Abstract

Introduction: Previous statistical analyses have demonstrated linkage disequilibrium of $\beta^E$-allele and surrounding in-cis sequences.

Objective: To confirm the linkage disequilibrium of $\beta^E$-allele and surrounding in-cis sequences.

Materials and Methods: Associations of the $\beta^E$-allele with the $\beta$-globin haplotype, (AT)$_N$(AT)$_m$ motif of $\beta$LCR-HS2, XmnI$^+_G$ polymorphism, pre-$\alpha^G$ framework, (TG)$_n$ motif in IVSII of $\alpha^\gamma$-globin gene and HBBP1 : rs2071348 polymorphism were analyzed in three Thai HbE/$\beta^O$-thalassemia families.

Results: The $\beta^E$-allele was always in-cis with the $\beta$-globin haplotype III, pre $\alpha^G$-framework 2, XmnI+, (TG)$_{13}$ motif within IVSII of $\alpha^\gamma$-globin gene and “C” at HBBP1 : rs2071348 locus, all previously shown to be associated with high HbF expression. Combination of the (AT)$_N$(AT)$_m$ motif of $\beta$LCR-HS2 and these cis-acting factors seemed to be required for maximum HbF-expression.

Conclusion: This family study substantially confirmed the linkage between the $\beta^E$ allele and cis-acting loci associated with high HbF phenotype. This study also highlighted the closed relationship between HbE and HbF as well as the need of combination of the analyzed cis-acting sequences in activating HbF expression in human.


Keywords: HbE/$\beta^O$-thalassemia, HbF, $\beta$-globin haplotype, $\beta$-globin mutations, cis-acting loci
Introduction

HbE (αβE) is abnormal hemoglobin arising from an assembly of mutated βE (A-G at codon 26 altering glutamic to lysine) and normal α-globin chains. This hemoglobin variant is common in Southeast Asia, attaining a frequency of about 60% among the population of Thailand, Laos and Cambodia. The βE-mRNA was shown to be ineffectively synthesized owing to competitive utilization of an alternative splice site at codon 25 and the silencing effect of the (AT)9 motif in the β-globin promoter. The reduction of βE-mRNA leads to a reduced βE-globin chain, which results in the β-thalassemia phenotype. Thus, compound heterozygote of the βE-globin and β-thalassemia alleles leads to a syndrome referred to as HbE/β-thalassemia.

Phenotypically, HbE/β-thalassemia presents in a broad range of clinical symptoms, ranging from transfusion-dependent β-thalassemia major to a mild form of β-thalassemia intermedia requiring only occasional blood transfusions. Several factors have been shown to be responsible for this clinical heterogeneity, including types of the counterpart β-thalassemia alleles, co-existence of α-thalassemia and co-inheritance of high HbF genetic variants. The Xmnl, site, a “T” substitution for “C” at nucleotide -158 in 3′-γ promoter, was shown to be a major cis-acting genetic factor associated with increased HbF. Recent statistical analyses in Thai HbE carriers, HbE homozygotes and HbE/β-thalassemia revealed association of the Xmnl-γ site and other surrounding cis-acting loci with βE-allele.

In this study, we attempted to confirm these previous findings by conducting family studies to determine associations of βE-allele and cis sequence variations, including the Xmnl, polymorphism, β-globin haplotypes, (AT)N, motif of LCR-HS2, pre-γ framework (PGγF), (TG)n motif in IVSII of 3′-γ-globin gene and HBBP1 : rs2071348 single nucleotide polymorphism. The information obtained from this study should confirm the previous linkage disequilibrium analyses and provide evidence of the close relationship between HbE and HbF.
**Materials and methods**

**Subjects**
This study was approved by the research ethics board of Faculty of Associated Medical Sciences, Chiang Mai University (Approval notice 205/2012). Three northern Thai families, each having one HbE/β°-thalassemia member, were recruited. EDTA blood samples were collected. Hb typing was performed by cation exchange HPLC (The Primus Variant System 99, Kansas City, MO, USA). The β-thalassemia mutations in the first and second families were identified by multiplex allele-specific PCR and that in the third family by automated fluorescent nucleotide sequencing. Based on the results of Hb typing and β-globin genotyping, all volunteers were diagnosed. The first family had four members: HbE/β°1146(A-T)-thalassemia proband, HbE-heterozygous father, β-thalassemia heterozygous mother and HbE-heterozygous brother. The second family had four members: HbЕ/β°1146(A-T)-thalassemia proband, HbE-heterozygous mother, HbE-heterozygous brother and β-thalassemia heterozygous father. The third family had four members: HbE/β°IVS1-nt1(G-T)-thalassemia proband, HbE/β°IVS1-nt1(G-T)-thalassemia sister, HbE heterozygous father and β-thalassemia heterozygous mother.

