RESEARCH ARTICLE

A candidate transacting modulator of fetal hemoglobin gene expression in the Arab—Indian haplotype of sickle cell anemia

Vinod Vathipadiekal,[1] John J. Farrell,[1] Shuai Wang,[2] Heather L. Edward,[1] Heather Shappell,[2] A.M. Al-Rubaish,[3] Fahad Al-Muhanna,[3] Z. Naserullah,[4,5] A. Alsuliman,[6] Hatem Othman Qutub,[7] Irene Simkin,[1] Lindsay A. Farrer,[1] Zhihua Jiang,[1] Hong-Yuan Luo,[1] Shengwen Huang,[1] Gustavo Mostoslavsky,[1] George J. Murphy,[1] Pradeep K. Patra,[8] David H.K. Chui,[1] Abdulrahman Alsultan,[9] Amein K. Al-Ali,[10] Paola Sebastiani,[2] and Martin H. Steinberg[1]*

Fetal hemoglobin (HbF) levels are higher in the Arab–Indian (AI) β-globin gene haplotype of sickle cell anemia compared with African-origin haplotypes. To study genetic elements that effect HbF expression in the AI haplotype we completed whole genome sequencing in 14 Saudi AI haplotype sickle hemoglobin homozygotes—seven selected for low HbF (8.2% ± 1.3%) and seven selected for high HbF (23.5% ± 2.6%). An intronic single nucleotide polymorphism (SNP) in *ANTXR1,* an anthrax toxin receptor (chromosome 2p13), was associated with HbF. These results were replicated in two independent Saudi AI haplotype cohorts of 120 and 139 patients, but not in 76 Saudi Benin haplotype, 894 African origin haplotype and 44 AI haplotype patients of Indian origin, suggesting that this association is effective only in the Saudi AI haplotype background. *ANTXR1* variants explained 10% of the HbF variability compared with 8% for *BCL11A*. These two genes had independent, additive effects on HbF and together explained about 15% of HbF variability in Saudi AI sickle cell anemia patients. *ANTXR1* was expressed at mRNA and protein levels in erythroid progenitors derived from induced pluripotent stem cells (iPSCs) and $CD34^+$ cells. As $CD34^+$ cells matured and their HbF decreased *ANTXR1* expression increased; as iPSCs differentiated and their HbF increased, *ANTXR1* expression decreased. Along with elements in cis to the HbF genes, *ANTXR1* contributes to the variation in HbF in Saudi AI haplotype sickle cell anemia and is the first gene in trans to *HBB* that is associated with HbF only in carriers of the Saudi AI haplotype.
Am. J. Hematol. 91:1118–1122, 2016.

Introduction

Fetal hemoglobin (HbF) concentrations in sickle cell anemia are heritable, variable, and genetically controlled (reviewed in [1–4]). The HbF genes, *HBG2* and *HBG1*, are located within the *HBB* gene cluster (11p15.5). HbF levels vary according to the *HBB* gene haplotype suggesting an important element of *cis*-acting control over HbF gene expression [5]. Nevertheless, patients with the same *HBB* haplotype can have different HbF concentrations implying that trans-acting quantitative trait loci (QTL) also modulate HbF gene expression. Common variants in the trans-acting loci *BCL11A* (2p16) and *MYB*, and *cis*-acting elements within the *HBB* gene cluster explained 8–50% of HbF variation in patients with sickle cell anemia depending on the population examined [6–11]. Although a unique *cis*-acting subhaplotype containing putative functional elements was exclusive to the AI haplotype and postulated to have a major effect on expression of their HbF genes, especially *HBG2*, as with all other haplotypes these patients can have widely divergent HbF levels [9,12,13].

To search for novel variants that might modulate HbF expression in AI haplotype sickle cell anemia we examined by whole genome sequencing a highly selected group of these patients who had different HbF levels. Two single nucleotide polymorphisms (SNPs) in an anthrax toxin receptor gene, *ANTXR1,* appeared strongly associated with HbF. This observation was replicated in two independent cohorts of Saudi sickle cell anemia

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

[1]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts; [2]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts; [3]Department of Internal Medicine, College of Medicine, University of Dammam, Dammam, Kingdom of Saudi Arabia; [4]Al-Omran Scientific Chair for Hematological Diseases, King Faisal University, Al-Ahsa, Kingdom of Saudi Arabia; [5]Department of Pediatrics, Maternity and Child Hospital, Dammam, Kingdom of Saudi Arabia; [6]Alomran Scientific Chair, King Faisal University, King Fahd Hospital, Hafof, Al-Ahsa, Kingdom of Saudi Arabia; [7]Alomran Scientific Chair, King Faisal University, Al-Ahsa, Kingdom of Saudi Arabia; [8]Department of Biochemistry, Pt. J. N. M. Medical College, Raipur, Chattisgarh, India; [9]Sickle Cell Disease Research Center and Department of Pediatrics, College of Medicine, King Saud University, Riyadh, Saudi Arabia; [10]Center for Research and Medical Consultation, University of Dammam, Dammam, Kingdom of Saudi Arabia

