, present in the conserved 5' β-globin promoter region and 3' untranslated region, may alter the mRNA expression, leading to the moderate synthesis of the β-globin chain [2]. It has been observed that patients inheriting the same β-globin gene mutations, display phenotypic heterogeneity. Thus, the clinical variability observed among the patients prompted the search for the additional genetic modulators of these diseases.

Several modifier genes have been identified which influence the severity of hemoglobinopathies. The most common secondary modifier is the co-inheritance of α-thalassemia and the elevated levels of fetal hemoglobin (HbF), both of which directly reduce the globin chain imbalance. An interacting combination of α-thalassemia with hemoglobinopathy has been shown to alleviate the severity by reducing the intracellular precipitation of free alpha-globin chains in β-thalassemia and by reducing HbS concentration in SCA patients [3,4].

The possibility that the severity of hemoglobinopathies could be reduced by induction of HbF was realized as the symptoms of hemoglobinopathies are observed after 6 months of age after the birth when there is a gradual reduction in the HbF levels. Thus, the study of other secondary mechanisms related to the sustained production of HbF in adult life may be important. With the advent of genome-wide association studies, recently several unlinked genetic factors have been identified which elevate the HbF levels [5]. Four major quantitative trait loci (QTLs): the HBB locus, the BCL11A gene, the HBSB1L-cMYB inter-region along with KLF1 gene variations bring about 20−50 % variation in HbF levels in hemoglobinopathy patients [5]. Thus, in this study, we have screened for the presence of primary and secondary modifiers in hemoglobinopathy patients and have analysed the cumulative effect of these modifiers on the phenotypic variability in the patient group.

2. Materials And Methods::

2.1 Selection of patients and healthy controls:

The study was approved by the National Institute of Immunohaematology-Institutional Ethics Committee **and all methods were performed in accordance with the relevant guidelines and regulations.** Unrelated 100 β-thalassemia homozygous and 100 SCA patients were selected for this study. 50 unrelated healthy adult blood donors with normal hematological indices, with no transfusion history and medication, were randomly selected. The blood samples were collected in K2 EDTA vacutainers after informed consent.

2.2 Primary screening and molecular analysis:

Primary screening involved complete blood count analysis and hemoglobin electrophoresis on BioRad Variant II high-performance liquid chromatography to evaluate the concentration of different hemoglobin fractions.

Molecular analysis of the β-globin gene was first carried out to confirm the hemoglobinopathy status in the patient samples by CRDB, AMRS PCR, or by direct DNA sequencing. α-globin gene deletions were detected by multiplex PCR. The γ-globin promoter region was screened by direct DNA sequencing. Five BCL11A polymorphisms which showed the highest correlation with HbF levels [rs11886868 (C→T), rs7557939 (A→G), rs4671393 (A→G), rs1427407 (G→T) and rs7606173 (G→C)] were analyzed by real-time SNP genotyping and by ARMS PCR. 2 polymorphisms in HBS1L-MYB intergenic region [rs66650371 (Intact 'TAC'→ Deletion 'TAC'), rs9399137 (T→C)] were screened by ARMS PCR and by DNA sequencing respectively.

The patients were clinically evaluated and the disease severity score was calculated based on the detailed clinical history of the patient [6,7].

2.3 Statistical analysis:

Statistical analysis of the data was performed using GraphPad version 6.01 software (Graph Pad Prism Inc, California, U.S.A). The hematological indices among different patient groups and normal controls are represented as Mean ± Standard Deviation (SD). Fischer extract test was used to compare the polymorphism distribution among the patients and the control groups. The comparison of the quantitative variables among the groups and between differing genotypes was carried out by unpaired non-parametric Mann-Whitney U test. The P-value ≤ 0.05 was considered to be statistically significant. Generalized Multifactor Dimensionality Reduction (GMDR) software version beta 0.9 was used to analyse the interaction among the SNPs in different patient groups. The Kaplan Meier survival curve analysis was performed to determine the age of presentation by considering the transfusion free survival among the patient groups.

3. Results:

On the basis of clinical history, the β-thalassemia patients were classified into a severe group (50 Thalassemia major: TM) and milder group (50 Thalassemia Intermedia: TI).

3 parameters were considered for clinical analysis in both the hemoglobinopathy patient group. These included the age of presentation, frequency of blood transfusion, and organomegaly. As expected, the patients in the β-thalassemia major group had an early age of presentation (9.2 ± 2.7 months) and recurrent blood transfusion requirement (14.8 ± 4.0 times/ year) as compared to the milder β-thalassemia intermedia group (mean age: 4.5 ± 3.3 years and transfusion frequency: 2.5 ± 3.3/ year). In the SCA, the mean age of presentation was found to be 6.3 ± 5.2 years. Further to know the contribution of HbF levels in the clinical presentation of the SCA patients, they were divided into two groups considering the median HbF level of 17.4 %. It was observed that SCA patients with HbF levels ≤ 17.4 % showed higher mean transfusion frequency (3.3 ± 5.5 times) as compared to patients with HbF > 17.4 % (1.7 ± 5.4 times).

Hepatosplenomegaly was pronounced in β-thalassemia intermedia group as against thalassemia major (**P: 0.001**). Also, in SCA patients, hepatosplenomegaly was observed. The detailed clinical analysis among the patient groups is shown in Table 1. The hematological parameters of the patient groups are shown in Table 2.

Table 1
Clinical characteristics among the patient groups

Parameters	β-thalassemia Homozygous		Sickle cell anemia patients	
	β-thalassemia Major n = 50	β -thalassemia intermedia n = 50	HbF level : ≤ 17.4 % n = 50	HbF level : >17.4 % n = 50
Mean ± SD				
Age of presentation	9.2 ± 2.7 months	4.5 ± 3.3 years	5.2 ± 4.9	7.1 ± 5.5
Blood transfusion per year (Times)	14.8 ± 4.0	2.5 ± 3.3	3.3 ± 5.5	1.7 ± 5.4
Hepatomegaly (cms)*	2.7 ± 2.9	4.6 ± 3.5	5.3 ± 3.1	
Splenomegaly (cms)#	4.7 ± 1.7	6.5 ± 3.6	6.8 ± 5.1	
n: Number of patients				
*: Hepatomegaly pronounced in thalassemia intermedia and Sickle cell anemia patients				
#: Splenomegaly in thalassemia intermedia and Sickle cell anemia patients				

