

DANIELA MARIA RIBEIRO

**INFLUÊNCIA DOS POLIMORFISMOS DO ELEMENTO
REGULATÓRIO MAIOR DOS GENES DO CLUSTER DA
GLOBINA α HUMANA NA EXPRESSÃO GÊNICA *IN VITRO***

CAMPINAS

2007

DANIELA MARIA RIBEIRO

**INFLUÊNCIA DOS POLIMORFISMOS DO ELEMENTO
REGULATÓRIO MAIOR DOS GENES DO CLUSTER DA
GLOBINA α HUMANA NA EXPRESSÃO GÊNICA *IN VITRO***

*Tese de Doutorado apresentada à Pós-Graduação da
Faculdade de Ciências Médicas da Universidade Estadual de
Campinas para obtenção do título de Doutor em Ciências
Médicas, área de concentração em Ciências Biomédicas.*

ORIENTADORA: PROFA. DRA. MARIA DE FÁTIMA SONATI

CAMPINAS

2007

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA FACULDADE DE CIÊNCIAS MÉDICAS DA UNICAMP**

Bibliotecário: Sandra Lúcia Pereira – CRB-8ª / 6044

R354i Ribeiro, Daniela Maria
Influência dos polimorfismos do elemento regulatório maior dos genes do cluster da globina α humana na expressão gênica *in vitro* / Daniela Maria Ribeiro. Campinas, SP : [s.n.], 2007.

Orientador : Maria de Fátima Sonati
Tese (Doutorado) Universidade Estadual de Campinas. Faculdade de Ciências Médicas.

1. Haplótipos. 2. Globina. I. Sonati, Maria de Fátima. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Título em inglês : Influence of the polymorphisms of the alpha-major regulatory element (α -MRE) on *in vitro* gene expression

Keywords: • Haplotypes
• Globins

Titulação: Doutor em Ciências Médicas
Área de concentração: Ciências Biomédicas

Banca examinadora: Profa. Dra Maria de Fátima Sonati
Prof Dr Francisco Mauro Salzano
Prof Dr Marcio José da Silva
Profa. Dra. Maricilda Palandi de Mello
Profa. Dra. Mônica Barbosa de Melo

Data da defesa: 24 - 08 - 2007

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA FACULDADE DE CIÊNCIAS MÉDICAS DA UNICAMP**

Bibliotecário: Sandra Lúcia Pereira – CRB-8ª / 6044

UNIDADE BC

Nº CHAMADA:

T/UNICAMP R354i

V. _____ EX. _____

TOMBO BCCL 75133

PROC 16.145-07

C _____ D X

PREÇO 11,00

DATA 05/12/07

BIB-ID 417874

R354i

Ribeiro, Daniela Maria

Influência dos polimorfismos do elemento regulatório maior dos genes do cluster da globina α humana na expressão gênica *in vitro* / Daniela Maria Ribeiro. Campinas, SP : [s.n.], 2007.

Orientador : Maria de Fátima Sonati

Tese (Doutorado) Universidade Estadual de Campinas. Faculdade de Ciências Médicas.

1. Haplótipos. 2. Globina. I. Sonati, Maria de Fátima. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Título em inglês : Influence of the polymorphisms of the alpha-major regulatory element (α -MRE) on *in vitro* gene expression

Keywords: • Haplotypes

• Globins

Titulação: Doutor em Ciências Médicas

Área de concentração: Ciências Biomédicas

Banca examinadora: Profa. Dra Maria de Fátima Sonati

Prof Dr Francisco Mauro Salzano

Prof Dr Marcio José da Silva

Profa. Dra. Maricilda Palandi de Mello

Profa. Dra. Mônica Barbosa de Melo

Data da defesa: 24 - 08 - 2007

Banca examinadora da tese de Doutorado

Orientador(a): Prof(a). Dr(a). Maria de Fátima Sonati

Membros:

1. Profa. Dra. Maria de Fátima Sonati – 

2. Prof. Dr. Francisco Mauro Salzano – 

3. Prof. Dr. Márcio José da Silva – 

4. Profa. Dra. Mônica Barbosa de Melo – 

5. Profa. Dra. Maricilda Palandi de Mello – 

Curso de pós-graduação em Ciências Médicas da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

Data: 24/08/2007

17854811

DEDICATÓRIA

*À minha mãe, que sempre me mostrou a importância do saber
e é meu exemplo de vida, de amor e de fé.*

Ao meu namorado, pelo amor, compreensão e constante apoio.

AGRADECIMENTOS

A Deus, minha luz e fortaleza na realização deste trabalho.

À Profa. Dra. Maria de Fátima Sonati, pela confiança, carinho e amizade.

Aos professores Dr. Francisco Mauro Salzano, Dr. Marcio José da Silva, Dra. Maricilda Palandi de Mello e Dra. Mônica Barbosa de Melo, pela pronta disponibilidade em participar de minha banca e pelas precisas observações e sugestões.

Ao Prof. Dr. Fernando Ferreira Costa, pela participação no Projeto Temático.

Aos amigos do laboratório Andréia, Denise, Cíntia, Elza, Magnun, Márcia, Susan, Tânia e Vânia, pela grande amizade.

Em especial, aos amigos Denise, Fernanda, Magnun, Márcia, Priscila, Sula e Tânia, pelo sorriso e alegria que tornaram meus dias no laboratório muito felizes. À Tânia, por estar sempre presente nos momentos que precisei e pela sintonia que sempre tivemos. A Márcia, por ser tão companheira e dividir comigo as suas experiências. Ao Magnun, pela pronta disponibilidade em ajudar e por me fazer rir. A Denise, pela alegria e bondade. A Fernanda, Priscila e Sula, pelo carinho e longas conversas animadas.

Às minhas sempre amigas Denise Faustino, Nádia, Simone Bordignon, e Simone Sant'Anna que, apesar da distância, sempre estiveram presentes na minha vida.

À Alessandra, Ana Cristina, Fernando, Isa, Margareth, Rômulo e Rônei, pela amizade.

Aos meus queridos amigos Ariana, Cleber, Christian, Elaine, Eliane, Fabiane, Flávio, Guigo, Janaína, João, Mariângela, Melissa, Paulo, Priscila, Rodrigo, Waldemir, Xandão, e toda a minha turma, pela amizade, incentivo e momentos de muita alegria.

Em especial, as minhas amiguinhas do coração Ariana, Fabiane, Mariângela e Melissa, por serem minhas amigas de toda hora.

A minha mais nova amiga Luciana Miraldi, pelas longas conversas que tanto me ajudam.

Ao Márcio, Célia e sua família, pelo grande carinho e amizade em todos os momentos.

Ao meu namorado Carlos, que esteve ao meu lado durante toda a realização deste trabalho, sempre com uma palavra de ânimo e incentivo.

Aos meus cunhados e sobrinhos, pela alegria.

A minha mãe, por ter um coração tão grande.

Ao meu pai, que sempre torceu e, tenho certeza, que continua torcendo por mim.

A minha querida família, que soube entender os meus períodos de ausência e sempre me apoiou.

A Márcia e Carolina, do CIPOI/UNICAMP.

A Cleide Silva, do Serviço de Estatística da FCM/UNICAMP.

À FAPESP e CNPq, pelo suporte financeiro.

*“Aqueles que sonham durante o dia estão cientes de muitas coisas
que escapam àqueles que sonham apenas durante a noite”.*

Edgar Allan Poe

	<i>PÁG.</i>
RESUMO	<i>xi</i>
ABSTRACT	<i>xiii</i>
1- INTRODUÇÃO	15
1.1- As hemoglobinas humanas	16
1.2- Os genes das globinas	17
1.3- Regulação da Expressão dos Genes das Globinas	19
1.4- Polimorfismos do α-MRE	24
1.5- Polimorfismos de DNA	28
2- CAPÍTULOS	31
CAPÍTULO 1- POLYMORPHISMS OF THE MAJOR REGULATORY ELEMENT OF THE ALPHA-GLOBIN GENE CLUSTER (α-MRE) AFFECT IN VITRO GENE EXPRESSION	32
CAPÍTULO 2- REGULATION OF THE HUMAN ALPHA-GLOBIN GENES AND ALPHA-THALASSEMIA	49
3- DISCUSSÃO E CONCLUSÃO	64
4- REFERÊNCIAS BIBLIOGRÁFICAS	73
5- ANEXO	81

LISTA DE TABELAS

	<i>PÁG.</i>
Tabela 1- Haplótipos do α -MRE.....	26
Tabela 2- Frequência dos haplótipos do α -MRE nos grupos populacionais estudados.....	26
Tabela 3- Resultados dos experimentos de transfecção transiente em células K562.....	66

		PÁG.
Figura 1-	Molécula de Hemoglobina Humana.....	16
Figura 2-	<i>Cluster</i> β	18
Figura 3-	<i>Cluster</i> α	18
Figura 4-	Estruturas bi-dimensionais da <i>active chromatin hub</i>	23
Figura 5-	Sítios de ligação para proteínas regulatórias do α -MRE.....	25
Figura 6-	Vetor de expressão pGL2-Promoter.....	82
Figura 7-	Construções gênicas.....	83
Figura 8-	Vetor de expressão pRL-TK.....	84
Figura 9-	a) Vetor de expressão pGL2 – Basic.....	85
	b) Vetor de expressão pGL2 – Control.....	85
Figura 10-	Atividade da luciferase das construções gênicas em células K562.....	67
Figura 11-	Atividade relativa da luciferase das construções gênicas em células K562.....	68

RESUMO



A expressão dos genes do *cluster* da globina α em humanos é regulada pelo α -MRE (*α -Major Regulatory Element*), um elemento localizado 40 kb à montante do respectivo *cluster* na região telomérica do braço curto do cromossomo 16. O α -MRE é geneticamente polimórfico e seis diferentes haplótipos, nomeados de A a F, foram determinados em alguns grupos populacionais da África, Europa e Ásia e em indivíduos pertencentes a duas populações indígenas brasileiras. As substituições de base que resultaram nestes haplótipos estão localizadas entre sítios de ligação para fatores nucleares ou em um sítio não ocupado *in vivo*, exceto no caso do haplótipo D, em que o polimorfismo altera a sequência de um dos sítios ocupado *in vivo* pelo fator NF-E2. Não há, de nosso conhecimento, nenhum estudo experimental feito para avaliar se essa variabilidade pode influenciar a expressão gênica. Assim, no presente trabalho foi analisada e comparada a expressão do gene repórter da luciferase em células K562 transientemente transfectadas com construções que tiveram como *enhancers* os diferentes haplótipos do α -MRE, além de 3 dos polimorfismos isoladamente (+130, +199 e +209). Os resultados revelaram redução da expressão do gene da luciferase com todas as construções em relação à do haplótipo selvagem A: os haplótipos B e C corresponderam a 19% do grau de expressão do haplótipo A, o D a 21%, o E a 15%, o F a 3%, o polimorfismo +130 a 24%, o +199 a 32% e o +209 a 3%. Em seu conjunto, eles demonstram que os polimorfismos responsáveis pelos diferentes haplótipos do α -MRE, em sua maioria situados nas sequências flanqueadoras dos sítios de ligação para proteínas regulatórias, reduzem a expressão gênica *in vitro*.

ABSTRACT



The expression of human α -like globin genes is regulated by the α -MRE (α -Major Regulatory Element), an element located 40 kb upstream of the α cluster in the short arm of chromosome 16. The α -MRE is genetically polymorphic and six different haplotypes, named A to F, have been identified in some population groups from Europe, Africa and Asia and in native Indians from two Brazilian Indian tribes. The base substitutions that resulted in these haplotypes are located between binding sites for nuclear factors or in a site that is considered not to be active *in vivo*, with the exception of haplotype D, in which the polymorphism changes the first binding site for the NF-E2 factor occupied *in vivo*. To our knowledge, there are no experimental studies evaluating whether this variability may influence gene expression. Thus, the present work analyzed and compared the expression of the luciferase reporter gene in K562 cells transiently transfected with constructs that have, as enhancers, the different α -MRE haplotypes, besides three isolated polymorphisms (+130, +199 and +209). The results demonstrated a reduction in luciferase gene expression with all the constructs compared with the wild type α -MRE (A haplotype): the B and C haplotypes corresponded to 19% of the A haplotype expression, the D to 21%, the E to 15%, the F to 3%, the polymorphism +130 to 24%, the +199 to 32% and the +209 to 3%. They demonstrate that the polymorphisms responsible for the α -MRE haplotypes, most located in the flanking sequences of the regulatory protein binding sites, decrease *in vitro* gene expression.

1- INTRODUÇÃO

1.1- As hemoglobinas humanas

As hemoglobinas (Hb) humanas são proteínas tetraméricas compostas por duas cadeias polipeptídicas do "tipo α " (α ou ζ) e duas cadeias do "tipo β " (β , δ , γ ou ϵ). Cada cadeia globínica está associada a um grupo prostético heme (Figura 1) que possui um átomo de ferro divalente capaz de se ligar de modo reversível à molécula de oxigênio, transportando-a dos pulmões aos tecidos (Perutz et al., 1960; Bunn e Forget, 1986). Na fase embrionária, são produzidas as hemoglobinas Gower I ($\zeta_2\epsilon_2$), Gower II ($\alpha_2\epsilon_2$) e Portland I ($\zeta_2\gamma_2$). Durante o período fetal, a produção dessas hemoglobinas dá lugar à síntese de Hb F ($\alpha_2\gamma_2$), que, por sua vez, é substituída pelas Hbs A ($\alpha_2\beta_2$) e A₂ ($\alpha_2\delta_2$) no estágio adulto. Após os seis meses de vida, os indivíduos normais apresentam cerca de 95% de Hb A, 2-3% de Hb A₂ e 0-2% de Hb F (Bunn e Forget, 1986; Stamatoyannopoulos et al., 1994).

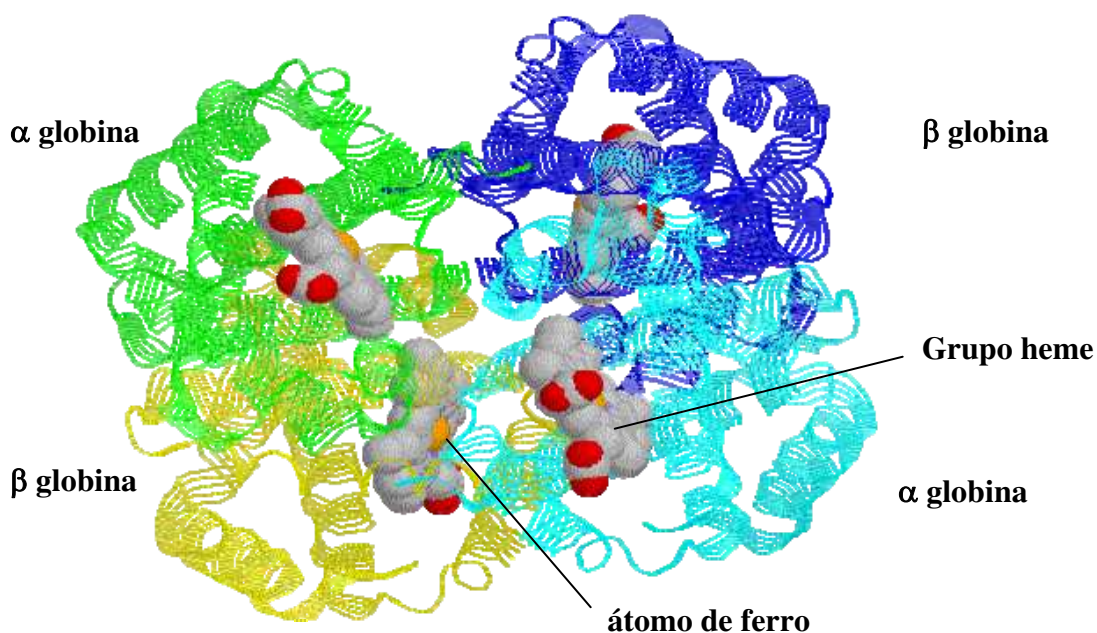


Figura 1- Molécula de Hemoglobina Humana. Cada globina está associada a um grupo prostético heme que possui um átomo de ferro central (figura adaptada do sítio <http://globin.cse.psu.edu/html/huisman/variants/image.html>).

1.2- Os genes das globinas

Os genes responsáveis pela codificação das cadeias globínicas são compactos, com cerca de 1 a 2 kilobases (kb) de DNA, formados por 3 exons e 2 introns e estão localizados em dois agrupamentos, ou *clusters*, denominados de α e β , onde se encontram dispostos na ordem em que são expressos durante o desenvolvimento do indivíduo.

O *cluster* β é composto pelo pseudogene β , pelo gene embrionário ϵ , pelos genes $^G\gamma$ e $^A\gamma$ e pelos genes δ e β , e está localizado no braço curto do cromossomo 11 (11p15.5) (Fritsch et al., 1980) (Figura 2).

O *cluster* α é formado pelo gene embrionário ζ , os pseudogenes $\psi\zeta$ e $\psi\alpha_1$, os genes da globina α duplicados (α_2 e α_1) e os genes θ e α^D (também conhecido como μ), e está localizado no braço curto do cromossomo 16 (16p13.3) (Higgs, 1993; Hugges et al., 2005) (Figura 3).

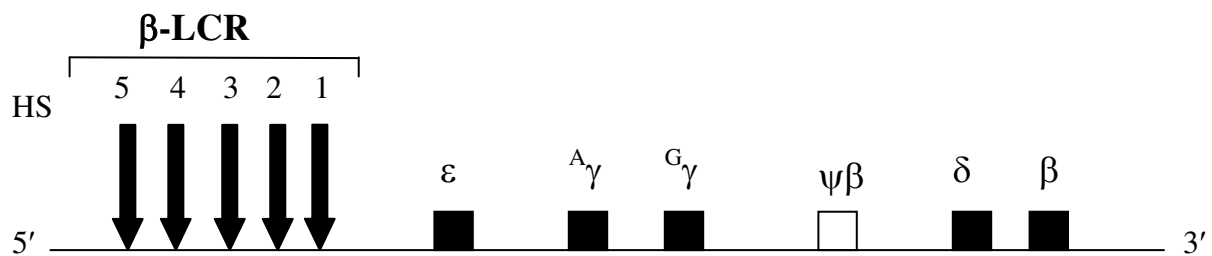


Figura 2- *Cluster beta*. Os quadrados pretos representam os genes funcionais e o branco o pseudogene, as setas representam as seqüências regulatórias associadas a cinco sítios hipersensíveis à DNase I (HS 1-5), o segmento de DNA que compreende os sítios é denominado Região de Controle do Locus β (β -LCR) (figura adaptada de Beutler et al., 2001).