Conflict of interest: Nothing to report.
***Correspondence to:** Martin H. Steinberg, MD, Professor of Medicine, Pediatrics, Pathology and Laboratory Medicine, Boston University School of Medicine, 72 E. Concord St. Boston, MA 02118. Tel.: 617-414-1020. Fax: 617-414-1021. E-mail: mhsteinb@bu.edu
Contract grant sponsor: University of Dammam; Contract grant number: SP 11/2011.
Contract grant sponsor: NIH; Contract grant numbers: U01 HL107443; R01 HL 87681; T32 HL007501; T32 GM74905; KL2RR025770.
Contract grant sponsor: Office of Collaboration and Knowledge Exchange, University of Dammam.
Received for publication: 27 July 2016; **Revised:** 2 August 2016; **Accepted:** 3 August 2016
Am. J. Hematol. 91:1118–1122, 2016.
Published online: 7 August 2016 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/ajh.24527

TABLE I. Patients Selected for Whole Genome Sequencing and in Follow-up Patient Cohorts

	N (cases)	Age (yrs.); SD	HbF (%); SD
Whole genome sequencing			
Saudi AI Haplotype			
(low HbF)	7	34.1; 10.3	8.2; 1.3
(high HbF)	7	25.9; 6.8	23.5; 2.6
TaqMan assays			
CSSCD	894	13.6; 11.3	5.2; 5.6
Saudi AI Haplotype Cohort 1	120	26.7; 10.1	18.0; 7.0
Saudi AI Haplotype Cohort 2	139	28.9; 12.6	16.2; 7.5
Indian AI Haplotype	44	14.6; 4.7	23.0; 4.8
Saudi Benin Haplotype	76	15.9; 9.9	11.4; 5.7

Forty-seven percent of African American patients (CSSCD) were Benin haplotype homozygotes, 27% Benin/Bantu compound heterozygotes, 10% Benin/Senegal compound heterozygote and 16% had other haplotypes including 8 Senegal haplotype homozygotes with HbF levels of 9.2, 2.3, 4.6, 16.0, 16.3, 11, 5.9, and 3.4% and ages of 24, 30, 16, 6, 24, 27, 44, and 23 years, respectively.

patients with the AI haplotype totaling 259 individuals but not in 76 Saudi Benin haplotype, 894 African origin *HBB* haplotype and 44 AI haplotype sickle cell anemia patients of Indian origin. The association of this locus explained about 10% of the HbF variance and was independent of the effect of *BCL11A*. *ANTXR1* was expressed in erythroid progenitor cells and, like *BCL11A* and *MYB*, appeared to be a suppressor of HbF gene expression [14–16].

Methods

Patients, haplotypes, and HbF. Fourteen unrelated patients with sickle cell anemia (homozygosity for rs334, GAG-GTG, E6V) from the indigenous population of the Al-Qatif and Al-Hasa regions along the Western shore of the Persian Gulf in the Eastern Province of the Kingdom of Saudi Arabia were selected for whole genome sequencing. These patients were homozygous for the AI haplotype.

To validate and replicate the results of whole genome sequencing in the Saudi AI haplotype we studied two independent cohorts of similar Saudi sickle cell anemia patients with the AI haplotype. Cohort 1 consisted of 120 Saudi AI haplotype patients; ascertained before 2014 and Saudi Cohort 2 had 139 patients ascertained between 2014 and 2015. Both cohorts were recruited from the Al-Qatif and Al-Hasa regions in the Eastern Province of the Kingdom of Saudi Arabia. We also studied 76 Saudi Benin haplotype patients, 44 Indian AI haplotype patients and 894 African American patients from the Cooperative Study of Sickle Cell Disease (CSSCD) with diverse *HBB* haplotypes of African origin (Table I). Ascertainment of the AI haplotype and HbS homozygosity was previously described and the *HBB* haplotypes associated with the HbS gene is shown in Supporting Information Fig. S1 [9]. These studies were approved by the IRB of all participating institutions.

HbF was measured in Saudi Arabia, India and Boston using high performance liquid chromatography. HbF in the CSSCD and in the patient homozygous for the Senegal haplotype studied by whole genome sequencing was measured by alkali denaturation. Patients were studied after age 5 years when HbF levels had stabilized and they were not prescribed hydroxyurea.

Whole genome sequencing. DNA was randomly sheared to fragments with a median size range of 300 bp. The extracted DNA was amplified by ligation-mediated PCR and paired-end 100 bp reads were sequenced at 40× coverage on the Illumina HiSeq 2000 platform. High-coverage whole genome reads were aligned to the Human Reference (GRCh37) using the Burrows-Wheeler Aligner version 0.5.9 which achieved an average alignment of 97%. The genome analysis toolkit (GATK) was used to provide duplicate removal, base quality score recalibration, and sample level local realignment around indels, genotyping, and variant recalibration. Further details are given in the Supporting Information Methods. The scheme used to filter these samples is shown in Supporting Information Fig. S2.

