Frontiers in Bioscience Scholar

Front. Biosci. (Schol Ed) **2025**; 17(2): 38388
https://doi.org/10.31083/FBS38388

Original Research

Genetic Modifiers of HbF in HbAA and HbAS Women From São Tomé e Príncipe: An Association Study of Common Genetic Variants in *BCL11A*, *MYB*, *HBG2*, and *BGLT3*

Licínio Manco[1,2,*], Afonso Marques Morais[1], Sara Miguel Almeida[1], Inês Salgado[1], Luís Relvas[3], Celdidy Monteiro[4], Guilherme Queiroz[1,5], Celeste Bento[1,2]

[1]Centro de Investigação em Antropologia e Saúde (CIAS), Universidade de Coimbra, 3000 Coimbra, Portugal
[2]Departmento de Ciências da Vida, Universidade de Coimbra, 3000 Coimbra, Portugal
[3]Departmento de Hematologia, Centro Hospitalar e Universitário de Coimbra (CHUC), 3000 Coimbra, Portugal
[4]Hospital Dr. Ayres de Menezes, São Tomé, República Democrática de São Tomé e Príncipe
[5]Unidade de Saúde Pública, Unidade Local de Saúde da Região de Aveiro, 3810 Aveiro, Portugal
*Correspondence: lmanco@antrop.uc.pt (Licínio Manco)
Academic Editor: Gustavo Caetano-Anollés
Submitted: 25 February 2025 Revised: 7 May 2025 Accepted: 28 May 2025 Published: 25 June 2025

Abstract

Background: While an increase in fetal hemoglobin (HbF) has no consequences in healthy adults, clinical benefits can be promoted in sickle cell disease (SCD) and β-thalassemia patients. Single-nucleotide polymorphisms (SNPs) in three genomic regions: the *HBB* gene cluster, the *BCL11A* gene, and the *HBS1L-MYB* (*HMIP*) intergenic region, have been associated with HbF regulation. Therefore, the present study aimed to examine the potential association of SNPs in *BCL11A* (rs11886868 and rs1427407), *HMIP* (rs66650371 and rs4895441), *HBG2* (rs7482144), and *BGLT3* (rs7924684) with HbF levels in an adult population sample from São Tomé e Príncipe (Central Africa). **Methods**: A total of 145 women aged 18 to 49 years were involved in this study, comprising 98 women with the normal hemoglobin (Hb) genotype (HbAA) and 47 with sickle cell trait (HbAS). From the HbAA individuals, we selected a control group of 60 subjects with normal HbF levels, ranging from 0.2% to 1.4% (mean: 0.75%), and a case group of 38 subjects with elevated HbF levels, ranging from 1.8% to 3.7% (mean: 2.35%). In the group of HbAS individuals, the HbF levels ranged from 0.4% to 3.7% (mean: 1.56%). SNP genotyping was conducted using standard molecular methods. **Results**: Logistic regression, in the additive model, revealed significant associations with increased levels of HbF for the minor alleles of the two *BCL11A* SNPs, rs11886868 [C] and rs1427407 [T], in HbAA women ($p = 0.00018$ and $p = 0.00076$, respectively). When comparisons of HbF levels were conducted among genotypes in the HbAA women, significant differences were observed for *BCL11A* SNPs rs11886868 and rs1427407, as well as for the *HBG2* rs7482144 and *BGLT3* rs7924684 variants. We found no association between HbF levels and the two *HMIP* variants rs66650371 and rs4895441 in the HbAA women. Among the HbAS women, no statistically significant associations were observed between the six analyzed polymorphisms and HbF levels ($p > 0.05$). **Conclusions**: We successfully replicated the association between the two well-known *BCL11A* SNPs, rs11886868 and rs1427407, with HbF levels in women with the normal HbAA genotype from São Tomé e Príncipe. Other signals of association with HbF levels were identified for the SNPs *HBG2* (rs7482144) and *BGLT3* (rs7924684).

Keywords: HbF regulation; *BCL11A*; *HMIP*; *HBG2*-XmnI; *BGLT3*; São Tomé e Príncipe

1. Introduction

Fetal hemoglobin (HbF) is a tetrameric molecule composed of two alpha-globin and two gamma-globin chains ($\alpha 2\gamma 2$) typically expressed during fetal development. It starts to decrease just before birth when it is replaced by adult hemoglobin A (HbA; $\alpha 2\beta 2$), which has a lower affinity for oxygen [1]. In healthy adults, red blood cells predominantly contain HbA, along with 2.5% to 3.5% HbA2 ($\alpha 2\delta 2$) and less than 1% HbF, which is unevenly distributed among erythrocytes. However, about 10% to 15% of normal adults have elevated levels of HbF that can reach up to 5.0%. This condition, known as hereditary persistence of fetal hemoglobin (HPFH), can be caused by deletions within the *HBB* gene cluster on chromosome 11, point mutations in the promoters of the HbF genes *HBG1* and *HBG2*, or variations in major repressors of the HbF genes, such as *BCL11A* and *MYB* [2,3].

In adults of African descent with sickle cell disease (SCD), HbF typically comprises 4% to 10% of the total hemoglobin levels, although it rarely exceeds 30% [2]. In contrast, African Americans with sickle cell trait (HbAS) have a mean HbF level of 1.4% [4]. While the increase in HbF levels has no consequences in healthy adults, it can provide clinical benefits in patients with SCD and β-thalassemia [5]. Therefore, understanding and managing the biological pathways associated with HbF expression could lead to clinical interventions for individuals with severe β-hemoglobinopathies by genetic modification of HbF expression [6,7].