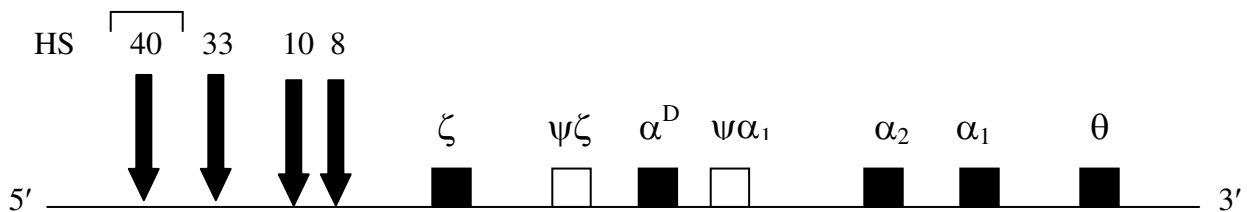


Figura 3- *Cluster alpha*. Os quadrados pretos representam os genes funcionais e os brancos os pseudogenes, as setas representam as seqüências regulatórias associadas a quatro sítios hipersensíveis à DNase I (HS 8, 10, 33 e 40), o sítio HS-40 é denominado Elemento Regulatório Maior dos genes do *cluster* da globina α (α -MRE) (figura adaptada de Beutler et al., 2001).

1.3- Regulação da expressão dos genes das globinas

A expressão dos genes globínicos em ambos os *clusters* é regulada de acordo com o estágio de desenvolvimento humano e de maneira tecido-específica. Esta complexa regulação depende da interação de proteínas nucleares com seqüências de DNA específicas que podem estar adjacentes aos genes nas regiões promotoras ou a vários kilobases de distância de cada *cluster* gênico nas seqüências regulatórias (Forrester et al., 1987; Grosveld et al., 1987; Higgs et al., 1990).

Na região promotora do gene se associam os fatores gerais de transcrição e a RNA polimerase II. A característica mais importante do promotor é o TATA *box*, uma seqüência curta de bases timina (T) e adenina (A), que é reconhecida pelo fator geral de transcrição TFIID, responsável pelo posicionamento da RNA polimerase II no sítio de início da transcrição. Este sítio está localizado aproximadamente 25 pares de bases (pb) após o TATA *box*. As seqüências regulatórias conferem altos níveis de transcrição gênica e possuem vários sítios de ligação para proteínas regulatórias (Tjian, 1995).

No *cluster* β , seqüências regulatórias estão associadas a cinco sítios hipersensíveis à DNase I (HS 1-5), específicos de células eritróides e localizados 4 a 20 kb à montante do gene ϵ . O segmento de DNA que compreende os sítios é denominado Região de Controle do Locus β (β -LCR) (Forrester et al., 1987; Grosveld et al., 1987) (Figura 2).

Similarmente, no *cluster* α foram identificados quatro sítios hipersensíveis à DNase I, localizados 8 (HS-8), 10 (HS-10), 33 (HS-33) e 40 (HS-40) kb à montante do gene ζ , em direção ao telômero (Higgs et al., 1990) (Figura 3). Estudos envolvendo a caracterização de deleções naturais, análises de híbridos interespecíficos, experimentos com transfecções estáveis e estudos em camundongos transgênicos indicam que somente a seqüência regulatória localizada 40 kb à montante do *cluster* (HS-40) exerce efeito significativo na expressão dos genes das globinas “tipo α ” (ζ e α) (Hatton et al., 1990; Liebhaber et al., 1990; Wilkie et al., 1990; Romao et al., 1991, 1992; Sharpe et al., 1992; Flint et al., 1994, 1996; Higgs et al., 1998; Viprakasit et al., 2003, 2006; Hartevelde et al., 2005). Esta seqüência foi denominada posteriormente Elemento Regulatório Maior dos genes do *cluster* da globina α (α -MRE) (Jarman et al., 1991).

Cada seqüência regulatória, associada a um sítio hipersensível à DNase I, contém um conjunto de sítios de ligação para proteínas regulatórias específicas de células eritróides e não eritróides (Philipsen et al., 1990; Talbot et al., 1990; Jarman et al., 1991; Pruzina et al., 1991; Talbot e Grosveld, 1991; Ikuta e Kan, 1991; Reddy e Shen, 1991; Strauss et al., 1992; Strauss e Orkin, 1992; Zhang et al., 1993).

Estudos vêm sendo realizados com o objetivo de compreender o funcionamento das seqüências regulatórias dos *clusters* α e β , ou seja, como estas seqüências conferem altos níveis de expressão, de maneira tecido-específica e de acordo com o estágio de desenvolvimento humano, a genes localizados a grandes distâncias. Devido à ancestralidade comum e à similar organização entre ambos os *clusters*, assumiu-se durante algum tempo que a regulação da expressão dos genes globínicos pelo β -LCR e α -MRE ocorria de maneira similar. Porém, estudos *in vivo* e *in vitro* sugeriram um papel diferente para estes elementos regulatórios na regulação gênica de seus respectivos *clusters* (Vyas et al., 1992; Craddock et al., 1995).

Além de atuar como um elemento *enhancer*, ativando a expressão dos genes do *cluster* β , o β -LCR tem a função adicional de alterar a estrutura da cromatina em que estes genes estão inseridos, para que eles possam ser transcritos. O *cluster* β está localizado em uma região do genoma rica nas bases A-T, de cromatina altamente condensada, onde se localizam genes ativos em poucos tipos celulares. Em células não eritróides, o domínio cromatínico onde o *cluster* β está contido é insensível à DNase I, replica-se tardiamente (durante a segunda metade da fase S do ciclo celular) e está transcricionalmente inativo. Em células eritróides, este domínio torna-se sensível à DNase I, replica-se logo no início da fase S e a transcrição gênica é ativada. As mudanças tecido-específicas na estrutura da cromatina são atribuídas ao β -LCR, que provoca sua descondensação, facilitando a interação dos fatores de transcrição ao DNA e possibilitando a expressão dos genes do *cluster* β (Craddock et al., 1995).

Diferentemente do β -LCR, o α -MRE não altera a estrutura da cromatina onde o *cluster* α está contido. Ele está localizado em uma região do genoma rica em guanina (G) e citosina (C), de cromatina descondensada tanto em células eritróides quanto em não

eritróides, onde se localizam genes ativos em todas as células, que é replicada logo no início da fase S do ciclo celular e transcricionalmente ativa. Assim sendo, a função principal do α -MRE é atuar como um elemento *enhancer*, ativando a expressão dos genes do *cluster* α (Vyas et al., 1992).

Nos últimos anos, diversos estudos vêm demonstrando que a estrutura da cromatina desempenha um papel crítico na expressão gênica. *In vivo*, os genes estão associados a proteínas histonas para formar a cromatina e os nucleossomas, estruturas repetitivas da cromatina formadas pelas histonas, que geralmente inibem a transcrição gênica. A ativação da transcrição envolve modificações na estrutura e/ou na posição dos nucleossomas nas regiões promotoras dos genes pela ação de grandes complexos de proteínas envolvidos na modulação da estrutura da cromatina (Vignali et al., 2000; Berger, 2002; Dean, 2004). Os complexos de proteínas remodeladoras da cromatina são recrutados para as regiões promotoras dos genes por proteínas nucleares que se associam às seqüências regulatórias. Estes complexos alteram, então, a estrutura cromatínica dos promotores, permitindo, assim, a ligação da maquinaria de transcrição (Elefant et al., 2000; Schubeler et al., 2000; Gui e Dean, 2001; Dean, 2004).

Sendo assim, o contato entre a região promotora e a seqüência regulatória é essencial para a ativação da transcrição. A formação de uma alça (*loop*) tem sido proposta para explicar como seqüências regulatórias distantemente posicionadas podem atuar (Bulger e Groudine, 1999). Na ativação da transcrição dos genes do *cluster* β em células eritróides, o β -LCR interage, via alça (*loop*), com os promotores dos genes ativos formando uma estrutura cromatínica única, conhecida como *active chromatin hub* (*ACH*) (Tolhuis et al., 2002; Patrinos et al., 2004) (Figura 4 - A). A formação desta estrutura cromatínica eritróide específica é também regulada de acordo com os diferentes estágios de desenvolvimento do indivíduo, uma vez que, em cada estágio de desenvolvimento, somente um gene interage com o β -LCR para formar uma *ACH* (Patrinos et al., 2004).

Diferentemente do *cluster* β , localizado em uma região do cromossomo onde se encontram genes ativos em poucos tipos celulares, o *cluster* α é cercado por genes ativos em todas as células; sendo assim, uma importante questão em relação a este *cluster* é

compreender como o elemento regulatório controla a expressão dos genes “tipo α ” sem a interferência dos genes vizinhos (Zhou et al., 2006). Durante algum tempo, pensou-se que haveria a formação de uma estrutura cromatínica diferente daquela formada no *cluster* β ; no entanto, estudos recentes de conformação da cromatina do *cluster* α de camundongos demonstraram que os genes vizinhos a este *cluster* mantêm no núcleo uma estrutura cromatínica permanente, similar àquela formada no *cluster* β , devido a sua expressão constitutiva. O recrutamento dos promotores dos genes ativos do *cluster* α e do elemento regulatório para esta *ACH* resulta na expressão dos genes globínicos em células eritróides murinas (Figura 4 - B), já a exclusão dos genes das globinas “tipo α ” e de seu elemento regulatório desta estrutura cromatínica em células não eritróides murinas, inativa a expressão gênica (Figura 4 - C) (Zhou et al., 2006).

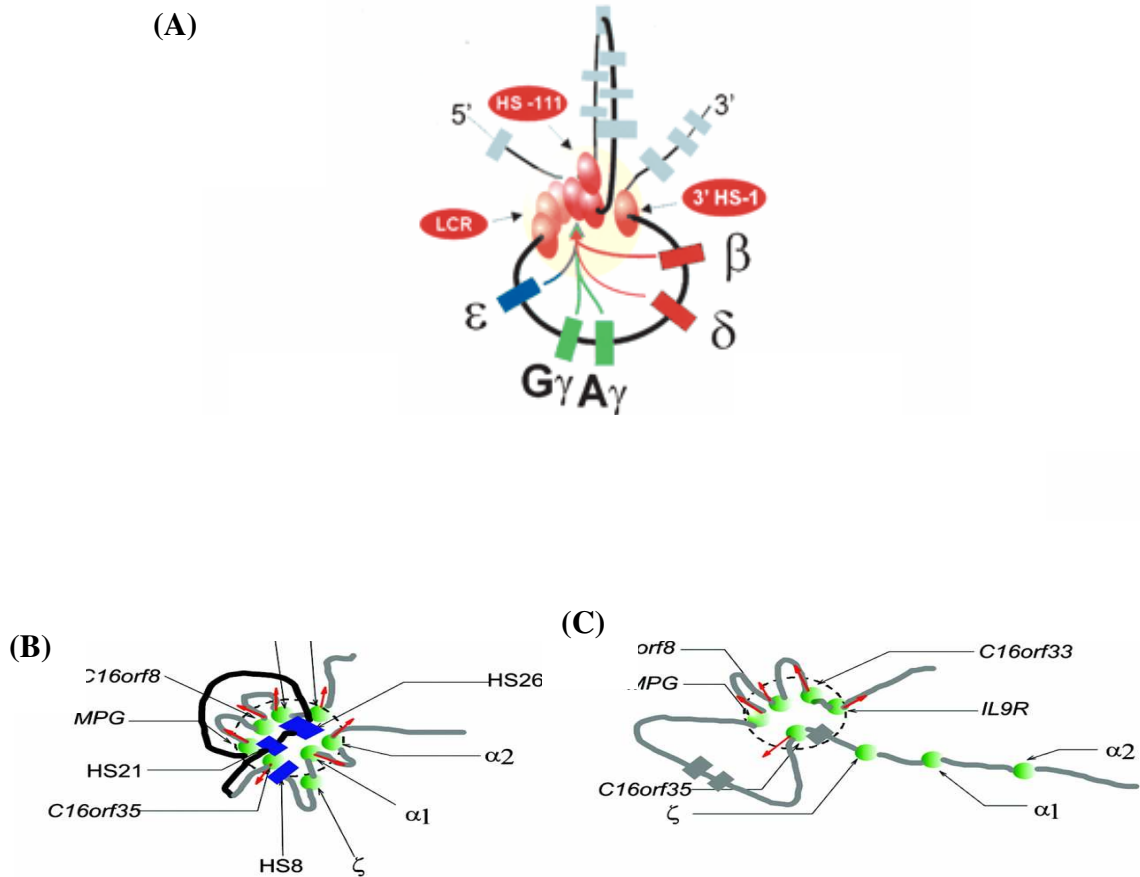


Figura 4- Estruturas bi-dimensionais da *active chromatin hub* (ACH). (A) Representação da β -ACH em células eritróides humanas. O β -LCR (representado por elipses vermelhas) interage com os promotores ativos dos genes do *cluster* β (indicados por retângulos azuis, verdes e vermelhos), formando a β -ACH (representada pela esfera amarela) (Patrinos et al., 2004); (B) Representação da α -ACH em células eritróides murinas; (C) Representação da α -ACH em células não eritróides murinas. No painel B, os promotores ativos dos genes do *cluster* α (representados pelas pequenas esferas verdes) interagem com os sítios HS (indicados por retângulos azuis), formando a α -ACH (representada pela esfera tracejada). A α -ACH é compartilhada pelos genes vizinhos ao *cluster* α (representados também por esferas verdes). No painel C, os genes do *cluster* α (esferas verdes) e os sítios HS (indicados por retângulos cinzas) são excluídos da α -ACH (Zhou et al., 2006).

1.4- Polimorfismos do α -MRE

A atividade do elemento regulatório α -MRE está restrita a um fragmento central de DNA de aproximadamente 350 pb, onde vários sítios de ligação para proteínas regulatórias têm sido identificados. Estes incluem quatro potenciais sítios para o fator de transcrição eritróide-específico GATA-1, nomeados de A a D, dois para o fator eritróide-específico NF-E2, denominados 5' NF-E2 e 3' NF-E2, e quatro *boxes* CACC para a proteína SP1, nomeados de I a IV (Jarman et al., 1991) (Figura 5). Estudos *in vivo* demonstraram a ocupação de três dos quatro sítios para o fator GATA-1 (B, C e D), de ambos os sítios NF-E2 (5' e 3') e de um *box* CACC (II) em células eritróides (Strauss et al., 1992). Estes sítios de ligação foram bem conservados durante a evolução, enquanto as seqüências externas aos mesmos mostraram uma menor conservação, onde polimorfismos são freqüentemente encontrados (Kielman et al., 1994; Kielman, 1996).

GATA-1 (A)

5' - TCGACCCTCTGGAAC CTATCA GGGACCACAGTCAGCCAGGCAAGC
 3' - AGCTGGGAGACCTTGGATAGT CCCTGGTGTTCAG TCGGTCCGT TCG

CACC (I) +96

ACATCTGCCCAAGCCAAGG GTGG AGGCATGCAGCTGTGGGGGTCTG
 TGTAGACGGGTTTCGGTT CC CACC TCC GTACGTCGACACCCCCAGAC

+130 GATA-1 (B) 5'-NF-E2 +158

AAACACTTGAGGGAGC AGATAA CT GGGCC AACCA TGACT CAGT GC
 TTT GTGAACTCCC TCG TC TAT T GA CCCGGTT GGT ACTGA GTCA CG

3'-NF-E2 +199 CACC (II) CACC (III) +212

GGAGGCCAACAGGACT GCTGAGTCA TCCTGT **GGGG** GTGG AGGTGG
 CC TCCGG TTGTCCTGA CGACTCAGT AGGACACCCC CACC TC CACC

CACC (IV) GATA-1 (C) +209

GGACAAGGGAAAGGG GTGG AATGGTACTGC TGAT TA CAACCTCTGG
 CCT GTTCCC TT T CCC CACC T TACCATGACG ACTAAT GTTGGAGACC

GATA-1 (D)

TGCT GCCT CCCCCTCC T GT T TATCT GAGAGGGAA -3'
 ACGACGGAGGGGGAGGACA AATAGA CT CTCCC TT -5'

Figura 5- Sítios de ligação para proteínas regulatórias do α -MRE. As seqüências de consenso dos sítios estão sublinhadas, os polimorfismos identificados por Hartevelde et al. (2002) estão em negrito (figura adaptada de Zhang et al., 1993).

Os polimorfismos do elemento regulatório α -MRE foram primeiramente estudados por Hartevelde et al. (2002) em 7 diferentes grupos populacionais, entre africanos da língua Bantu, pigmeus africanos, holandeses, italianos, chineses, indianos e indonésios. Seis diferentes haplótipos do α -MRE, denominados de A a F (Tabela 1), foram encontrados. Estes haplótipos foram também estudados em indivíduos pertencentes a duas populações indígenas brasileiras: Parakanã e Xikrin (Ribeiro et al., 2003). As frequências dos haplótipos do α -MRE nos grupos populacionais estudados por Hartevelde et al. (2002) e Ribeiro et al. (2003) estão apresentadas na Tabela 2.

Tabela 1- Haplótipos do α -MRE.

	+96 ^a	+130	+158	+199	+209	+212
Haplótipo A	C	G	C	G	G	G
Haplótipo B	C	A ^b	C	A	G	G
Haplótipo C	C	A	C	A	A	G
Haplótipo D	C	G	T	G	G	G
Haplótipo E	C	G	C	G	G	C
Haplótipo F	A	G	C	G	G	G

^a Posições dos sítios polimórficos do α -MRE em relação ao sítio CAP do gene ζ

^b As letras em negrito indicam as substituições nucleotídicas em comparação ao haplótipo A

Fonte: Hartevelde et al., 2002

Tabela 2- Frequência dos haplótipos do α -MRE nos grupos populacionais estudados.

