Blood Cells, Molecules and Diseases 54 (2015) 224–230

Contents lists available at ScienceDirect

Blood Cells, Molecules and Diseases

journal homepage: www.elsevier.com/locate/bcmd

BCL11A enhancer haplotypes and fetal hemoglobin in sickle cell anemia

P. Sebastiani [a,*], J.J. Farrell [b], A. Alsultan [c], S. Wang [a], H.L. Edward [b], H. Shappell [a], H. Bae [d], J.N. Milton [a], C.T. Baldwin [b], A.M. Al-Rubaish [e], Z. Naserullah [f], F. Al-Muhanna [e], A. Alsuliman [g], P.K. Patra [h], L.A. Farrer [b], D. Ngo [b], V. Vathipadiekal [b], D.H.K. Chui [b], A.K. Al-Ali [i], M.H. Steinberg [b]

[a] *Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States*
[b] *Department of Medicine, Boston University School of Medicine, Boston, MA, United States*
[c] *Sickle Cell Disease Research Center and Department of Pediatrics, College of Medicine, King Saud University, Riyadh, Saudi Arabia*
[d] *College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States*
[e] *Department of Internal Medicine, College of Medicine, University of Dammam, Dammam, Saudi Arabia*
[f] *Department of Pediatrics, Maternity & Child Hospital, Dammam, Saudi Arabia*
[g] *Department of Hematology, King Fahd Hospital, Hafof, Al-Ahsa, Saudi Arabia*
[h] *Deptartment of Biochemistry, Pt. J.N.M. Medical College, Raipur, Chattisgarh, India*
[i] *Prince Mohammed Center for Research & Consultation Studies, University of Dammam, Dammam, Saudi Arabia*

ARTICLE INFO

Article history:
Submitted 7 January 2015
Accepted 7 January 2015
Available online 30 January 2015

(Communicated by M. Narla, DSc,
07 January 2015)

Keywords:
Sickle cell disease
Hemoglobin
Genetic association
Haplotype analysis

ABSTRACT

Background: : Fetal hemoglobin (HbF) levels in sickle cell anemia patients vary. We genotyped polymorphisms in the erythroid-specific enhancer of *BCL11A* to see if they might account for the very high HbF associated with the Arab–Indian (AI) haplotype and Benin haplotype of sickle cell anemia.
Methods and results: : Six *BCL112A* enhancer SNPs and their haplotypes were studied in Saudi Arabs from the Eastern Province and Indian patients with AI haplotype (HbF ~20%), African Americans (HbF ~7%), and Saudi Arabs from the Southwestern Province (HbF ~12%). Four SNPs (rs1427407, rs6706648, rs6738440, and rs7606173) and their haplotypes were consistently associated with HbF levels. The distributions of haplotypes differ in the 3 cohorts but not their genetic effects: the haplotype TCAG was associated with the lowest HbF level and the haplotype GTAC was associated with the highest HbF level and differences in HbF levels between carriers of these haplotypes in all cohorts were approximately 6%.
Conclusions: : Common HbF *BCL11A* enhancer haplotypes in patients with African origin and AI sickle cell anemia have similar effects on HbF but they do not explain their differences in HbF.

Introduction

Fetal hemoglobin (HbF) is the predominant modulator of the phenotype of sickle cell anemia. By inhibiting sickle hemoglobin (HbS) polymerization it reduces the tissue injury and hemolytic anemia exemplifying this disease [1–4]. HbF levels and the distribution of HbF concentrations in sickle erythrocytes are highly variable. The genetic regulation of HbF was first associated with the haplotype of the β-globin gene (*HBB*) cluster suggesting the importance of cis-acting regulation [5,6]. In the Middle East and in India the HbS gene is often on an autochthonous Arab-Indian (AI) *HBB* haplotype that is associated with HbF levels more than twice as high as those found with African origin haplotypes; see Fig. 1 [7,8]. This is clinically important as the youngest individuals with the AI haplotype have the mildest phenotype of all sickle cell anemia patients, although when their HbF level falls from about 30% in children to 15–20% in adults the disease becomes more severe [9–14]. Saudi Arabs with the African origin Benin haplotype have HbF levels nearly twice that of non-Arab patients with this haplotype [15]. Within each haplotype group there is considerable heterogeneity of HbF [5,6,16], suggesting that trans-acting elements also effect γ-globin gene (*HBG*) expression.