Table 2
Hematological analysis among the patient and the control groups

Parameters	Controls	β-thalassemia Homozygous			Sickle cell anemia patients		
	Mean ± SD	TM, n: 27 Mean ± SD #	TI, n: 50 Mean ± SD	*P value 95 % CI	HbF ≤ 17.4% n: 50 Mean ± SD	HbF > 17.4% n: 50 Mean ± SD	**P value 95 % CI
RBCs(10^6/µl)	4.6 ± 0.6	2.8 ± 0.70	3.5 ± 0.69	**< 0.0001**	3.1 ± 0.9	3.1 ± 0.9	0.5
Hb (g/dL)	13.5 ± 2.04	5.4 ± 1.17	7.8 ± 1.4	**< 0.0001**	8.1 ± 2.1	9.1 ± 2.5	**0.03**
MCV (fL)	83.9 ± 4.2	75.1 ± 17.3	70.7 ± 7.8	0.92	79.4 ± 12.0	84.0 ± 10.6	**0.04**
MCH (pg)	28.9 ± 4.2	24.4 ± 5.2	22.8 ± 3.4	0.28	26.4 ± 4.4	29.2 ± 3.8	**0.008**
MCHC (g/dL)	34.5 ± 1.5	31.6 ± 3.5	32.2 ± 1.9	0.18	33.2 ± 2.1	34.8 ± 2.5	**0.0003**
RDW (%)	13.85 ± 1.5	32.8 ± 9.8	30.8 ± 5.9	0.36	22.8 ± 4.6	20.8 ± 6.3	**0.0016**
Platelets (10^3/µl)	291.7 ± 38.0	333 ± 246.8	370.8 ± 269	0.23	349.7 ± 211.1	221 ± 115.6	**0.0002**
HbA$_2$ (%)	2.6 ± 0.2	3.6 ± 1.5	3.4 ± 1.9	0.08	3.0 ± 0.5	2.9 ± 1.1	**0.007**
HbF (%)	0.31 ± 0.2	65.6 ± 36.3	66.3 ± 34.2	0.46	12.7 ± 3.7	26.4 ± 7.1	**< 0.0001**
F-cells (%)	2.4 ± 0.5	67.1 ± 26.01	68.0 ± 24.1	0.81	31.8 ± 18.6	43.2 ± 18.2	0.06
HbS (%)	-	-	-	-	74.1 ± 11.7	67.5 ± 9.4	**< 0.0001**

: At the time of testing 23 TM patients were on regular transfusions, hence these patients were excluded from hematological analysis.

* : P value calculated among the β-thalassemia homozygous groups, ** : P value calculated among the SCA patient groups. The P value was calculated using unpaired non parametric Man-Whitney U test.

Molecular analysis of the β-globin gene showed that IVS 1−5 (G→C) was the commonest mutation encountered in both the β-thalassemia homozygous group (TM: 32 %, TI: 28 %). The six common Indian mutations [IVS 1−5 (G→C), 619 bp deletion, IVS 1−1 (G→T), codons 8/9 (+ G), codon 15 (G→A), codons 41/42 (-CTTT)] accounted for 82.5% of molecular lesions (TM: 44.5 %, TI: 38 %). It was observed that the

overall prevalence of milder mutations was higher in TI patients (8.0 %) as compared to TM patients (1.0 %, P: 0.07). (**Suppl. Table 1**)

As co-inheritance of α-thalassemia is a well-known disease modifier of β-thalassemia and SCA, the presence of α-globin-gene deletions was screened in the patient groups. A much higher prevalence of single alpha globin gene deletions was observed in SCA patients (51%). Among the β-thalassemia homozygotes, the β-thalassemia intermedia showed a higher prevalence (26%) of α-globin gene deletions as compared to β-thalassemia major (20%). (P:0.47) (**Suppl. Table 2**)

The second powerful modifier of disease severity in hemoglobinopathy patients is elevated HbF levels. Hence the polymorphisms located in the three loci linked to raised HbF levels: γ-globin promoter region, BCL11A and HBSL1-MYB intergenic region were analysed in this study.

In the ^{G}γ globin promoter region, the XmnI polymorphism residing in the − 158 (C→T) position was only detected. In the β-thalassemia homozygous group, the homozygosity for the mutant T allele [T/T, Xmn I +/+] was significantly higher in TI (44 %) as compared to TM (28%), (**P: 0.01**).Similarly, in SCA patients, 94% of the patients were homozygous for the T allele. As thalassemia major patients were on recurrent transfusion, the genotypes could not be compared with the HbF levels, however, in thalassemia intermediates, the TT [XmnI: +/+] genotype was found to be significantly associated with raised HbF levels (79.1% ± 29.0, p: 0.04) as compared to the CC [XmnI: -/-] genotype (53.08 % ± 35.9). [Figure 1A]

In the ^{A}γ-globin promoter region, + 25 (G→A) variation was detected and A allele was found to be the variant allele. The A allele was found to be significantly higher in the TI group as compared to TM (P: 0.005). Also, the A allele in β-thalassemia intermediates, was significantly associated with increased HbF levels (79.9 % ±28.6, P: 0.03). In SCA patients, 94% of the patients were homozygous for the A allele, an observation similar to the XmnI polymorphism. [Figure 1B].

Among the 5 intronic polymorphisms in the BCL11A gene screened, the mutant T allele of rs1427407 (G→T) polymorphism, was significantly higher in the β-thalassemia intermedia group as compared to the thalassemia major group (P: 0.002, OR: 5.6, 1.84−17.22). In SCA patients the T allele was found to be significantly associated with raised HbF levels (HBF > 17.4 %, (P: 0.003, OR: 3.14, 1.46−6.75) as compared to the other group. The T allele was also found to be significantly associated with HbF levels in both the patient groups. (P < 0.05). [Figure 1C]

In the sickle cell anemia patients, the C allele of rs11886868 C→ T polymorphism was found to be significantly associated with increased HbF levels (P: 0.02, HbF: 20.9 % ± 8.8). [Figure 1D] Among the HBS1L and MYB gene polymorphisms, the deletional allele of rs66650371 (Intact TAC → Deletion –- 'TAC') polymorphism and the C allele of rs9399137 (T→C) were found to be significantly present in the milder β-hemoglobinopathy patients. As reported in earlier studies these 2 polymorphisms were found to be in complete linkage disequilibrium. The minor alleles of these polymorphisms were found to be significantly associated with the HbF levels. [Figure 1E **and F**] Table 3 gives a detailed analysis of the allelic frequency of these polymorphisms determined among the patient and the control group.

Table 3
Genotypic and allelic frequency of the polymorphisms in patient and control group

Allelic frequencies (N: 100)		Controls n:50	β-thalassemia Homozygous		Sickle cell anemia patients		*P value/ Odds Ratio 95 % CI	**P value/ Odds Ratio 95 % CI
			TM n:50	TI n:50	HbF ≤ 17.4% n:50	HbF > 17.4% n:50		
-158 (C→T)	C	81	59	41	5	2	0.01/	0.29/
	T	19	41	59	95	98	2.07 (1.1–3.6)	2.4 (0.4–12.7)
+25 (G→A)	G	81	60	40	5	2	0.005/	0.29/
	A	19	40	60	95	98	2.2 (1.27–3.96)	2.4 (0.4–12.7)
rs11886868(C→T)	C	60	55	59	47	57	0.56/	0.11 /
	T	40	45	41	53	43	1.17(0.67–2.06)	1.56(0.89–2.7)
rs7557939 (A→G)	A	38	42	43	47	36	0.88/	0.11/
	G	62	58	57	53	64	1.04 (0.5–1.8)	1.57 (0.9–2.7)
rs4671393 (A→G)	A	26	12	18	30	31	0.2/	0.95 /
	G	74	88	82	70	69	1.6 (0.73–3.5)	1.01 (0.55–1.85)
rs1427407(G→T)	G	95	96	81	89	72	0.002/	0.003/
	T	5	4	19	11	28	5.6 (1.84–17.2)	3.14 (1.46–6.75)
rs7606173(G→C)	G	93	93	95	80	78	0.55 /	0.72 /
	C	07	7	5	20	22	0.69 (0.21–2.28)	1.1 (0.57–2.22)
rs66650371	TAC	85	86	75	94	79	0.06/	0.007 /

n: Number of patients in each group, N: Total number of alleles in each group

*: Thalassemia patients

**: SCA patients

		Controls n:50	β-thalassemia Homozygous		Sickle cell anemia patients			
TAC→ 3 base Deletion	—	15	14	25	6	21	2.0 (0.96–4.12)	3.76 (1.4–9.8)
rs9399137(T→C)	T	85	86	75	94	79	0.06/	0.007 /
	C	15	14	25	6	21	2.0 (0.96–4.12)	3.76 (1.4–9.8)
n: Number of patients in each group, N: Total number of alleles in each group								
*: Thalassemia patients								
**: SCA patients								