Genotyping and imputation. One-hundred twenty-eight SNPs associated with HbF in the whole genome sequencing analysis were genotyped in the Saudi AI Cohort 1, in the Indian AI haplotype, and in African American sickle cell anemia cases using the KASP reagent (Supporting Information Methods). Only rs4527238 was genotyped in the 139 cases of Saudi AI Cohort 2. For additional analysis, 76 Saudi Benin haplotype patients and 894 African Americans were imputed to the 1000 Genomes reference panel using IMPUTE2. A subset of SNPs was further validated by Sanger sequencing.

Functional annotation of candidate variants. Variants in *ANTRX1* were annotated using the GTEX portal (http://gtexportal.orgi) [17] and RegulomeDB (http://www.regulomedb.org/). To identify SNPs associated with regulatory protein or transcription factors we used SCAN (http://www.scandb.org/newinterface/about.html), a large-scale genomic database for functional and physical-based SNP annotation [18,19], Hembase (http://hembase.niddk.nih.gov), a database for human erythroid gene activity [20] and ErythronDB (http://www.cbil.upenn.edu/ErythronDB/) that provides transcript level information on expression patterns of genes involved in the mammalian erythropoiesis [21]. Splice-site analysis for the *ANTXR1* intronic SNPs was conducted using the Human Splice Finder. STRING 9.1 (http://string-db.org) was used to model predicted protein interactions of ANTXR1 [22].

Induced pluripotent stem cell (iPSC) and CD34$^+$ cell culture. iPSCs were reprogrammed from lymphocytes from the blood of two normal donors and CD34$^+$ cells were isolated from the blood of these same individual for studies using RNA sequencing. iPSCs cells from blood of patients with sickle cell anemia [23] were also reprogramed from peripheral blood lymphocytes and driven toward erythroid differentiation. Both iPSC and CD34$^+$ cells were differentiated into erythroid lineage cells as previously described [24,25], (Supporting Information Methods).

Gene expression. RNA isolated from iPSCs and CD34$^+$ cells was sequenced using the Illumina HiSeq-2000 platform. Library preparation, RNA sequencing and data analysis was as described [26,27] (Supporting Information Methods). For normalizing read depth and length, data were expressed as read counts per million (RPKM); raw counts are also displayed to show expression of protein coding splice variants.

Immunofluorescence studies. Sickle iPSCs at days 15, 20, and 26 of erythroid differentiation were fixed with methanol and incubated in 1% BSA-0.1% Tween20 phosphate buffer for 1 hr to permeabilize cells and block non-specific binding. Cells were incubated with antibodies to ANTXR1 (1:50 dilution) and HbF-FITC for 2 hr at room temperature. After washing, the 2nd antibody Rhodamine-conjugated AffinityPure F (ab') 2 Fragment Donkey anti-Rabbit IgG (H + L) was used at 1:200 dilution for 0.5 hr. Unstained controls without the 1st antibody added were carried out simultaneously in every immunostaining experiment for each cell type at different times during the cultures and all were negative (data not shown). As ANTXR1 showed specific expression patterns in the cultured CD34$^+$ and iPSCs, isotype controls were not done. Prior to immunostaining the 2nd antibody was optimized as to concentration, incubation time, whole antibody vs. *F*(ab')2 fragment and antibody competitions when costaining with anti-HbF Slides were mounted in DAPI medium for and examined under a fluorescent microscope.

Statistical analysis. Association of SNPs with HbF in the whole genome sequencing experiment was tested using simple linear regression with an additive genetic model for common SNPs (MAF > 0.05), and the Sequence Kernel Association Test (SKAT) [28] using sliding windows of 5 SNPs for rare SNPs. PLINK [29] was employed to remove all SNPs in high linkage disequilibrium (LD; r2 >0.80). The association of the 128 most significant SNPs with HbF was tested using data imputed to the 1000 Genomes reference panel in the African American cohort and genotyped data in the other cohorts. Cubic root transformation of HbF was used as in [30] and a regression model of the allele dosages was used, adjusted for age and sex. The data from 76 Saudi Benin were analyzed using a linear mixed effect model to adjust for relatedness as in [31]. To determine the variation in HbF jointly explained by *ANTXR1* and *BCL11A* in Saudi AI haplotype sickle cell anemia and their additive effects, these three regression models were fitted: (1) HbF with rs1427407 in *BCL11A*, (2) HbF with rs4527238 in *ANTXR1*, and (3) HbF with both rs4527238 and rs1427407 (all models adjusted by age and sex). All analyses were conducted using the statistical program R v.3.2.

Results

Patients

The *HBB* haplotype, number of cases in each patient cohort, age and HbF levels are shown in Table I. AI haplotype carriers had higher HbF levels than other haplotypes.