POPULAÇÕES	HAPLÓTIPOS					
	A	B	C	D	E	F
Holandesa (35)^a	0,43	0,57				
Italiana (53)	0,56	0,43	0,01			
Indiana (39)	0,67	0,32	0,01			
Chinesa (23)	0,74	0,26				
Indonésiana (34)	0,78	0,22				
Africana (32)	0,73	0,11		0,16		
Pigmeus (60)	0,76	0,16		0,04	0,02	0,02
Parakanã (70)	0,70	0,30				
Xikrin (95)	0,87	0,13				

^a Número de indivíduos estudados

Fonte: Hartevelde et al., 2002; Ribeiro et al., 2003.

O haplótipo A é o haplótipo selvagem e deriva da seqüência publicada por Jarman et al. (1991). O haplótipo B difere do A em duas substituições de pares de bases, nas posições +130 (G→A) e +199 (G→A). Os demais haplótipos derivaram dos haplótipos A ou B por uma simples substituição de base. O haplótipo C derivou do B por uma única substituição G→A na posição +209, enquanto os haplótipos D, E e F derivaram de A por substituições nas posições +158 (C→T), +212 (G→C) e +96 (C→A), respectivamente (Harteveld et al., 2002).

Os haplótipos A e B são freqüentes em holandeses e italianos (43%: 57% e 56%: 43%, respectivamente). Nas populações indiana, africana, chinesa, indonesiana, Parakanã e Xikrin, o haplótipo A é o predominante (67% a 87%), seguido do haplótipo B (11% a 32%). O haplótipo C é relativamente raro em italianos e indianos, com uma freqüência de 1%, e não foi encontrado nas demais populações estudadas. O haplótipo D foi observado somente entre africanos, com uma freqüência relativamente elevada (4 a 16%). Os haplótipos E e F foram observados exclusivamente em pigmeus africanos, apresentando uma freqüência relativamente baixa (2% cada) ao lado dos haplótipos A, B e D, mais predominantes (Harteveld et al., 2002; Ribeiro et al., 2003) (Tabela 2).

Os haplótipos A, B e D encontram-se associados a determinantes da talassemia α (3 deleções e 3 mutações de ponto). As deleções $-^{SEA}$ e $-^{MED}$ estão ligadas ao haplótipo A e a deleção $-\alpha^{3.7}$ está associada aos haplótipos A, B e D. Já as mutações de ponto α^{HphI} e α^{Cs} estão ligadas ao haplótipo A e a mutação α^{NcoI} está associada ao haplótipo B (Harteveld et al., 2002).

Espera-se que os polimorfismos do α -MRE não interfiram na regulação dos genes do *cluster* α , uma vez que os mesmos estão localizados entre sítios de ligação para fatores nucleares (polimorfismos +96, +130 e +199) ou em um sítio não ocupado *in vivo* (polimorfismos +209 e +212), exceto o polimorfismo +158, que constitui o haplótipo D, onde a substituição nucleotídica altera a seqüência de consenso do sítio 5'-NF-E2 de TGACTCAGT para TGACTTAGT (Figura 5) (Andrews et al., 1993).

1.5- Polimorfismos de DNA

Grande parte das variações genômicas é constituída de polimorfismos de nucleotídeos únicos (conhecidos em inglês como *SNPs* ou *Single Nucleotide Polymorphisms*), com frequências entre 1 e 99% e responsáveis pela diversidade genética entre os indivíduos. Os *SNPs* são resultantes de mutações pontuais, que podem estar presentes tanto em regiões codificadoras (*cSNPs*) como em introns e regiões reguladoras (*rSNPs*), são bialélicos e apresentam uma baixa taxa de mutação, tornando-se, portanto, bons indicadores para estudos evolutivos (Collins et al., 1998). Estima-se que, no genoma humano haja cerca de 1 *SNP* a cada 500-1.000 pares de bases e 6 *SNPs* por região codificadora por gene (Collins et al., 1998; Sherry et al., 1999; The International SNP Map Working Group, 2001).

O conhecimento da variabilidade genética entre os indivíduos pode ter um importante papel na elucidação da história e organização do genoma humano, bem como em estudos de associação visando correlacionar diferentes genótipos com susceptibilidade ou proteção a doenças e resposta a tratamentos específicos (Collins et al., 1998; Chakravarti, 2001; Wang e Moulton, 2001).

A identificação de *SNPs* na região codificadora de genes humanos pode auxiliar no entendimento das bases moleculares de doenças de característica genética multifatorial, como o câncer, aterosclerose, doenças autoimunes e neuropsiquiátricas. A presença de *SNPs* em regiões reguladoras pode alterar a afinidade de ligação de proteínas regulatórias aos seus respectivos sítios e, conseqüentemente, modificar a expressão gênica (Collins et al., 1998; Wang e Moulton, 2001; Knight, 2005; Wang et al., 2005).

A influência dos *SNPs* na expressão gênica foi recentemente ressaltada por um estudo demonstrando um mecanismo adicional causador de talassemia α : um *SNP* regulatório (*rSNP*), localizado em uma região não codificante entre os genes α e o elemento regulatório α -MRE, constitui um novo promotor que interfere na expressão normal dos genes α . A simples substituição de uma timina (T) por uma citosina (C) criou um novo sítio de ligação para o fator de transcrição GATA-1, eritróide-específico. A inserção deste novo elemento promotor altera a expressão dos genes α devido,

provavelmente, à sua interação preferencial com o elemento regulatório α -MRE, competindo com os promotores endógenos dos genes da globina α (Gobbi et al., 2006).

O interesse pelo estudo dos polimorfismos de DNA aumentou bastante nos últimos anos, como podemos observar pelo número crescente de *SNPs* depositados nos bancos públicos, ultrapassando os 9 milhões de *SNPs* (Wang et al., 2005). No entanto, ainda existem poucos trabalhos pesquisando o papel dos polimorfismos e haplótipos nas regiões reguladoras dos genes humanos e identificando aqueles que alteram a expressão gênica. Para demonstrar que um determinado polimorfismo modifica a expressão gênica são necessários experimentos *in vitro* e *in vivo* que comprovem essa influência. A grande maioria dos trabalhos baseia-se na análise dos efeitos dos polimorfismos regulatórios em linhagens celulares através de ensaios com genes repórteres, não os associando, porém, a variações da expressão gênica *in vivo* (Rockman e Wray, 2002; Buckland et al., 2004, 2005).

Os sistemas de gene repórter vêm sendo empregados com a finalidade de aumentar a eficiência e a sensibilidade dos experimentos *in vitro* de análise da expressão gênica. Nesses sistemas, seqüências promotoras ou regulatórias são clonadas em vetores de expressão e transfectadas em células somáticas. Os vetores de expressão contêm um gene codificador para uma proteína cuja expressão é determinada, geralmente, através de um ensaio enzimático; sua atividade corresponderá à transcrição do gene e, conseqüentemente, a força do promotor e/ou elemento regulatório estudado (Li et al., 2005).

Dentre os diversos sistemas de gene repórter, o que utiliza o gene da luciferase apresenta as vantagens de ser um método simples, rápido e bastante sensível. A luciferase é uma proteína de 62 kDa, estruturalmente simples (sem grupos prostéticos ou ligações com metais) que não requer nenhuma modificação pós-transcricional para exercer sua atividade enzimática, podendo ser expressa na grande maioria das células (Wood, 1984; de Went, 1985). Esta enzima catalisa uma reação de quimioluminescência produzindo luz, medida através de um luminômetro. Como os organismos geralmente não possuem luminescência endógena, não há interferência de sinal, e a medida da luz produzida reflete a produção de luciferase. Este sistema pode detectar menos que 10^{-20} moles de luciferase, o

que representa uma sensibilidade cerca de 100 vezes maior que a obtida utilizando outros sistemas de gene repórter como o CAT (*chloramphenicol acetyltransferase*), por exemplo (Williams et al., 1989).

Outra vantagem do sistema de gene repórter da luciferase é a possibilidade de se utilizar dois vetores de expressão simultaneamente. Há dois tipos de luciferase: a Firefly luciferase e a Renilla luciferase, a primeira é derivada do organismo *Photinus pyralis* e a segunda do *Renilla reniformis*, sendo enzimas diferentes que utilizam substratos diferentes, podendo, portanto, serem usadas conjuntamente. Os genes codificadores para as duas enzimas podem estar no mesmo vetor, ligados a promotores diferentes, ou em vetores diferentes que podem ser co-transfectados em uma mesma célula (Promega, 2005).

Tendo em vista que os polimorfismos do α -MRE foram muito pouco investigados a nível mundial e, de nosso conhecimento, não há nenhum estudo experimental realizado para investigação do efeito de sua variabilidade na expressão de genes sob seu controle, o objetivo do presente estudo foi analisar a influência dos polimorfismos do elemento regulatório maior dos genes do *cluster* da globina α humana na expressão gênica *in vitro*, através do sistema de gene repórter da luciferase, em células da linhagem eritroleucêmica humana K562 (Capítulo 1).

E, considerando que a hemoglobina e os genes de globina são, de longa data, importantes modelos para estudo e determinação da estrutura, da função e dos mecanismos de regulação de outros genes e proteínas, concomitantemente aos experimentos de expressão e frente às várias descrições de deleções nesta região regulatória que resultam em talassemia α , julgamos pertinente a elaboração de um artigo de revisão sobre o tema (Capítulo 2).

2- CAPÍTULOS

CAPÍTULO 1

POLYMORPHISMS OF THE MAJOR REGULATORY ELEMENT OF THE ALPHA-GLOBIN GENE CLUSTER (α -MRE) AFFECT *IN VITRO* GENE EXPRESSION

DANIELA M. RIBEIRO; TANIA R. ZACCARIOTTO; MAGNUN N. N. SANTOS;
FERNANDO F. COSTA and MARIA F. SONATI

From the Department of Clinical Pathology (D.M.R.; T.R.Z.; M.N.N.S.; M.F.S.) and
Department of Clinical Medicine (F.F.C.), School of Medical Sciences, State University of
Campinas-UNICAMP, Campinas (SP), Brazil

Address for correspondence: Prof. Maria F. Sonati, Department of Clinical Pathology -
School of Medical Sciences - State University of Campinas – UNICAMP – P.O. Box 6111
– ZIP Code 13083-970 – Campinas – State of São Paulo (SP) – Brazil – Phone (55-19)
3521-9451 – Fax (55-19) 3521-9434 – E-mail: sonati@fcm.unicamp.br (preferred) /
sonati_mf@yahoo.com.br (alternative)

KEYWORDS: α -MRE haplotypes, α -globin genes, genetic polymorphisms, gene
expression

ABSTRACT

The α -MRE is the major regulatory element responsible for the expression of human α -like globin genes. It is genetically polymorphic, and different haplotypes have been identified in some population groups from Europe, Africa and Asia and in native Indians from two Brazilian Indian tribes. Most of the α -MRE polymorphisms are located between binding sites for nuclear factors or in a site that is considered not to be active *in vivo*. To our knowledge, there are no experimental studies evaluating whether this variability may influence gene expression. The aim of the present work was to analyze and compare the expression of the luciferase reporter gene in K562 cells transiently transfected with constructs that have the different α -MRE haplotypes as enhancers in addition to the SV40 promoter. We observed a reduction in luciferase gene expression with all the constructs compared with the wild type α -MRE (A haplotype), thus demonstrating that these polymorphisms, most located in the flanking sequences of the nuclear factor binding sites, decrease *in vitro* gene expression.

INTRODUCTION

Human hemoglobin (Hb) is a globular protein that consists of two α -like and two β -like globin polypeptide chains. The α -like globin gene family is clustered on chromosome 16, and the transcriptionally active members of this cluster are arranged in the order of their expression during erythroid development: 5'- ζ (embryonic) - α_2 - α_1 - (fetal/adult) - 3'.¹

The appropriate erythroid lineage-specific and developmental stage-specific expression of the α -like globin genes is controlled by an element located 40 kb upstream of the ζ -globin gene; this erythroid-specific DNase I-hypersensitive site is known as HS-40 or the α -Major Regulatory Element (α -MRE).^{2,3} Experimental studies have demonstrated that the α -MRE confers high levels of expression of the α -like globin genes and behaves as an enhancer for the promoter activity of these genes.³⁻¹⁰

A core 350-base pair (bp) fragment retains most of the activity of the α -MRE and contains a high density of well-conserved GATA-1 and NF-E2 nuclear factor binding sites, as well as several CACC motifs.³ *In vivo* binding studies demonstrated that six binding sites are occupied with nuclear factors in an erythroid lineage-specific and developmental stage-specific manner: three GATA-1 sites (B, C and D), two NF-E2 sites (5' and 3') and one CACC box (II) (figure 1).¹¹

The α -MRE is genetically polymorphic, and its variability was studied by Harteveld et al. in seven population groups from Africa (Bantu-speaking and Pygmies), Europe (Dutch and Italians) and Asia (Chinese, East Indians and Indonesians).¹² Six different haplotypes, named A to F, were found (table 1). Two Brazilian Indian tribes (Parakanã and Xikrin) were also investigated,¹³ revealing the presence of the A and B haplotypes. These polymorphic sites are not expected to interfere with regulation of the α cluster, since, with the exception of haplotype D, they are located either between binding sites for nuclear factors (polymorphisms +96, +130 and +199) or in a site considered not to be active *in vivo* (polymorphisms +209 and +212). In haplotype D, the base substitution at position +158 changes the first consensus binding site for the NF-E2 factor (Figure 1).¹⁴ To our knowledge, however, there are no experimental studies that evaluate whether the polymorphisms responsible for the different α -MRE haplotypes are able to influence its enhancer activity. This was the aim of the present study.

MATERIAL AND METHODS

Plasmid constructs

Vectors for transient DNA transfection of human K562 erythroleukemia cell lines were constructed according to the following procedure.

The α -MRE core fragment (350 bp) was first amplified by polymerase chain reaction (PCR) as described elsewhere,¹³ using genomic DNA from two controls: one with the wild type α -MRE (A haplotype) and the other with the B haplotype. These fragments were used to generate the other α -MRE haplotypes (C, D, E and F) and three other

elements containing the polymorphisms +130, +199, and +209 separately by site-directed mutagenesis in a two-step PCR.¹⁵ In the first step, primer pairs A/D and B/C were used to amplify overlapping halves of the mutant segment; the two PCR products were combined and the full-length mutant sequence was amplified with primer pair A/B. The products were sequenced to verify if any additional mutations had been introduced during amplification. The sequences of the oligonucleotides used are shown in table 2.

The different amplified elements were cloned by blunt-end ligation into the *Sma* I cloning site of plasmid pGL2-Promoter (Promega, USA), which consists of an SV40 promoter linked to the Firefly luciferase reporter gene. The enhancer activity of the α -MRE on the SV40 promoter is similar to its activity on α -globin gene promoters.⁷

Plasmid DNA was grown in DH5 α bacteria and purified using Qiagen Plasmid Maxiprep (Qiagen, Germany). The correct insertion and sequences of the constructs were confirmed by sequencing them. Only those constructs in which the α -MRE was in the same genomic orientation as the SV40 promoter were used in the transfection assays.

Nine constructs were obtained: one with the A haplotype, five with the other α -MRE haplotypes (B, C, D, E and F haplotypes) and three with the polymorphisms +130, +199 and +209 separately.

Cell culture and DNA transfection assay

The K562 cells were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum and penicillin-streptomycin (all from GIBCO, USA) at 37°C and 5% CO₂ at a density of 1 x 10⁶ cells per mL.¹⁶ Approximately 8 x 10⁶ cells were transfected in suspension by electroporation (250 volts and 975 μ F for 9 ms) with 20 μ g of each construct and 1 μ g of pRL-TK plasmid (Promega, USA).¹⁷ The vector pRL-TK, which expresses Renilla luciferase, was used as an internal control for electroporation.

Luciferase assays

After transfection the cells were cultivated, and harvested after 24 hours. The Firefly and Renilla luciferase activities were measured sequentially in the cell lysates using a luminometer (Turner Designs, USA) and the Dual Luciferase Reporter Assay (Promega,

USA). The light output was measured for 15 seconds. The vectors pGL2-Basic and pGL2-Control (both from Promega, USA) were used as negative and positive controls, respectively, for the expression of luciferase (data not shown).

Data from the Firefly luciferase activity, expressed from the constructs, were normalized against the Renilla luciferase activity, expressed from co-transfected pRL-TK vector. All electroporations were performed in triplicate, and each experiment was repeated three times to obtain the average expression level for each construct. The results are given in relative luminescence unit (RLU).

RESULTS

To assess whether the α -MRE polymorphisms modify the luciferase gene expression, wild type and mutant constructs were co-transfected with a control vector into K562 cells. The results of the transient transfection experiments are summarized in table 3 and in figure 2.

To compare the enhancer activities of the different α -MRE elements, the median values of the luciferase activity of the constructs are shown as a percentage of the activity of construct containing the wild type α -MRE (A haplotype), defined as 100% (figure 3). Median was chosen here because it better reflects the variability of these experiments.

A noticeable decrease in luciferase activity was seen with all the enhancers, but the magnitude of the decrease varied with the type of α -MRE present in the construct. With the B haplotype, luciferase activity was reduced by 81% compared with that for the wild type α -MRE. The B haplotype differs from the A one in two polymorphisms (+130 and +199), both located near the nuclear factor-binding sites (GATA-1 and SP1) occupied *in vivo*. Analyses of the constructions in which the polymorphisms +130 and +199 were present separately revealed a reduction in expression of 76% and 68%, respectively.

The construction with the C haplotype, which differs from the B haplotype in one polymorphism (at position +209), also reduced the luciferase activity by 81% compared with the wild type α -MRE. However, the construct containing only the polymorphism +209, located at an SP1 protein site believed not to be active *in vivo*, repressed the enhancer function of the α -MRE by 97%.

Haplotype D, in which the base substitution at position +158 changes the site 5'-NF-E2 occupied *in vivo*, and haplotype E, which contains the polymorphism +212 located at an SP1 protein site not occupied *in vivo*, caused a decrease in luciferase activity of 79% and 85%, respectively, compared with the A haplotype.

Very little luciferase activity was also produced by the construct with the F haplotype. The polymorphism at position +96, near an SP1 protein binding site that is not active *in vivo*, reduced the α -MRE enhancer function in K562 cells by 97%.