The linkage disequilibrium plot showed that the + 25 (G→A) polymorphism in ^{A}γ-globin gene and the XmnI polymorphism in the ^{G}γ-globin are highly linked (Linkage Disequilibrium coefficient D': 93). Also, the SNPs rs11886868 (C→T) and rs7557939 (A→G) in the BCL11A intronic region, are strongly linked (D': 88) with each other. [Figure 2]

Further, the best SNP models accompanied by the lowest prediction error (Testing balance accuracy), the highest CVC, and the P-value of significant level were calculated. The results revealed a cumulative effect of the mutant alleles of the 3 SNPs – 158(C→T), rs11886868 (C→T), and rs1427407 (G→T) significantly higher in β-thalassemia intermedia patients, and as the best SNP model with testing balance accuracy of 74.9% and cross-validation consistency of 9/10. Further gene-gene interaction studies showed a synergistic effect may coexist among these SNPs in elevating the HbF level [Figure 3A] Similarly among the SCA patients, gene-gene interaction between the mutant alleles of rs66650371 and rs1427407 were found to be significantly higher in the sickle cell anemia patients with HbF levels > 17.4 % with a testing balance accuracy 66.0 % and cross-validation consistency 10/10. [Figure 3B] The generation of GMDR models for determining the most influential SNPs among the 9 SNPs studied in the patient groups is shown in **supplement** Table 3.

The presence of ameliorating alleles may significantly delay the age of presentation and transfusion requirement in β-hemoglobinopathy patients. Hence, for the analysis we included both primary modifiers [the type of β-globin gene mutation in β-thalassemia patients] and secondary modifiers: α-globin genotype and the HbF modulators [γ-globin promoter variations, BCL11A, MYB and KLF1 variations (the KLF1 data from our previous published paper) [8]. Among the SCA patients, a strong negative correlation was observed between the HbF levels and the disease severity score. (Pearson correlation coefficient r: -0.7, P < 0.00001) The patients inheriting the higher numbers of modulating allele showed significantly elevated HbF levels (mean HbF: 21.9% ±9.8) as compared to patients with a less number of disease severity modulating alleles (mean HbF:16.5 %±7.1). Also showed a significant delay in the age of first transfusion as compared to the other group. [Figure 4A and 4B] The β-thalassemia intermedia patients inheriting more number of the disease ameliorating alleles showed elevated HbF levels (mean HbF: 75.1 % ± 29.9), with reduced disease severity score (mean DSS :5.6) as compared to patients with lower numbers of disease

severity modulating alleles, who had lower HbF levels (mean HbF: 54.1% ± 36.9). Further, when compared to the age of first transfusion, however, no significant difference was observed among the 2 groups. [Figure 4C and 4D] The **supplementary table 4** shows the median transfusion free survival and hazard ratio in both the patient groups. The number of modulating alleles, transfusion free survival ratio was inversely associated with the hazard ratio.

4. Discussion:

Though β-thalassemia and sickle cell disease are single-gene disorders with prototypical Mendalian inheritance patterns, both the disorders display a wide spectrum of clinical phenotypes. Thus, the search for the genetic modifiers was triggered, as 5–10 % of β-thalassemia homozygous patients with the same β-globin gene mutation and sickle cell anemia patients showed a variable pattern of clinical expression [9].

In this study, we first classified the β -thalassemia patients according to the clinical severity and then studied the influence of the genetic modifiers. **Modell and Berdukas 1984**, reported that 60 % of β-thalassemia homozygous patients presented in the first year of life, these patients were segregated as β-thalassemia major and 9 % of the β-thalassemia homozygous patients who presented after 2 years of age, with intermediate clinical severity were classified as β- thalassemia intermedia [10,11]. A similar observation was made in our study, in which the β-thalassemia major patients presented early by 9.2 ± 2.7 months and the patients in the β-thalassemia intermedia group had a delayed age of presentation mean of 4.3 ± 3.3 years. The β-thalassemia intermedia patients also showed a significantly higher mean baseline hemoglobin of 7.8 ± 1.4 g/dL as compared to thalassemia major patients. Similarly, another study showed that in 63 β-thalassemia intermedia patients, the hemoglobin values ranged between 7–9 g/dL with occasional transfusion regimen and splenomegaly [12]. In our study as well, pronounced hepatosplenomegaly was observed in β-thalassemia intermedia patients as compared to β-thalassemia major. Mpalampa L et al., 2012 considering the mean HbF cut-off as 10 %, in 216 sickle cell anemia patients observed a strong negative correlation of HbF levels with the total number of transfusions ($r = -0.181$, P: 0.004), hospitalisations rate ($r = -0.173$, P : 0.006), and significant positive correlation with the age at diagnosis ($r = 0.151$, P : 0.013) [13]. In the Indian context, Nayak et al., 2018 studied 60 sickle cell anemia patients and observed fewer episodes of painful crises in children with high baseline HbF level as compared to children with low HbF level [14]. Correspondingly in our study as well, the mean age of diagnosis among 100 SCA patients was found to be 6.3 ± 5.2 years which is very much delayed as compared to the patient cohort studied by Mpalampa L et al., 2012 This observation could be due to inherently elevated HbF levels in Indian patients mainly due to Arab-Indian haplotype which is a major determinant of HbF levels in Indian SCA patients [15]. Further, it was observed that the patients with higher HbF level had a delayed age of presentation (7.1 ± 5.5 years) with less transfusion requirement and sporadic painful crisis compared to patients with HbF level ≤ 17.4 % (age of presentation: 5.2 ± 4.9 years).

As the β-thalassemia alleles inherited by the patient act as a primary modulator of the disease severity in β-thalassemia, Colah et al., 2004 observed that the milder mutations are prevalent in β-thalassemia

intermedia group as compared to severe β-thalassemia major patients [16]. Similarly in our study, the presence of milder β-thalassemia alleles were significantly higher in β-thalassemia intermedia as compared to β-thalassemia major patients [P: 0.004, Odds Ratio: 8.6 (95 % CI: 1.9–37.9)]. However, Garewal et al.,2007 described that, in majority of the Indian patients, the beta genotype alone cannot predict the clinical phenotype of the patients [17]. A similar observation was seen in our patients, which suggested the presence of other genetic factors that may play a synergistic role in modifying the disease severity of β-thalassemia. In a previous study by Nadkarni et al., 2001, the associated α-thalassemia was found to be significantly higher in the thalassemia intermedia group (37%) as compared to β-thalassemia major group (5%) (P < 0.025) [18]. A study by Pandey et al., 2011 revealed 32 % sickle cell anemia patients with co-existing α-globin gene deletion, showed a relatively milder clinical course with improved hematological indices and reduced transfusion history [19]. Similarly, Rumaney et al., 2014 observed that in Cameroon sickle cell disease patients, co-inheritance of α-thalassemia showed improved hematological indices with a better survival rate [20]. Similarly in this study, we observed that the coinheritance of α-thalassemia was higher in the milder β-thalassemia patient group as compared to the other group. 51 % of SCA patients also showed presence of α-thalassemia.