ANTXR1, a candidate gene for HbF regulation in Saudi AI haplotype sickle cell anemia

Single SNP analysis or SKAT-based SNP-set analysis of the 14 Saudi AI haplotype HbS homozygotes subjected to whole genome sequencing did not identify any SNPs with genome-wide significance (P value < 10E-8). We identified 465 SNPs with $P \leq 10^{-4}$ for association with HbF that we investigated further (Supporting Information Table SI). One-hundred twenty of these variants were selected for follow up analysis in the replication cohorts based on lack of pairwise correlation, and 99 of these SNPs were successfully genotyped (Supporting Information Table SII). Following analysis of these SNPs, rs4527238 and rs35685045 ($D' = 1$) in intron 9 of *ANTXR1* (anthrax

toxin receptor 1, 2p13.1) remained associated with HbF after correction for multiple comparisons ($P = 0.00011$, 3.34E-05, Table II). As these SNPs were in perfect linkage disequilibrium (LD) the remaining studies focus on rs4527238.

Figure 1 shows the distribution of HbF by genotype of rs4527238 in 120 Saudi AI haplotype patients (Cohort 1), an additional independent set of 139 Saudi AI haplotype patients (Cohort 2), 76 Saudi Benin haplotype patients, 44 Indian AI haplotype HbS homozygotes, and 894 African Americans. The C allele of rs4527238 was associated with lower HbF only in the Saudi AI sickle cell patients, but not in the Indian AI, Saudi Benin, or African American patient cohorts. There was no LD between SNPs in *ANTXR1* and 6 SNPs in *BCL11A* including the sentinel SNP rs1427407 (D' 0.006–0.152) in Saudi Cohort 1. *BCL11A* variants were not genotyped in Saudi Cohort 2.

In joint analysis, both rs4527238 (*ANTXR1*) and rs1427407 (*BCL11A*) were significantly associated with HbF in AI haplotype patients. The regression model using rs1427407 in *BCL11A* explained about 8% of HbF variability. The regression model of HbF with rs4527238 in *ANTXR1* explained about 10% of HbF variability. A model including both *BCL11A* and *ANTXR1* SNPs explained about 15% of HbF variability. *BCL11A* rs1427407 and *ANTXR1* rs4527238 have independent, additive effects on HbF as shown by the lack of difference between the SNP effect when we compared the unadjusted and adjusted model, and the genetic effect of rs4527238 (2.60) is slightly larger than that of rs1427407 (2.23) (Supporting Information Table SIII).

TABLE II. Association of rs4527238 with HbF in AI Haplotype Saudi Sickle Cell Anemia and Other Haplotypes

	Coded allele	Coded allele frequency	Beta (SE)	*P* value
Saudi AI (Cohort 1)	T	0.52	3.23 (0.84)	0.00011
Saudi AI (Cohort 2)	T	0.52	1.54 (0.87)	0.0795
Saudi Benin	T	0.66	−1.42 (1.03)	0.16819
Indian AI	T	0.6	0.92 (0.95)	0.33435
CSSCD	T	0.66	0.03 (0.07)	0.72168

There was >95% power to detect a genetic effect if present. CSSCD—Cooperative Study of Sickle Cell Disease, an observational study of African American sickle cell disease.

ANTXR1 is conserved in vertebrates, expressed in iPSC and $CD34^+$ derived erythroid progenitors and its protein levels vary inversely with HbF expression. Functional annotation using the GTEX portal showed that rs4527238 is a significant eQTL for *ANTXR1* in whole blood, and *ANTXR1* expression increased with each copy of the its C allele of rs4527238 ($P = 0.04$; Supporting Information Fig. S3) that was associated with significantly lower level of HbF. SNP rs4527238 had no effect on the expression of *BCL11A* ($P = 0.3$; data not shown).

ANTXR1 is widely expressed with highest levels in bone marrow (Supporting Information Fig. S4). Rs35685045, which is in perfect LD with rs4527238, is conserved in vertebrates [32] and is located on a promoter-flanking region with possible regulatory feature (Ensemble Regulation (ENSR00001543322, Chr2: 69096821-69100420). RegulomeDB analysis predicted binding sites for TBP and MEF-2. Further, this SNP is located within the 5 kb region upstream of miR-3126 (Chr2: 69,103,682-69,103,755).

Expression of *ANTXR1* as RPKM and its splice variants in iPSCs and $CD34^+$ cells are shown in Supporting Information Fig. S5. In $CD34^+$ and iPSCs isolated from two different normal subjects, expression increased over time in $CD34^+$ erythroid cells and decreased over time in iPSC-derived erythroid progenitors; only the long splice variant of *ANTXR1* was expressed.