**Molecular analysis**
Genomic DNA was prepared from leukocytes by the direct Chelex extraction method. The XmnI-G°γ-polymorphism was determined by XmnI digestion of PCR products covering the position -158 on G°γ-promoter, following the procedure described previously. Determination of A/C polymorphism of HBBP1: rs2071348 was performed by the tetra-primer ARMS-PCR, established previously in our laboratory. The RFLP haplotypes (HindII-E, HindIII-5'β, HindIII-5'γ, HindIII-3'β, HindIII-3'γ) were generated following the procedure described previously. Sanger’s direct nucleotide sequencing was employed to genotype the (AT)_N (AT)_y motif of LCR-HS2, PG1/F, (TG)n motif in IVSII of β°γ-globin gene.

**Results**

**Association of β°-allele and sequence variations on β-globin gene cluster**

Five β-globin haplotypes (HindII-E, HindIII-5'β, HindIII-5'γ, HindIII-3'β, HindIII-3'γ, Avall-β and Hinfl-β) were observed in these three families, including haplotypes I (+ - - - -++), II (+ + + + + -), III (+ + + + + +), IV (- + - - - +) and V (+ - - - - +) (Table 1). The β°-allele was in-cis with only haplotype III.

Five (AT)xNz(A) motifs of βLCR-HS2 were found. These included (AT)_N (AT)_y, (AT)_N (AT)_y, (AT)_N (AT)_y, (AT)_N (AT)_y, (AT)_N (AT)_y, (AT)_N (AT)_y, (AT)_N (AT)_y, (AT)_N (AT)_y and (AT)_N (AT)_y (Figure 1 and Table 1). The β° allele was predominately in-cis with the (AT)_N (AT)_y motif.

According to Pissard & Beuzard, the combination of SNPs at positions -1450, -1280 and -1225 related to G°γ-globin gene cap site formed three pre-G°γ-frameworks (PG°γF). These included PG°γ1F (TGA), PG°γ2F (TAG) and PG°γ4F (TGG). Interestingly, combinations of SNPs at -1067, -807, -369 and -309 formed two sub-PG°γFs (TCCA and AACAA) and were shown to be PG°γ-specific. The “TCCA” was specific to PG°γF2, whereas the “AACAA” was always in-cis to the PG°γFs 1, 4. The β°-allele was in-cis with only PG°γF2 (Table 1).

The XmnI+ and XmnI- sites were equally distributed in these three families. The β°1146(A-T), β°1146(A-T)-TCCT and β°IVS1-nt1(G-T) were in-cis with the XmnI+ site. The βA-allele was in-cis with both the XmnI+ and XmnI- sites. However, the β°-allele was in-cis with only the XmnI+ site (Table 1).

Two (TG)n motifs within IVS2 of β°γ-globin gene were observed, including (TG)_13 and (TG)_8 (CG)_6 (TG)_8 motifs (Figure 2). Combination of single nucleotide polymorphisms (SNPs) loci rs 2187608, rs 28379094, rs 28440105 and rs 3841756 formed three sub-haplotypes within this region, including “GGGA”, “CAGA” and “CAG dela”. The (TG)_13 motif was in-cis with only the “GGGA” sub-haplotype. The (TG)_8 (CG)_6 (TG)_8 motif was in-cis with either the “CAGA” or “CAG dela” sub-haplotypes. The β°-allele was in-cis with only the (TG)_13 motif within this region (Figure 2 and Table 1).
In cases of A-C polymorphism of the HBBP1: rs 2071348 locus, the polymorphic “C” allele was less frequent than wild type “A” allele. However, the \( \beta^\text{E} \)-allele was \textit{in-cis} with only the polymorphic “C” allele (Table 1).