The effect of the genetic modifiers of fetal hemoglobin was also analysed in this study. A study in the Egyptian β-thalassemia patients showed that 83.3 % of β-thalassemia intermedia cases were heterozygous for XmnI polymorphism as compared to β-thalassemia major (57.6 %) and that β-thalassemia intermedia with single T allele of XmnI showed delayed age of diagnosis, raised HbF levels and milder disease phenotype as compared patients negative for the XmnI polymorphism [21]. In another study, it was also determined that the patients with homozygosity for the mutant T allele of XmnI polymorphism significantly showed higher mean HbF levels (85.5 ± 6.8 %) as compared to the thalassemia intermedia patients homozygous for XmnI CC genotype (19.5 % ± 29.3) [22]. A similar result was observed in our patient group where in the β-thalassemia intermedia patients homozygous for variant allele T showed significantly higher HbF level.

+ 25 G→A polymorphism in ^{A}γ-globin promoter region was found to be significantly associated with elevated HbF levels in the β-thalassemia intermedia group. This polymorphism was first reported by Bianchi et al., 2016 and a strong linkage of this polymorphism with the − 158 C→T (XmnI polymorphism) was observed in their study as well [23]. It has been reported that + 25 G→A polymorphism reduces the binding efficacy of LYAR transcription factor (repressor of γ-globin gene expression) and abolishes the binding of 2 negative epigenetic regulators [DNA methyltransferase 3 alpha (DNMT3A) and protein arginine methyltransferase 5 (PRMT5)] to this promoter region [23, 24]. Thus, it could be speculated that there could be a cumulative effect of mutant alleles of both XmnI polymorphism (T allele) and + 25 G→A polymorphism (A allele) in synergistically elevating the HbF levels.

The association of BCL11A polymorphisms with elevated HbF levels and their effect on amelioration of the disease phenotype was studied by Uda et al., 2008 in Sardinian β-thalassemia homozygous patients [25]. They showed that the mutant C allele of rs11886868 (C→T) formed the major allele in Sardinian population and was significantly associated with elevated HbF levels in β-thalassemia intermedia patient

group. Similarly, in Indian patients Dadheech et al., 2016, determined that the C allele was significantly associated with the raised HbF levels and delayed the age of presentation in both thalassemia homozygous and SCA groups [26]. In our study, the mutant CC genotype was found to be significantly associated with HbF levels only in the sickle cell anemia patients.

Similarly, in Indonesian HbE-β-thalassemia patients inheriting variant alleles of rs11886868, rs766432 in the *BCL11A* gene, showed higher HbF levels and reduced disease severity as compared to patients with wild type alleles [27]. The second SNP that was found to be significantly associated with the HbF levels is rs1427407 (G→T) polymorphism in the *BCL11A* gene. Our results were found to be consistent with the earlier report by Bhanushali et al., 2014, who showed a similar distribution of allelic frequency of rs1427407 in Indian SCA patients [28]. Studies have demonstrated that the patients with the mutant T allele of rs1427407 (G→T) showed significantly higher HbF level, the results of which are concordant with our study, thus suggesting a crucial role of this SNP in modulating the HbF levels [28, 29].

Similarly, Chaouch et al., 2016 observed that co inheritance of the mutant C allele of the rs11886868 and the mutant A allele of the rs46713939 ameliorated the clinical phenotype of SCA patients [30]. In our study, though the A allele of rs46713939 was found to be higher in the TI group, no significant difference in the allelic frequencies among the milder and severe groups could be observed. Studies have identified a restricted 14 kb region in BCL11A intron 2 to be associated with H3-acetylation, RNA pol II activity as well as a strong GATA-I, TAL-1 binding site, all of which indicated the presence of a regulatory sequence in this region [31, 32]. Thus, suggesting that the presence of a polymorphism could potentially alter the recruitment of transcription factors to this region.

Similarly, in our population, a 100 % linkage was observed between rs66650371 and rs9399137 polymorphism. A similar observation was seen in the Tanzanian SCD patients where, both these polymorphisms in HMIP 2A block were strongly associated with HbF levels and showed a strong linkage [33]. Similarly, Lai et al., 2017 in β-thalassemia intermedia patients showed that the mutant alleles of rs9376090, rs7776054, rs9399137, rs9389268, rs9402685 in the HbS1L-MYB intergenic region and rs189984760 in the BCL11A locus, showed significant association with high HbF level [34]. Bioinformatic characterization of the 3 bp deletion polymorphism showed that this region acts as a binding site for 4 transcription factors TAL1/SCL, E47, GATA-2 and RUNX1/ AML1 all of which are important erythroid differentiation, erythropoiesis and the presence of the mutant allele may also disrupt the MYB gene expression, which is a negative regulator of the γ-globin gene [35, 36].

In our study, it was observed that in SCA patients, the prevalence of the T allele of rs1427407 (G→T) and the 3 bp deletional allele of rs66650371 both were significantly higher in the SCA patient group possessing higher HbF levels (HbF > 17.4 %) as compared to the SCA patients, with lower HbF level (HbF ≤ 17.4 %). A similar result was shown by Adeyemo et al., 2018, where in patients with the mutant alleles of these 2 polymorphisms had a milder form of the disease, with improved hemoglobin levels [37].

Similarly, in β-thalassemia homozygous patients, the cumulative effect of 3 HbF associated mutant alleles of the SNPs – 158(C→T), rs11886868 (C→T) and rs1427407 (G→T), were observed significantly in the β-

thalassemia intermedia group as compared to β-thalassemia major group. A comparable result was reported by Allawi et al., 2015, where they determined the main factors leading to milder phenotypes were the attenuated β-thalassemia alleles, the T allele of XmnI polymorphism and the minor allele of BCL11A rs10189857 [38]. Cardoso et al., 2014, studied the influence of three known major loci on the HbF trait (HBG2, rs748214; BCL11A, rs4671393; and HBS1L-MYB, rs28384513, rs489544 and rs9399137) in north Brazilian SCA patients and they showed that the raised HbF trait was primarily influenced by mutant alleles of BCL11A [39].

Further to predict the disease severity in presence of these genetic modifiers, Badens et al., 2011 studied 5 genetic modifiers of β-thalassemia. By regression analysis, all 5 types of favorable allele were found to be significantly associated with thalassemia intermedia phenotype. The β-globin gene mutations and XmnI polymorphism were the most influential modifiers of the disease severity [5]. A similar observation was reported by Danjou et al., 2012, wherein they further evaluated the age of the first transfusion with respect to the inheritance of HbF boosting alleles and observed that the age of transfusion was found to be delayed in presence of more number of HbF inducing alleles [40]. Similarly, in our study, it was observed that, 8% of thalassemia intermedia patients (HbF: 82.7 % ± 22.6) had inherited more than 10 disease ameliorating alleles as compared to none in thalassemia major patients. Thalassemia major patients (64 %) showed less number of disease modifying alleles as compared to the thalassemia intermedia patients (36%). Similarly, in SCA patients, the ameliorating alleles was found to be higher in the raised HbF group (HbF > 17.4) (31.3 %) (mean HbF: 26.2%±8.3) as compared to patients in low HbF group (HbF ≤ 17.4%) (4 %) (mean HbF: 12.1 %± 3.5).

These observations suggest that the presence of increased number of an ameliorating allele, may help in reducing the disease severity in hemoglobinopathy patients mainly by restoring the globin chain imbalance. The precise identification of the polymorphisms associated with elevated HbF levels may help in developing a molecular chip that may assist in predicting the disease severity.

5. Conclusions:

The present study expands the knowledge of the frequency of the genetic modifiers (primary and secondary modifiers) and the independent effect of individual predictor genes on HbF levels in hemoglobinopathy patients. The analysis of the cumulative effect of the HbF modulators may help in identifying the strongest response gene to the HbF level in both β-thalassemia and sickle cell anemia patients in the population. The predictions based on genetic modifiers thus can foresee the severity of β-thalassemia and SCA. This study may assist the clinicians, to predict the clinical phenotype of hemoglobinopathy patients at an early stage and thus may help in the efficient management of the disease. This may contribute towards molecular mechanisms of HbF regulation and the development of therapeutic approaches for β- hemoglobinopathies.