DISCUSSION

We analyzed the effects of α -MRE polymorphisms on *in vitro* gene expression. This kind of study allowed the simultaneous comparison of all polymorphisms described so far. In the transient expression assay employed here, the polymorphisms, most of which are located between nuclear protein binding sites, substantially decreased luciferase reporter gene expression.

DNA polymorphisms are genomic variations present in more than 1% of a population and are associated with genetic diversity.¹⁸ Although the majority are functionally neutral, some have been associated with susceptibility to diseases and individual responses to therapeutic treatments.^{19,20} Greater emphasis has been placed on the polymorphisms occurring in gene regulatory regions such as promoters, enhancers, silencers and introns, in view of their important role in modulating gene expression.²¹ A recent study identified a regulatory polymorphism in a non-genic region between the α -

globin genes and the α -MRE that results in α -thalassemia by creating a new promoter-like element that interferes with normal activation of the downstream α -like globin genes.²²

Gene expression in eukaryotes involves many strictly controlled steps. However, its major regulation is achieved at the level of transcription, when gene promoters and regulatory elements are bound by regulatory proteins, which recruit the transcriptional machinery.^{20,23} These protein-DNA complexes are important in the formation of the looping structures proposed in a model explaining how the distant α -MRE element activates the α -like globin genes;²⁴ they also appear to have a role in the modulation of the promoter chromatin structure, which is important for gene expression during development and differentiation.²⁴⁻²⁶

Mutations that affect DNA-binding and transcriptional activation properties of one or more of these protein complexes may alter the configuration and function of the regulatory elements and, consequently, gene expression.^{9,16,20} Little is known about sequence variation in the binding sites of regulatory proteins, and far less is known about variations in the flanking sequences of these binding sites.^{16,20}

Most of the polymorphisms analyzed in this study are located in the flanking sequences of the protein binding sites that constitute the functional core of the α -MRE element, which is responsible for interacting with the gene promoters during the final step of the transcriptional activation in K562 cells. In the D haplotype, the base substitution occurs in the 5'-NF-E2 site. The transient expression assay employed reflects the structure-function relationship during the final stage of transcriptional activation. In this step, gene promoters and regulatory elements are already physically associated.²⁷

Reporter gene assays are a powerful tool for analyzing the effects of regulatory polymorphisms on gene expression; they are sensitive, reproducible and specific. Some caution, however, is necessary when interpreting their results since they do not always represent exactly what is happening *in vivo*.²⁸ The DNA sequences are not in the native chromatin configuration and the results have been found to be dependent on the choice of cell type to transfect, the mode and efficiency of transfection and the plasmids used in the constructs.¹⁹

If the *in vivo* expression of the human α -globin genes were also influenced by the different α -MRE polymorphisms, α -thalassemia would be presented by the B to F haplotype carriers. Hartevelde et al. analyzed the hematological parameters of two homozygous for both the $-\alpha^{3.7}$ rightward deletion and the D haplotype and did not find any significant differences from the expected values.¹² However, only a small number of individuals and populations have been investigated to date.

CONCLUSION

In summary, our results represent the first analysis of the influence of α -MRE polymorphisms on *in vitro* gene expression and demonstrate that the simple base substitutions responsible for its haplotypes decrease gene control even outside of the nuclear factor binding sites. Further studies, particularly involving populations under malarial selective pressure, and experiments aimed at identifying DNA-protein interactions at these polymorphic sites should help to understand how these polymorphisms influence gene regulation.

ACKNOWLEDGEMENTS

We are grateful to Ms. Marcia C. Fornazim and Ms. Maria Carolina Spago of CIPOI/UNICAMP for the K562 cells; Ms. Angela M. Assis, Mr. Anderson F. Cunha and Ms. Ana Flavia Brugneroto of the HEMOCENTRO/UNICAMP for their technical assistance; Ms. Cleide Silva of the Statistics Advisory Service of UNICAMP Medical School for statistical analysis and Mr. Marcio J. Silva of CBMEG/UNICAMP for the luminometer. Funding was provided by FAPESP/Brazil (fellowship no 03/07412-0 and grants no 02/13801-7) and CNPq/ Brazil (grants no 475481/2006-2).

REFERENCES

1. Higgs DR, Vickers MA, Wilkie AO, Pretorius IM, Jarman AP, Weatherall DJ (1989) A review of the molecular genetics of the human α -globin gene cluster. *Blood* 73: 1081-1104.
2. Higgs DR, Wood WG, Jarman AP, Sharpe J, Lida J, Pretorius IM, Ayyub H (1990) A major positive regulatory region located far upstream of the human α -globin gene locus. *Genes and Dev* 4: 1588-1601.
3. Jarman AP, Wood WG, Sharpe JA, Gourdon G, Ayyub H, Higgs DR (1991) Characterization of the major regulatory element upstream of the human α -globin gene cluster. *Mol Cell Biol* 11: 4679-4689.
4. Pondel MD, George M, Proudfoot NJ (1992) The LCR-like α -globin regulatory element functions as an enhancer in transiently transfected cells during erythroid differentiation. *Nucleic Acids Res* 20: 237-243.
5. Sharpe JA, Chan-Thomas PS, Lida J, Ayyub H, Wood WG, Higgs DR (1992) Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. *EMBO J* 11: 4565-4572.
6. Vyas P, Vickers MA, Simmons DL, Ayyub H, Craddock CF, Higgs DR (1992) Cis-acting sequences regulating expression of the human alpha-globin cluster lie within constitutively open chromatin. *Cell* 69: 781-793.
7. Ren S, Luo X, Atweh GF (1993) The major regulatory element upstream of the α -globin gene has classical and inducible enhancer activity. *Blood*, 81 (4): 1058-1066.
8. Sharpe JA, Summerhill RJ, Vyas P, Gourdon G, Higgs DR, Wood WG (1993) Role of upstream DNase I hypersensitive sites in the regulation of human α -globin gene expression. *Blood* 82: 1666-1671.
9. Zhang Q, Reddy PM, Yu CY, Bastiani C, Higgs DR, Stamatoyannopoulos G, Papayannopoulou T, Shen CK (1993) Transcriptional activation of human ζ_2 globin promoter by the α globin regulatory element (HS-40): functional role of specific nuclear factor-DNA complexes. *Mol Cell Biol* 13: 2298-2308.

10. Gourdon G, Sharpe JA, Wells D, Wood WG, Higgs DR (1994) Analysis of a 70 kb segment of DNA containing the human ζ and α -globin genes linked to their regulatory element (HS-40) in transgenic mice. *Nucleic Acids Res* 22: 4139-4147.
11. Strauss EC, Andrews NC, Higgs DR, Orkin SH (1992) In vivo footprinting of the human α -globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol Cell Biol* 12: 2135-2142.
12. Harteveld CL, Muglia M, Passarino G, Kielman MF, Bernini LF (2002) Genetic polymorphism of the major regulatory element HS-40, upstream of the human α -globin gene cluster. *Br J Haematol* 119 (3): 848-854.
13. Ribeiro DM, Figueiredo MS, Costa FF, Sonati MF (2003) Haplotypes of alpha-globin gene regulatory element in two Brazilian native populations. *Am J Phys Anthropol* 121 (1): 58-62.
14. Andrews NC, Erdjument-Bromage H, Davidson MB, Tempst P, Orkin SH (1993) Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 362: 722-728.
15. Sabath DE, Koehler KM, Yang WQ, Patton K, Stamatoyannopoulos G (1995) Identification of a major positive regulatory element located 5' to the human ζ -globin gene. *Blood* 85: 2587-2597.
16. Wen SC, Roder K, Hu KY, Rombel I, Gavva NR, Daftari P, Kuo YY, Wang C, Shen CKJ (2000) Loading of DNA-binding factors to an erythroid enhancer. *Mol Cell Biol* 20: 1993-2003.
17. Takahashi T, Schreiber R, Krieger JE, Saad ST, Costa FF (2003) Analysis of the mechanism of action of the Brazilian type of hereditary persistence of fetal hemoglobin. *Eur J Haematol* 71: 418-424.
18. Collins FS, Brooks LD, Chakravati A (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 8: 1229-1231.
19. Knight JC (2005) Regulatory polymorphisms underlying complex disease traits. *J Mol Med* 83: 97-109.

20. Wang X, Tomso DJ, Liu X, Bell DA (2005) Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes. *Toxicology and Applied Pharmacology* 207: 84-90.
21. Ponomarenko JV, Orlova GV, Merkulova TI, Gorshkova EV, Fokin ON, Vasiliev GV, Frolov AS, Ponomarenko MP (2002) rSNP guide: an integrated database tools system for studying SNPs and site-directed mutations in transcription factor binding sites. *Hum Mutat* 20: 239-248.
22. Gobbi MD, Viprakasit V, Hughes JR, Fisher C, Buckle VJ, Ayyub H, Gibbons RJ, Vernimmen D, Yoshinaga Y, Jong P, Cheng JF, Rubin EM, Wood WG, Bowden D, Higgs DR (2006) A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 312: 1215-1217.
23. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA (1998) Dissecting the regulatory circuitry of a eucaryotic genome. *Cell* 95: 717-728.
24. Zhou GL, Xin L, Song W, Di LJ, Liu G, Wu XS, Liu DP, Liang CC (2006) Active chromatin hub of the mouse α -globin locus forms in a transcriptional factory of clustered housekeeping genes. *Mol Cell Biol* 26: 5096-5105.
25. Dean A (2004) Chromatin remodeling and the interactions between enhancers and promoters in the β -globin locus. *Briefings in Functional Genomics and Proteomics* 2: 344-354.
26. Patrinos GP, Krom M, Boer E, Langeveld A, Ali Imam AM, Strouboulis J, Laat W, Grosveld FG (2004) Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes and Dev* 18: 1495-1509.
27. Zhang Q, Rombel I, Reddy N, Gang JB, Shen CKJ (1995) Functional roles of in vivo footprinted DNA motifs within an α globin enhancer. *J Biol Chem* 270: 8501-8505.
28. Li Q, Emery DW, Han H, Sun J, Yu M and Stamatoyannopoulos G (2005) Differences of globin transgene expression in stably transfected cell lines and transgenic mice. *Blood* 105 (8): 3346-3352.

Table 1. α -MRE haplotypes.

	+96 ^a	+130	+158	+199	+209	+212
Haplotype A	C	G	C	G	G	G
Haplotype B	C	A ^b	C	A	G	G
Haplotype C	C	A	C	A	A	G
Haplotype D	C	G	T	G	G	G
Haplotype E	C	G	C	G	G	C
Haplotype F	A	G	C	G	G	G

Note - Source: Hartevelde et al.¹²

^a Positions of the α -MRE polymorphic sites in relation to the CAP-site of the ζ gene

^b Bold letters indicate the nucleotide substitutions in relation to the A haplotype

Table 2. Primers used in site-directed mutagenesis.

Primers	Sequence
A ^a : HM13- α MRE	5'- GCC GGG AGG CTC TCA GGA ACA AGA-3'
B ^b : HBIO- α MRE	5'- CCG CAG CCT GGC TGT GAA CAC TTT -3'
C ^a : P96s	5'- AAG GGT GGA GGA ^c ATG CAG CTG TGG-3'
D ^b : P96as	5'- CCA CAG CTG CAT TCC ACC CTT-3'
C: P130s	5'- AAA CAC TTG AGA GAG CAG ATA ACT-3'
D: P130as	5'- AGT TAT CTG CTC TCT CAA GTG TTT-3'
C: P199s	5'- AGT CAT CCT GTA GGG GTG GAG GTG-3'
D: P199as	5'- CAC CTC CAC CCC TAC AGG ATG ACT-3'
C: P209s	5'- TGG GGG TGG AGA TGG GAC AAG GGA-3'
D: P209as	5'- TCC CTT GTC CCA TCT CCA CCC CCA-3'
C: P212s	5'- GGG TGG AGG TGC GAC AAG GGA AAG-3'
D: P212as	5'- CTT TCC CTT GTC GCA CCT CCA CCC-3'

Note - Source: Hartevelde et al.¹²

^a A and C indicate the forward primers

^b B and D indicate the reverse primers

^c Bold letters indicate the nucleotide substitutions in relation to the wild-type sequence.

Table 3. Results of the transient transfection experiments in K562 cells.

Genic Constructions	N	Mean ^a	Standard Deviations ^a	Minimum ^a	Median ^a	Maximum ^a
A	3	7.998	3.794	5.756	5.859	12.378
B	3	2.676	3.190	0.555	1.129	6.344
C	3	1.037	0.543	0.450	1.140	1.521
D	3	1.339	1.116	0.282	1.230	2.506
E	3	1.110	0.982	0.249	0.901	2.180
F	3	0.244	0.205	0.069	0.193	0.469
+130	3	1.677	1.090	0.748	1.404	2.877
+199	3	1.459	1.244	0.056	1.892	2.430
+209	3	0.198	0.072	0.154	0.158	0.281

Note

^a Values are given in relative luminescence unit (RLU)

GATA-1 (A)

5'- TCGACCCT CT GGAAC CTATCA GGGACCACAGTCAGCCAGGCAAGC
 3'- AGCTGGGAGACC TTG GATAGT CCCT GGTGTCAGTCGGTCC GTTCG

CACC (I) +96

ACATCTGCCCAAGCCAAGG GTGG **AGGCATGC**AGCTGTGGGGGTCTG
 TGTAGACGGGTTCGGTT CC CACC TCC **GTACGTCGACAC**CCCCAGAC

+130 GATA-1 (B) 5'-NF-E2 +158

AAACACTTGAG**GG**GAGC AGATAA CTGGGCCAACCA TGACT CAGT GC
 TTT GTGAACTCCC TCG TCTATT GACCCGGTT GGT ACTGAGTCA CG

3'-NF-E2 +199 CACC (II) CACC (III) +212

GGAGGCCAACAGGACT GCTGAGTCA TCCTGT **GGGG** GTGG **AGGTGG**
 CCTCCGG TT GTCCTGA CGACTCAGT AGGACACCCC CACC TC CACC

CACC (IV) GATA-1 (C) +209

GGACAAGGGAAAGGG GTGG **AATGGTACTGC** TG ATTACAACCT CTGG
 CC TGTTCCC TT TCCC CACC T TACCAT GACGACTAAT GTTGGAGACC

GATA-1 (D)

TGCT GCCT CCCCCTCC TG T TT ATCT GAGAGGGAA -3'
 ACGACGGAGGGGGAGGACA AATAGA CTC TCCC TT -5'

Figure 1. Binding sites for the nuclear factors GATA-1, NF-E2 and CACC boxes (underlined) and the α -MRE polymorphisms (in bold letters).

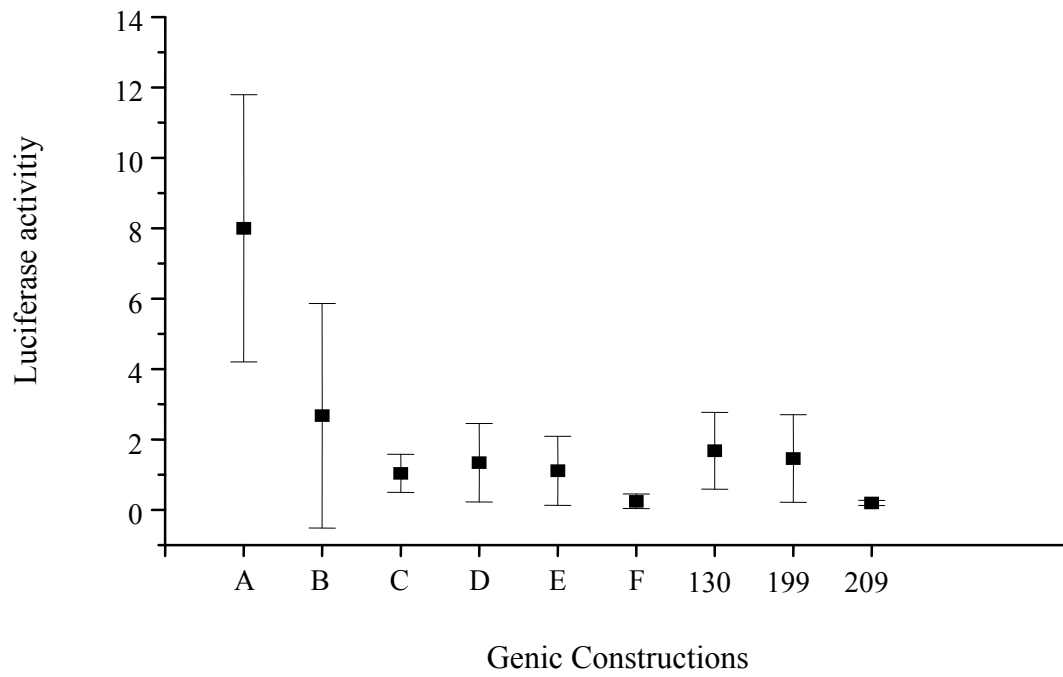


Figure 2. Luciferase activity of the different genic constructions in K562 cells. Means and Standard Deviations are given in relative luminescence unit (RLU).

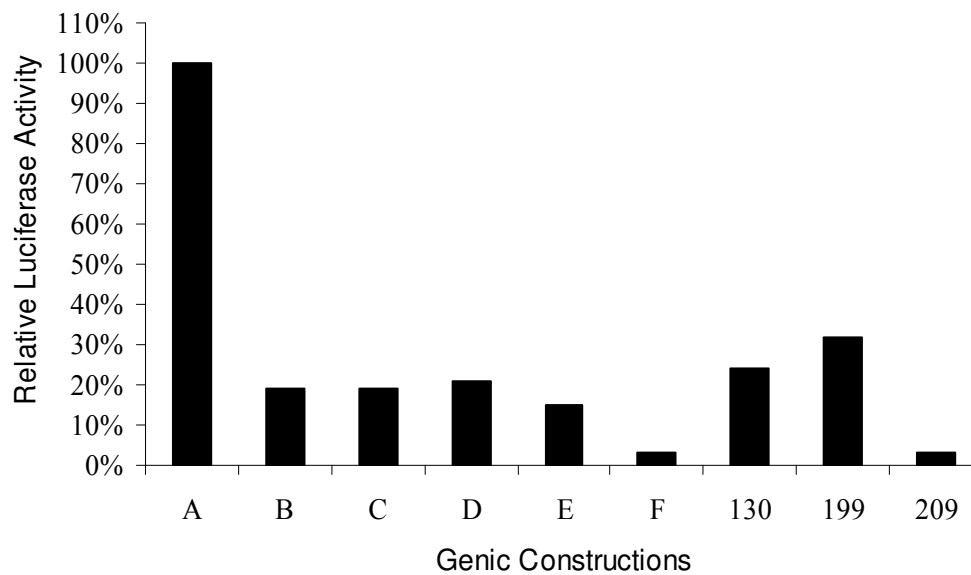


Figure 3. Relative luciferase activity of the different genic constructions in K562 cells. The medians of the luciferase activity of the genic constructions are shown as a percentage of the activity of the construction containing the wild type α -MRE (A haplotype), defined as 100%.