Research has shown that variations in HbF levels in adults are significantly influenced by genetic factors [5,8]. One of the most studied variants is the XmnI restriction site polymorphism located in the promoter region of the *HBG2* gene, upstream of the transcription start site (–158C>T) (rs7482144). This polymorphism is known to impact HbF levels, particularly under conditions of increased erythropoietic stress, such as in patients with β-thalassemia and SCD patients [7,8]. Some patients with SCD have exceptionally high levels of HbF associated with the Senegal and Saudi-Indian haplotypes, both characterized by the presence of the rs7482144 [T] allele [5]. It is known that only the *HBG2* expression is affected by the –158 polymorphism; however, the transcription factors binding to the –158 *HBG2* motif remain uncharacterized [8].

Additionally, minor alleles of several single nucleotide polymorphisms (SNPs) in two other quantitative trait loci (QTL), the *BCL11A* gene on chromosome 2 and the *HBS1L-MYB* (*HMIP*) intergenic region on chromosome 6, are also strongly associated with HbF regulation [5]. Both genes *BCL11A* and *MYB* encode repressors of γ-globin genes [8]. The transcription factor BCL11A binds TGACCA motifs in the *HBB* gene cluster, preferentially to position –115 in the *HBG2* and *HBG1* gene promoters [8]. Functional variants of the *BCL11A* gene are marked by the intron 2 SNP rs1427407 which marks an erythroid-specific enhancer [8,9]. The *HMIP* intergenic variants affect regulatory elements that are occupied by key erythroid transcription factors within this region, affecting long-range interactions with *MYB* and in consequence the *MYB* expression levels [10]. The most significant functional motif accounting for *HMIP* modulation of HbF is rs66650371, a 3-bp deletion polymorphism, which is in complete LD with the SNP rs9399137 shown by several Genome-Wide Association Studies to be most significantly associated with HbF in Europeans, Asians, and African Americans [4,9].

Variants in these three QTLs have been reported to contribute up to 50% of HbF variation [11]. The increase of HbF remaining unexplained could be a result of rare alleles or polymorphic variations in other genes coding for protein factors involved in the regulation of HbF production, such as *KLF1* that activates expression of the *BCL11A* gene [12] or other still uncharacterized HbF repressors. Moreover, the impact of these variants can be different from one population to another, depending on how common a risk allele is in a population. For example, in European populations, a strong HbF influence is known for variants at *HMIP*, which are rare in the Sub-Saharan populations and hence have a low impact on HbF levels in such populations [11].

The high prevalence of SCD in Africa justifies genetic studies linked to HbF regulation in African populations. Although many HbF association studies were implemented in SCD patients of African ancestry [13], a large proportion of the HbF variance remains unexplained. Additional studies in healthy populations could highlight different genetic factors or pathophysiological pathways associated with HbF regulation. Therefore, the present study examines for the first time an adult population sample of the São Tomé e Príncipe archipelago in Central Africa for the potential association with HbF levels of several SNPs in the *HBB* gene cluster, *BCL11A* gene, and *HMIP* intergenic region.

2. Material and Methods

2.1 Study Sample and Hematological Data

The archipelago of São Tomé e Príncipe is located on the Equator in the Gulf of Guinea (Central Africa). A total of 145 women (18–49 years; mean 33.06) living in São Tomé and Príncipe were involved in this study, 98 having the normal hemoglobin (Hb) genotype (HbAA) and 47 with sickle cell trait (HbAS). This study population is a subsample of that previous analyzed by Queiroz *et al*. [14] for the prevalence of sickle cell trait and other hemoglobin variants in women of reproductive age from São Tomé e Príncipe. In the group of HbAA women, we selected a control group consisting of 60 subjects who had normal HbF values ranging from 0.2% to 1.4% (mean 0.75%) and a case group of 38 subjects with elevated HbF levels ranging from 1.8% to 3.7% (mean 2.35%). Among the 47 individuals with HbAS, the HbF levels range from 0.4% to 3.7% (mean 1.56%). Table 1 presents the demographic and hematological parameters of the study population. The measurement of HbF and other Hb variants were performed with the Automatic Glycohemoglobin Analyzer point-of-care device Lab001 (ARKRAY Inc., Kyoto, Japan) or, when this was not possible, by high performance liquid chromatography (HPLC) on the VARIANT II™ Hemoglobin Testing System (Bio-Rad Laboratories Inc., Hercules, CA, USA) in the laboratory context, as previously described in Queiroz *et al*. [14]. The HbAS condition was defined as a proportion of HbS between 20% and 45% and the HbAA condition was defined as a proportion of HbA above 70% showing no other Hb variants of interest [14].

2.2 Genotyping

Blood samples were collected on Guthrie filter paper (blood spots) from each individual and DNA extraction was carried out with the innuPREP Forensic Kit (LOT 021-23) (Innuscreen, GmbH, Berlin, Germany) according to the manufacturer's instructions.

The SNPs *BCL11A* rs1427407, *HMIP* rs4895441, *HBG2* rs7482144 (XmnI) and *BGLT3* rs7924684 were genotyped throughout the technique polymerase chain reaction - restriction fragment length polymorphism (PCR-RFLP). The PCR was performed with the Qiagen Multiplex PCR kit (Qiagen GmBH, Hilden, Germany) in a 12.5 µL reaction volume containing 10–50 ng of genomic DNA, 1× PCR Master Mix (Qiagen), and 5 pmol of direct and reverse primers. The PCR cycling conditions included an initial heat activation step of 15 min at 95 °C followed by 35 cycles of 95 °C denaturation for 30 s, an-

Table 1. Characteristics of women with a normal HbAA genotype and with HbAS from S. Tomé e Príncipe.