Declarations

References

1. Balgir, R. The burden of haemoglobinopathies in India and the challenges ahead. *Curr Sci*. 79, 1536–47 (2000). https://www.jstor.org/stable/24104846

2. Colah, R. *et al*. Impact of beta globin gene mutations on the clinical phenotype of beta thalassemia in India. *Blood Cells Mol Dis*. **33**, 153–157 https://doi.org/10.1016/j.bcmd.2004.05.002 (2004).

3. Garewal, G., Das, R., Awasthi, A., Ahluwalia, J. & Marwaha, R. The clinical significance of the spectrum of interactions of CAP + 1 (A→C), a silent beta-globin gene mutation, with other beta-thalassemia mutations and globin gene modifiers in north Indians. *Eur J Haematol*. **79**, 417–421 https://doi.org/10.1111/j.1600-0609.2007.00958.x (2007).

4. Nadkarni, A. *et al*. Molecular pathogenesis and clinical variability of beta-thalassaemia syndromes among Indians. *Am J Hematol*. **68**, 75–80 https://doi.org/10.1002/ajh.1156 (2001).

5. Pandey, S., Pandey, S., Mishra, R., Sharma, M. & Saxena, R. Genotypic influence of α-deletions on the phenotype of Indian sickle cell anemia patients. *Korean J Hematol*. **46**, 192–195

https://doi.org/10.5045/kjh.2011.46.3.192 (2011).

6. Rumaney, M. *et al.* The co-inheritance of alpha-thalassemia and sickle cell anemia is associated with better hematological indices and lower consultations rate in Cameroonian patients and could improve their survival. *PLoS One.* **9**, e100516 https://doi.org/10.1371/journal.pone.0100516 (2014).

7. Said, F. & Abdel-Salam, A. XmnI polymorphism: Relation to β thalassemia phenotype and genotype in Egyptian Children. *Egypt J Med Hum Genet.* **16**, 123–127 https://doi.org/10.1016/j.ejmhg.2014.12.005 (2015).

8. Miri-Moghaddam, E., Bahrami, S., Naderi, M., Bazi, A. & Karimipoor, M. Xmn1-158 γ G Variant in β-Thalassemia Intermediate Patients in South-East of Iran. Int J Hematol Oncol Stem Cell Res.11,165–171(2017). PMID: 28875012

9. Bianchi, N. *et al.* Structural and Functional Insights on an Uncharacterized Aγ-Globin-Gene Polymorphism Present in Four β0-Thalassemia Families with High Fetal Hemoglobin Levels. *Mol Diagn Ther.* **20**, 161–173 https://doi.org/10.1007/s40291-016-0187-2 (2016).

10. Chen, D. *et al.* A Genetic Variant Ameliorates β-Thalassemia Severity by Epigenetic-Mediated Elevation of Human Fetal Hemoglobin Expression. *Am J Hum Genet.* **101**, 130–138 https://doi.org/10.1016/j.ajhg.2017.05.012 (2017).

11. Uda, M. *et al.* Genome-wide association study shows BCL11A associated with persistent foetal haemoglobin and amelioration of the phenotype of β-thalassaemia. *Proc Natl Acad Sci U S A.* **105**, 1620–1625 https://doi.org/10.1073/pnas.0711566105 (2008).

12. Dadheech, S. *et al.* Association of BCL11A genetic variant (rs11886868) with severity in β-thalassaemia major & sickle cell anaemia. *Indian J Med Res.* **143**, 449–454 https://doi.org/10.4103/0971-5916.184285 (2016).

13. Rujito, L. *et al.* Modifying effect of XmnI, BCL11A, and HBS1L-MYB on clinical appearances: A study on β-thalassemia and hemoglobin E/β-thalassemia patients in Indonesia. *Hematol Oncol Stem Cell Ther.* **9**, 55–63 https://doi.org/10.1016/j.hemonc.2016.02.003 (2016).

14. Bhanushali, A., Patra, P., Nair, D., Verma, H. & Das, B. Genetic variant in the BCL11A (rs1427407), but not HBS1-MYB (rs6934903) loci associate with fetal hemoglobin levels in Indian sickle cell disease patients. *Blood Cells Mol Dis.* **54**, 4–8 https://doi.org/10.1016/j.bcmd.2014.10.003 (2015).

15. Sebastiani, P. *et al.* BCL11A enhancer haplotypes and fetal hemoglobin in sickle cell anemia. *Blood Cells Mol Dis.* **54**, 224–230 https://doi.org/10.1016/j.bcmd.2015.01.001 (2015).

16. Chaouch, L. *et al.* rs11886868 and rs4671393 of BCL11A associated with HbF level variation and modulate clinical events among sickle cell anemia patients. *Hematology.* **3**, 1–5 https://doi.org/10.1179/1607845415Y.0000000026 (2016).

17. Jawaid, K., Wahlberg, K., Thein, S. & Best, S. Binding patterns of BCL11A in the globin and GATA1 loci and characterization of the BCL11A fetal hemoglobin locus. *Blood Cells Mol Dis.* **45**, 140–146 https://doi.org/10.1016/j.bcmd.2010.05.006 (2010).

18. Liu, L. *et al.* Original Research: A case-control genome-wide association study identifies genetic modifiers of fetal hemoglobin in sickle cell disease. *Exp Biol Med (Maywood).* **241**, 706–718

https://doi.org/10.1177/1535370216642047 (2016).

19. Mtatiro, S. *et al.* Genetic association of fetal-hemoglobin levels in individuals with sickle cell disease in Tanzania maps to conserved regulatory elements within the MYB core enhancer. *BMC Med Genet.* **16**, 4 https://doi.org/10.1186/s12881-015-0148-3 (2015).

20. Lai, Y. *et al.* Genetic variants at BCL11A and HBS1L-MYB loci influence Hb F levels in Chinese Zhuang β-thalassemia intermedia patients. *Hemoglobin.* **40**, 405–410 https://doi.org/10.1016/j.bcmd.2017.01.011 (2016).

21. Farrell, J. *et al.* A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. *Blood.* **117**, 4935–4945 https://doi.org/10.1182/blood-2010-11-317081 (2011).

22. Stadhouders, R. *et al.* HBS1L-MYB intergenic variants modulate foetal haemoglobin via long-range MYB enhancers. *J Clin Invest.* **124**, 1699–1710 https://doi.org/10.1172/JCI71520 (2014).

23. Adeyemo, T. *et al.* A survey of genetic fetal-haemoglobin modifiers in Nigerian patients with sickle cell anaemia. *PLoS One.* **13**, e0197927 https://doi.org/10.1371/journal.pone.0197927 (2018).

24. Al-Allawi, N., Puehringer, H., Raheem, R. & Oberkanins, C. Genetic Modifiers in β-Thalassemia Intermedia: A Study on 102 Iraqi Arab Patients. *Genet Test Mol Biomarkers.* **19**, 242–247 https://doi.org/10.1089/gtmb.2014.0310 (2015).

25. Cardoso, G. *et al.* DNA polymorphisms at BCL11A, HBS1L-MYB and Xmn1-HBG2 site loci associated with fetal hemoglobin levels in sickle cell anemia patients from Northern Brazil. *Blood Cells Mol Dis.* **53**, 176–179 https://doi.org/10.1016/j.bcmd.2014.07.006 (2014).

