American Journal of Hematology 69:77–79 (2002)

Study of the Single Nucleotide Polymorphism (SNP) at the Palindromic Sequence of Hypersensitive Site (HS)4 of the Human β-Globin Locus Control Region (LCR) in Indian Population

Ritushree Kukreti,[1] Chandrika B-Rao,[1] Swapan Kr Das,[2] Madhusnata De,[2] Geeta Talukder,[2] Flavian Vaz,[3] I.C. Verma,[4] and Samir K. Brahmachari[1]*

[1]Functional Genomics Unit, Centre for Biochemical Technology (CSIR), Delhi University Campus, Delhi, India
[2]Thalassemia Counselling Department, Vivekananda Institute of Medical Sciences, Ramakrishna Mission Seva Pratisthan, Calcutta, India
[3]B.J. Wadia Hospital for Children and Institute of Child Health Research Society, Acharya Donde Marg, Parel, Mumbai, India
[4]Sir Ganga Ram Hospital, Rajinder Nagar, New Delhi, India

LCR, a genetic regulatory element, was examined in β-thalassemia patients who do not show any mutation in the β-globin genes. We sequenced LCR-HS2, HS3, and HS4 in samples from 16 such patients from the Indian population and found only one SNP A-G in the inverted repeat in HS4. A significant association was observed between the G allele and occurrence of β-thalassemia by Fisher's exact test. The AG and GG genotypes showed higher relative risk as compared to the AA genotype. We also observed linkage disequilibrium between the A/G polymorphism and the AT-rich motif of the LCR HS2 region, suggesting that the G allele could be an evolutionarily new mutation in the study population. Am. J. Hematol. 69:77–79, 2002. © 2002 Wiley-Liss, Inc.

Key words: β-thalassemia β-globin gene; locus control region; hypersensitive site; single nucleotide polymorphism

INTRODUCTION

β-Thalassemia is a highly prevalent autosomal recessive disorder characterised by the complete absence of, or some defect in, the β-globin genes, leading to an imbalance of the α- and β-globin chains [1]. The β-globin LCR, a major regulatory element necessary for high-level transcription of the β-globin genes located approximately 5−20 kb upstream of the ε gene [2]. LCR consists of five DNase I −HS1-5. Genetic variation in the binding sites of the core fragments of the β-LCR HS 2, 3, and 4 may play a role in cases of β-thalassemia (both major and carrier phenotypes) in which none of the known mutations is found in the β-globin gene. In this study, we sequenced the 1.8-kb region of the β-globin and β-LCR-HS 2, 3, and 4 of the uncharacterized patients in our sample in an effort to discover new polymorphisms that may be associated with the disease and aid in carrier detection.

SUBJECTS AND METHODS

Subjects

1,800 β-thalassemia patients with varying disease severity were recruited from three centres in northern, eastern, and western India after diagnosis. Control blood samples were collected from different communities with informed consent.

Phenotype Analysis

Genomic DNA was isolated from whole blood in anticoagulant (EDTA) by using proteinase K−phenol−chloroform standard procedure [3]. Hemoglobin

*Correspondence to: Dr. Samir K. Brahmachari, Functional Genomics Unit, Centre for Biochemical Technology (CSIR), Delhi University Campus, Mall Road, Delhi 110 007, India. E-mail: skb@cbt.res.in

Received for publication 15 June 2001; Accepted 15 August 2001

DOI 10.1002/ajh.10026

variants were analysed by agarose gel electrophoresis at pH 8.6; HbA_2 was determined by gel elution and HbF by alkali denaturation, respectively [4].

Mutation Analysis and Examination of β-LCR HSs

ARMS-PCR method was used to detect known 23 β-thalassemia mutations [3,5]. β-globin gene cluster locus control 5′ HS 2, 3, and 4 (~460 bp, ~446 bp, and ~442 bp from GeneBank coordinates 8757–9217, 4397–4991, and 668–1109) were sequenced by the ABI Prism 377 automated DNA sequencer using dye terminators chemistry. Sequences were aligned with the corresponding wild-type sequences using Factura and Sequence Navigator program. Statistical analysis was done using Fisher's exact test, genotype relative risk (GRR), and χ^2 test.

RESULTS AND DISCUSSION

Of the 1,800 β-thalassemia patients screened by ARMS-PCR 27 did not show any of the 23 known common mutations in the β-globin gene in the Indian population. By direct sequencing of the 1.8-kb region of β-globin gene of these 27 samples, 11 were found to have rare mutations (2 had codon 29(C→T), 2 had codon 39(C→T), and 6 had codon 110 (T→C)), while 16 showed no mutation. Analysis of the core fragment of LCR HS2, 3, and 4 was carried out in these 16 samples in search of novel mutations associated with the disease phenotype.