Parameters	Total	Group 1 (HbF <1.5%)	Group 2 (HbF >1.5%)
Women with HbAA			
Number	98	60	38
Age (years): Mean (SD)	33.00 (8.15)	33.58 (8.77)	32.08 (7.07)
Age (years): Median (IQR)	33 (11)	33.00 (17)	33 (10)
Age (years): Range	18–49	18–49	18–45
HbF (%): Mean (SD)	1.37 (0.88)	0.75 (0.3)	2.35 (0.52)
HbF (%): Median (IQR)	1 (1.4)	0.8 (0.4)	2.2 (0.83)
HbF (%): Range	0.20–3.7	0.2–1.4	1.8–3.7
Women with HbAS			
Number	47		
Age (years): Mean (SD)	33.19 (8.11)		
Age (years): Median (IQR)	33 (11)		
Age (years): Range	18–49		
HbF (%): Mean (SD)	1.56 (0.77)		
HbF (%): Median (IQR)	1.4 (0.80)		
HbF (%): Range	0.4–3.7		

Abbreviations: SD, standard deviation; IQR, interquartile range; HbF, fetal hemoglobin; HbAA, normal hemoglobin; HbAS, sickle cell trait.

nealing at 58 °C for 45 s and extension at 72 °C for 45 s. The primer sequences, detailed in **Supplementary Table 1**, were previously described in other studies [15,16] or newly designed using the online software Primer3web version 4.0.1 (https://primer3.ut.ee/). The genotype discrimination was performed using the restriction enzymes HpyCH4III, RsaI, XmnI and PuvII, as detailed in **Supplementary Table 2**. Sanger DNA sequencing was performed in about 10% of samples to confirm genotypes of the six analyzed SNPs using the ABI PRISM® Big Dye® Terminator v1.1 kit (Applied Biosystems, Foster City, CA, USA) in a SeqStudio Genetic Analyzer (Applied Biosystems). SNP *BCL11A* rs11886868 was genotyped by real-time PCR with a pre-designed TaqMan SNP Genotyping Assay (C__11363852_10; Applied Biosystems). The polymorphism *HMIP* rs66650371, which consists in a 3-bp TAC indel in the *HBS1L-MYB* intergenic region, was genotyped with two sets of primers as described in Farrell *et al*. [16] (see **Supplementary Table 1** for details in allelic discrimination).

2.3 Statistical Analysis

Allele frequencies for the different SNPs were calculated by direct counting and the departure from Hardy-Weinberg equilibrium (HWE) was calculated by an exact test. In the HbAA women, the association was tested by logistic regression in additive and dominant models, crude and adjusted for age, splitting participants in two groups: subjects with increased HbF levels *vs*. subjects with normal HbF levels, using 1.5% HbF as cutoff. Among individuals with HbAS, the association between individual SNPs and levels of HbF was explored by linear regression models after logarithmic transformation of HbF values to normalize the quantitative trait distribution. Haplotype association was tested using a case-control approach in HbAA individuals and by linear regression in HbAS individuals. Haplotype associations were tested using the --hap-assoc command in Plink which estimates haplotype frequencies via the expectation-maximization algorithm. All these tests were performed using the software Plink v1.07 (https://zzz.bwh.harvard.edu/plink/download.shtml) [17].

The normality of the data was evaluated using the Shapiro-Wilk and Kolmogorov-Smirnov tests. To compare HbF levels between genotypes in HbAA women, we employed the non-parametric Kruskal-Wallis test or the Mann-Whitney U test, depending on the number of categorical independent groups involved. Graphical analyses and data normality tests were performed with IBM® SPSS® Statistics, version 27 (IBM-SPSS Statistics, Chicago, IL, USA).

3. Results

3.1 HbAA Women

The allelic and genotype frequencies of the six individual SNPs analyzed in the HbAA women are presented in Table 2. The minor allele frequencies (MAF) were: 0.306 and 0.276 for the *BCL11A* SNPs rs11886868 [C] and rs1427407 [T], respectively; 0.036 and 0.046 for the *HMIP* SNPs rs66650371 [3 bp del] and rs4895441 [G], respectively; and 0.143 and 0.184 for the *HBG2* rs7482144 [T] and *BGLT3* rs7924684 [T] SNPs, respectively. All genotype distributions were in accordance with the HWE ($p > 0.05$), with exception for rs1427407 ($p = 0.023$).

Table 2. Allelic and genotype frequencies and association with HbF levels for the six analyzed SNPs in 98 women with a HbAA genotype.