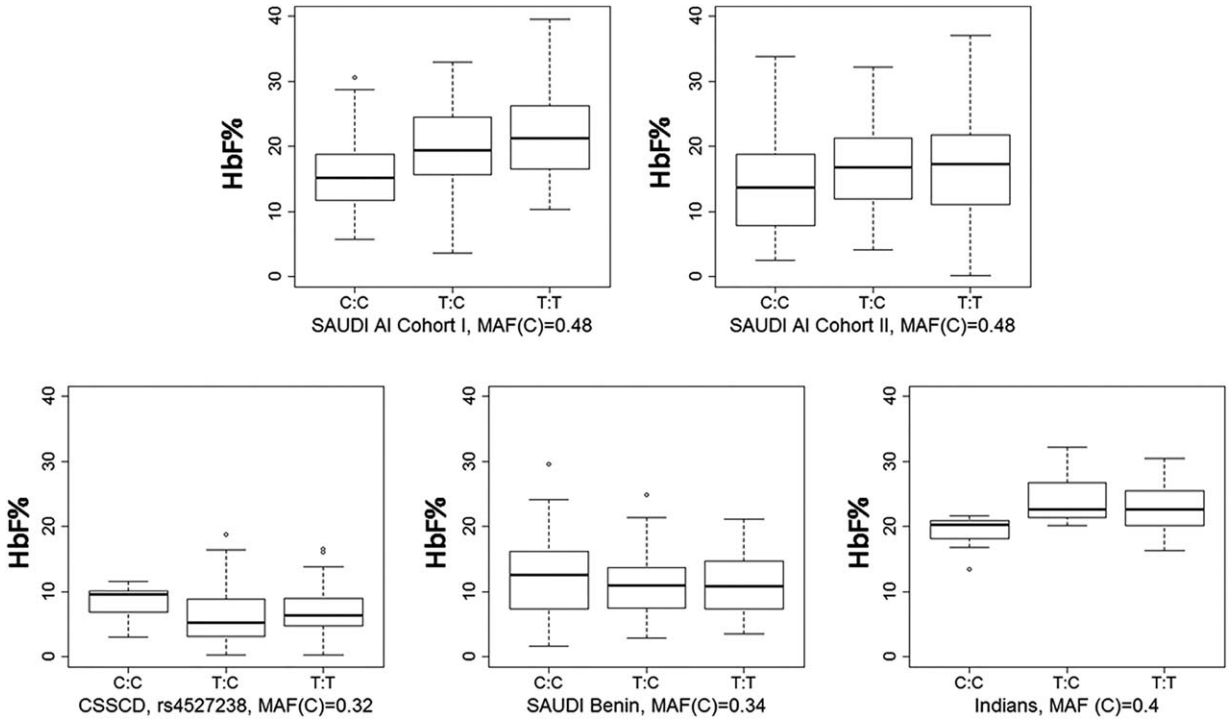

Figure 1. Association of rs4527238 in *ANTXR1* with HbF expression in Saudi AI Cohort I ($n = 120$), Saudi AI Cohort II ($n = 139$), Indian AI ($n = 44$), Saudi Benin ($n = 76$), and CSSCD ($n = 894$) datasets. The data for Saudi AI Cohort 1, Saudi Benin and CSSCD were based on genome-wide SNP analysis imputed to the 1000 Genomes reference panel while rs452728 was genotyped directly in the Indian AI cohort. The CSSCD sample included 47% Benin, 27% Benin/Bantu, 10% Benin/Senegal and 16% other haplotypes. The bar in each rectangle is the median value for the HbF in percentage scale.

 doi:10.1002/ajh.24527

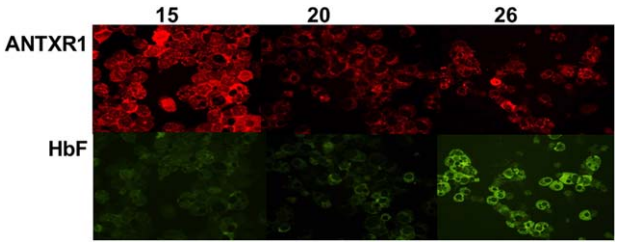

Figure 2. Immunofluorescence co-staining studies of ANTXR1 and HbF in sickle iPSCs. As ANTXR1 staining falls from culture days 15–20 and 26 (top row), HbF staining increases (bottom row). This reflects increased HbF synthesis as iPSCs mature and decreased HbF synthesis as $CD34^+$ cells differentiate. A similar pattern was seen for the ϵ-globin chain (data not shown). Under our culture conditions established for erythroid differentiation of human $CD34^+$ cells 80–90% of cells are erythrocytes on day 22, based on FACS analysis and Wright–Giemsa-staining (data not show). When the iPSCs were induced toward erythroid differentiation, >80% of cells are erythrocytes on day 15 of culture. Occasional bright ANTXR1 staining cells are nonerythroid based on Wright–Giemsa staining. [Color figure can be viewed at wileyonlinelibrary.com]

Figure 2 shows immunofluorescence costaining for ANTXR1 and HbF during the erythroid differentiation of sickle iPSCs. ANTXR1 showed specific expression patterns in the cultured $CD34^+$ cells and iPSCs, which cannot be explained as nonspecific binding. As hemoglobin expression increased fetal (γ) and embryonic (ϵ) globin chains (not shown) are the predominant non-α globin. iPSCs displayed increasing expression of γ-globin chains as ANTXR1 immunofluorescence fell.