26. Danjou, F. *et al.* Genetic modifiers of β-thalassemia and clinical severity as assessed by age at first transfusion. *Haematologica.* **97**, 989–993 https://doi.org/10.3324/haematol.2011.053504 (2012).

27. 27. Rujito, L., Basalamah, M., Siswandari, W., Setyono, J., Wulandari, G., Mulatsih, S., et al. Modifying effect of XmnI, BCL11A, and HBS1L-MYB on clinical appearances: A study on β-thalassemia and hemoglobin E/β-thalassemia patients in Indonesia. *Hematol Oncol Stem Cell Ther*. 9, 55–63 (2016). doi: 10.1016/j.hemonc.2016.02.003.

28. 28. Bhanushali, A., Patra, P., Nair, D., Verma, H., and Das, B. Genetic variant in the BCL11A (rs1427407), but not HBS1-MYB (rs6934903) loci associate with fetal hemoglobin levels in Indian sickle cell disease patients. *Blood Cells Mol Dis.* 54, 4–8 (2015). doi: 10.1016/j.bcmd.2014.10.003.

29. 29. Sebastiani, P., Farrell, J., Alsultan, A., Wang, S., Edward, H., Shappell, H, et al. BCL11A enhancer haplotypes and fetal hemoglobin in sickle cell anemia. *Blood Cells Mol Dis.* 54, 224–30 (2015). doi: 10.1016/j.bcmd.2015.01.001.

30. 30. Chaouch, L., Moumni, I., Ouragini, H., Darragi, I., Kalai, M., Chaouachi, D., et al. rs11886868 and rs4671393 of BCL11A associated with HbF level variation and modulate clinical events among sickle cell anemia patients. *Hematology.* 3, 1–5 (2016). doi: 10.1179/1607845415Y.0000000026.

31. 31. Jawaid, K., Wahlberg, K., Thein, S., and Best, S. Binding patterns of BCL11A in the globin and GATA1 loci and characterization of the BCL11A fetal hemoglobin locus. *Blood Cells Mol Dis*. 45,140-6 (2010). doi: 10.1016/j.bcmd.2010.05.006.

32. 32. Liu, L., Pertsemlidis, A., Ding, L., Story, M., Steinberg, M., Sebastiani, P., Hoppe, C., et al. Original Research: A case-control genome-wide association study identifies genetic modifiers of fetal hemoglobin in sickle cell disease. *Exp Biol Med (Maywood).* 241, 706 – 18 (2016). doi: 10.1177/1535370216642047.

33. 33. Mtatiro, S., Mgaya, J., Singh, T., Mariki, H., Rooks, H., Soka, D., et al. Genetic association of fetal-hemoglobin levels in individuals with sickle cell disease in Tanzania maps to conserved regulatory elements within the MYB core enhancer. *BMC Med Genet.* 16, 4 (2015). DOI 10.1186/s12881-015-0148-3

34. 34. Lai, Y., Chen, Y., Chen, B., Zheng, H., Yi, S., Li, G., et al. Genetic variants at BCL11A and HBS1L-MYB loci influence Hb F levels in Chinese Zhuang β-thalassemia intermedia patients. *Hemoglobin.* 40, 405 – 10 (2016). DOI: 10.1016/j.bcmd.2017.01.011

35. 35. Farrell, J., Sherva, R., Chen, Z., Luo, H., Chu, B., Ha, S., et al. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. *Blood.*117, 4935-45 (2011).DOI: 10.1182/blood-2010-11-317081

36. 36. Stadhouders, R., Aktuna, S., Thongjuea, S., Aghajanirefah, A., Pourfarzad, F., van IJcken, W., et al. HBS1L-MYB intergenic variants modulate foetal haemoglobin via long-range MYB enhancers. *J Clin Invest.* 124, 1699 – 710 (2014). doi:10.1172/JCI71520.

37. 37. Adeyemo, T., Ojewunmi, O., Oyetunji, I., Rooks, H., Rees, D., Akinsulie, A., et al. A survey of genetic fetal-haemoglobin modifiers in Nigerian patients with sickle cell anaemia. *PLoS One.* 13, e0197927 (2018). doi: 10.1371/journal.pone.0197927

38. 38. Al-Allawi, N., Puehringer, H., Raheem, R., and Oberkanins, C. Genetic Modifiers in β-Thalassemia Intermedia: A Study on 102 Iraqi Arab Patients. *Genet Test Mol Biomarkers.* 19, 242-7 (2015). doi: 10.1089/gtmb.2014.0310.

39. 39. Cardoso, G., Diniz, I., da Silva, A., Cunha, D., da Silva Junior, J., Uchôa, C., et al. DNA polymorphisms at BCL11A, HBS1L-MYB and Xmn1-HBG2 site loci associated with fetal hemoglobin levels in sickle cell anemia patients from Northern Brazil. *Blood Cells Mol Dis.* 53, 176 – 79 (2014). doi: 10.1016/j.bcmd.2014.07.006.