Chr:position (hg38)	SNP ID	Gene	Alleles 1:2	MAF (n = 98)	p-HWE	MAF (11/12/22) HbF <1.5 (n = 60)	MAF (11/12/22) HbF >1.5 (n = 38)	OR (95% CI) *p*-value (*)	OR (95% CI) *p*-value (**)
2:60493111	rs11886868	*BCL11A*	C:T	0.306	0.093	0.192 (0/23/37)	0.487 (13/11/14)	3.53 (1.82–6.83) p = **0.00018**	2.76 (1.19–6.39) p = **0.018**
2:60490908	rs1427407	*BCL11A*	T:G	0.276	0.023	0.175 (0/21/39)	0.434 (12/9/17)	3 (1.58–5.69) p = **0.00076**	2.29 (0.99–5.26) p = 0.05
6:135097495…499	rs66650371	*HMIP*	Del:TAC	0.036	1	0.025 (0/3/57)	0.053 (0/4/34)	2.24 (0.47–10.6) p = 0.311	2.24 (0.47–10.6) p = 0.311
6:135105435	rs4895441	*HMIP*	G:A	0.046	1	0.042 (0/5/55)	0.053 (0/4/34)	1.29 (0.33–5.16) p = 0.715	1.29 (0.33–5.16) p = 0.715
11:5254939	rs7482144	*HBG2*	T:C	0.143	0.405	0.117 (0/14/46)	0.184 (3/8/27)	1.65 (0.75–3.60) p = 0.212	1.34 (0.53–3.36) p = 0.535
11:5245498	rs7924684	*BGLT3*	T:C	0.184	0.305	0.225 (4/19/37)	0.118 (1/7/30)	0.49 (0.22–1.09) p = 0.082	0.43 (0.17–1.09) p = 0.077

Abbreviations: Del, allele with TAC deletion; OR, odds ratio; CI, confidence interval; Alleles, 1-minor, 2-major; MAF, minor allele frequency; Genotypes, 11-homozygous for the minor allele, 12-heterozygous, 22-homozygous for the major allele; n, number of samples; p-HWE, *p*-value for Hardy–Weinberg Equilibrium.

The *p*-value for association was obtained by binary logistic regression under the additive (*) and dominant (**) models. Significant association *p*-values are in bold.

The logistic regression, under the additive model, showed that the minor allele of the two *BCL11A* polymorphisms rs11886868 [C] and rs1427407 [T] present a significant association with increased levels of HbF ($p = 0.00018$ and $p = 0.00076$, respectively). The remaining four polymorphisms showed no significant association ($p > 0.05$) with HbF levels (Table 2). Moreover, logistic regression under the dominant model reveals similar patterns of significance (Table 2).

Consistent with these results, when comparisons of continuous HbF levels were made among the three genotypes using the Kruskal-Wallis test, the homozygous individuals for the minor allele of the two BCL11A SNPs rs11886868 and rs1427407 exhibited significantly higher HbF levels ($p < 0.05$) (Table 3; Fig. 1A,B). Additionally, the comparison of HbF levels between homozygous individuals for the major allele and the grouped heterozygous and homozygous individuals for the minor allele, using the Mann-Whitney U test, also revealed significant differences ($p < 0.05$) (Table 3).

Of note, for the SNP rs7482144 (XmnI), homozygous individuals for the minor allele [T] present higher HbF levels in comparison with heterozygous and homozygous for the major allele [C] (**Supplementary Fig. 1**), with significant differences observed in the Kruskal-Wallis test ($p = 0.031$); however, this significance does not hold for the Mann-Whitney U test comparing homozygous for the major allele vs. heterozygous and homozygous for the minor allele grouped together (Table 3). Otherwise, for the *BGLT3* SNP rs7924684, homozygous individuals for the major allele [C] show higher levels of HbF (**Supplementary Fig. 2**), with a significant difference ($p = 0.032$) in the Mann-Whitney U test when comparing homozygous for the major allele vs. heterozygous and homozygous for the minor allele grouped together (Table 3).

The haplotype association test showed the *BCL11A* rs11886868|rs1427407 [CT] haplotype with the two minor alleles significantly more frequent in individuals with high HbF levels (0.434) in relation to individuals with lower HbF levels (0.158) ($p = 2.02 \times 10^{-5}$). The inverse relationship (0.513 vs. 0.792) was observed for the complementary haplotype [TG], also statistically significant ($p = 4.4 \times 10^{-5}$). No other significant associations were observed for the remaining haplotypes, addressing also the two other sets of SNPs in *HMIP* and *HBB* gene cluster (**Supplementary Table 3**).

Table 3. Comparison of HbF levels between genotypes in 98 women with a HbAA genotype.

Chr:posição (hg38)	Gene	SNP ID	Genotypes (n)	Mean (SD)	Median (IQR)	*p*-value (*)	*p*-value (**)
2:60493111	*BCL11A*	rs11886868	T/T (51)	1.13 (0.77)	0.90 (1.30)	**0.000013**	**0.006**
			T/C (34)	1.27 (0.75)	0.90 (1.13)		
			C/C (13)	2.56 (0.63)	2.30 (1.20)		
2:60490908	*BCL11A*	rs1427407	G/G (56)	1.18 (0.79)	0.90 (1.18)	**0.000016**	**0.01**
			G/T (30)	1.23 (0.70)	0.95 (1.05)		
			T/T (12)	2.62 (0.61)	2.40 (1.00)		
6:135097495	*HMIP*	rs66650371	TAC/TAC (91)	1.35 (0.88)	0.90 (1.30)	N/A	0.336
			TAC/Del (7)	1.64 (0.89)	1.80 (1.30)		
6:135105435	*HMIP*	rs4895441	A/A (89)	1.34 (0.88)	0.90 (1.35)	N/A	0.224
			A/G (9)	1.61 (0.80)	1.40 (1.50)		
11:5254939	*HBG2*	rs7482144	C/C (73)	1.30 (0.87)	1.50 (1.00)	**0.031**	0.132
			C/T (22)	1.39 (0.80)	1.05 (1.23)		
			T/T (3)	2.83 (0.23)	2.70 (-)		
11:5245498	*BGLT3*	rs7924684	C/C (67)	1.50 (0.91)	1.20 (1.40)	0.099	**0.032**
			C/T (26)	1.09 (0.75)	0.90 (1.18)		
			T/T (5)	1.00 (0.60)	0.90 (1.15)		

Abbreviations: Del, allele with TAC deletion; n, number of samples; N/A, not applicable; SD, standard deviation; IQR, interquartile range. The *p*-values were obtained with the Kruskal-Wallis (*) test or the Mann-Whitney U (**) test. In the Mann-Whitney U test for SNPs rs11886868, rs1427407, rs7482144, and rs7924684, heterozygous and homozygous genotypes for the minor allele were grouped. Significant *p*-values (<0.05) are in bold.