\textbf{Figure 1} Electropherograms of the (AT)xNz(AT)y motif of \( \beta \)-LCR-HS2. Boundary of (AT)x, Nz and (AT)y repeats are shown as the lines above peaks within each electropherogram.
Figure 1 Electropherograms of the (AT)xNz(AT)y motif of βLCR-HS2. Boundery of (AT)x, Nz and (AT)y repeats are shown as the lines above peaks within each electropherogram. (continued)
Figure 1. Electropherograms of the (AT)xNz(AT)y motif of β-LCR-HS2. Boundery of (AT)x, Nz and (AT)y repeats are shown as the lines above peaks within each electropherogram. (continued)

Figure 2. (TG)n motif within IVS2 of β1-globin gene. A, B are the (TG)13/(TG)13 and (TG)9(CG)5/(TG)8/(TG)9(CG)5(TG)8, respectively.

Table 1. β-globin haplotypes and sequence variations of β-globin cluster in three Thai HbE/β0-thalassemia families.

<table>
<thead>
<tr>
<th>Family</th>
<th>AT(NG)ATF</th>
<th>Haplotype</th>
<th>Hind III</th>
<th>Hind II</th>
<th>Hind III</th>
<th>Hind II</th>
<th>Hind III</th>
<th>Hind II</th>
<th>Hind III</th>
<th>Hind II</th>
<th>Hind III</th>
<th>Hind II</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-I</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>C</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>I-II</td>
<td></td>
<td>+</td>
<td>1</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>II-I*</td>
<td>V</td>
<td>+</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>II-III</td>
<td>V</td>
<td>+</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

(AT)xNz(AT)y is the repeat motif of β-LCR-HS2. β-haplotypes are the RFLP-haplotypes of β-thalassemia chromosome. PGF stands for θγ framework. rs2071348 is the SNP (C/A) in HBBP1 locus. Numbers preceded by the minus symbol indicate locations of SNPs in promoter region of β1-globin gene. -158 is the XmnII-θγ site. *+* indicates presence of cutting sites and *-* indicates absence of cutting sites for the restriction enzymes.
Table 1. β-globin haplotypes and sequence variations of β-globin cluster in three Thai HbE/β^0-thalassemia families.

(continued)

<table>
<thead>
<tr>
<th>Family II</th>
<th>(AT)xNz(AT)y</th>
<th>β-haplotypes</th>
<th>Hind II</th>
<th>Hind III</th>
<th>Hind II</th>
<th>Hind II</th>
<th>Hind II</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-I</td>
<td></td>
<td>VII</td>
<td>+</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9-12-31</td>
<td>+</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>I-II</td>
<td></td>
<td>III</td>
<td>-</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9-12-30</td>
<td>-</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>II-I</td>
<td></td>
<td>III</td>
<td>-</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>II-II*</td>
<td></td>
<td>VII</td>
<td>+</td>
<td>1</td>
<td>A</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9-12-30</td>
<td>-</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

[Diagram of β-LCR and HSB2]
Extended β-globin haplotypes and HbF synthesis in Thai HbE/β⁰-thalassemia

Extended haplotypes of the β-globin gene cluster in three HbE/β-thalassemia; two with Hbs EF and another with Hb EE, were analyzed and compared. The first patient (II-I in family I, as shown in Table 1 and A in Figure 3), with 66.6% HbE and 20.9% HbF, was compound heterozygous for βE and β17(A-T) alleles. The second patient (II-II in family II, as shown in Table 1 and B in Figure 3), with 61.3% HbE, 22.1% HbF, was compound heterozygote for βE and β17(A-T). The third patient (II-II in family III, as shown in Table 1 and C in Figure 3), with 1.7% HbF and 89.4% HbE, was heterozygous for βE and βIVS1-nt1 (G-T) (Figure 3). As shown in Figure 3, the βE-allele in the two patients (“A” and “B”) with high HbF levels was in-cis with the (AT)₉N₁₂(AT)₁₀ haplotype III, PG₁:F2, (TG)₁₃ and “C” at the HBBP₁: rs 2071348 locus. In contrast, the βE-allele in the patient (“C”) with a lower HbF level than patients “A”, “B” was also in-cis with haplotype III, PG₁:F2, (TG)₁₃ and “C” at the HBBP₁: rs 2071348 locus. However, the (AT)xNz(AT)y motif of the β₁LCR-HS2 in patient “C” was (AT)₈N₁₂(AT)₁₁.