One trans-acting element is *BCL11A*, a repressor of γ-globin gene expression [10,17–21]. Functional studies have shown that *BCL11A* expression is regulated by erythroid-specific enhancers in its 2nd intron. The enhancer elements contain 3 DNase hypersensitive sites (DHS) lcated +62, +58 and +55 kb from the transcription initiation site [10]. Two SNP haplotypes of the enhancer elements were associated with HbF levels in African American patients with sickle cell anemia. The strongest association with HbF levels in African Americans with sickle cell anemia was with rs1427407 in DHS +62.

The cause of high HbF in the AI haplotype and the Saudi Benin haplotype is unexplained and could be due to increased *HBG* expression mediated by cis- or trans-acting regulators. We focused on *BCL11A* enhancer polymorphisms and examined the association of their haplotypes with HbF levels in Saudi and Indian patients with the AI haplotype

* Corresponding author at: Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Boston, MA 02118, United States.
E-mail address: sebas@bu.edu (P. Sebastiani).

http://dx.doi.org/10.1016/j.bcmd.2015.01.001
1079-9796/

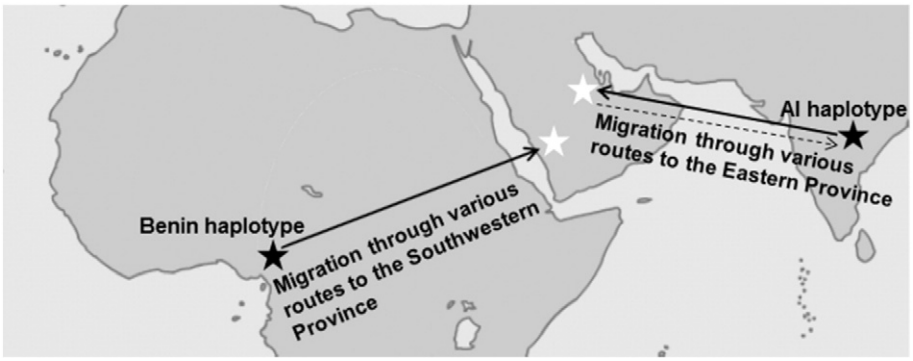

Fig. 1. Origins of the HbS mutation in Africa and India and migration to the Arabian Peninsula. The primary HbS gene-associated haplotype in the Southwestern Province of Saudi Arabia is Benin, which was introduced from Africa. Eastern Province patients have the Arab–Indian (AI) haplotype that might have originated in India (solid arrow) or alternatively, originated in the Middle East and migrated to India (dashed arrow).

and Saudi HbS homozygotes with the Benin haplotype, and compared these results to those of African American patients with sickle cell anemia.

Materials and methods

Study populations and HbF measurement

SNPs in *BCL11A* enhancers were genotyped directly or imputed from genome-wide SNP analysis in the following cohorts (Table 1):

1. 894 African American HbS homozygotes, diverse haplotypes, from the Cooperative Study of Sickle Cell Disease (CSSCD), aged >5 years [22].
2. 96 Saudi HbS homozygotes mostly with the Benin haplotype from the Southwestern Province of Saudi Arabia (Saudi W), aged 4–55 years, not taking hydroxyurea. (Saudi W)
3. 110 Saudi HbS homozygotes all with AI haplotype from the Eastern Province of Saudi Arabia, aged 11–59 years, not taking hydroxyurea (Saudi E) [16].
4. 44 Indian HbS homozygotes all with the AI haplotype, age 10–32 years, not taking hydroxyurea.