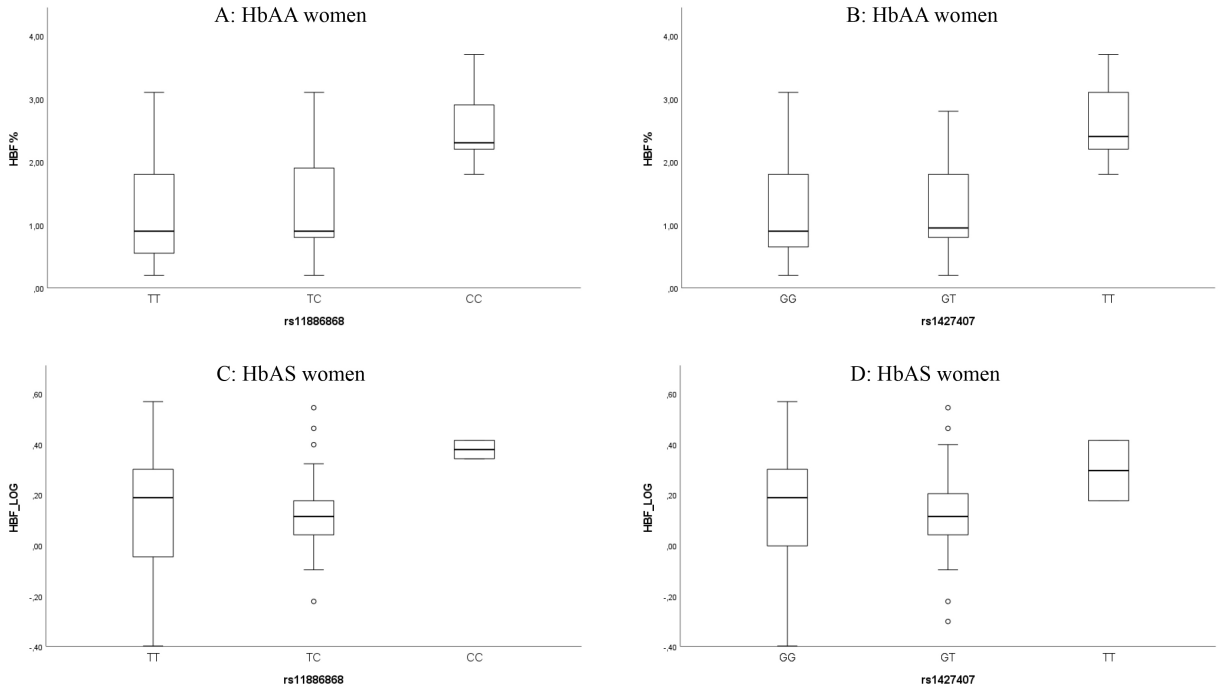

Fig. 1. Box plots showing the distribution of HbF levels within genotypes of the two *BCL11A* SNPs rs11886868 and rs1427407 in HbAA women (A,B), and in women with HbAS after logarithmic transformation, base 10 (C,D). Each rectangle represents the data between the 25th and 75th quartiles, and the bar within each rectangle is the median value for HbF.

3.2 HbAS Women

The genetic and association data of the six individual SNPs analyzed in women with HbAS were displayed in Table 4. The MAF for the analyzed polymorphisms were: 0.309 for *BCL11A* rs11886868 [C] and rs1427407 [T], 0.032 and 0.064 for *HMIP* rs66650371 [3 bp del] and

Table 4. Allelic and genotype frequencies and association results with HbF levels for the six analyzed SNPs in 47 women with HbAS.

Chr:position (hg38)	SNP ID	Gene	Alleles (1:2)	Genotypes (11/12/22)	MAF	p-HWE	β (SE) p-value (*)	β (SE) p-value (**)
2:60493111	rs11886868	*BCL11A*	C:T	2/25/20	0.309	0.168	0.045 (0.057) $p = 0.433$	0.018 (0.066) $p = 0.789$
2:60490908	rs1427407		T:G	2/25/20	0.309	0.168	0.009 (0.057) $p = 0.879$	–0.015 (0.066) $p = 0.815$
6:135097495-135097499	rs66650371	*HMIP*	Del:TAC	0/3/44	0.032	1	0.003 (0.133) $p = 0.983$	0.003 (0.133) $p = 0.983$
6:135105435	rs4895441		G:A	0/6/41	0.064	1	–0.023 (0.097) $p = 0.815$	–0.023 (0.097) $p = 0.815$
11:5254939	rs7482144	*HBG2*	T:C	0/9/38	0.096	1	–0.113 (0.081) $p = 0.167$	–0.113 (0.081) $p = 0.167$
11:5245498	rs7924684	*BGLT3*	T:C	0/13/34	0.138	0.574	–0.079 (0.072) $p = 0.278$	–0.079 (0.072) $p = 0.278$

Abbreviations: Del, allele with TAC deletion; Chr, Chromosome; MAF, Minor allele frequency; Alleles, 1-minor, 2-major; Genotypes, 11-homozygous for the minor allele, 12-heterozygous, 22-homozygous for the major allele; p-HWE, p-value for Hardy-Weinberg Equilibrium; β, regression coefficient beta; SE, standard error.