CAPÍTULO 2

Minireview

REGULATION OF THE HUMAN ALPHA-GLOBIN GENES AND ALPHA-THALASSEMIA

Daniela Maria Ribeiro¹ and Maria de Fátima Sonati¹

¹Universidade Estadual de Campinas, Departamento de Patologia Clínica, Campinas, SP, Brasil

Correspondence should be addressed to Maria de Fátima Sonati. Departamento de Patologia Clínica da Faculdade de Ciências Médicas da Universidade Estadual de Campinas, Campinas, SP, Brasil – Caixa Postal 6111 – CEP 13083-970 — E-mail: sonati@fcm.unicamp.br or sonati_mf@yahoo.com.br

KEYWORDS: alpha-globin genes, alpha-Major Regulatory Element, alpha-thalassemia

Abstract

Human α -globin synthesis is encoded by two adjacent genes (α_2 and α_1) clustered on chromosome 16. Their expression is controlled mainly by a regulatory element located 40 kb upstream on the same chromosome, the α -Major Regulatory Element (α -MRE), whose activity is restricted to a core fragment of 350 base pairs (bp) in which several regulatory protein binding sites have been found. The α -MRE is genetically polymorphic, and different haplotypes have been identified in some population groups from Europe, Africa and Asia and in native Indians from two Brazilian Indian tribes, the Parakanã and the Xikrin. Deletions involving the α -MRE constitute a particular category of α -thalassemia determinant in which the α -globin genes are physically intact but functionally inactive. The present article summarizes the information available to date about this important regulatory element and the relationship between alterations in this element and α -thalassemia.

Hemoglobin and Globin Genes

Human hemoglobin is a globular protein that consists of two α -like and two β -like globin polypeptide chains. Each globin chain is associated with a heme group, which is capable of binding reversibly one molecule of oxygen and transporting it from the lungs to the peripheral tissues (Perutz et al., 1960; Bunn and Forget, 1986). The cluster of the genes coding for the β -like globin chains, on the short arm of chromosome 11 (11p15.5), includes the embryonic ε -globin gene, the fetal globin genes $^G\gamma$ and $^A\gamma$, the pseudogene $\psi\beta$ and the adult δ - and β -globin genes (Fritsch et al., 1980) (Figure 1-A). The α -globin gene cluster, located in the telomeric region of the short arm of chromosome 16 (16p13.3), includes an embryonic gene (ζ), two minor α -like genes (α^D , also called μ , and θ), two pseudogenes ($\psi\alpha_1$ and $\psi\zeta$) and two α genes (α_2 and α_1) (Higgs, 1993; Hugges et al., 2005) (Figure 1-B).

Regulation of Globin Gene Expression

The globin genes are regulated in a tissue- and developmental stage-specific manner to produce different hemoglobins. This complex expression pattern of two physically separate loci depends on local cis-acting sequences, like the promoter sequences,

and remote cis-acting sequences, like the Locus Control Region (LCR) of the β -globin gene cluster and the Major Regulatory Element (MRE) of the α -globin gene cluster (Grosveld et al., 1987; Forrester et al., 1987; Higgs et al., 1990; Jarman et al., 1991).

In the β cluster, the regulatory sequences are associated with five erythroid-specific DNase I hypersensitive sites (named HS-1 to HS-5) distributed over a region 4 – 20 kilobases (kb) upstream of the ϵ -gene. The entire segment of DNA that includes these sites is referred to as the β -Locus Control Region (β -LCR) (Grosveld et al., 1987; Forrester et al., 1987) (Figure 1-A).

Similarly, in the α globin cluster, four erythroid-specific DNase I hypersensitive sites, located 8 (HS-8), 10 (HS-10), 33 (HS-33) and 40 (HS-40) kb upstream of the ζ -globin mRNA cap site, were identified (Higgs et al., 1990) (Figure 1-B). Characterization of natural deletions, analysis of interspecific hybrids and stable transfectants, and studies of transgenic mice indicate that only the regulatory sequence located 40 kb upstream of the α cluster (HS-40) has a significant effect on α -gene expression (Hatton et al., 1990; Liebhaber et al., 1990; Wilkie et al., 1990; Romao et al., 1991, 1992; Sharpe et al., 1992; Flint et al., 1994, 1996; Higgs et al., 1998; Viprakasit et al., 2003, 2006; Harteveld et al., 2005). This sequence was later termed the α -Major Regulatory Element (α -MRE) (Jarman et al., 1991).

Because the α and β clusters have a common ancestry, coordinated expression and similar organization, it was assumed that regulation of globin gene expression by the β -LCR and α -MRE was similar. However, experimental evidence from *in vivo* and *in vitro* studies suggests different roles for the β -LCR and α -MRE in β -globin and α -globin gene expression, respectively (Vyas et al., 1992; Craddock et al., 1995). Transgenic mice carrying constructs with the β -globin gene under the control of the β -LCR show expression levels proportional to the number of gene copies integrated, irrespective of the positions at which they are integrated into the mouse genome (Grosveld et al., 1987). In contrast to the β -LCR, the α -MRE does not confer position independence or copy-number dependence on transgenic α -globin gene expression (Sharpe et al., 1992; Gourdon et al., 1994).

These differences in the regulation of gene expression could be explained by chromatin effects at the integration site. It has been suggested that the β -LCR, in addition to its action as an enhancer, mediates a transition from “closed” chromatin to “open” transcriptionally active chromatin in a tissue-specific manner. Even though the α -MRE has some ability to form “open” chromatin in transgenic mice, it does not appear to be required for the organization of the chromatin structure of the α -globin locus, which is located in a constitutively “open” chromatin environment (Vyas et al., 1992; Craddock et al., 1995).

Recent studies of the structure of chromatin in the β -cluster showed that the β -LCR interacts with the promoters of the active β -like genes through looping, forming an active chromatin hub (ACH) to regulate gene expression. The formation of this erythroid cell-specific chromatin structure is also developmentally regulated (Tolhuis et al., 2002; Patrinos et al., 2004).

With regard to the α cluster, studies of the mouse α -globin locus identified an ACH-like structure similar to that formed in the β cluster. In erythroid cells, the upstream regulatory element interacts with the promoters of the active α -globin genes through looping, forming an α -ACH that is shared with the co-expressed housekeeping genes. In non-erythroid cells, the housekeeping gene promoters still colocalize, but the α -globin genes are excluded from the α -ACH (Zhou et al., 2006).

The α -Major Regulatory Element (α -MRE)

The regulatory element α -MRE behaves as a classical enhancer; its main function in the normal chromosomal environment is to activate, and enhance expression from, the ζ -globin as well as the α -globin promoters (Zhang et al., 1993).

The functional domain of this element is restricted to a 350-bp core fragment in which several well-conserved nuclear protein binding sites have been identified. These include four potential binding sites for the erythroid-specific factor GATA-1, four CACC boxes and two binding sites for the erythroid-factor NF-E2 (Jarman et al., 1991) that are occupied *in vivo* in an erythroid lineage-specific and developmental stage-specific manner.

In vivo, three out of four GATA-1 sites, both NF-E2 sites and one CACC box are occupied in erythroid cells, but not in non-erythroid cells (Strauss et al., 1992).

The α -MRE element is genetically polymorphic, and this polymorphism was first studied by Hartevelde et al. (2002) in seven population groups from Africa, Europe and Asia. Six different α -MRE haplotypes, named A to F, were found (Table 1). These haplotypes were also studied in native Indians from two non-miscegenated tribes (the Parakanã and Xikrin) in Amazonia in the north of Brazil (Ribeiro et al., 2003). The α -MRE haplotype frequencies in the population groups studied by Hartevelde et al. (2002) and Ribeiro et al. (2003) are shown in Table 2.

Most of these polymorphisms are not expected to interfere with regulation of α -gene expression, since they are in between binding sites for nuclear factors or in sites considered not to be active *in vivo*, except for haplotype D, in which the base substitution at position 158 changes the first consensus binding site of the factor NF-E2. However, an *in vitro* study, recently performed in our laboratory, employing constructs with the different α -MRE haplotypes in linkage to the luciferase reporter gene, demonstrated that the α -MRE polymorphisms caused a noticeable decrease in luciferase gene expression in transiently transfected K562 cells (*unpublished data*).

Alpha-Thalassemia

Thalassemias are inherited hemoglobin disorders characterized by a quantitative reduction of the α - or β -globin chains (Weatherall and Clegg, 1981). Alpha-thalassemias result from underproduction of the α -globin chains of fetal and adult hemoglobin. Down-regulation of one or two of the four α -globin genes leads to mild alterations in red blood cells; when three α -genes are involved, excess β -globin chains form tetramers (β_4 – Hb H), causing a moderate to severe hemolytic anemia called Hb H disease. Inactivation of the four α -genes results in tetramerization of γ -chains (γ_4 – Hb Barts) and in Hb Barts hydrops fetalis syndrome (Weatherall and Clegg, 1981). The majority of α -thalassemia determinants are deletions that remove the α -genes; inactivating point

mutations are a less common cause of this type of thalassemia but they may occur at high frequencies in certain areas under selective pressure by malaria (Higgs, 1993).

Deletions involving the regulatory element α -MRE are sporadic but characterize a particular category of α -thalassemia determinant in which the α -globin genes are physically intact but functionally inactive (Romao et al., 1991). At present, about fifteen deletions that remove the α -MRE and cause α -thalassemia have been found, eleven of which have been fully characterized (Hatton et al., 1990; Liebhaber et al., 1990; Wilkie et al., 1990; Romao et al., 1991, 1992; Flint et al., 1994, 1996; Viprakasit et al., 2003, 2006; Hartevelde et al., 2005) (Figure 2). The patients that were studied inherited chromosomes with deletions that removed from 35 to 160 kb of the region upstream of the α cluster, while the linked α genes remained intact, with entirely normal DNA sequences. All these patients have α -thalassemia and their phenotypes are consistent with severe down-regulation (< 1% – 2%) of α -gene expression from the affected chromosome (Anguita et al., 2002). The hematological findings in the patients show that simple heterozygotes for an α -MRE deletion [$(\alpha\alpha)/\alpha\alpha$] have low mean corpuscular volume (<80 fl) and low mean corpuscular hemoglobin (<25 pg) and are indistinguishable from patients with only two α genes ($--/\alpha\alpha$). Compound heterozygotes [$(\alpha\alpha)/-\alpha$] have the clinical phenotype of Hb H disease and are indistinguishable from patients who inherited a single functional α gene ($--/-\alpha$) (Viprakasit et al., 2003, 2006).

A recent study demonstrated the existence of an additional mechanism causing α -thalassemia, namely, a regulatory single nucleotide polymorphism (rSNP) in a non-genic region between the α -globin genes and their regulatory element. This polymorphism creates a new promoter-like element that interferes with normal activation of all downstream α -like genes and disrupts α -gene expression, probably as a result of its preferential interaction with the regulatory element α -MRE, outcompeting the endogenous α -globin promoters (Gobbi et al., 2006). These findings demonstrate the important role of SNPs in gene expression, especially SNPs located in regulatory regions.

Deletions of the α -MRE are also present in the Brazilian population. In 2002, Wenning et. al described a patient with Hb H disease resulting from a combination of the $-\alpha^{3.7}$ rightward deletion (the most common α -thalassemic determinant) with the $(\alpha\alpha)^{MM}$

deletion, which removes the regulatory element α -MRE, abolishing the α -gene expression from the affected chromosome. The $(\alpha\alpha)^{MM}$ deletion was first characterized by Romao et al. in 1991, in a patient from the Azores Islands, Portugal, and the Brazilian case represented the first of this type of deletion in an American population. A number of other cases have yet to be characterized. An awareness of the presence of these deletions in the populations is important because the heterozygotes present with reduced red-cell indices in the absence of iron deficiency, β -thalassemia or α -globin gene alterations.

Regulation of α -globin gene expression by the α -MRE is an important model for understanding the control systems of other eukaryotic genes and investigating the interplay between gene transcription and modifications to chromatin structure related to gene functioning.

Acknowledgements

Funding was provided by FAPESP/Brazil (fellowship no 03/07412-0) and CNPq/ Brazil (grants no 475481/2006-2).

References

- Anguita E, Sharpe JA, Stanley JAS, Tufarelli C, Higgs DR and Wood WG (2002) Deletion of the mouse α -globin regulatory element (HS-26) has an unexpectedly mild phenotype. *Blood* 100: 3450-3456.
- Beutler E, Coller BS, Lichtman MA, Kipps TJ and Seligsohn U (2001) *Williams Hematology*. 6^a ed. McGraw-Hill Companies, New York.
- Bunn HF and Forget BG (1986) *Hemoglobin: Molecular, Genetic and Clinical Aspects*. Philadelphia, W B Saunders Company.
- Craddock CF, Vyas P, Sharpe JA, Ayyub H, Wood WG and Higgs DR (1995) Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. *EMBO J* 14: 1718-1726.
- Flint J, Craddock CF, Villegas A, Bentley DP, Williams HJ, Galanello R, Cao A, Wood WG, Ayyub H and Higgs DR (1994) Healing of broken human chromosomes by the addition of telomeric repeats. *Am. J Hum Genet* 55: 505-512.
- Flint J, Rochette J, Craddock CF, Dode C, Vignes B, Horsley SW, Kearney L, Buckle VJ, Ayyub H and Higgs DR (1996) Chromosomal stabilization by a subtelomeric rearrangement involving two closely related *Alu* elements. *Hum Mol Genet* 5: 1163-1169.
- Forrester WC, Takegawa S, Papayannopoulou T, Stamatoyannopoulos G and Groudine M (1987) Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res* 15: 10159-10177.
- Fritsch EF, Lawn RM and Maniats T (1980) Molecular cloning and characterization of the human β -like globin gene cluster. *Cell* 19: 959-972.
- Gobbi MD, Viprakasit V, Hughes JR, Fisher C, Buckle VJ, Ayyub H, Gibbons RJ, Vernimmen D, Yoshinaga Y, Jong P, Cheng JF, Rubin EM, Wood WG, Bowden D and Higgs DR (2006) A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 312: 1215-1217.

- Gourdon G, Sharpe JA, Wells D, Wood WG and Higgs DR (1994) Analysis of a 70 kb segment of DNA containing the human ζ and α -globin genes linked to their regulatory element (HS-40) in transgenic mice. *Nucleic Acids Res* 22: 4139-4147.
- Grosveld F, Blom van Assendelft G, Greaves DR and Kollias G (1987) Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* 51: 975-985.
- Harteveld CL, Muglia M, Passarino G, Kielman MF and Bernini LF (2002) Genetic polymorphism of the major regulatory element HS-40, upstream of the human α -globin gene cluster. *Br J Haematol* 119 (3): 848-854.
- Harteveld CL, Voskamp A, Phylipsen M, Akkermans N, Dunnen J, White SJ and Giordano PC (2005) Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha- and beta-thalassemia characterized by high resolution multiple ligation-dependent probe amplification. *J Med Genet* 42: 922-931.
- Hatton CSR, Wilkie AOM, Drysdale HC, Wood WG, Vickers MA, Sharpe J, Ayyub H, Pretorius IM, Buckle VJ and Higgs DR (1990) Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human α globin gene cluster. *Blood* 76: 221-227.
- Higgs DR, Wood WG, Jarman AP, Sharpe J, Lida J, Pretorius IM and Ayyub H (1990) A major positive regulatory region located far upstream of the human α -globin gene locus. *Genes and Dev* 4: 1588-1601.
- Higgs DR (1993) α -Thalassemia. In: Higgs DR, Weatherall DJ (eds) *Baillière's Clinical Haematology*, vol 6, nr. 1. Baillière Tindall, London, pp 117-150.
- Higgs DR, Sharpe JA and Wood WG (1998) Understanding alpha globin gene expression: a step towards effective gene therapy. *Semin Hematol* 35: 93-104.
- Hughes JR, Cheng JF, Ventress N, Prabhakar S, Clark K, Anguita E, Gobbi M, Jong P, Rubin E and Higgs DR (2005) Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. *Proc Natl Acad Sci USA* 102: 9830-9835.

- Jarman AP, Wood WG, Sharpe JA, Gourdon G, Ayyub H and Higgs DR (1991) Characterization of the major regulatory element upstream of the human α -globin gene cluster. *Mol Cell Biol* 11: 4679-4689.
- Liebhaber SA, Griese EU, Weiss I, Cash FE, Ayyub H, Higgs DR and Horst J (1990) Inactivation of human α -globin gene expression by a *de novo* deletion located upstream of the α -globin gene cluster. *Proc Natl Acad Sci USA* 87: 9431-9435.
- Patrinos GP, Krom M, Boer E, Langeveld A, Ali Imam AM, Strouboulis J, Laat W and Grosveld FG (2004) Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev* 18: 1495-1509.
- Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G and North ACT (1960) Structure of hemoglobin. *Nature* 185: 416-422.
- Ribeiro DM, Figueiredo MS, Costa FF and Sonati MF (2003) Haplotypes of alpha-globin gene regulatory element in two Brazilian native populations. *Am J Phys Anthropol* 121 (1): 58-62.
- Romao L, Osorio-Almeida L, Higgs DR, Lavinha J and Liebhaber SA (1991) Alpha-thalassemia resulting from deletion of the regulatory sequences far upstream of the α -globin structural genes. *Blood* 78: 1589-1595.
- Romao L, Cash F, Weiss I, Liebhaber S, Piratsu M, Galanello R, Loi A, Paglietti E, Ioannou P and Cao A (1992) Human α -globin gene expression is silenced by terminal truncation of chromosome 16p beginning immediately 3' of the ζ -globin gene. *Hum Genet* 89: 323-328.
- Sharpe JA, Chan-Thomas PS, Lida J, Ayyub H, Wood WG and Higgs DR (1992) Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. *EMBO J* 11: 4565-4572.
- Strauss EC, Andrews NC, Higgs DR and Orkin SH (1992) In vivo footprinting of the human α -globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol and Cell Biol* 12: 2135-2142.