Predicted protein interactions with ANTXR1

The predicted protein interactions of ANTXR1 using String 9.1 are shown in Supporting Information Fig. S6. ANTXR1 interacted directly with LRP6 [33], HDAC2 [34], BCLX and BRAC1. The highest level of confidence of these interactions was with LRP6 (0.95). Nine other proteins had scores of 0.551 and higher, including HDAC2 (0.61).

Discussion

High HbF levels are associated with reduced morbidity and mortality in sickle cell disease [35,36]. In the Middle East and in India the HbS gene is often on the AI *HBB* haplotype [37]. Amongst these patients mean HbF is two to four times higher than in individuals with African-origin haplotypes, but it can have a wide range of HbF concentrations [9,38].

A polymorphism in intron 9 of *ANTXR1*, a type 1 transmembrane protein and receptor for anthrax toxin was associated with HbF only in Saudi patients with the AI haplotype. The intronic SNPs associated with HbF are found in all four coding variants of the seven splice variants of *ANTXR1*. Only the longest isoform of the coding variants was expressed in iPSC and $CD34^+$ derived erythroid cells (Supporting Information Fig. S5) and rs4527238 is at a potential acceptor splice site (Supporting Information Fig. S7). *ANTXR1* has not been previously associated with a hematologic phenotype. Loss of function mutations in *ANTXR1* that has been linked to cell adhesion and migration and angiogenesis causes GAPO syndrome, a rare recessive disorder associated with growth retardation, alopecia, pseudoanodontia, and progressive optic atrophy that is characterized by excessive extracellular matrix deposition [39,40].

MicroRNAs have been shown to modulate hematopoietic differentiation, proliferation and activity of hematopoietic cells [41–43]. MiR-3126 is embedded within the intron of *ANTXR1* that is the site of the HbF associated variants and might have independent intronic promoters [41]. MiR-221 and miR-222 can regulate HbF expression in erythroid cells by modulating the Kit receptor [44]. Mir-96 regulated HbF expression by direct post-transcriptional inhibition of γ-globin mRNA during adult erythropoiesis, and MiR-15a and miR-16-1 modulated *MYB* expression in human erythroid cells increasing fetal and embryonic hemoglobin gene expression [45,46]. Whether or not MiR-3126 has any effect on HbF expression is unknown.

As cultured $CD34^+$ erythroid cells mature HbF synthesis wanes and this is associated with increased *ANTXR1* expression and decreased HbF/HbA and HbF mRNA/HbA mRNA ratios. Conversely, as erythroid progenitors derived from iPSCs matured in culture they begin to produce HbF as these cells have an embryonic/fetal globin gene expression profile. As HbF production increased *ANTXR1* expression declined (Supporting Information Fig. S5). These data suggest that *ANTXR1* is a suppressor of HbF gene expression.

How might *ANTXR1* impact HbF gene expression? Bioinformatic analysis suggests two possibilities based on its putative role in hematopoiesis and Wnt signaling and interaction with HDAC2 and globin gene expression. A model for the effect of *ANTXR1* on HbF regulation is shown in Supporting Information Fig. S8. Wnt signaling is differentially activated during hematopoiesis [47]. ANTXR1 was predicted to interact with LRP6, a Wnt coreceptor, and with FZD1, another Wnt signaling-related protein (Supporting Information Fig. S6) [33]. These interactions might occur at the extracellular von Willebrand Factor A domain of ANTXR1 that is a protein–protein interaction module [33,48–50]. *ANTXR1* knockdown decreased the expression of Wnt pathway-associated proteins including LRP6 [51]. Low *ANTXR1* expression might reduce Wnt signaling, and through effects on hematopoietic stem cells, indirectly increase HbF as stress erythropoiesis is associated with increased HbF production [52–54].

ANTXR1 might also modulate HbF gene expression directly by interactions with HDAC2 [55]. *HDAC2* inhibition increased HbF gene expression and was associated with reduced expression of *ANTXR1*. *MYB* has similar dual effects on HbF concentration affecting both erythropoiesis and transcription within the *HBB* gene cluster [16].

In Saudi sickle cell anemia with the AI haplotype *ANTXR1* is a candidate for an additional trans-acting HbF modulator. Along with cis-acting HbF modulatory elements, *ANTXR1* and other trans-acting modulators like *BCL11A*, might account for the Gaussian distribution of HbF levels in the AI haplotype [9]. Whether targeting the *ANTXR1* pathway is a therapeutic approach to HbF induction will require functional and mechanistic studies.