Linear regression values were obtained under the additive (*) and dominant (**) models.

rs4895441 [G], respectively; and 0.096 and 0.138 for *HBG2* rs7482144 (XmnI) [T] and *BGLT3* rs7924684 [T], respectively. All genotype distributions were in accordance with the HWE ($p > 0.05$).

Using HbF as a continuous variable, the simple linear regression either in additive, dominant, or recessive (data not shown) models, showed no statistically significant associations between the six analyzed polymorphisms and HbF levels ($p > 0.05$) in the group of HbAS individuals (Table 4). However, it should be noted that the two homozygous individuals for the minor allele in both SNPs present higher HbF levels in comparison with heterozygous and homozygous for the major allele (Fig. 1C,D). In the haplotype-based analysis, no statistically significant associations with HbF levels were observed for any of the haplotypes combining SNPs in the same chromosome (**Supplementary Table 4**).

4. Discussion

Characterizing genetic variants associated with HbF regulation is a paramount topic because it has been established that the induction of higher levels of HbF by genome-editing strategies or pharmacological inducer agents can improve the clinical and hematological features of severe hemoglobinopathies such as SCD and β-thalassemia [6].

In the present study, we replicated in a population sample of HbAA women from São Tomé e Príncipe the significant associations with HbF levels of the two well-known *BCL11A* variants rs11886868 T>C and rs1427407 G>T. The study of a population group of 98 HbAA women in a case-control framework using logistic regression revealed that the minor allele frequencies of the two *BCL11A* polymorphisms rs11886868 [C] and rs1427407 [T] were significantly higher in the group with higher levels of HbF. In concordance, homozygous individuals for the minor allele present significantly higher HbF levels compared with heterozygous and homozygous for the major allele. Despite the higher levels of HbF in two homozygous individuals for the minor allele in the HbAS women, no significant association was found between the minor allele of the two *BCL11A* SNPs and high HbF levels using linear regression in the additive, dominant, or recessive models.

SNPs in the intronic *BCL11A* enhancer were implicated in HbF regulation by many genome-wide and candidate association studies, addressing mainly SCD and β-thalassemia patients [18–20], but also β-thalassemia carriers [16,21,22] or even normal healthy Caucasian individuals with common forms of hereditary persistence of fetal hemoglobin (HPFH) [21,23]. The transcription factor BCL11A is a major repressor of γ-globin gene expression and switching from fetal to adult hemoglobin in erythroid cells of mammals by directly binding the fetal globin (*HBG1/2*) gene promoters and other sites in the β-globin gene cluster [24,25]. Increased levels of HbF are present with the downregulation of BCL11A in adult erythroblasts and cell lines expressing adult hemoglobin [8]. Moreover, the BCL11A binding site disruption in the γ-globin promoters increases γ-globin gene expression and HbF production [3]. Because of these results, the down-regulation or blockage of BCL11A has been a guide for new targeted HbF induction therapies [7,26].

The *HBG2*-XmnI polymorphism (–158C>T) (rs7482144) in HbAA women, showed significantly higher values of HbF in the homozygous genotypes for the minor

allele T in comparison with heterozygous and homozygous for the major allele. However, as only three homozygous TT individuals were found in this group, this result should be interpreted with caution. No signals of association between rs7482144 and HbF was observed in women with HbAS genotypes. It is well established that the *HBG2*-XmnI polymorphism is strongly associated with increased levels of HbF and can improve the phenotypes in response to stress erythropoiesis in conditions like β-thalassemia and sickle cell anemia [5]. The effect on HbF levels linked to the XmnI polymorphism is minimal or undetectable in healthy adults of Caucasian ancestry [21,23]. Consequently, the significant or near-significant association of the XmnI minor allele [T] with elevated HbF levels observed in HbAA individuals in São Tomé and Príncipe may indicate some form of erythropoietic stress.

In the association tests for the *BGLT3* SNP rs7924684 C>T, significant and near-significant results were found linking the major [C] allele to higher levels of HbF in women with the HbAA genotype. These findings align with previous studies that reported significant associations between the ancestral [C] allele of SNP rs7924684 and elevated levels of HbF among Portuguese β-thalassemia carriers [27] and Angolan children with SCD [28]. The *BGLT3* gene, which encodes a long non-coding RNA (lncRNA), serves as a positive regulator of γ-globin genes expression [29]. This regulation may occur through various mechanisms, including looping to the γ-globin genes and recruiting coregulators via the *BGLT3* transcript [29].

For the two remaining SNPs here analyzed, rs66650371 and rs4895441, located in the *HMIP* intergenic region, no significant associations were observed with HbF levels in the HbAA or HbAS groups. Several SNPs linked to increased HbF levels have been identified in this genomic region on chromosome 6q23, particularly within the contexts of β-thalassemia and SCD [5,30]. Variants in these loci have been reported to contribute to 20–50% of the variability in HbF levels among non-African populations; however, the impact of these variants varies from one population to another. As the *HMIP* variants, such as rs66650371 and rs4895441, are less frequent in Sub-Saharan African populations, as was also observed herein for the study sample of São Tomé e Príncipe showing lower MAFs of 0.036 and 0.046, respectively, in HbAA women, and 0.032 and 0.064, respectively, in the HbAS women, this can explain the limited impact on HbF levels in that populations.