- Tolhuis B, Palstra RJ, Splinter E, Grosveld F and Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell Biol* 10: 1453-1465.
- Viprakit V, Kidd AMJ, Ayyub H, Horsley S, Hughes J and Higgs DR (2003) *De novo* deletion within the telomeric region flanking the human α globin locus as a cause of α thalassemia. *Br J Haematol* 120: 867-875.
- Viprakit V, Harteveld CL, Ayyub H, Stanley JS, Giordano PC, Williams GW and Higgs DR (2006) A novel deletion causing α thalassemia clarifies the importance of the major human alpha globin regulatory element. *Blood* 107: 3811-3812.
- Vyas P, Vickers MA, Simmons DL, Ayyub H, Craddock CF and Higgs DR (1992) Cis-acting sequences regulating expression of the human alpha-globin cluster lie within constitutively open chromatin. *Cell* 69: 781-793.
- Weatherall DJ and Clegg JB (1981) *The thalassaemia syndromes*. 3^a ed. Blackwell Publishing, Oxford.
- Wenning MRSC, Harteveld CL, Giordano PC, Kimura EM, Saad STO, Costa FF and Sonati MF (2002) Hemoglobin H disease resulting from the $-\alpha 3.7$ rightward deletion and the $(\alpha\alpha)$ MM deletion in a Brazilian patient. *Eur J Haematol* 69 (3): 179-181.
- Wilkie AOM, Lamb J, Harris PC, Finney RD and Higgs DR (1990) A truncated human chromosome 16 associated with α -thalassemia is stabilized by addition of telomeric repeat $(TTAGGG)_n$. *Nature* 346: 868-871.
- Zhang Q, Reddy PM, Yu CY, Bastiani C, Higgs DR, Stamatoyannopoulos G, Papayannopoulou T and Shen CK (1993) Transcriptional activation of human ζ_2 globin promoter by the α globin regulatory element (HS-40): functional role of specific nuclear factor-DNA complexes. *Mol Cell Biol* 13: 2298-2308.
- Zhou GL, Xin L, Song W, Di LJ, Liu G, Wu XS, Liu DP and Liang CC (2006) Active chromatin hub of the mouse α globin locus forms in a transcription factory of clustered housekeeping genes. *Mol Cell Biol* 26: 5096-5105.

Table 1. α -MRE haplotypes.

	+96 ^a	+130	+158	+199	+209	+212
Haplotype A	C	G	C	G	G	G
Haplotype B	C	A ^b	C	A	G	G
Haplotype C	C	A	C	A	A	G
Haplotype D	C	G	T	G	G	G
Haplotype E	C	G	C	G	G	C
Haplotype F	A	G	C	G	G	G

Note - Source: Hartevelde et al., 2002.

^a Positions of the α -MRE polymorphic sites in relation to the CAP-site of the ζ gene

^b Bold letters indicate the nucleotide substitutions in relation to the A haplotype

Table 2. α -MRE haplotype frequencies in the population groups studied.

Populations	Haplotype					
	A	B	C	D	E	F
Dutch (35) ^a	0.43	0.57				
Italian (53)	0.56	0.43	0.01			
Indian (39)	0.67	0.32	0.01			
Chinese (23)	0.74	0.26				
Indonesian (34)	0.78	0.22				
African (32)	0.73	0.11		0.16		
Pygmy (60)	0.76	0.16		0.04	0.02	0.02
Parakanã (70)	0.70	0.30				
Xikrin (95)	0.87	0.13				

Note - Source: Hartevelde et al., 2002; Ribeiro et al., 2003.

^a Number of individuals studied

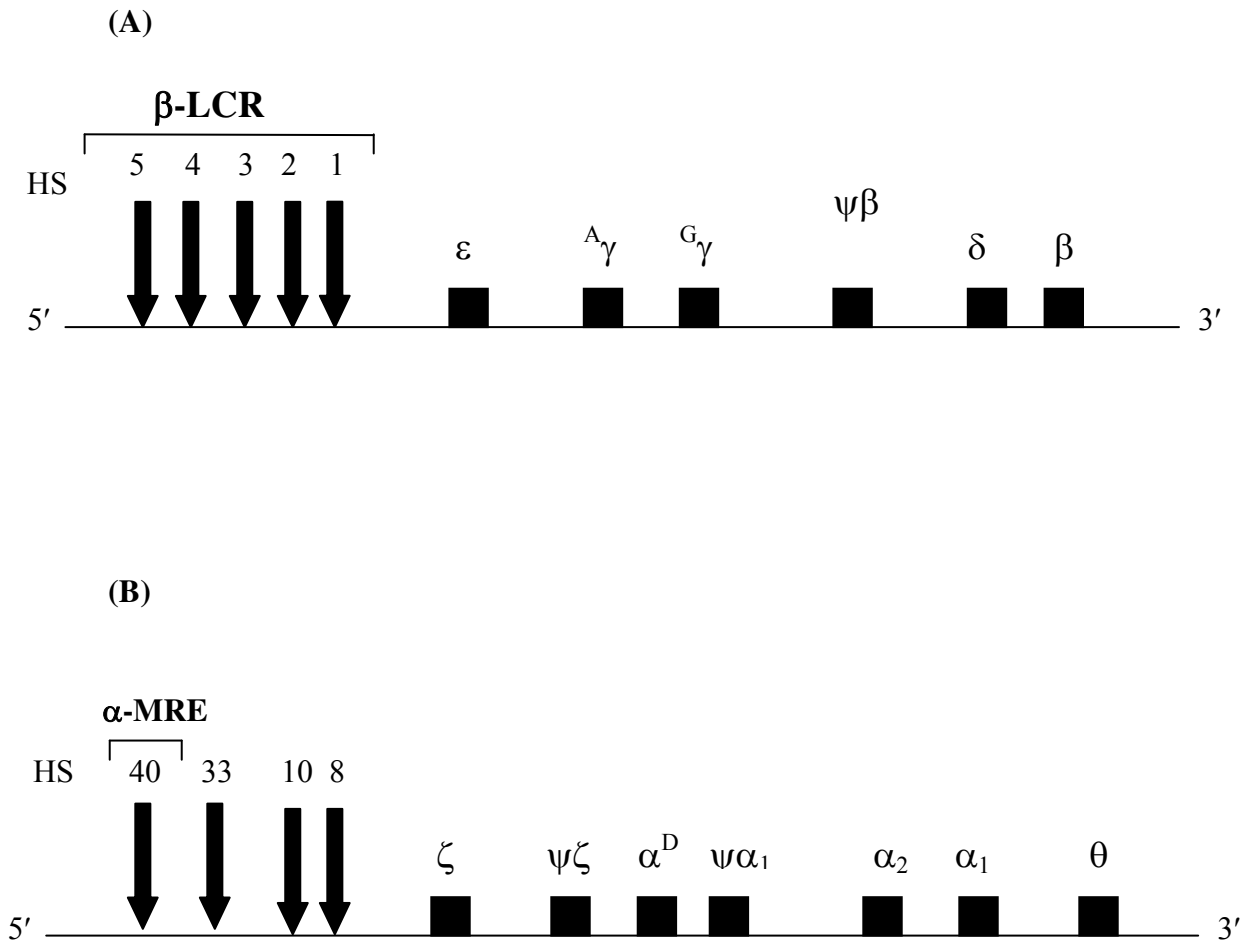


Figure 1. (A) β cluster and its regulatory sequences on the short arm of chromosome 11. (B) α cluster and its regulatory sequences on the short arm of chromosome 16. The black boxes represent the genes, and the arrows represent the regulatory sequences (adapted from Beutler et al., 2001).

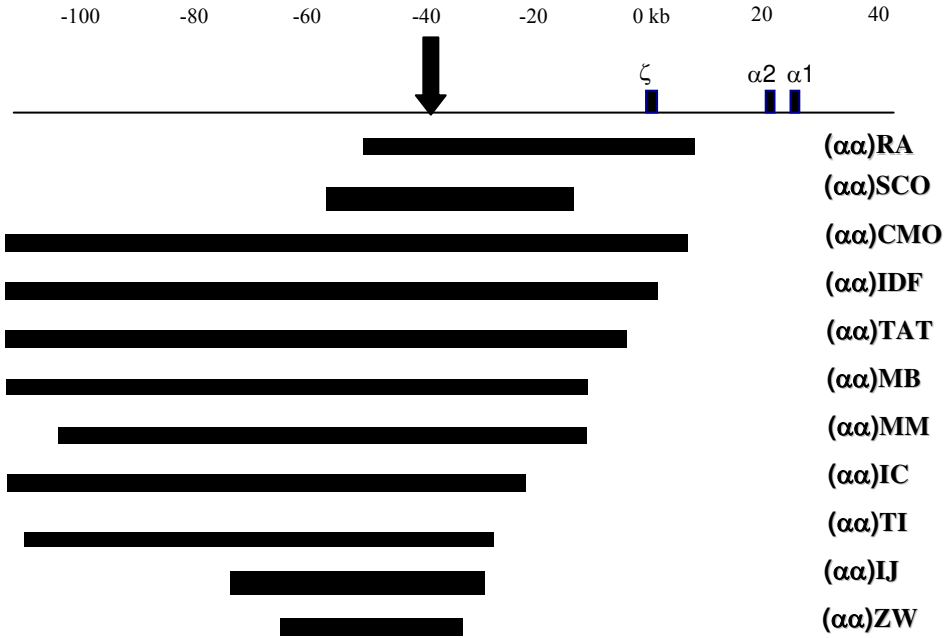


Figure 2. Diagram of upstream deletions of the α cluster. The black boxes above the line represent the genes, and the α -MRE is shown as an arrow. The black boxes below the line represent the deletions (adapted from Viprakasit et al., 2003 and Harteveld et al., 2005).

3- DISCUSSÃO E CONCLUSÃO

Para a análise da influência dos polimorfismos do α -MRE na expressão gênica *in vitro*, experimentos de transfecção transiente foram realizados em células da linhagem K562, utilizando-se o sistema do gene repórter da luciferase. As diferentes seqüências polimórficas do α -MRE foram clonadas no vetor de expressão pGL2-Promoter (Promega) (Figura 6 - Anexo), que contém o promotor SV40 junto ao gene da Firefly luciferase, uma vez que, o elemento regulatório do *cluster* α pode ativar, de modo similar, tanto o promotor de seu gene alvo endógeno como o promotor SV40 (Ren et al., 1993).

Nove construções gênicas foram totalizadas, compostas pelo gene da Firefly luciferase associada ao promotor SV40 e aos diferentes haplótipos do elemento α -MRE (A, B, C, D, E e F) e aos polimorfismos +130, +199 e +209 isolados (Figura 7 - Anexo).

Tais construções gênicas foram transfectadas, por eletroporação, em células da linhagem K562, células eritroleucêmicas humanas que crescem em suspensão e são capazes de expressar os genes das globinas, exceto o da globina β (Donovan-Peluso et al., 1984). O vetor pRL-TK (Promega) (Figura 8 - Anexo), que expressa a enzima Renilla luciferase, foi co-transfectado juntamente com o vetor a ser estudado para controle interno do experimento de transfecção. Normalizando a atividade do vetor experimental com a atividade do vetor controle interno, minimizamos a variabilidade experimental causada por diferenças na viabilidade celular ou na eficiência da transfecção (Promega, 2005).

Os experimentos foram realizados em triplicata e repetidos três vezes. Os vetores pGL2 - Basic, que contém apenas o gene da luciferase sem promotor e *enhancer*, e pGL2 - Control, que possui um promotor e um *enhancer* (ambos da Promega) (Figura 9 - Anexo), foram utilizados, respectivamente, como controles negativo e positivo da expressão da luciferase.

Após a transfecção, as células K562 foram cultivadas para permitir a expressão gênica, sendo então lisadas e o extrato celular utilizado para a leitura da luminescência da Firefly luciferase, assim como a leitura da luminescência da Renilla luciferase. Os valores de luminescência obtidos com os vetores experimentais foram normalizados pelos valores obtidos com o vetor controle interno através da relação Firefly/Renilla.

Os resultados dos experimentos de transfecção transiente, expressos em unidade relativa de luz (URL), estão apresentados na Tabela 3. Para a comparação da relação Firefly/Renilla entre as diferentes construções gênicas foi utilizado o teste ANOVA (Análise de Variância) para medidas repetidas com transformação por postos nas medidas, devido à variabilidade da amostra. O valor de p calculado ($p = 0,0025$) indica a existência de diferença significativa das médias entre as diferentes construções gênicas.

Tabela 3- Resultados dos experimentos de transfecção transiente em células K562.

Construções gênicas	N	Média ^a	Desvio padrão ^a	Mínimo ^a	Mediana ^a	Máximo ^a
A	3	7,998	3,794	5,756	5,859	12,378
B	3	2,676	3,190	0,555	1,129	6,344
C	3	1,037	0,543	0,450	1,140	1,521
D	3	1,339	1,116	0,282	1,230	2,506
E	3	1,110	0,982	0,249	0,901	2,180
F	3	0,244	0,205	0,069	0,193	0,469
+130	3	1,677	1,090	0,748	1,404	2,877
+199	3	1,459	1,244	0,056	1,892	2,430
+209	3	0,198	0,072	0,154	0,158	0,281

^a resultados em unidade relativa de luz (URL)

$p = 0,0025$ (ANOVA para medidas repetidas com transformação por postos)

As médias e os respectivos desvios-padrão de cada construção gênica estão apresentados na Figura 10.

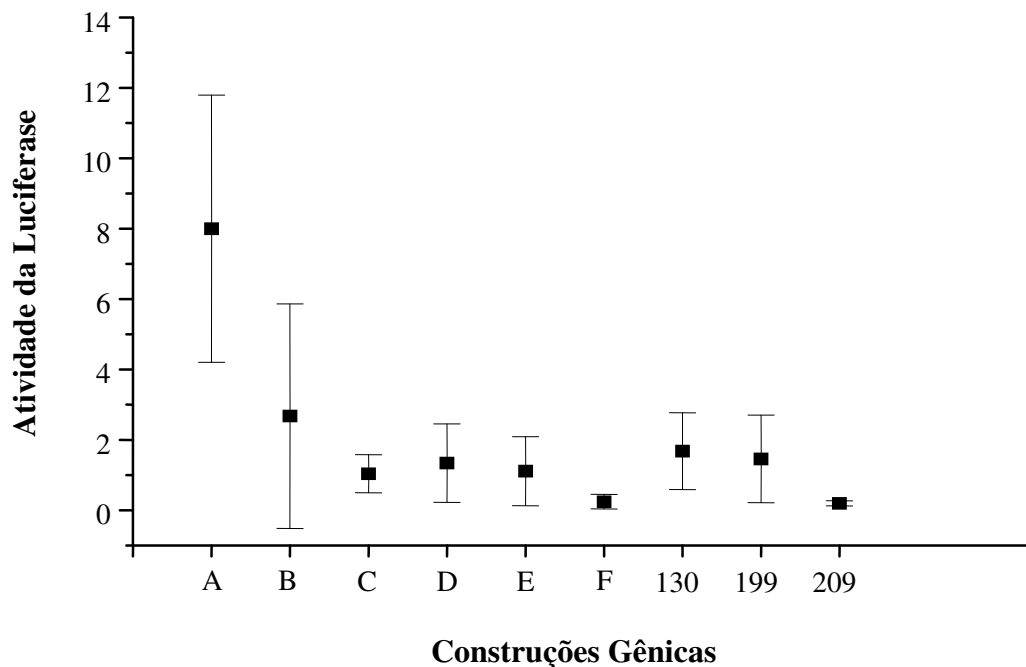


Figura 10- Atividade da luciferase das construções gênicas em células K562. Médias e desvios-padrão em unidade relativa de luz (URL).

Para comparar a atividade *enhancer* dos diferentes elementos α -MRE e, desse modo, analisar a influência dos polimorfismos deste elemento regulatório na expressão dos genes por ele controlados, os valores de mediana da atividade da luciferase das diferentes construções gênicas foram expressos como porcentagem da atividade da construção contendo o haplótipo selvagem A do elemento regulatório, definida como tendo 100% de atividade (Figura 11). A mediana foi escolhida aqui por refletir melhor a variabilidade dos experimentos.

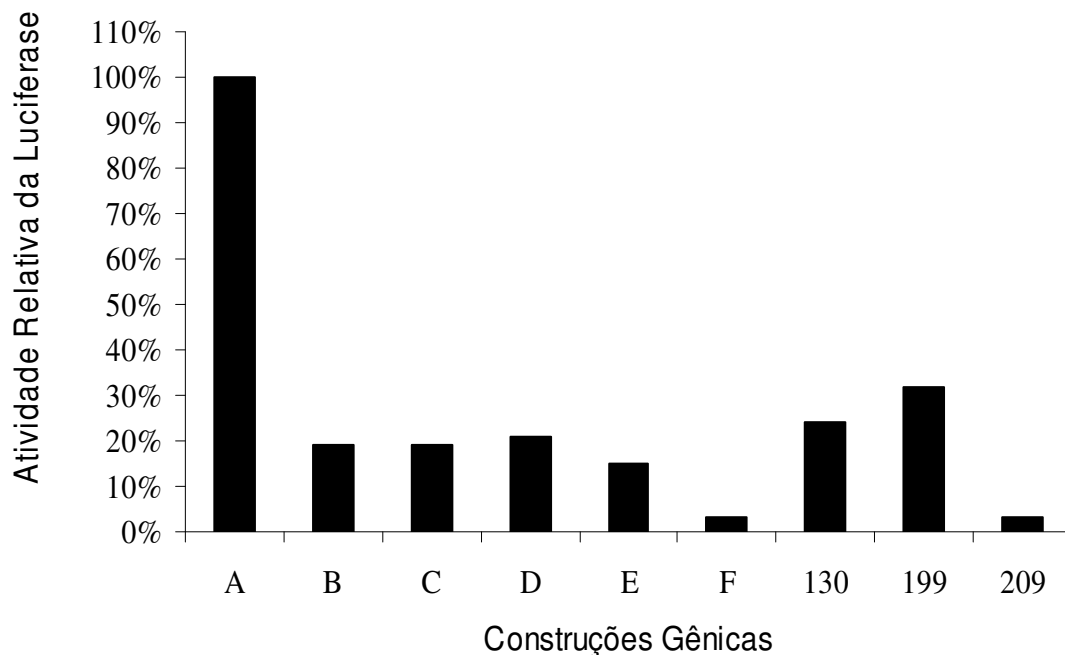


Figura 11- Atividade relativa da luciferase das construções gênicas em células K562. As medianas da atividade da luciferase das diferentes construções gênicas são expressas como porcentagem da atividade da construção contendo o haplótipo A do α -MRE, definida como 100% de atividade.