![Figure 3](image-url)

**Figure 3** Hb typing by HPLC (top panel) and cis-acting sequence variations (lower panel) in HbE/β⁰-thalassemia patients “A”, “B” and “C” as mentioned in the text. β₁⁷ represents β17(A-T), β17/42 represents β17/42(-TTCT), βIVS1-nt represents βIVS1-nt1(G-T), β-haplo represents β-haplotypes and PG₁:F represents pre-α₁-framework.

<table>
<thead>
<tr>
<th>β⁺-alleles</th>
<th>(AT)xNz(AT)y</th>
<th>β⁻-haplo</th>
<th>XmnL⁺⁻γ</th>
<th>γIVS2-(TG)n</th>
<th>HBBP₁-rs2071348</th>
</tr>
</thead>
<tbody>
<tr>
<td>A βE</td>
<td>9-12-10</td>
<td>III</td>
<td>2</td>
<td>(TG)₁₃</td>
<td>C</td>
</tr>
<tr>
<td>β¹⁷</td>
<td>9-12-10</td>
<td>VII</td>
<td>1</td>
<td>(TG)₁₉(CC)₁₀(TG)₁₀</td>
<td>A</td>
</tr>
<tr>
<td>B βE</td>
<td>9-12-10</td>
<td>III</td>
<td>2</td>
<td>(TG)₁₃</td>
<td>C</td>
</tr>
<tr>
<td>β¹⁷</td>
<td>11-12-10</td>
<td>VII</td>
<td>1</td>
<td>(TG)₁₉(CC)₁₀(TG)₁₀</td>
<td>A</td>
</tr>
<tr>
<td>C βE</td>
<td>8-12-11</td>
<td>III</td>
<td>2</td>
<td>(TG)₁₃</td>
<td>C</td>
</tr>
<tr>
<td>βIVS1nt</td>
<td>9-12-11</td>
<td>I</td>
<td>1</td>
<td>(TG)₁₉(CC)₁₀(TG)₁₀</td>
<td>A</td>
</tr>
</tbody>
</table>
The βE-allele mutates from the normal βA-allele by missense mutation of codon 26 of the β-globin gene, altering the codon for glutamic acid to that for lysine. This mutation is common in Southeast Asia. Recent pairwise LD analyses in Thai HbE/β-thalassemia by Ma et al. and in Thai homozygous HbE (EE), heterozygous HbE (AE) and normal (A2A) individuals by Ohashi et al. indicated significant linkage disequilibrium of the βE-allele and surrounding in-cis sequence variations. These studies assumed that malarial selection was involved. The family studies undertaken in this study found that the βE-allele was always in-cis with the (AT)9N12(AT)10 motif, β-globin haplotype III, PG F2, (TG)13 motif within IVS2 of α-globin gene, XmnI+ and “C” allele of HBBP1 : rs2071348, confirming those previous statistical assumptions. Interestingly, all of these in-cis sequence variations have been shown to be associated with increased HbF expression. Hence, closed relation between HbE and cis-acting genetic factors associated with increased HbF expression observed in this study was most likely a result of positive selection against malaria, which is endemic in Thailand. Red blood cells containing Hbs E and F retarded the growth of Plasmodium falciparum, probably due to increased oxidative stress. Thus, individuals inheriting both HbE and HbF-activating loci are likely to survive malarial infection.