HbF: HbF was measured in all Saudi samples using high performance liquid chromatography (HPLC) or capillary electrophoresis. HbF in the CSSCD was measured by alkali denaturation [23]. HbF in Indian patients was measured by HPLC both in India and at Boston University.

Genotyping

HbS mutation and HBB haplotype

Homozygosity for the HbS gene was confirmed using amplification refractory mutation system analysis [24]. HbS homozygosity in the CSSCD cohort was based on clinical and hematologic studies. The AI haplotype was ascertained by analysis of rs7482144 (Xmn1 C-T restriction site 158 bp 5′ to *HBG2*), rs3834466 (Hinc2 restriction site 5′ to *HBE1* and the C-T SNP 68 bp 5′ to *HBD* [25].

SNPs: Targeted genotyping of *BCL11A* enhancer SNPs was done with tetra-primer ARMS-PCR, TaqMan assays and Sanger sequencing.

Table 1
Cohorts studied. Saudi E and Indian patients are homozygous for the AI haplotype and were not taking hydroxyurea. CSSCD patients have different African haplotypes. Cooperative Study of Sickle Cell Disease (CSSCD). Illumina/Imputation signifies imputation of SNPs not included in the Illumina SNP arrays.

	N	Age ± SD (yrs.)	HbF ± SD (%)	Genotyping
CSSCD	894	13.6 ± 11.3	5.2 ± 5.6	Illumina/Imputation
Saudi W	96	17.7 ± 9.84	11.4 ± 6.0	Illumina/Imputation
Saudi E: AI Haplotype	110	26.7 ± 10.1	18.0 ± 7.0	PCR/Illumina/Imputation
Indian: AI Haplotype	44	14.6 ± 4.6	23.0 ± 4.8	PCR

Imputation of genotypes

To derive haplotypes of these SNPs in cohorts where direct genotyping was not done we imputed to 1000 Genomes level data from genome-wide SNP data obtained using Illumina technology (Supplementary material).

Data analysis

Data are described by mean and standard deviation. Single SNP associations were estimated using sex and age adjusted linear regression with additive genetic effects, and the B allele in the forward strand was the coded allele (Table 2). A mixed effect model with kinship coefficients implemented in the coxme package of the R statistical software was used to analyze the association between HbF levels and SNPs of Saudi W samples, since some subjects were related. HbF levels were approximately normally distributed in the Saudi E and Indian samples (see Supplement Fig. S1), and a cubic root transformation was used for normalizing the HbF levels of CSSCD samples as in [22]. HbF levels of Saudi W samples were also normalized using a cubic root transformation. Linkage disequilibrium was evaluated using the program HaploView. Haplotype analysis was conducted using the haplo.stats package in R software [26]. The program uses a two-step EM algorithm to iteratively update the probability of subjects haplotypes, based on coefficient of the age and sex adjusted regression model, and to update the coefficients of the regression model based on the posterior probability of a subject haplotype. The analysis was conducted in the combined Saudi E and Indian samples and, separately, in the Saudi W and CSSCD samples using linear regression adjusted for age and sex. Haplotype pairs for each subject were inferred based on the most likely haplotypes and the distribution of HbF was displayed using boxplots. To test if the effect of haplotype pairs changed substantially in the different cohorts, a multivariable regression model of the HbF levels versus haplotype pairs was fitted, using age, sex, indicators for the haplotype pairs, an indicator of the cohort type (CSSCD, Saudi W, and Saudi E + Indians), and an interaction between indicator variables of the haplotype pairs and study cohorts. Lack of a statistically significant interaction indicated no change of haplotype pairs effects in the 3 cohorts. All analyses were conducted using the statistical software R.

These studies were approved by the Institutional Review Boards of the participation institutions.