5. Conclusions

In conclusion, we successfully replicated in HbAA women from São Tomé e Príncipe (Central Africa), the known associations with HbF variation of the two *BCL11A* variants rs11886868 and rs1427407. No evidence of association was found between levels of HbF and the two *HMIP* variants rs66650371 and rs4895441, which can be explained by the lower MAF from these variants in the study population. For the variants located at the *HBB* gene cluster, the *BGLT3* SNP rs7924684 C>T showed signs of an association between the major C allele and higher levels of HbF in women with the HbAA genotype. A signal of association with increased levels of HbF was observed for the minor allele of the XmnI polymorphism rs7482144 in HbAA women. In the group of women with HbAS, no significant associations were found between HbF levels and the six analyzed SNPs in *BCL11A*, *HMIP*, and *HBB* gene cluster. This lack of association may be attributed, in part, to the small sample size of only 47 individuals, in particular given the low frequency of homozygous individuals for the minor alleles of the two *BCL11A* variants rs11886868 and rs1427407.

Availability of Data and Materials

The authors confirm that the findings of this study are available within the article and its supplementary materials. The data supporting the findings of this study are available from the first author upon reasonable request.

Author Contributions

CB, GQ and CM collected samples, performed blood analyses, and interpreted the hematological data. LR performed laboratory blood analysis. AMM, SMA and IS performed molecular and statistical analysis. LM designed the research study, interpreted the molecular and statistical data, wrote, and edited the manuscript. All authors read and approved the final manuscript. All authors contributed to editorial changes in the manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee of São Tomé e Príncipe, identified as CESIC Case PC022_2022. All participants signed an informed consent form before participating in the study. All data were anonymized with respect to confidentiality and processed according to the ethical principles outlined in the Declaration of Helsinki.

Acknowledgment

Not applicable.

Funding

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under the institutional grant UIDB/00283/2020 and by Forum Hematologico (CHUC).

Conflict of Interest

The authors declare no conflict of interest. Given his role as the Editorial Board member, Licínio Manco had no

involvement in the peer-review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Gustavo Caetano-Anollés.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/FBS38388.

References

[1] Sankaran VG, Orkin SH. The switch from fetal to adult hemoglobin. Cold Spring Harbor Perspectives in Medicine. 2013; 3: a011643. https://doi.org/10.1101/cshperspect.a011643.

[2] Thein SL, Craig JE. Genetics of Hb F/F cell variance in adults and heterocellular hereditary persistence of fetal hemoglobin. Hemoglobin. 1998; 22: 401–414. https://doi.org/10.3109/03630269809071538.

[3] Steinberg MH. Fetal Hemoglobin in Sickle Hemoglobinopathies: High HbF Genotypes and Phenotypes. Journal of Clinical Medicine. 2020; 9: 3782. https://doi.org/10.3390/jcm9113782.

[4] Akinsheye I, Alsultan A, Solovieff N, Ngo D, Baldwin CT, Sebastiani P, *et al*. Fetal hemoglobin in sickle cell anemia. Blood. 2011; 118: 19–27. https://doi.org/10.1182/blood-2011-03-325258.

[5] Thein SL. Genetic Basis and Genetic Modifiers of β-Thalassemia and Sickle Cell Disease. Advances in Experimental Medicine and Biology. 2017; 1013: 27–57. https://doi.org/10.1007/978-1-4939-7299-9_2.

[6] Bou-Fakhredin R, De Franceschi L, Motta I, Cappellini MD, Taher AT. Pharmacological Induction of Fetal Hemoglobin in β-Thalassemia and Sickle Cell Disease: An Updated Perspective. Pharmaceuticals (Basel, Switzerland). 2022; 15: 753. https://doi.org/10.3390/ph15060753.

[7] Gambari R, Waziri AD, Goonasekera H, Peprah E. Pharmacogenomics of Drugs Used in β-Thalassemia and Sickle-Cell Disease: From Basic Research to Clinical Applications. International Journal of Molecular Sciences. 2024; 25: 4263. https://doi.org/10.3390/ijms25084263.

[8] Steinberg MH. Fetal hemoglobin in sickle cell anemia. Blood. 2020; 136: 2392–2400. https://doi.org/10.1182/blood.2020007645.

[9] Menzel S, Thein SL. Genetic Modifiers of Fetal Haemoglobin in Sickle Cell Disease. Molecular Diagnosis & Therapy. 2019; 23: 235–244. https://doi.org/10.1007/s40291-018-0370-8.

[10] Stadhouders R, Aktuna S, Thongjuea S, Aghajanirefah A, Pourfarzad F, van Ijcken W, *et al*. HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. The Journal of Clinical Investigation. 2014; 124: 1699–1710. https://doi.org/10.1172/JCI71520.

[11] Nkya S, Mwita L, Mgaya J, Kumburu H, van Zwetselaar M, Menzel S, *et al*. Identifying genetic variants and pathways associated with extreme levels of fetal hemoglobin in sickle cell disease in Tanzania. BMC Medical Genetics. 2020; 21: 125. https://doi.org/10.1186/s12881-020-01059-1.