Analysis in 3 HbE/β-thalassemia patients having typical EF (“A”, “B” in Figure 3) and atypical EE (“C” in Figure 3) phenotypes highlighted the importance of combination of the (AT)9N12(AT)10 motif, β-globin haplotype III, PG F2, (TG)13 motif within IVS2 of α-globin gene, XmnI+ and “C” allele of HBBP1 : rs2071348 in activating HbF production. The β17(A-T), β14/42 (-TTCT) and βIVS1-nt1(G-T) alleles in these three patients were in-cis to the β-globin haplotype I & VII, PG F1, (TG)9(CG)1(TG)9 motif within IVS2 of α-globin gene, XmnI- and “A” allele of HBBP1 : rs2071348 previously shown to be involved with low HbF expression. Thus, HbF expression in these three patients was likely the result of increased expression of the α-globin gene on the βE-chromosome. However, in patient “C” (Figure 3), the (AT)9N12(AT)10 motif in the βE-chromosome was (AT)9N12(AT)11, which has been previously shown to be involved in normal HbF expression and should be responsible for the atypically low HbF in patient “C”. The (AT)9N12(AT)10 motif of βLCR-HS2 has been shown to be a classical enhancer for α-globin gene expression via opening chromosomal domain, thus increasing accessibility of RNA polymerase. The (AT)9N12(AT)10 motif might have less impact on the chromosome opening activity at α-globin promoter, resulting in minimal α-globin gene expression. This study was, however, the first to describe this unusual phenomenon.

In conclusion, this family study confirmed the linkage of the βE allele and cis-acting loci associated with high HbF phenotype. The close relationship between HbE and HbF as well as the need of combination of the analyzed cis-acting sequences in activating HbF expression was highlighted.

Acknowledgements

This work was financially supported by a Research Chair Grant from The National Science and Technology Development Agency of Thailand. The authors would like to thank the Research Administration Center, Chiang Mai University for proofreading and correcting English language. LT was supported by the Post Graduate Study Fund, Faculty of Associated Medical Sciences, Chiang Mai University.

Conflict of interest

None
allele of + and "C" allele of altering the codon for glutamic acid to that for lysine. Thus, HbF expression in these three haplotype I & VII, PG F I, (TG) This mutation is common in Southeast Asia. This high expression. 15,33,34  Thus, HbF expression in these three patients was likely the result of increased expression of HbE (EE), heterozygous HbE (AE) and normal (A 2A) association with increased HbF expression observed in this study. The family studies undertaken in this study found that the selection was involved. The close relationship between HbE and -thalassemia by Ma et al. 8  and in Thai homozygous -globin gene expression. This study was, however, the first to describe this unusual phenomenon. The motif within IVS2 of A -globin gene, XmnI + and "C" motif, (CG)5(TG)8 motif within -globin chain. Int J Hematol. 2012; 95(4): 386-93.

Discussion

The alleles in these three patients were -thalassemia by Ma et al. 8  and in Thai homozygous -globin gene, XmnI and in-cis IVS1-nt1(G-T) motif within IVS1-nt1(G-T) motif, cis-acting genetic factors in-cis in-activating motif within -globin gene, XmnI and "C" motif, (CG)5(TG)8 motif within -globin gene, XmnI and XmnI+ motifs, confirming those previous findings. The motif, (AT)y motif in the -globin chain. Int J Hematol. 2012; 95(4): 386-93.

The close relationship between HbE and -thalassemia by Ma et al. 8  and in Thai homozygous -globin gene expression. This study was, however, the first to describe this unusual phenomenon. The motif within IVS2 of A -globin gene, XmnI and "C" motif, (CG)5(TG)8 motif within -globin gene, XmnI and XmnI+ motifs, confirming those previous findings. The motif, (AT)y motif in the -globin chain. Int J Hematol. 2012; 95(4): 386-93.

Acknowledgements

The authors would like to thank the Research Administration Center, Chiang Mai University for proofreading and reviewing the manuscript. They would also like to thank the Department of Hematology, Chiang Mai University for providing free laboratory services. They would like to thank the patients and their families for their cooperation. This study was supported by a grant from the National Research Council of Thailand. The authors would like to thank the Research Administration Center, Chiang Mai University for proofreading and reviewing the manuscript. They would also like to thank the Department of Hematology, Chiang Mai University for providing free laboratory services. They would like to thank the patients and their families for their cooperation. This study was supported by a grant from the National Research Council of Thailand.

References


28. Lanclos KD, Oner C, Dimovski AJ, Gu YC, Huisman THJ. Sequence variations in the 5’ flanking and IVS II regions of \(^0\gamma\) and \(^\gamma\) genes of \( \beta^6 \) chromosomes with five different haplotypes. Blood. 1991; 77(11): 2488-96.


38. Li Q, Harju S, Peterson KR. Locus control regions; coming of age at a decade plus. TIG. 1999: 403-8.