40. 40. Danjou, F., Anni, F., Perseu, L., Satta, S., Dessì, C., Lai, M., et al. Genetic modifiers of β-thalassemia and clinical severity as assessed by age at first transfusion. *Haematologica.* 97, 989 – 93 (2012). https://doi.org/10.3324/haematol.2011.053504
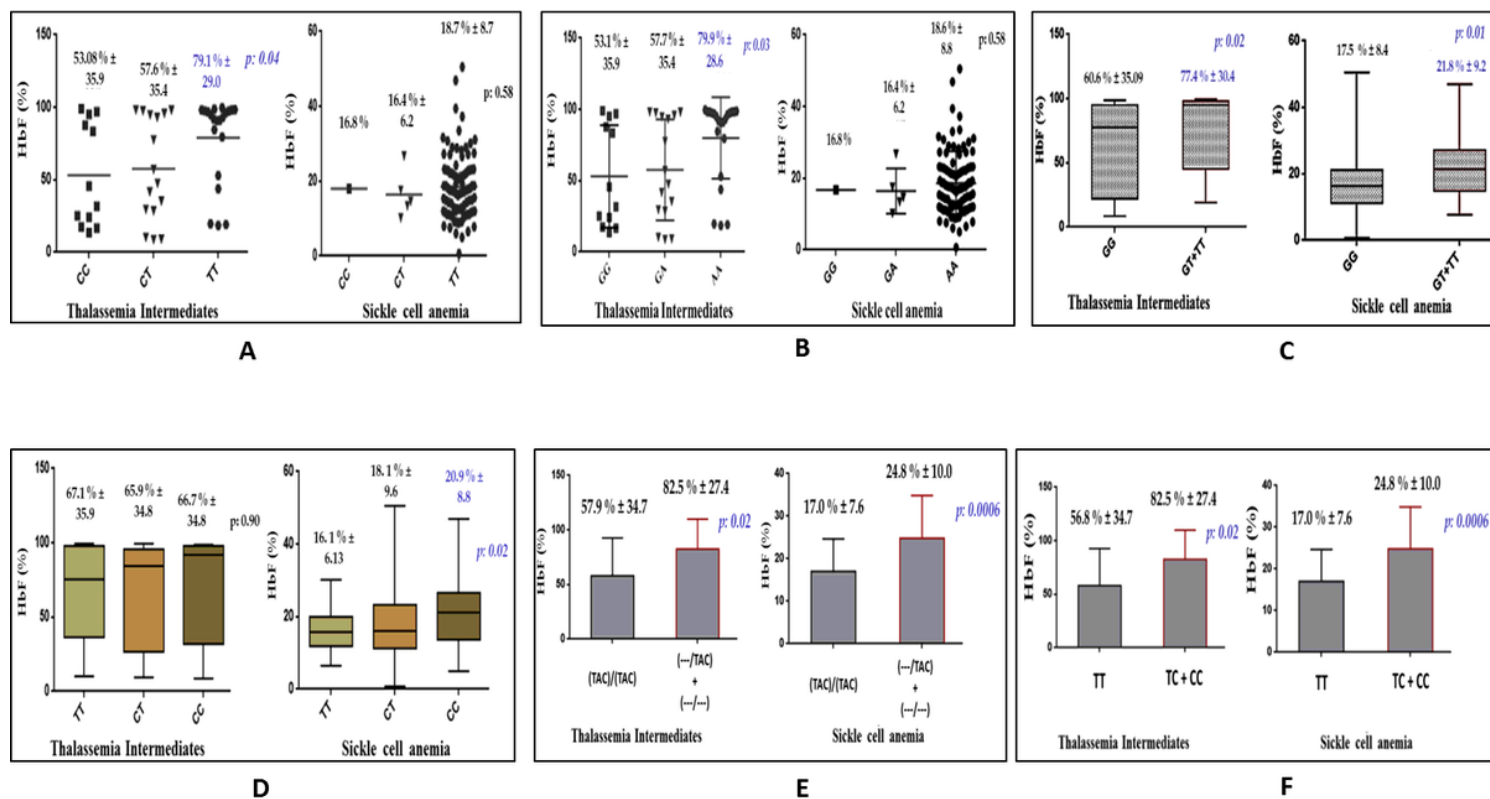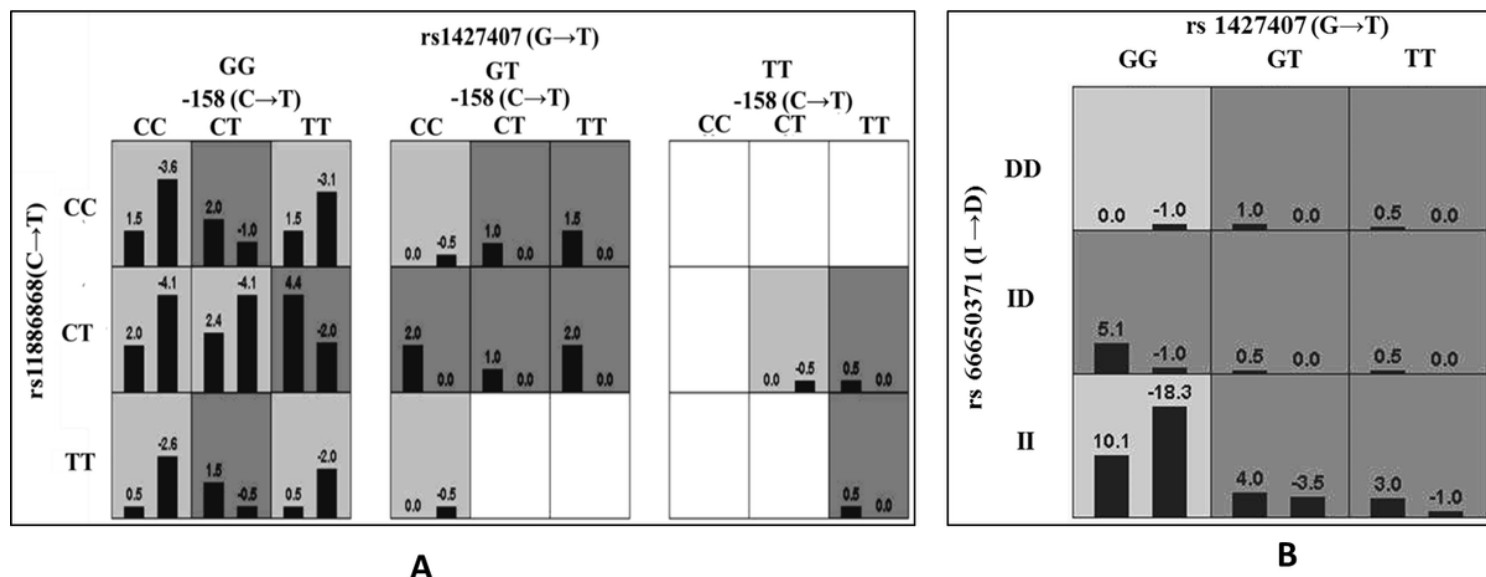
Figures