[12] Siatecka M, Bieker JJ. The multifunctional role of EKLF/KLF1 during erythropoiesis. Blood. 2011; 118: 2044–2054. https://doi.org/10.1182/blood-2011-03-331371.

[13] Kirkham JK, Estepp JH, Weiss MJ, Rashkin SR. Genetic Variation and Sickle Cell Disease Severity: A Systematic Review and Meta-Analysis. JAMA Network Open. 2023; 6: e2337484. https://doi.org/10.1001/jamanetworkopen.2023.37484.

[14] Queiroz G, Monteiro C, Manco L, Relvas L, Trovoada MDJ, Leite A, *et al*. Sickle cell trait in São Tomé e Príncipe: a population-based prevalence study in women of reproductive age. BMC Public Health. 2024; 24: 850. https://doi.org/10.1186/s12889-024-17761-1.

[15] Sutton M, Bouhassira EE, Nagel RL. Polymerase chain reaction amplification applied to the determination of beta-like globin gene cluster haplotypes. American Journal of Hematology. 1989; 32: 66–69. https://doi.org/10.1002/ajh.2830320113.

[16] Farrell JJ, Sherva RM, Chen ZY, Luo HY, Chu BF, Ha SY, *et al*. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood. 2011; 117: 4935–4945. https://doi.org/10.1182/blood-2010-11-317081.

[17] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics. 2007; 81: 559–575. https://doi.org/10.1086/519795.

[18] Uda M, Galanello R, Sanna S, Lettre G, Sankaran VG, Chen W, *et al*. Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105: 1620–1625. https://doi.org/10.1073/pnas.0711566105.

[19] Galanello R, Sanna S, Perseu L, Sollaino MC, Satta S, Lai ME, *et al*. Amelioration of Sardinian beta0 thalassemia by genetic modifiers. Blood. 2009; 114: 3935–3937. https://doi.org/10.1182/blood-2009-04-217901.

[20] 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 (Amsterdam, Netherlands). 2016; 21: 425–429. https://doi.org/10.1080/10245332.2015.1107275.

[21] Pereira C, Relvas L, Bento C, Abade A, Ribeiro ML, Manco L. Polymorphic variations influencing fetal hemoglobin levels: association study in beta-thalassemia carriers and in normal individuals of Portuguese origin. Blood Cells, Molecules & Diseases. 2015; 54: 315–320. https://doi.org/10.1016/j.bcmd.2015.02.001.

[22] Chan NCN, Lau KM, Cheng KCK, Chan NPH, Ng MHL. A Multi-locus Approach to Characterization of Major Quantitative Trait Loci Influencing Hb F Regulation in Chinese β-thalassemia Carriers. Hemoglobin. 2016; 40: 400–404. https://doi.org/10.1080/03630269.2016.1245198.

[23] Guida V, Cappabianca MP, Colosimo A, Rafanelli F, Amato A, Dallapiccola B. Influence of Ggamma-158C –> and beta-(AT)x(T)y globin gene polymorphisms on HbF levels in Italian beta-thalassemia carriers and wild-type subjects. Haematologica. 2006; 91: 1275–1276.

[24] Sankaran VG, Xu J, Byron R, Greisman HA, Fisher C, Weatherall DJ, *et al*. A functional element necessary for fetal hemoglobin silencing. The New England Journal of Medicine. 2011; 365: 807–814. https://doi.org/10.1056/NEJMoa1103070.

[25] Huang P, Keller CA, Giardine B, Grevet JD, Davies JOJ, Hughes JR, *et al*. Comparative analysis of three-dimensional chromosomal architecture identifies a novel fetal hemoglobin regulatory element. Genes & Development. 2017; 31: 1704–1713. https://doi.org/10.1101/gad.303461.117.

[26] Lessard S, Rimmelé P, Ling H, Moran K, Vieira B, Lin YD, *et al*. Zinc finger nuclease-mediated gene editing in hematopoietic stem cells results in reactivation of fetal hemoglobin in sickle cell disease. Scientific Reports. 2024; 14: 24298. https://doi.org/10.1038/s41598-024-74716-7.

[27] Manco L, Santos R, Rocha C, Relvas L, Bento C, Maia T, *et al*.

Hb F Levels in β-Thalassemia Carriers and Normal Individuals: Known and Unknown Quantitative Trait Loci in the β-Globin Gene Cluster. Hemoglobin. 2022; 46: 168–175. https://doi.org/10.1080/03630269.2022.2070498.

[28] Delgadinho M, Ginete C, Santos B, Miranda A, Brito M. Genotypic Diversity among Angolan Children with Sickle Cell Anemia. International Journal of Environmental Research and Public Health. 2021; 18: 5417. https://doi.org/10.3390/ijerph18105417.

[29] Ivaldi MS, Diaz LF, Chakalova L, Lee J, Krivega I, Dean A. Fetal γ-globin genes are regulated by the *BGLT3* long noncoding RNA locus. Blood. 2018; 132: 1963–1973. https://doi.org/10.1182/blood-2018-07-862003.

[30] Mohammad SNNA, Iberahim S, Wan Ab Rahman WS, Hassan MN, Edinur HA, Azlan M, *et al*. Single Nucleotide Polymorphisms in *XMN1-HBG2*, *HBS1L-MYB*, and *BCL11A* and Their Relation to High Fetal Hemoglobin Levels That Alleviate Anemia. Diagnostics (Basel, Switzerland). 2022; 12: 1374. https://doi.org/10.3390/diagnostics12061374.