Disclosure Statement

VV wrote the article and performed bioinformatic and genetic analysis, JJF, HS, PS and HW analyzed and annotated whole genome sequences and did staistical anlysis, AA, H-YL, HE and IS performed genetic, genomic analysis and immunofluoresct studies and edited the article, AKA-A, AMA-R, FA-M, ZN, AAl, HOQ, PKP contributed patients and patient data and edited the article, GM and GJM generated iPSCs and CD34 cells and prepared samples for RNA-seq, DHKC edited the article, MHS supervised and conceived of the research, wrote and edited the article. The authors declare no competing interests.

Acknowledgments

Charles Jahnke provided technical assistance with the use of the Boston University Medical Campus Linux Clusters for Genetic Analysis computing resource. Whole genome sequencing results are available on request from the University of Dammam and Boston University.

References

1. Akinsheye I, Alsultan A, Solovieff N, et al. Fetal hemoglobin in sickle cell anemia. Blood 2011; 118:19–27.
2. Bauer DE, Orkin SH. Update on fetal hemoglobin gene regulation in hemoglobinopathies. Curr Opin Pediatr 2011;23:1–8.
3. Sankaran VG, Xu J, Orkin SH. Advances in the understanding of haemoglobin switching. Br J Haematol 2010;149:181–194.
4. Wilber A, Nienhuis AW, Persons DA. Transcriptional regulation of fetal to adult hemoglobin switching: New therapeutic opportunities. Blood 2011;117:3945–3953.
5. Nagel RL, Fabry ME, Pagnier J, et al. Hematologically and genetically distinct forms of sickle cell anemia in Africa. The Senegal type and the Benin type. N Engl J Med 1985;312:880–884.
6. Bae H, Baldwin CT, Sebastiani P, et al. Meta-analysis of 2040 sickle cell anemia patients: *BCL11A* and HMIP are the major genetic modifiers of HbF in African Americans. Blood 2012; 120:1961–1962.
7. Bauer DE, Kamran SC, Lessard S, et al. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science 2013;342:253–257.
8. Lettre G, Sankaran VG, Bezerra MA, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc Natl Acad Sci USA 2008;105:11869–11874.
9. Ngo D, Bae H, Steinberg MH, et al. Fetal hemoglobin in sickle cell anemia: Genetic studies of the Arab–Indian haplotype. Blood Cells Mol Dis 2013;51:22–26.
10. Galarneau G, Palmer CD, Sankaran VG, et al. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. Nat Genet 2010;42:1049–1051.
11. Sedgewick A, Timofeev N, Sebastiani P, et al. BCL11A (2p16). is a major HbF quantitative trait locus in three different populations. Blood Cells Mol Dis 2008;41:255–258.
12. Vathipadiekal V, Alsultan A, Baltrusaitis K, et al. Homozygosity for a haplotype in the HBG2-OR51B4 region is exclusive to Arab–Indian haplotype sickle cell anemia. Am J Hematol 2016;91:E308–E311.
13. Shaikho EM, Habara AH, Alsultan A, et al. Variants of *ZBTB7A* (LRF) and its β-globin gene cluster binding motifs in sickle cell anemia. Blood Cells Mol Dis 2016;59:49–45.
14. Sankaran VG, Menne TF, Xu J, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science 2008;322:1839–1842.
15. 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;117:4935–4945.
16. Stadhouders R, Aktuna S, Thongjuea S, et al. HBS1L-MYB intergenic variants modulate fetal hemoglobin via long-range MYB enhancers. J Clin Invest 2014;124:1699–1710.
17. Human Genomics. The genotype–tissue expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science 2015;348:648–660.
18. Gamazon ER, Zhang W, Konkashbaev A, et al. SCAN: SNP and copy number annotation. Bioinformatics 2010;26:259–262.
19. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012; 22:1790–1797.
20. Goh SH, Lee YT, Bouffard GG, et al. Hembase: browser and genome portal for hematology and erythroid biology. Nucl Acids Res 2004;32: D572–D574.
21. Kingsley PD, Greenfest-Allen E, Frame JM, et al. Ontogeny of erythroid gene expression. Blood 2013;121:e5–e13.
22. Jensen LJ, Kuhn M, Stark M, et al. STRING 8—A global view on proteins and their functional interactions in 630 organisms. Nucl Acids Res 2009;37:D412–D416.
23. Park S, Gianotti-Sommer A, Molina-Estevez FJ, et al. A comprehensive ethnically diverse library of sickle cell disease-specific induced pluripotent stem cells. Stem Cell Rep 2016 (submitted).
24. Smith BW, Rozelle SS, Ubellacker J, et al. The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. Blood 2013;122:376–385.
25. Sommer AG, Rozelle SS, Sullivan S, et al. Generation of human induced pluripotent stem cells from peripheral blood using the STEMCCA lentiviral vector. J Vis Exp 2012;(68). pii: 4327. doi: 10.3791/4327.