Uma vez que a atividade da enzima luciferase corresponde à expressão do seu gene, observamos no gráfico acima que a construção gênica contendo o haplótipo B do elemento α -MRE apresentou uma redução de 81% na expressão do gene da luciferase quando comparado à construção contendo o haplótipo selvagem A. O haplótipo B é constituído pelos polimorfismos +130 e +199, ambos localizados próximos a sítios de ligação para fatores nucleares (GATA-1 e SP1) ocupados *in vivo*. Analisando as construções contendo os polimorfismos +130 e +199 isoladamente, observamos uma redução de 76% na expressão da luciferase quando somente o polimorfismo +130 está presente no elemento α -MRE e uma redução de 68% quando temos somente o polimorfismo +199.

O haplótipo C, que difere do haplótipo B por um único polimorfismo na posição +209, apresentou também uma redução de 81% na expressão do gene da luciferase quando comparado ao haplótipo A. No entanto, a construção gênica contendo apenas o polimorfismo +209, localizado em um sítio de ligação para a proteína SP1 não ocupado *in vivo*, reduziu a atividade *enhancer* do elemento α -MRE em 97%.

O haplótipo D, constituído pelo polimorfismo +158 que altera a seqüência de consenso do sítio 5'-NF-E2 ocupado *in vivo*, e o haplótipo E, formado pelo polimorfismo +212 localizado em um sítio para a proteína SP1 não ocupado *in vivo*, apresentaram reduções acentuadas (79% e 85%, respectivamente) na expressão do gene da luciferase quando comparados ao haplótipo A.

A maior redução na expressão da luciferase entre os haplótipos do α -MRE foi obtida com o haplótipo F (97%), que é constituído pelo polimorfismo +96, localizado próximo a um sítio de ligação para a proteína SP1 não ocupado *in vivo*.

Os resultados apresentados acima permitem observar que os polimorfismos do elemento regulatório α -MRE reduzem a expressão do gene repórter da luciferase em células K562. Estes polimorfismos estão, em sua maioria, situados nas seqüências flanqueadoras dos sítios de ligação para proteínas regulatórias que constituem o núcleo funcional do elemento α -MRE, responsável por interagir com os promotores dos genes durante a etapa final da ativação transcricional em células K562. No haplótipo D, a substituição de base ocorre no sítio 5'-NF-E2. O experimento de transfecção transiente empregado neste trabalho reflete o que acontece no estágio final da ativação da transcrição, quando os promotores dos genes e o elemento regulatório já estão fisicamente associados (Zhang et al., 1995).

Estudos anteriores de mutagênese sítio-dirigida e transfecção transiente demonstraram que os sítios GATA-1 (B, C e D), ambos os sítios NF-E2 (5' e 3') e o sítio CACC (II) regulam positivamente a expressão dos genes do *cluster* α em células K562 (Rombel et al., 1995; Zhang et al., 1995). Mutações em cada um destes sítios reduziram a expressão gênica em 50 a 90%. Duas mutações em particular, nos sítios 3'-NF-E2 e CACC

(II), praticamente aboliram a expressão dos genes do *cluster* α por desestabilizarem a ligação de proteínas regulatórias aos demais sítios do elemento α -MRE (Wen et al., 2000).

Mutações que afetam a capacidade de ligação das proteínas regulatórias ao DNA e, conseqüentemente, suas propriedades ativadoras da transcrição, podem alterar a configuração e a função dos elementos regulatórios e, assim, modificar a expressão gênica (Zhang et al., 1993; Wen et al., 2000; Wang et al., 2005). Variações nas seqüências de consenso dos sítios de ligação para estas proteínas ainda são pouco conhecidas, porém, menos ainda se sabe sobre polimorfismos nas seqüências flanqueadoras destes sítios (Wen et al., 2000; Wang et al., 2005).

Os polimorfismos do α -MRE modificam a expressão do gene da luciferase em células K562 provavelmente por alterarem a afinidade de ligação das proteínas regulatórias aos seus respectivos sítios. Devido a suas localizações, eles podem estar causando um remodelamento da estrutura cromatínica dos sítios de ligação, alterando, assim, a dinâmica e a configuração do elemento regulatório α -MRE e, conseqüentemente, sua função.

Se a expressão *in vivo* dos genes da globina α humana for também influenciada pelos polimorfismos do α -MRE, os portadores dos diferentes haplótipos devem apresentar fenótipos α -talassêmicos variados, a depender do genótipo final de cada indivíduo (se heterozigoto ou homozigoto) e de possíveis mecanismos compensatórios existentes nos organismos mais complexos. Hartevelde et al. (2002) analisaram os parâmetros hematológicos de dois indivíduos homozigotos da deleção $-\alpha^{3.7}$ e do haplótipo D do α -MRE e não observaram nenhuma diferença significativa em relação aos valores esperados. Um número maior de indivíduos e de populações, no entanto, ainda precisa ser investigado, particularmente em regiões sob pressão seletiva pela malária, onde os haplótipos mais raros podem eventualmente atingir uma freqüência maior.

A interpretação dos resultados dos experimentos *in vitro* deve ser bastante cuidadosa, uma vez que eles nem sempre representam exatamente o que está acontecendo *in vivo*. A transfecção transiente é uma técnica bastante utilizada devido a sua facilidade e rapidez, porém sofre maior interferência que a transfecção estável por analisar seqüências

não integradas ao DNA celular (Li et al., 2005). A transfecção estável, por sua vez, utiliza vetores que conferem resistência a antibióticos às células transfectadas, permitindo, assim, selecionar células resistentes, ou seja, as que integraram a seqüência a ser estudada ao seu material genético; é um método mais trabalhoso e demorado, principalmente quando se utiliza células em suspensão, pela dificuldade em selecionar as linhagens resistentes, mas que oferece maior fidelidade de resultados uma vez que as seqüências estão inseridas no contexto da cromatina intacta, constituindo então um modelo melhor para o estudo do que acontece *in vivo* (Li et al., 2005).

Os experimentos com animais transgênicos representam o método mais adequado de análise do que ocorre realmente *in vivo*, pois possibilita o estudo da expressão gênica em seu ambiente regulatório natural. Desse modo, podemos analisar não somente a expressão dos genes individualmente, mas também em conjunto, assim como, a ativação e inativação gênica nos diferentes períodos de desenvolvimento do indivíduo (Li et al., 2005). A obtenção destes modelos animais é, no entanto, difícil e cara, limitando sua utilização na análise de seqüências regulatórias. Além disso, tanto os experimentos de transfecção estável em linhagens celulares quanto os estudos com animais transgênicos estão sujeitos aos efeitos de integração na cromatina que podem gerar resultados controversos (Anguita et al., 2002; Li et al., 2005). No estudo da regulação dos genes do *cluster α* , o camundongo é o modelo animal mais estudado, porém parece não ser ainda o ideal por apresentar algumas diferenças no controle da expressão gênica em relação ao que ocorre em humanos (Anguita et al., 2002).

Os ensaios utilizando genes repórteres, por sua vez, são importantes ferramentas para a análise dos efeitos dos polimorfismos regulatórios na expressão gênica, demonstrando boa sensibilidade, especificidade e reprodutibilidade. Entretanto, além de constituírem ensaios *in vitro*, os resultados destes experimentos dependem também do tipo celular a ser transfectado, do método de transfecção empregado e de sua eficiência, e das construções gênicas utilizadas (Knight, 2005).

As células da linhagem K562 aqui utilizadas são muito úteis no estudo da transcrição de genes com especificidade eritrocitária; são, no entanto, células que se encontram no estágio embrionário/fetal de desenvolvimento. Haveria, então, a possibilidade

de utilização de células que mimetizam o ambiente eritróide adulto, como a célula MEL, que é uma célula eritroleucêmica de camundongo, capaz de expressar as globinas α e β presentes na fase adulta (Marks e Rifkind, 1978). Acredita-se que os fatores nucleares destas células e os complexos formados entre eles possam diferir daqueles presentes nas células K562 (Zhang et al., 1995). As células MEL são, no entanto, células de um outro organismo e, possivelmente, também apresentarão divergências com o que ocorre em células humanas.

Estudos com o objetivo de identificar as interações DNA-proteína entre as regiões polimórficas do elemento e proteínas nucleares de células eritróides, poderão auxiliar na compreensão de como os polimorfismos interferem no mecanismo de regulação exercido pelo α -MRE e em sua interação com os genes sob seu controle.

Este é o primeiro estudo avaliando a influência dos diferentes polimorfismos do elemento α -MRE sobre a expressão gênica. Os resultados obtidos indicam que a variabilidade na sequência de DNA deste elemento regulatório, ainda que não afetando diretamente os sítios de ligação para proteínas regulatórias, reduz, ao menos *in vitro*, o grau de expressão dos genes por ele regulados.

4- REFERÊNCIAS BIBLIOGRÁFICAS

Andrews NC, Erdjument-Bromage H, Davidson MB, Tempst P, Orkin SH. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 1993; 362: 722-28.

Anguita E, Sharpe JA, Stanley JAS, Tufarelli C, Higgs DR, Wood WG. Deletion of the mouse α -globin regulatory element (HS-26) has an unexpectedly mild phenotype. *Blood* 2002; 100: 3450-56.

Berger SL. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* 2002; 12: 142-48.

Beutler E, Coller BS, Lichtman MA, Kipps TJ, Seligsohn U. *Williams Hematology*. New York, McGraw-Hill Companies; 2001.

Buckland PR, Hoogendoorn B, Coleman SL, Smith SK, Buxbaum JD, Haroutunian V, Donovan MC. A high proportion of polymorphisms in the promoters of brain expressed genes influences transcriptional activity. *Biochim Biophys Acta* 2004; 1690: 238-49.

Buckland PR, Hoogendoorn B, Coleman SL, Guy CA, Smith SK, Donovan MC. A high proportion of chromosome 21 promoter polymorphisms influence transcriptional activity. *Gene Expr* 2005; 11: 233-39.

Bulger M and Groudine M. Looping versus linking: toward a model for long-distance gene activation. *Genes and Dev* 1999; 13: 2465-77.

Bunn HF and Forget BG. *Hemoglobin: Molecular, Genetic and Clinical Aspects*. Philadelphia, W B Saunders Company; 1986.

Chakravati A. To a future of genetic medicine. *Nature* 2001; 409: 822-23.

Collins FS, Brooks LD and Chakravati A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 1998; 8: 1229-31.

Craddock CF, Vyas P, Sharpe JA, Ayyub H, Wood WG, Higgs DR. Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. *EMBO J* 1995; 14: 1718-26.

De Went JR. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc Natl Acad Sci USA* 1985; 82: 7870-73.

Dean A. Chromatin remodeling and the interactions between enhancers and promoters in the β -globin locus. *Briefings in Functional Genomics and Proteomics* 2004; 2: 344-54.

Donovan-Peluso M, Young K, Dobkin C, Bank A. Erythroleukemia (K562) cells contain a functional β -globin gene. *Mol Cell Biol* 1984; 4: 2553-57.

Elefant F, Cooke NE, Liebhaber SA. Targeted recruitment of histone acetyltransferase activity to a locus control region. *J Biol Chem* 2000; 275: 13827-834.

Flint J, Craddock CF, Villegas A, Bentley DP, Williams HJ, Galanello R, Cao A, Wood WG, Ayyub H, Higgs DR. Healing of broken human chromosomes by the addition of telomeric repeats. *Am J Hum Genet* 1994; 55: 505-12.

Flint J, Rochette J, Craddock CF, Dode C, Vignes B, Horsley SW, Kearney L, Buckle VJ, Ayyub H, Higgs DR. Chromosomal stabilization by a subtelomeric rearrangement involving two closely related *Alu* elements. *Hum Mol Genet* 1996; 5: 1163-69.

Forrester WC, Takegawa S, Papayannopoulou T, Stamatoyannopoulos G, Groudine M. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res* 1987; 15: 10159-177.

Fritsch EF, Lawn RM, Maniats T. Molecular cloning and characterization of the human β -like globin gene cluster. *Cell* 1980; 19: 959-72.

Gobbi MD, Viprakasit V, Hughes JR, Fisher C, Buckle VJ, Ayyub H, Gibbons RJ, Vernimmen D, Yoshinaga Y, Jong P, Cheng JF, Rubin EM, Wood WG, Bowden D, Higgs DR. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* 2006; 312: 1215-17.

Grosveld F, Blom van Assendelft G, Greaves DR, Kollias G. Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* 1987; 51: 975-85.

Gui CY and Dean A. Acetylation of a specific promoter nucleosome accompanies activation of the epsilon-globin gene by beta-globin locus control region HS2. *Mol Cell Biol* 2001; 21: 1155-63.

Harteveld CL, Muglia M, Passarino G, Kielman MF, Bernini LF. Genetic polymorphism of the major regulatory element HS-40, upstream of the human α -globin gene cluster. *Br J Haematol* 2002; 119 (3): 848-54.

Harteveld CL, Voskamp A, Phylipsen M, Akkermans N, Dunnen J, White SJ, Giordano PC. Nine unknown rearrangements in 16p13.3 and 11p15.4 causing alpha- and beta-thalassemia characterized by high resolution multiple ligation-dependent probe amplification. *J Med Genet* 2005; 42: 922-31.

Hatton CSR, Wilkie AOM, Drysdale HC, Wood WG, Vickers MA, Sharpe J, Ayyub H, Pretorius IM, Buckle VJ, Higgs DR. Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human α globin gene cluster. *Blood* 1990; 76: 221-27.

Higgs DR, Wood WG, Jarman AP, Sharpe J, Lida J, Pretorius IM, Ayyub H. A major positive regulatory region located far upstream of the human α -globin gene locus. *Genes and Dev* 1990; 4: 1588-1601.

Higgs DR. α -Thalassemia. In: Higgs DR, Weatherall DJ. *Baillière's Clinical Haematology*. London, Baillière Tindall; 1993. v 6 (1).

Higgs DR, Sharpe JA, Wood WG. Understanding alpha globin gene expression: a step towards effective gene therapy. *Semin Hematol* 1998; 35: 93-104.

[Http://globin.cse.psu.edu/html/huisman/variants/image.html](http://globin.cse.psu.edu/html/huisman/variants/image.html)

Hughes JR, Cheng JF, Ventress N, Prabhakar S, Clark K, Anguita E, Gobbi M, Jong P, Rubin E, Higgs DR. Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. *Proc Natl Acad Sci USA* 2005; 102: 9830-35.

Ikuta T and Kan YW. In vivo protein-DNA interactions at the β -globin gene locus. *Proc Natl Acad Sci USA* 1991; 88: 10188-192.

Jarman AP, Wood WG, Sharpe JA, Gourdon G, Ayyub H, Higgs DR. Characterization of the major regulatory element upstream of the human α -globin gene cluster. *Mol Cell Biol* 1991; 11: 4679-89.

Kielman MF, Smits R, Bernini LF. Localization and characterization of the mouse α -globin locus control region. *Genomics* 1994; 21: 431-33.

Kielman MF. The α -globin domain of man and mouse: integration of erythroid regulatory elements with a conserved cluster of upstream flanking genes. Rijksuniversiteit te Leiden. PhD Thesis; 1996.

Knight JC. Regulatory polymorphisms underlying complex disease traits. *J Mol Med* 2005; 83: 97-109.

Li Q, Emery DW, Han H, Sun J, Yu M, Stamatoyannopoulos G. Differences of globin transgene expression in stably transfected cell lines and transgenic mice. *Blood* 2005; 105 (8): 3346-52.

Liebhaber SA, Griese EU, Weiss I, Cash FE, Ayyub H, Higgs DR, Horst J. Inactivation of human α -globin gene expression by a *de novo* deletion located upstream of the α -globin gene cluster. *Proc Natl Acad Sci USA* 1990; 87: 9431-35.

Marks PA and Rifkind RA. Erythroleukemic differentiation. *Ann Rev Biochem* 1978; 47: 419-48.

Patrinos GP, Krom M, Boer E, Langeveld A, Ali Imam AM, Strouboulis J, Laat W, Grosveld FG. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes and Dev* 2004; 18: 1495-1509.

Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North ACT. Structure of hemoglobin. *Nature* 1960; 185: 416-22.

Philipsen S, Talbot D, Fraser P, Grosveld F. The β -globin dominant control region: hypersensitive site 2. *EMBO J* 1990; 9: 2159-67.

Promega. Dual luciferase reporter assay system. Technical Manual. Madison, Promega Corporation; 2005.

Pruzina S, Hanscombe O, Whyatt D, Grosveld F, Philipsen S. Hypersensitive site 4 of the human β -globin locus control region. *Nucleic Acids Res* 1991; 19: 1413-19.

Reddy PM and Shen CK. Protein-DNA interactions in vivo of an erythroid-specific, human β -globin locus enhancer. *Proc Natl Acad Sci USA* 1991; 88: 8676-80.

Ren S, Luo X, Atweh GF. The major regulatory element upstream of the α -globin gene has classical and inducible enhancer activity. *Blood* 1993; 81 (4): 1058-66.

Ribeiro DM, Figueiredo MS, Costa FF, Sonati MF. Haplotypes of alpha-globin gene regulatory element in two Brazilian native populations. *Am J Phys Anthropol* 2003; 121 (1): 58-62.