DNA sequencing of HS2, 3, and 4 core sequences showed only one polymorphism, an A-G, in the palindromic sequence, TGGGGACCCCA, of LCR HS4, in some of the uncharacterised samples. Using samples from 86 normal control subjects and 47 patients (β-thalassemia major and carriers), we tested for association of this SNP with β-thalassemia. The allelic frequency of A was 0.76 in controls and 0.63 in patients. We found significant association between the G allele and occurrence of the disorder (Fisher's exact test, P = 0.015). The genotypes AA, AG, and GG occurred in 34.0%, 57.5% and 8.5%, respectively, of patients and 55.8%, 40.7%, and 3.5%, respectively, of control individuals. Although this SNP has been reported earlier by Sriroongrueng et al. [6], this is the first time that the GG homozygote has been detected in a population. The genotype relative risk (1.49 with 95% confidence interval = (1.09, 2.04)) for genotypes GG and AG suffering from β-thalassemia as compared to the AA genotype is higher, which confirms the expectation from the allelic association of G with disease.

TABLE I. Linkage Disequilibrium Between SNP at the Palindromic Sequence in LCR-HS4 and $(AT)_xN_y(AT)_z$ Length Polymorphism in LCR-HS2[a]

	SNP	
$(AT)_xN_y(AT)_z$ polymorphism	A	G
$(AT)_{10}N_{12}(AT)_{11}$	10 (33.33)	0 (0)
$(AT)_9N_{12}(AT)_{11}$	2 (6.67)	0 (0)
$(AT)_9N_{12}(AT)_{10}$	2 (6.67)	16 (53.33)

[a]Number of haplotypes at the two loci (percentage frequency of the haplotype) are shown in the table.

We observed G allele in the HS4 locus is associated with only one distinct pattern $(AT)_9N_{12}(AT)_{10}$ of HS2 of the β-LCR (Table I), showing linkage disequilibrium between the two loci and suggesting that the G allele could be an evolutionarily new mutation in the Indian population. The occurrence of A in the middle of a short palindromic sequence causes the formation of a hairpin structure [7]. Replacement of A by G decreases the stability of this structure [7], and it may be responsible for the observed low frequency of the G allele. However, it does not by itself explain how this may play a role in β-thalassemia. Although structural variations caused by the replacement of A by G might not have a direct casual effect, it is possible that it modifies gene expression [8,9] and thus the disease severity.

ACKNOWLEDGMENT

This work was supported by Department of Biotechnology, Government of India, and Council of Scientific and Industrial Research, Government of India.

REFERENCES

1. Weatherall DJ, Clegg JB. Thalassemia—a global public health problem. Nat Med 1996;2:847–849.
2. Li Q, Harju S, Peterson KR. Locus control regions coming of age at a decade plus. Trends Genet 1999;15:403–408.
3. Old JM, Varawalla NY, Weatherall DJ. The rapid detection and prenatal diagnosis of β-thalassemia in the Asian Indian and Cypriot and populations in the U.K. Lancet 1990;336:834–837.
4. De M, Chakraborty G, Das SK, Bhattacharya DK, Talukdar G. Molecular studies of haemoglobin-E in tribal populations of Tripura. Lancet 1997;349:1297.
5. Thakur C., Vaz F, Banerjee M, Kapadia C, Natarajan PG, Yagnik H, Gangal S. Prenatal diagnosis of β-thalassemia and other haemoglobinopathies in India. Prenat Diagn 2000;20:194–201.

6. Sriroongrueng W, Schleiemacher E, Panich V, Nopparatana C, Saechan V, Laosombat V, Pornpatkul M, Fukumaki Y. Analysis of β-thalassemia mutations and β-locus control region hypersensitive sites 2, 3 and 4 in southern Thailand. Southeast Asian J Trop Med Public Health 1997;28:120–127.
7. Yoshizawa S, Kawai G, Watanabe K, Miura K, Hirao I. GNA trinucleotide loop sequences producing extraordinarily stable DNA minihairpins. Biochemistry 1997;36:4761–4767.
8. Brahmachari SK, Sarkar PS, Balagurumoorthy P, Burma PK, Bagga R. Synthetic gene design to investigate the role of *cis*-acting DNA structural elements in regulation of gene expression in vivo. Nucleic Acids Res 1991;24:163–166.
9. Sarkar PS, Bagga R, Balagurumoorthy P, Brahmachari SK. A noval approach to design *cis*-acting DNA structural elements for regulation of gene expression in vivo. Curr Sci 1991;61: 586–591.