Figure 1

Association of the SNP genotypes with HbF levels in β-hemoglobinopathy patients A: XmnI polymorphism: -158 (C→T), B: +25 (G→A) polymorphism, C: rs1427407 (G→T) polymorphism, D: rs11886868 (C→T), E: rs66650371 (Intact TAC → 3 bp deletion) polymorphism, F: rs9399137 (T→C) polymorphism
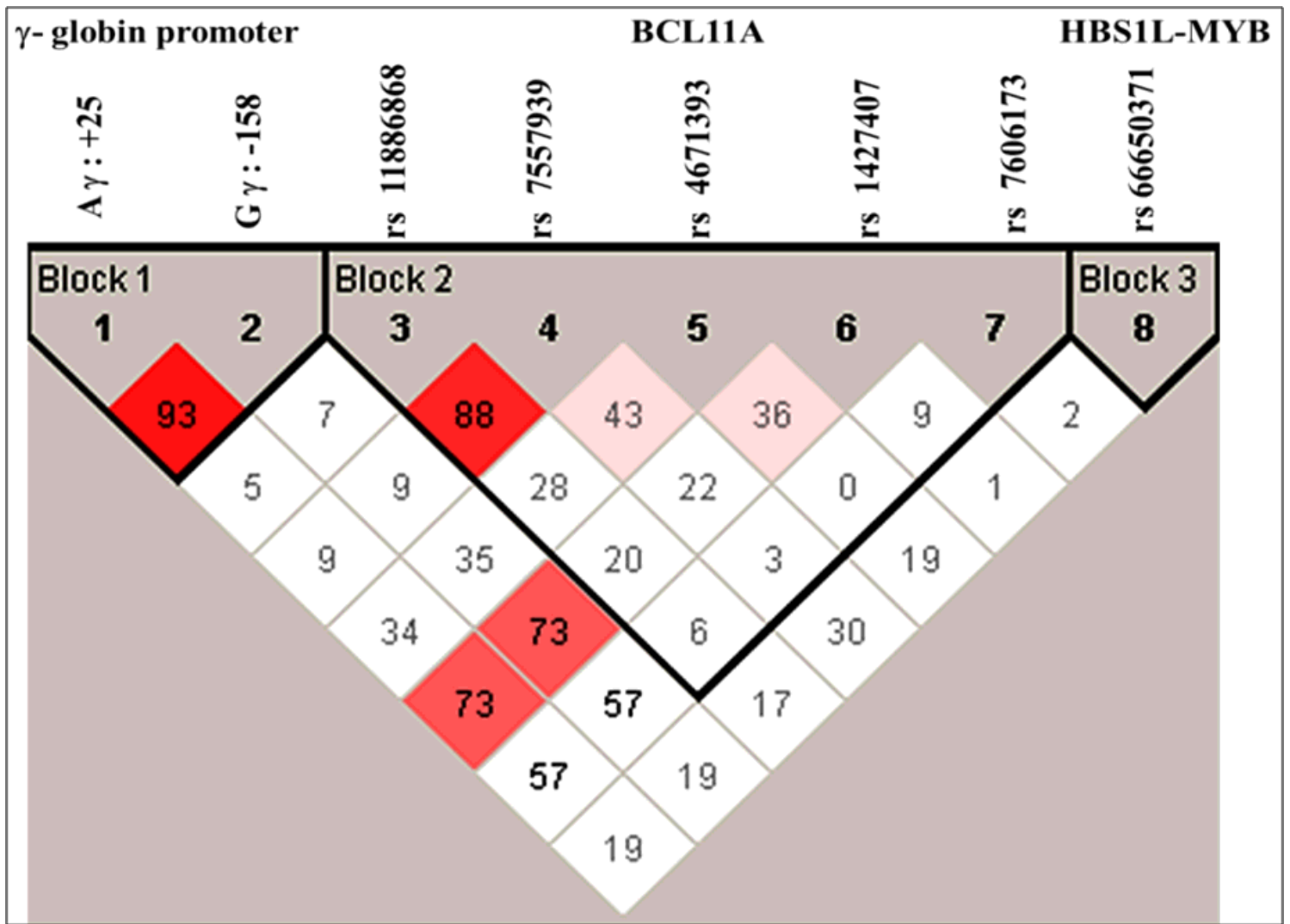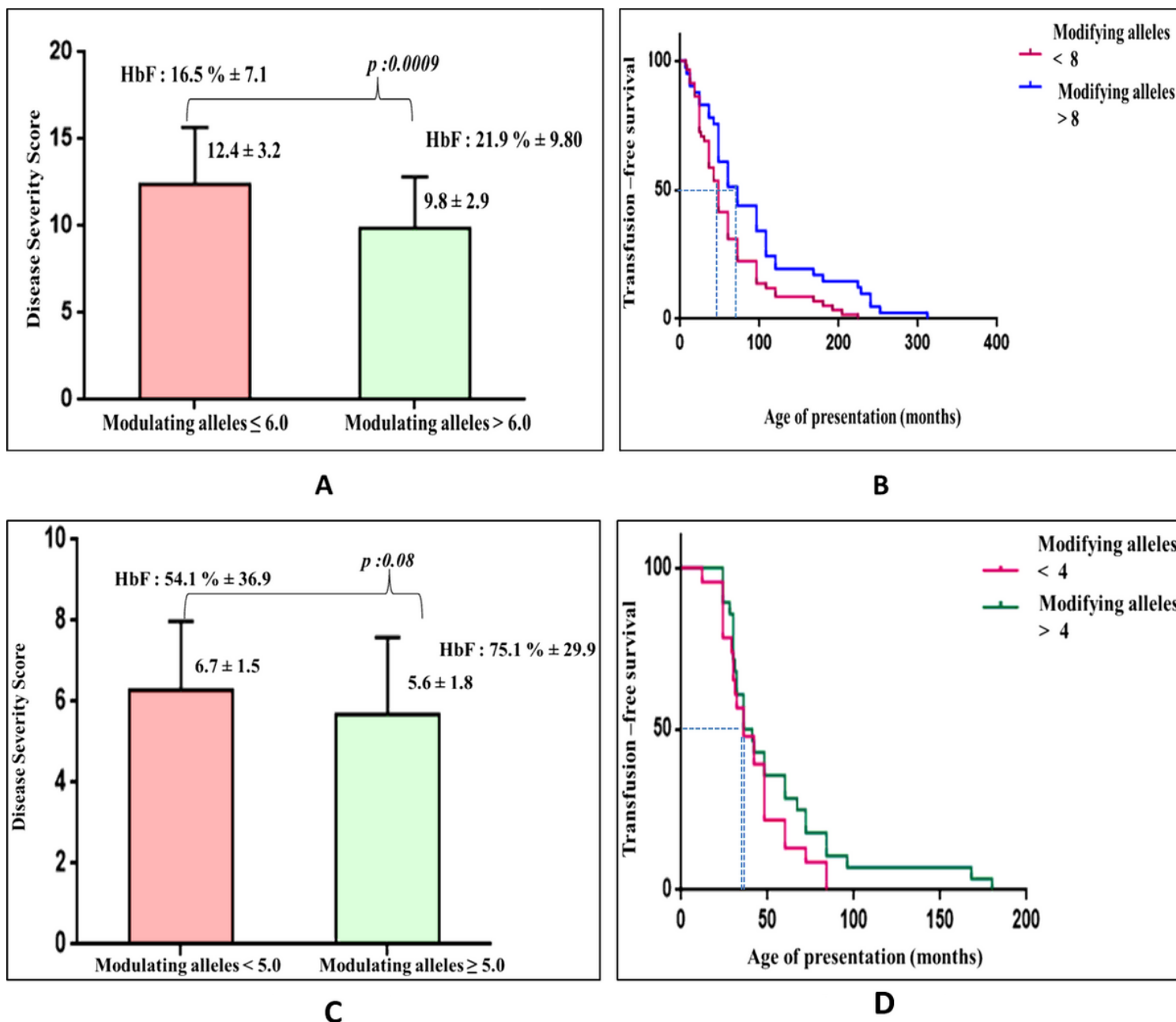
Figure 2

Linkage disequilibrium (LD) plot of the SNPs It could be observed that the +25 (G→A) polymorphism in Aγ-globin gene and the XmnI polymorphism (-158 C→T) in the Gγ-globin are highly linked (D': 93). Also the SNPs rs11886868 (C→T) and rs7557939 (A→G) in the BCL11A intronic region, are strongly linked (D': 88) with each other.

Figure 3

SNP tile analysis of gene−gene interaction in β-hemoglobinopathy patients The tiles are shaded according to the cumulative scores of both the bars. Here the dark grey tiles represent a positive score (presence of mutant HbF boosting allele in milder patient group) and the light grey tiles represent a negative score (presence of normal allele in the severe patient group) A: The mutant alleles of 3 SNPs (XmnI polymorphism, rs11886868(C→T) and rs1427407 (G→T) found to be significantly higher in β-thalassemia intermedia patients (the left bar) as compared to the β-thalassemia major patients (right bar). B: The mutant alleles 2 SNPs [rs66650371 (TAC (I) > deletion of 3 bp (D)) and rs1427407 (G→T)] found to be significantly higher in SCA patients with HbF levels > 17.4 % (left bar) as compared to the SCA patients with HbF ≤ 17.4 % (right bar).

Figure 4

Association of disease ameliorating alleles with disease severity in the patients A: Depicts the disease severity score against the median number of HbF modulating alleles in the SCA patients. The patients with higher number of modulating alleles showed significant positive correlation with the HbF levels (mean HbF: 21.9% ±9.8) as compared to the other patient group. B: Kaplan Meier transfusion free survival curve was plotted considering all the genetic modifiers (α-thalassemia and the HbF modifiers) and a delayed median age of presentation in patients with more number of modifying alleles: 72 months was observed as compared to the other group : 48 months. (P:0.02, Gehan Breslow Wilcoxon test) C: Depicts the disease severity score (DSS) against the number of HbF modulating alleles in the β-thalassemia intermedia patients. The patients inheriting more number of the disease ameliorating alleles showed elevated HbF levels (mean HbF: 75.1 % ± 29.9), with reduced disease severity score (mean DSS :5.6) as compared to the

patients in the other group (P: 0.08). D: Kaplan Meier transfusion free survival curve analysis (considering both primary and secondary modifiers) showing slight delay in median age of presentation in patients with more number of modifying alleles (38.5 months) as compared to group (36.0 months) (P: 0.36, Gehan Breslow Wilcoxon test).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.docx