Rockman MV and Wray GA. Abundant raw material for cis-regulatory evolutions in humans. *Mol Biol Evol* 2002; 19: 1991-2004.

Romao L, Osorio-Almeida L, Higgs DR, Lavinha J, Liebhaber SA. Alpha-thalassemia resulting from deletion of the regulatory sequences far upstream of the α -globin structural genes. *Blood* 1991; 78: 1589-95.

Romao L, Cash F, Weiss I, Liebhaber S, Piratsu M, Galanello R, Loi A, Paglietti E, Ioannou P, Cao A. Human α -globin gene expression is silenced by terminal truncation of chromosome 16p beginning immediately 3' of the ζ -globin gene. *Hum Genet* 1992; 89: 323-28.

Rombel I, Hu KY, Zhang Q, Papayannopoulou T, Stamatoyannopoulos G, Shen CKJ. Transcriptional activation of human adult α -globin genes by hypersensitive site HS-40 enhancer: function of nuclear factor binding motifs occupied in erythroid cells. *Proc Natl Acad Sci USA* 1995; 92: 6454-58.

Schubeler D, Francastel C, Cimborra DM. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes and Dev* 2000; 14: 940-50.

Sharpe JA, Chan-Thomas PS, Lida J, Ayyub H, Wood WG, Higgs DR. Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. *EMBO J* 1992; 11: 4565-72.

Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res* 1999; 9: 677-79.

Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varnus H. The molecular bases of blood diseases. Philadelphia, Saunders Company; 1994.

Strauss EC, Andrews NC, Higgs DR, Orkin SH. In vivo footprinting of the human α -globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol Cell Biol* 1992; 12: 2135-42.

Strauss EC and Orkin SH. In vivo protein-DNA interactions at hypersensitive site 3 of the human β -globin locus control region. *Proc Natl Acad Sci USA* 1992; 89: 5809-13.

Talbot D, Philipsen S, Fraser P, Grosveld F. Detailed analysis of the site 3 region of the human β -globin dominant control region. *EMBO J* 1990; 9: 2169-77.

Talbot D and Grosveld F. The 5' HS-2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBO J* 1991; 10: 1391-98.

The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001; 409: 928-33.

Tjian R. Molecular machines that control genes. *Sci American* 1995; February: 38-45.

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell Biol* 2002; 10: 1453-65.

Vignali M, Hassan AH, Neely KE, Workman JL. ATP-dependent chromatin-remodelling complexes. *Mol Cell Biol* 2000; 20: 1899-1910.

Viprakit V, Kidd AMJ, Ayyub H, Horsley S, Hughes J, Higgs DR. *De novo* deletion within the telomeric region flanking the human α globin locus as a cause of α thalassemia. *Br J Haematol* 2003; 120: 867-75.

Viprakit V, Harteveld CL, Ayyub H, Stanley JS, Giordano PC, Williams GW, Higgs DR. A novel deletion causing α thalassemia clarifies the importance of the major human alpha globin regulatory element. *Blood* 2006; 107: 3811-12.

Vyas P, Vickers MA, Simmons DL, Ayyub H, Craddock CF, Higgs DR. Cis-acting sequences regulating expression of the human alpha-globin cluster lie within constitutively open chromatin. *Cell* 1992; 69: 781-93.

Wang Z and Moulton J. SNPs, protein structure and disease. *Hum Mutation* 2001; 17: 263-70.

Wang X, Tomso DJ, Liu X, Bell DA. Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes. *Toxicology and Applied Pharmacology* 2005; 207: 84-90.

Wen SC, Roder K, Hu KY, Rombel I, Gavva NR, Daftari P, Kuo YY, Wang C, Shen CKJ. Loading of DNA-binding factors to an erythroid enhancer. *Mol Cell Biol* 2000; 20: 1993-2003.

Wilkie AOM, Lamb J, Harris PC, Finney RD, Higgs DR. A truncated human chromosome 16 associated with α -thalassemia is stabilized by addition of telomeric repeat (TTAGGG)_n. *Nature* 1990; 346: 868-71.

Williams TM, Burlein JE, Ogden S, Kricka LJ, Kant JA. Advantages of firefly luciferase as a reporter gene: application to the interleukin-2 gene promoter. *Anal Biochem* 1989; 176 (1): 28-32.

Wood KV. Bioluminescence and Chemiluminescence In: Stanley P and Kricka L. *Bioluminescence and Chemiluminescence: Current Status*. Chichester, John Wiley and Sons; 1984. v 11.

Zhang Q, Reddy PM, Yu CY, Bastiani C, Higgs DR, Stamatoyannopoulos G, Papayannopoulou T, Shen CK. Transcriptional activation of human ζ_2 globin promoter by the α globin regulatory element (HS-40): functional role of specific nuclear factor-DNA complexes. *Mol Cell Biol* 1993; 13: 2298-2308.

Zhang Q, Rombel I, Reddy N, Gang JB, Shen CKJ. Functional roles of in vivo footprinted DNA motifs within an α globin enhancer. *J Biol Chem* 1995; 270: 8501-05

Zhou GL, Xin L, Song W, Di LJ, Liu G, Wu XS, Liu DP, Liang CC. Active chromatin hub of the mouse α globin locus forms in a transcription factory of clustered housekeeping genes. *Mol Cell Biol* 2006; 26: 5096-5105.

5- ANEXO



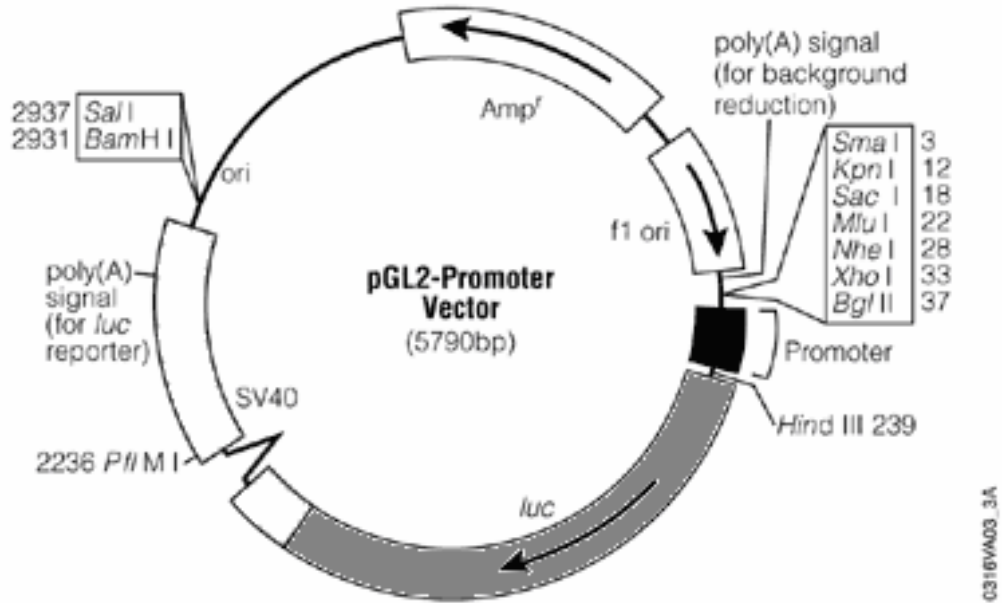


Figura 6- Vetor de expressão pGL2-Promoter. Em cinza, o gene da Firefly Luciferase e em preto, o promotor SV40. As diferentes seqüências polimórficas do α -MRE foram inseridas no sítio de restrição da enzima *Sma* I, localizado acima da região promotora.

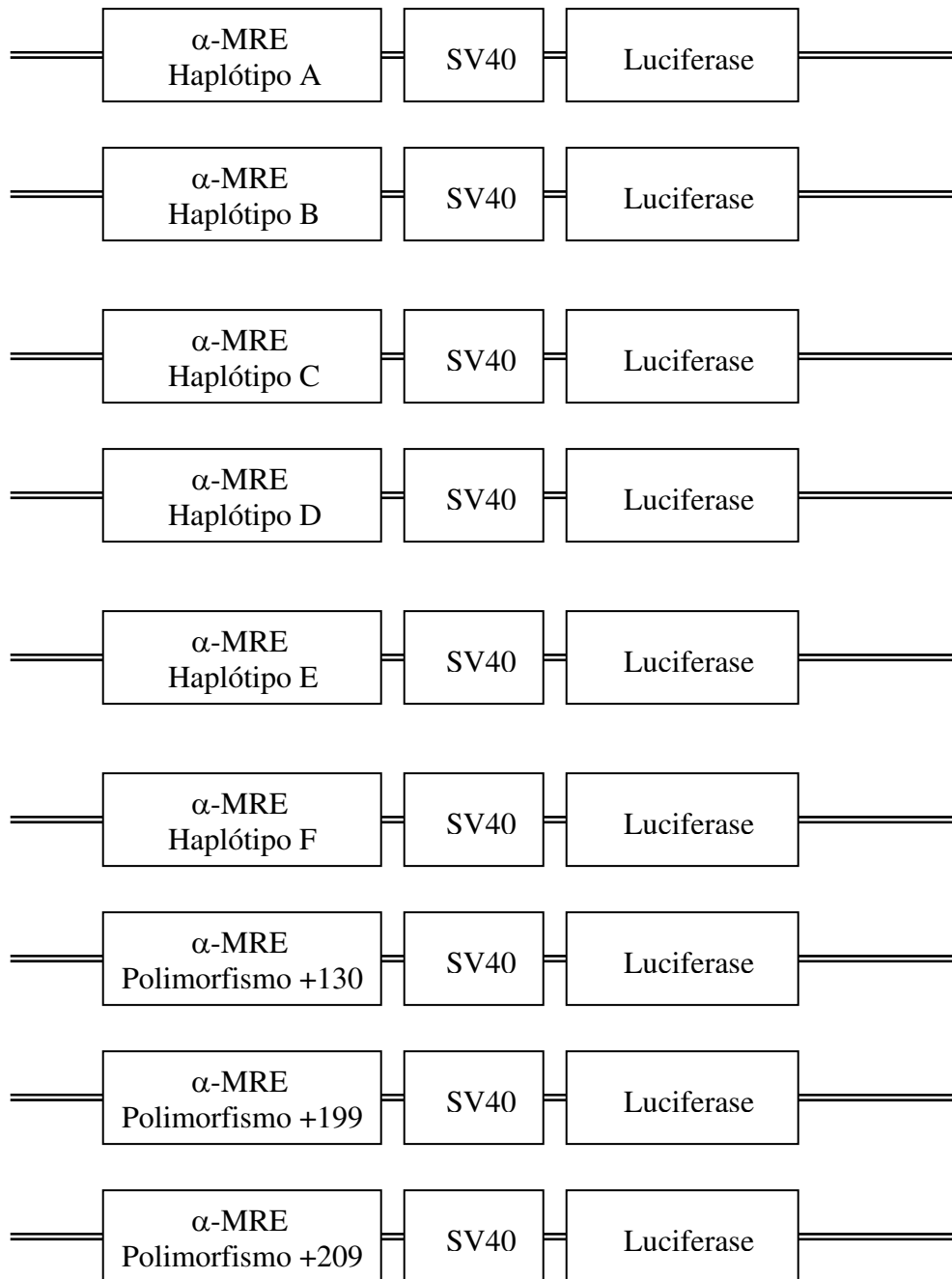


Figura 7- Construções gênicas. As construções são compostas pelo gene da Firefly luciferase associada ao promotor SV40 e aos diferentes haplótipos do elemento α -MRE (A, B, C, D, E e F) e aos polimorfismos +130, +199 e +209 isolados

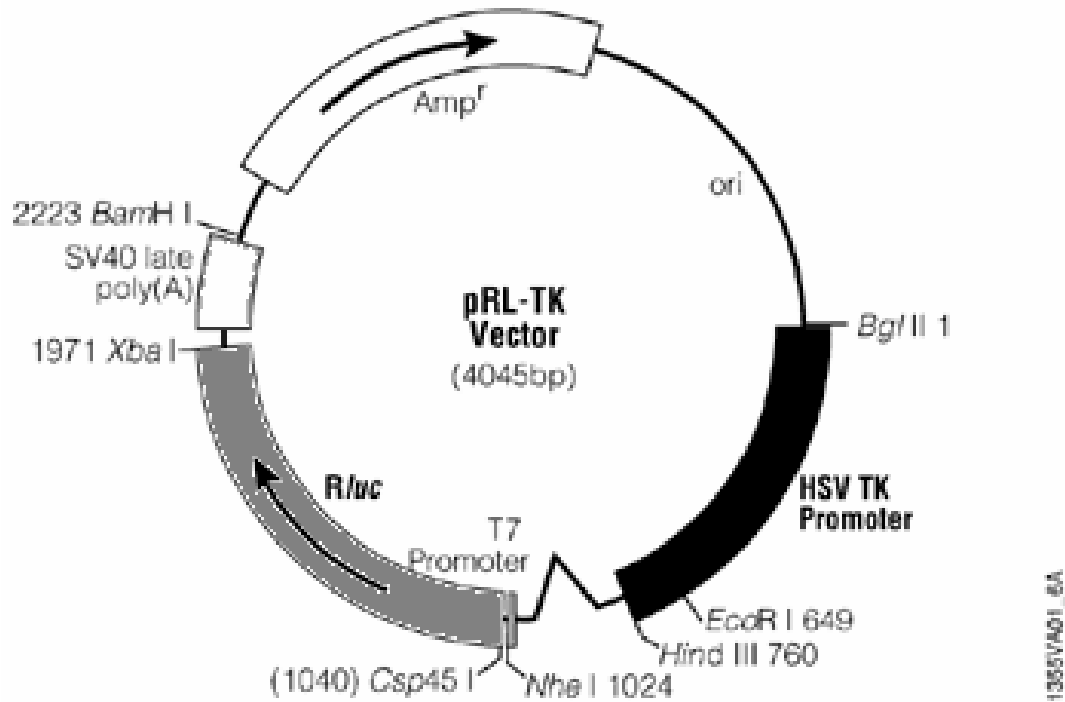


Figura 8- Vetor de expressão pRL-TK. Em cinza, o gene da Renilla Luciferase e em preto, o promotor HSV TK.

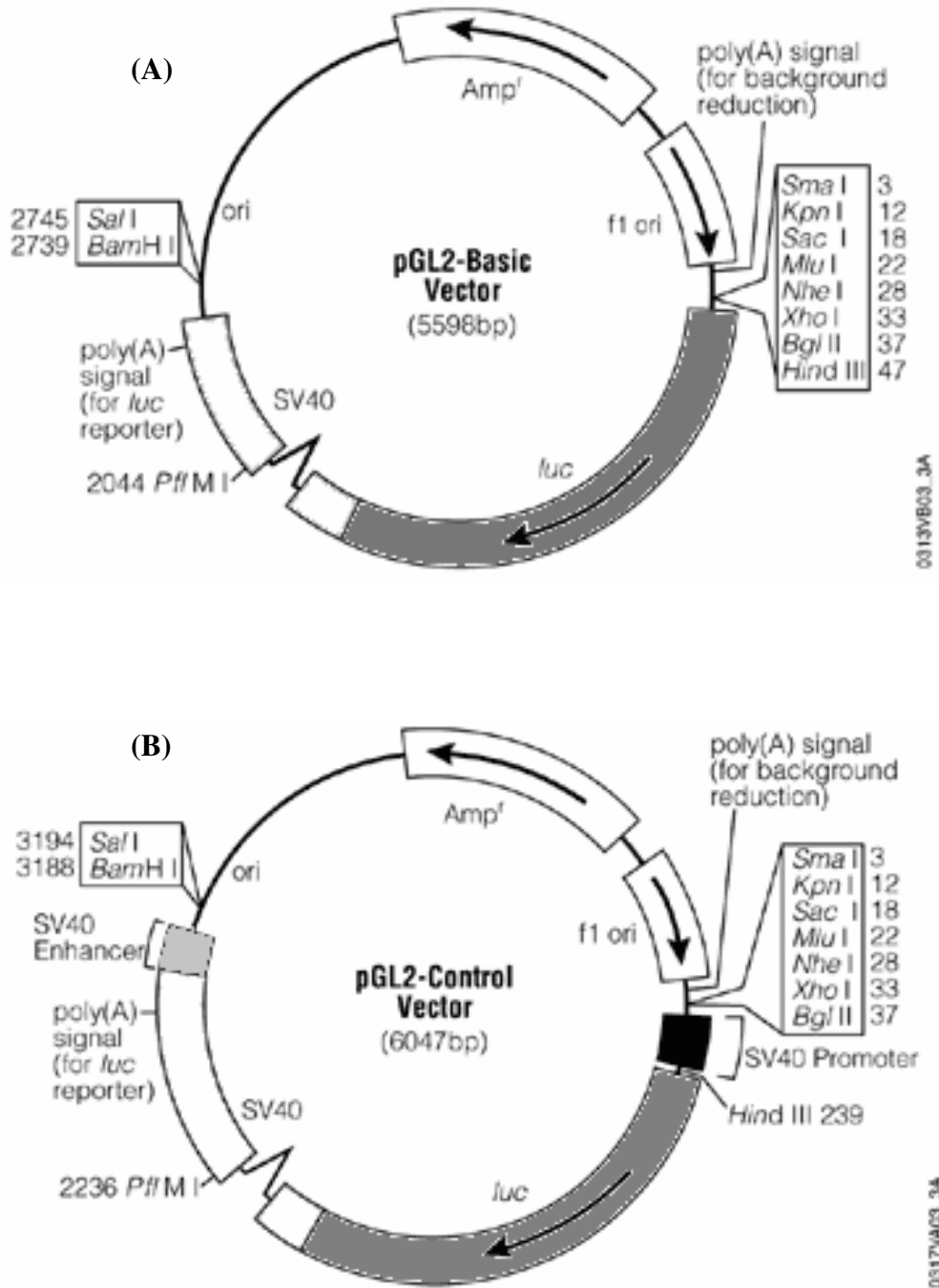


Figura 9- (A) Vetor de expressão pGL2-Basic. Em cinza, o gene da Firefly Luciferase, sem promotor e *enhancer*; (B) Vetor de expressão pGL2-Control. O promotor SV40 (em preto), seguido do gene da Firefly Luciferase (em cinza médio) e o *enhancer* SV40 (em cinza claro).