26. Wang L, Wang S, Li W. RSeQC: Quality control of RNA-seq experiments. Bioinformatics 2012; 28:2184–2185.
27. Liao Y, Smyth GK, Shi W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923–930.
28. Wu MC, Lee S, Cai T, et al. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 2011;89:82–93.
29. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genets 2007;81:559–575.
30. 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;115:1815–1822.
31. Sebastiani P, Farrell JJ, Alsultan A, et al. BCL11A enhancer haplotypes and fetal hemoglobin in sickle cell anemia. Blood Cells Mol Dis 2015;54:224–230.
32. Desmet FO, Hamroun D, Lalande M, et al. Human splicing finder: An online bioinformatics tool to predict splicing signals. Nucl Acids Res 2009;37:e67.
33. Wei W, Lu Q, Chaudry GJ, et al. The LDL receptor-related protein LRP6 mediates internalization and lethality of anthrax toxin. Cell 2006; 124:1141–1154.
34. Schmidt DR, Schreiber SL. Molecular association between ATR and two components of the nucleosome remodeling and deacetylating complex, HDAC2 and CHD4. Biochemistry 1999;38: 14711–14717.
35. Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. N Engl J Med 1994;330: 1639–1644.
36. Steinberg MH, Sebastiani P. Genetic modifiers of sickle cell disease. Am J Hematol 2012;87: 824–826.
37. Kulozik AE, Wainscoat JS, Serjeant GR, et al. Geographical survey of the β-globin gene haplotypes: Evidence for an independent Asian origin of the sickle-cell mutation. Am J Hum Genet 1986;39:239–244.
38. Alsultan A, Alabdulaali MK, Griffin PJ, et al. Sickle cell disease in Saudi Arabia: The phenotype in adults with the Arab–Indian haplotype is not benign. Br J Haematol 2014;164:597–604.
39. Bradley KA, Mogridge J, Mourez M, et al. Identification of the cellular receptor for anthrax toxin. Nature 2001;414:225–229.
40. Stranecky V, Hoischen A, Hartmannova H, et al. Mutations in ANTXR1 cause GAPO syndrome. Am J Hum Genet 2013;92:792–799.
41. Monteys AM, Spengler RM, Wan J, et al. Structure and activity of putative intronic miRNA promoters. RNA 2010;16:495–505.
42. Garzon R, Croce CM. MicroRNAs in normal and malignant hematopoiesis. Curr Opin Hematol 2008;15:352–358.
43. Bissels U, Bosio A, Wagner W. MicroRNAs are shaping the hematopoietic landscape. Haematologica 2012;97:160–167.
44. Gabbianelli M, Testa U, Morsilli O, et al. Mechanism of human Hb switching: A possible role of the kit receptor/miR 221-222 complex. Haematologica 2010;95:1253–1260.
45. Azzouzi I, Moest H, Winkler J, et al. MicroRNA-96 directly inhibits gamma-globin expression in human erythropoiesis. PLoS One 2011;6: e22838.
46. Sankaran VG, Menne TF, Scepanovic D, et al. MicroRNA-15a and −16-1 act via MYB to elevate fetal hemoglobin expression in human trisomy 13. Proc Natl Acad Sci USA 2011;108: 1519–1524.
47. Luis TC, Naber BA, Roozen PP, et al. Canonical wnt signaling regulates hematopoiesis in a dosage-dependent fashion. Cell Stem Cell 2011; 9:345–356.
48. Staal FJ, Sen JM. The canonical Wnt signaling pathway plays an important role in lymphopoiesis and hematopoiesis. Eur J Immunol 2008;38: 1788–1794.
49. Verma K, Gu J, Werner E. Tumor endothelial marker 8 amplifies canonical Wnt signaling in blood vessels. PLoS One 2011;6:e22334.
50. Whittaker CA, Hynes RO. Distribution and evolution of von Willebrand/integrin A domains: Widely dispersed domains with roles in cell adhesion and elsewhere. Mol Biol Cell 2002;13: 3369–3387.
51. Chen D, Bhat-Nakshatri P, Goswami C, et al. ANTXR1, a stem cell-enriched functional biomarker, connects collagen signaling to cancer stem-like cells and metastasis in breast cancer. Cancer Res 2013;73:5821–5833.
52. Stamatoyannopoulos G, Veith R, Galanello R, et al. Hb F production in stressed erythropoiesis: Observations and kinetic models. Ann NY Acad Sci 1985;445:188–197.
53. Xiang J, Wu DC, Chen Y, et al. In vitro culture of stress erythroid progenitors identifies distinct progenitor populations and analogous human progenitors. Blood 2015;125:1803–1812.
54. Luck L, Zeng L, Hiti AL, et al. Human CD34(+) and CD34(+)CD38(−) hematopoietic progenitors in sickle cell disease differ phenotypically and functionally from normal and suggest distinct subpopulations that generate F cells. Exptl Hematol 2004;32:483–493.
55. Bradner JE, Mak R, Tanguturi SK, et al. Chemical genetic strategy identifies histone deacetylase 1 (HDAC1) and HDAC2 as therapeutic targets in sickle cell disease. Proc Natl Acad Sci USA 2010;107:12617–12622.